Image Enhancement By Nonlinear Extrapolation in Frequency Space

Hayit Greenspan, Charles H. Anderson, and Sofia Akber

***Abstract*—A technique for enhancing the perceptual sharpness of an image is described. The enhancement algorithm augments the frequency content of the image using shape-invariant properties of edges across scale by using a nonlinearity that generates phase-coherent higher harmonics. The procedure utilizes the Laplacian transform and the Laplacian pyramid image representation. Results are presented depicting the power-spectra augmentation and the visual enhancement of several images. Simplicity of computations and ease of implementation allow for real-time applications such as high-definition television (HDTV).**

***Index Terms*—HDTV, image enhancement, Laplacian pyramid, multiresolution representation, scale space.**

I. Introduction

We present a technique for enhancing the perceptual sharpness of an image. We stress that the goal of this work is not in complete reconstruction, but rather in augmenting the sharpness, as perceived by the human observer. This goal, which is intrinsically comprised of more subjective constraints, is of importance in a variety of application domains. One such scenario is when a given input image is blurred with no degradation model known. If a degradation model does in fact exist, restoration techniques can be applied, together with other frequency enhancement techniques (unsharp masking) present in the literature [8]. The enhancement scheme described in this paper can then be applied as an additional enhancement utility. A second application domain relates to expanding an image up by a factor of two in size (so called "zoom in"). This is desirable in many applications (e.g., HDTV, video-phone), but generally results in an image which appears blurred because there is no power in the highest spatial frequency band.

This work concentrates on creating new high-spatial frequencies and thus can augment existing (linear) high-frequency enhancement techniques available in the literature. The given frequency content is augmented using shape-invariant properties of edges across scale. The augmentation procedure is based on a multiresolution image representation [1], [3] and can be described using the scale-space formalism [13]–[15]. Image representation across scale is described in Section II. In Section III, we formalize the enhancement procedure. This procedure includes a simple extrapolation across scale representations using a nonlinearity that generates phase coherent higher harmonics. In Section IV we present an evaluation of the algorithm parameters. Experimental results depicting the power-spectra augmentation and the visual enhancement of several images are presented in Section V. Computational cost considerations and the comparison to related works follow in Sections VI and VII, respectively. Section VIII concludes this work with suggested future directions.

Manuscript received August 20, 1997; revised June 17, 1999. This work was supported in part by an Intel research grant. The early part of this work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Henri Maitre.

H. Greenspan is with the Department of Biomedical Engineering, Faculty of Engineering, Tel-Aviv University, Ramat-Aviv, Tel-Aviv 69978, Israel (e-mail: hayit@eng.tau.ac.il).

C. H. Anderson is with the Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, MO 63110 USA.

S. Akber is with the Department of GeoPhysics, University of California, Berkeley, CA 94720 USA.

Publisher Item Identifier S 1057-7149(00)04698-4.

II. Image Representation Across Scale

An image can be represented by a composition of its frequency components, as long as there is no aliasing and overlapping involved

$$\text{Low-resolution image} + \text{High-resolution components} = \text{Original image} \quad (1)$$

with the low-resolution and high-resolution components taken from the same image.

We focus on the edge representation of an image across different image resolutions, or image frequency components. Edges are an important characteristic of images, since they correspond to object boundaries or to changes in surface orientation or material properties. An edge can be characterized by a local peak in the first derivative of the image brightness function, or by a zero in the second derivative, the so called zero crossings (ZC) [8]. An ideal edge (a step function) is scale invariant in that no matter how much one increases the resolution, the edge appears the same (i.e., remains a step function). This property provides a means for identifying edges and a method for enhancing real edges.

The application of the Laplacian transform to an ideal edge transition results in a series of self-similar transient structures. An edge of finite resolution would produce a decrease in amplitude of these transients with increasing spatial frequency, with it going to zero at frequencies above the Nyquist limit. An edge of finite resolution can be created by starting with a low-resolution image (e.g., using a Gaussian blur) and then adding on all the bandpass transient structures. To create an edge with twice the resolution requires the creation of a self-similar transient at the next level, hereby referred to as L_{-1}. The most essential features of these transient structures is that they are of the same sign at the same position in space, hence their ZC's line up, and they all have roughly the same amplitude. The precise shape of the structures need not necessarily be maintained so long as their

Fig. 1. Multiscale sequence of edge maps. Presented from left to right are the Laplacian pyramid components: L_0, L_1, and L_2, respectively.

scaled spatial frequency response is similar. The simple procedure described in this paper creates localized transients for L_{-1} that satisfies all these constraints except for the maintenance of constant amplitude. While more complicated procedures could handle this case, it was found that sharpening the stronger value edges produces in itself visually pleasing results.

In its simplest formulation, what we point out is that a high-resolution image can be obtained from a lower resolution one by adding a bandpass component, such as the Laplacian [see (1)]. Moreover, a high-resolution Laplacian of an ideal edge can be predicted from a low-resolution one.

A. Multiresolution Pyramidal Representation

We are interested in viewing an image in a multiresolution framework for characterizing edges across scale. Any bandpass filtering operation is valid (e.g. Laplacian) in extracting the input image's frequency components. One particular multiresolution framework involves Gaussian- and Laplacian-pyramid image representations, as described by Burt [3] or Anderson (FSD pyramid) [1]. The Gaussian pyramid consists of lowpass filtered (LPF) versions of the input image, with each stage of the pyramid achieved by Gaussian filtering of the previous stage and corresponding sub-sampling of the filtered output. The Laplacian pyramid consists of bandpass filtered (BPF) versions of the input image, with each stage of the pyramid constructed by the subtraction of two corresponding adjacent levels of the Gaussian pyramid.

We have arbitrarily chosen to use the FSD pyramid[1] construction as the multiresolution representation framework. In the following we refer to the input image as G_0, the LPF versions are labeled G_1 thru G_N with decreasing resolutions and the corresponding edge maps are labeled L_0 thru L_N, respectively. A recursive procedure allows for the creation of the pyramids, as follows:

$$\begin{aligned} G^0_{n+1} &= W \times G_n \\ L_n &= G_n - G^0_{n+1} \\ G_{n+1} &= \text{Subsampled } G^0_{n+1} \\ n &= 0 \cdots (N-1). \end{aligned} \tag{2}$$

[1]We will not address here any theoretical issues regarding the pyramid construction and reconstruction. The interested reader can refer to [1] and [2].

The Burt and Anderson Laplacian pyramids differ in details of when the subsampling step is applied and have slightly different bandpass characteristics. Generally, the weighting function W is Gaussian in shape, with variance 1, and normalized to have the sum of its coefficients sum to 1. The values used in this work for the LPF, which is a 5-sample separable filter, are (1/16, 1/4, 3/8, 1/4, 1/16). Fig. 1 presents an example of a Laplacian pyramid for the Lenna image.

The pyramid representation can be viewed as a discrete version of the scale-space description of ZC which has been introduced in the literature [13]–[15]. The scale-space formalism gives the position of the ZC across a continuum of scales. One of the main theorems [14], [15] states that ZC of an image filtered through a Gaussian filter have nice scaling properties, one of which is that ZC are not created as the scale increases. If an edge appears at lower resolutions of the image it will consistently appear as we shift to higher resolutions. Although theoretically defined, not much work has yet taken advantage of the image representation across scale. In this work, we utilize the shape invariant properties of edges across scale in a multiresolution (pyramid) representation and in agreement with the consistency characteristic of the scale-space formalism.

III. Enhancement Scheme

The enhancement algorithm shares the basic structure of other high-frequency enhancement methods, except that the linear filter is replaced by a nonlinear filter operation, as schematically summarized in Fig. 2.

Our objective is to form the next higher harmonic of the given signal while maintaining phase. Fig. 3 illustrates a one-dimensional high-contrast edge scenario. The given input, G_0, is shown in (0), together with its pyramid components, L_0 and G_1, shown in (1) and (2), respectively. From the pyramid reconstruction process we know that adding the high-frequency component L_0 to the G_1 component can sharpen G_1 to produce the input G_0. Ideally, we would like to take this a step further. We would like to predict a higher-frequency component, L_{-1}, preserving the shape and phase of L_0, as shown in (3), so that we can use the reconstruction process to produce an even sharper edge, which is closer to the ideal-edge objective, as shown in (4). The L_{-1} component can not be created by a linear operation on the given L_0 component. We can thus never hope

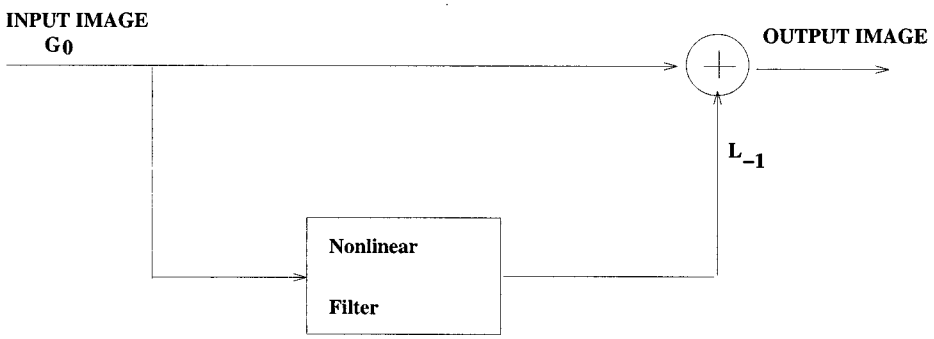

Fig. 2. Basic diagram of the image enhancement algorithm.

to create a higher-frequency output by a linear enhancement technique.

It remains to be shown how the L_{-1} component of the pyramid can be predicted. We extrapolate to the new resolution (L_{-1}) by preserving the Laplacian-filtering waveform shape, together with sharpening via a nonlinear operator. The waveform as in (5) is the result of bounding (or *clipping*) the L_0 response, multiplying the resultant waveform by a constant and then removing the low-frequencies present in order to extract a high-frequency response. The enhanced edge output is presented in (6).

Equation (3) formalizes the generation of L_{-1}

$$L_{-1} = BP(s \times (BOUND(L_0))) \tag{3}$$

where s is a scaling constant and $BOUND(x)$ is the following function:

$$BOUND(x) = \begin{cases} T, & \text{if } x > T \\ x, & \text{if } -T \le x \le T \\ -T, & \text{if } x < -T. \end{cases}$$

Here, $T = (1 - c) \times L_{0\max}$, with c a clipping constant ranging between 0 and 1. The result is high-passed (BP) in order to leave only the high frequencies of the signal.

Generating the new output image entails taking the L_{-1} image as the high-frequency component of the pyramid representation. The new output is generated next as the sum of the given input and L_{-1}

$$OutputImage = L_{-1} + InputImage. \tag{4}$$

IV. Parameter Estimation

In comparing enhanced edges with ideal edges two main deviations are present in the form of blurring (B) and ringing (R). These deviations correspond to the light and dark filled areas in Fig. 4, respectively. A tradeoff exists between the perceived ringing side-effects and the sharpness of the edges. The exact relationship, between the clipping and scaling parameters to the blurring and ringing deviations is complex. We know that the clipping parameter, c, provides us with the nonlinearity which augments the frequency content of the image. A strong scaling coefficient, s, provides for a sharp slope, thus reducing the blurring, yet augmenting the ringing side-effect. We start with a theoretical investigation into a special case of the algorithm. We then proceed to the more general case, in which we extract the parameters via error-function minimization.

A. Theoretical Evaluation—A Special Case

A theoretical investigation of the algorithm parameters can be pursued in the special case of an ideal step edge input, and ignoring the final BP operation (the second filtering operation). It was found experimentally that the final BP stage is not critical for achieving good enhancement results. We thus eliminate this stage and proceed with a theoretical evaluation of the algorithm, in which we can model both the input edge as well as the desired output edge.

Consider an arbitrary object. Optical imaging can be modeled by Gaussian low-pass filtering, G_0, with $\sigma_0 = 0.9$.[2] Nyquist then allows sampling with distance 1 to get the original input data. In order to get a zoom factor of two of the image, one should then low-pass filter, G_{-1}, with $\sigma_{-1} = 0.45$ and sample at a distance of 0.5.

Let us consider the special case of a step edge. In the implemented algorithm we use a low-pass filter, W, which is a 5-tap binomial approximation of a normalized Gaussian with a standard deviation of 1.0. As a result, the filtered low-pass image, G_1, has $\sigma_1 = 1.345$ $(\sigma_1^2 = \sigma_0^2 + \sigma_W^2)$. When applied to a unit step edge, $U(x)$, a normalized Gaussian filter produces an ErrorFunction. Thus,

$$\begin{aligned} L_0U &= G_0U - G_1U; \\ L_0U &= \text{Erf}(x/(\sigma_0)) - \text{Erf}(x/(\sigma_1)) \end{aligned} \tag{5}$$

where we use U instead of $U(x)$ for easier notation. The algorithm takes the maximum of L_0U, $L_{0\max}$, and clips, so as to get the maxima at $T = (1 - c) \times L_{0\max}$, where c is the clipping parameter. In the case of a step edge the maximum, $L_{0\max}$, can be derived analytically as follows: we wish to find $x = x_{\max}$ for which the derivative of L_0U is zero. Incorporating (5), we get

$$\text{Erf}'(x\text{max}/(\sigma_0)) - \text{Erf}'(x\text{max}/(\sigma_1)) = 0. \tag{6}$$

This is equivalent to the constraint

$$G(\sigma_0)_{x\text{max}} - G(\sigma_1)_{x\text{max}} = 0 \tag{7}$$

the solution of which leads to the following:

$$x\text{max} = sqrt(2\log(\sigma_1/\sigma_0)/(1/(\sigma_0)^2 - 1/(\sigma_1)^2)). \tag{8}$$

[2]We start with a set of typical values in order to carry out the analysis. The initial values may change depending on the particular imaging situation, in which case pursuing a similar analysis to the one shown here will enable extracting the adjusted set of parameters.

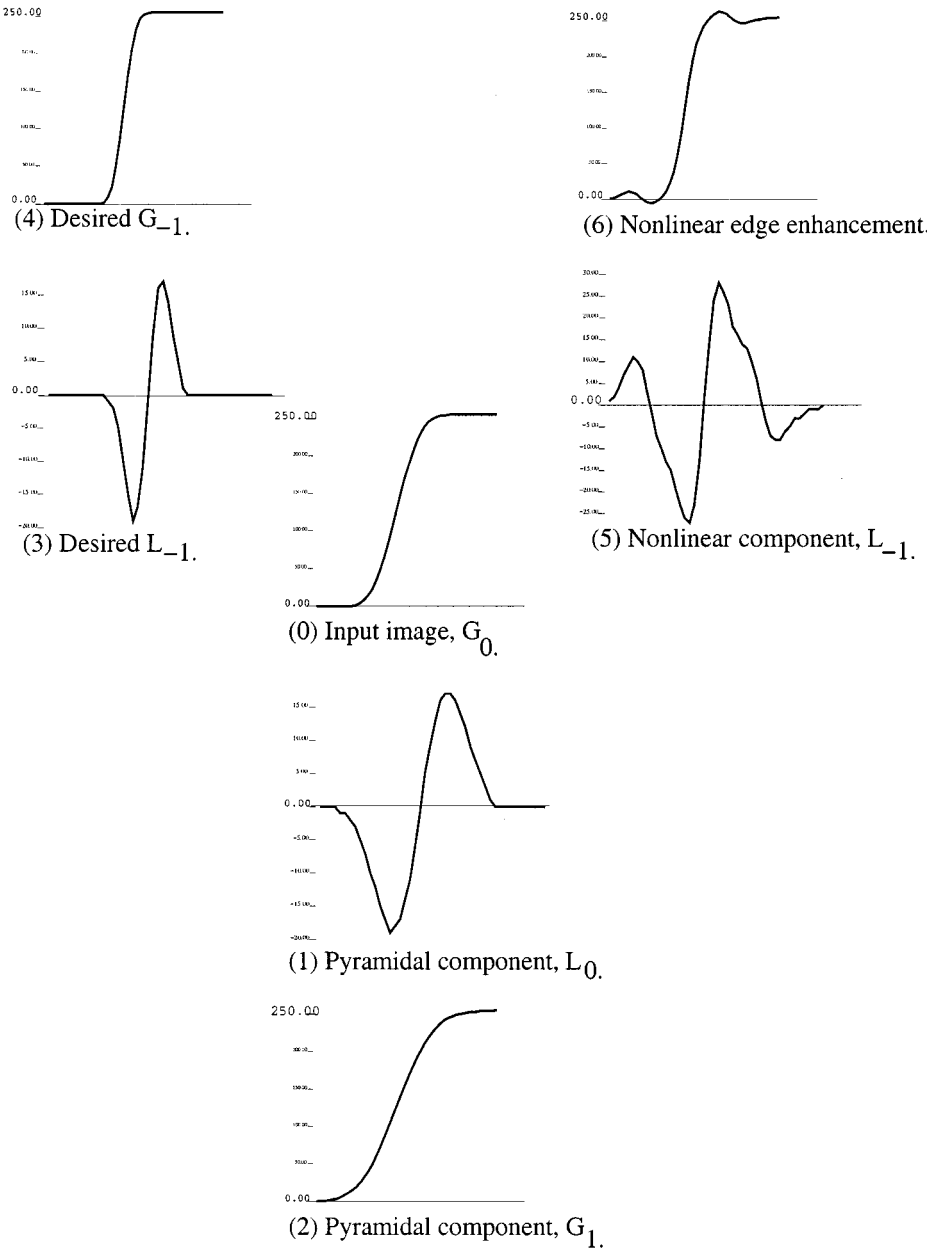

Fig. 3. One-dimensional ideal-edge scenario.

Here

$$x\max = sqrt(2\log(1.345/0.9)/(1/0.81 - 1/1.8)) = 1.085. \tag{9}$$

Hence

$$L_{0_{\max}} = \text{Erf}(1.085/0.9) - \text{Erf}(1.085/1.345) = 0.097. \tag{10}$$

The extremes of the *ideal signal*, L_{-1}, are found in a similar way, with σ_0, σ_{-1}, respectively,

$$\begin{aligned} x\max &= sqrt(2\log(0.9/0.45)/(4/0.81 - 1/0.81)) = 0.612, \\ L_{-1_{\max}} &= \text{Erf}(0.612/0.45) - \text{Erf}(0.612/0.9) \\ &= 0.91 - 0.75 = 0.16. \end{aligned} \tag{11}$$

We thus have our first constraint

$$s \times (1 - c) \times 0.097 = 0.16. \tag{12}$$

A second measure of comparison is the slope. The slope of the approximation at the zero-crossing position, $s \times BOUND(L_0U)$, is s times the slope of L_0U

$$\begin{aligned} L_0(x=0) &= G_0(x=0) - G_1(x=0) \\ &= (1/(0.9 \times \sqrt{2 \times \pi}) - 1/(1.345 \times \sqrt{2 \times \pi})) \\ &= 0.1467. \end{aligned}$$

The slope of the ideal $L_{-1}U$ at the zero-crossing position is

$$\begin{aligned} L_{-1}(x=0) &= G_{-1}(x=0) - G_0(x=0) \\ &= 1/(0.45 \times \sqrt{2 \times \pi}) - 1/(0.9 \times \sqrt{2 \times \pi})) = 0.44. \end{aligned}$$

We thus have our second constraint

$$s \times 0.1467 = 0.44. \tag{13}$$

The scale factor to achieve equal slope is $s = 3$.

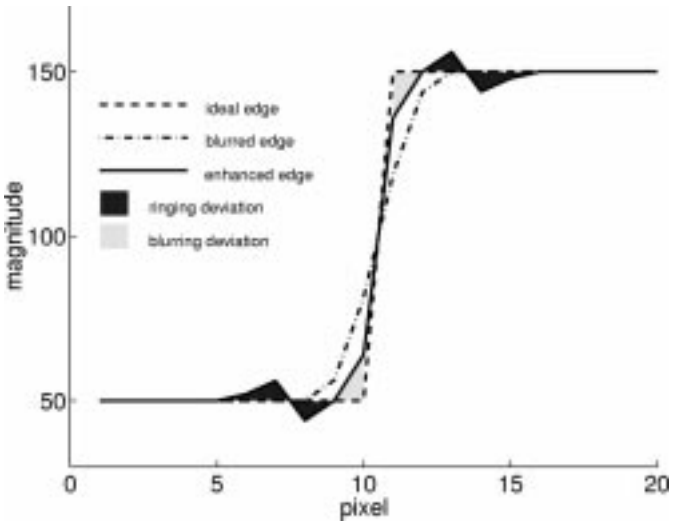

Fig. 4. Blurring and ringing deviations for an ideal edge blurred by a 5-tap Gaussian filter.

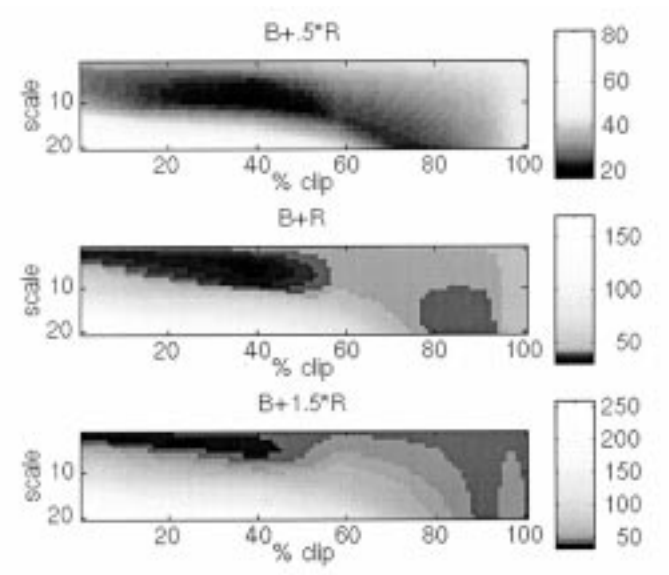

Fig. 5. Error surfaces for three different weighting constants. Black regions indicate low error with an increase in error at the brighter regions.

Revisiting the first constraint [(12)], we can now extract the clipping parameter as well

$$s \times (1 - c) \times 0.097 = 0.16.$$

Thus, $c = 0.45$.

In summary, the theoretical analysis for an ideal step edge scenario, not including the final BP filtering operation, indicates the following set of enhancement parameters:

$$c = 0.45; \quad s = 3.$$

We note that this analysis is based on the set of filter parameters: σ_0, σ_1, σ_{-1}, and can be generalized to other filter parameters than have been discussed here. Different standard-deviations can be used, additional levels of the multiresolution pyramid, and so forth. The exact set of enhancement parameters extracted depend on the details of the imaging scenario at hand as well as on the output human perception desired. With the analysis presented here we have modeled one such enhancement scenario, while providing more detailed insight into the enhancement procedure itself.

B. Parameter Estimation via Error Minimization

We next look at extracting the set of enhancement parameters via minimizing a cost function which gives a tradeoff between the two deviations of ringing (R) and blurring (B). In this scenario, we are able to account for the enhancement algorithm in its entirety, including the second BP filtering stage.

The energy-function is chosen as a weighted sum of the deviations

$$E = B + const \times R. \tag{14}$$

Fig. 5 shows three error surfaces, E, as a function of the enhancement parameters, for the step (edge) scenario of Fig. 4. The three surfaces (top to bottom) correspond to weighting constants of 0.5, 1, and 1.5, respectively. Here, black indicates low errors with an increase of the error in the brighter regions. Several things can be noted from the error surfaces. First, the bottom left corner, with the maximum error, represents the case of no clipping and maximum scaling. With too much scaling, the ringing deviations increase substantially and thus the increase in error. Second, we note the shift of the minima (black region) from mid-scale range (top plot) upwards toward lower-scales. This shift upwards corresponds with weighting the ringing deviations more heavily in the cost function. A consistent range for the clipping parameter across the three maps can be found between 30% and 50% clipping.

The different weighting constants in E reflect the subjective preference to less ringing side-effects versus less blur. This has to be set *a priori* by the user. We choose here a conservative set for the enhancement parameters

$$c = 0.4; \quad s = 5.$$

These are chosen as the minima of the center error map and are consistently in the black region across all three maps. Our algorithm thus entails clipping 40% of L_0, then scaling by 5, prior to adding the generated L_{-1} to the input image [see (3) and (4)].

The effect of choosing parameter sets from different regions of the (center) error surface is shown in Fig. 6. We have the following set of parameters (left to right): $c = 0.8$; $s = 2$ (from a light gray region), $c = 0.4$; $s = 5$ (black) and $c = 0.1$; $s = 10$ (white region). The corresponding enhanced results show a blurry output on the left (not enhanced enough) and an overly sharpened (too much ringing) on the right. The center result, based on our selected parameter set, shows a sharpened output with no (or minimal) ringing.

In addition to extracting the enhancement parameter set, the error surfaces of Fig. 5 provide us with sensitivity maps for these parameters.

V. Simulation Results

In this section, we show experimental results which indicate that the enhancement routine augments the frequency content of an input image achieving a visually enhanced output. We

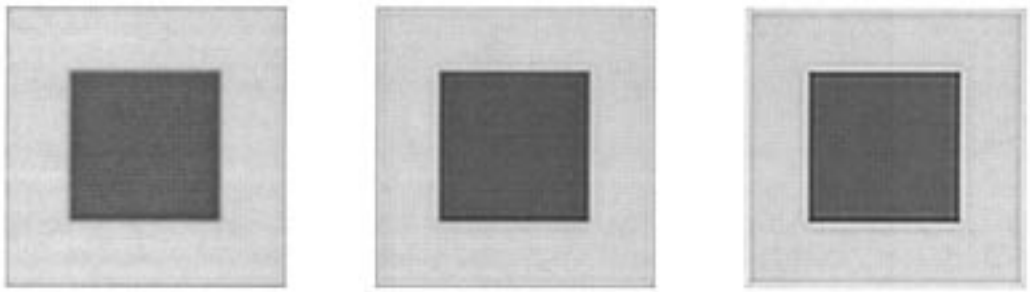

Fig. 6. Different parameter sets in energy-minimization scheme. We have the following set of parameters (left to right): $c = 0.8$; $s = 2$, $c = 0.4$; $s = 5$ and $c = 0.1$; and $s = 10$. The corresponding enhanced results show a blurry output on the left and an overly sharpened on the right. The center result, based on our selected parameter set, shows a sharpened output with minimal ringing.

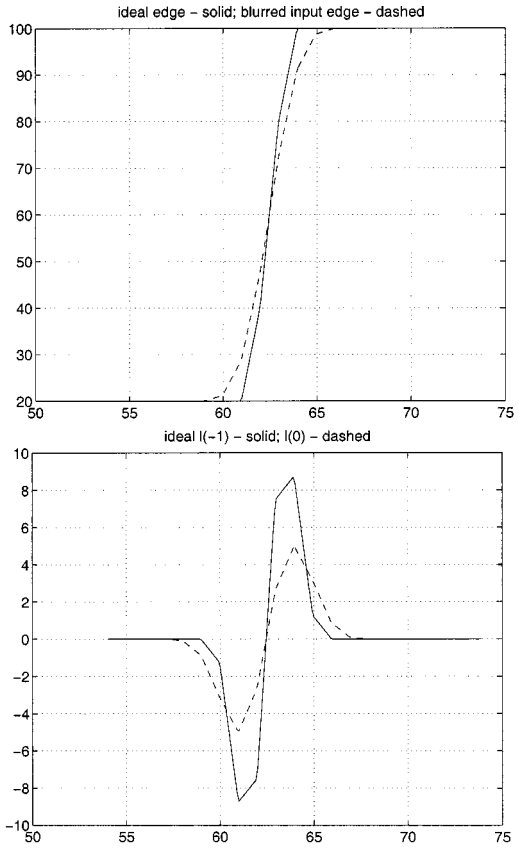

Fig. 7. Top: Input curves, ideal step edge (solid) and input blurred edge (dashed). Bottom: Corresponding Laplacians, ideal L_{-1} (solid), and extracted L_0 (dashed).

have extracted two parameter sets from the parameter estimation analysis discussed above. Both relate to an ideal edge scenario. In the following, we first demonstrate the performance of the parameter sets in synthetic examples, and then focus on the generalization of the extracted parameter sets to real-world images.

We start with a synthetic example. We compare the first, theoretically derived parameter set, using the enhancement routine without the second BP filtering stage, to the second set of parameters, extracted via the energy minimization scheme—including the BP filtering stage. Figs. 7 and 8 demonstrate the effect of the two parameter sets in a step-edge scenario. Fig. 7 presents the ideal edge and the given input edge (top plot), with corresponding L_{-1} and L_0 curves (bottom plot). The top plot in Fig. 8 presents a comparison of the generated L_{-1} curves via each of the parameter sets and its correspondence with the ideal L_{-1} curve. If we examine the first parameter set with $s = 3$ (circles), we note that it in fact accomplishes desired criteria of matching slope and height with the ideal L_{-1} curve. As no constraint was imposed on the width of the curve lobes, the main lobe is wider than the ideal lobe. The second set of parameters, with $s = 5$ and an additional BP stage (stars), has an exact slope match as well as maximum peak, and in addition has a closer match in the overall lobe shape (with the BP operation "pulling down" the height and "pulling in" the side lobe). The BP step seems to be beneficial in preserving higher frequencies of the response, yet one notes the cost of ringing, as the generated curve wiggles around the ideal one. The results presented in this figure demonstrate the validity of the two parameter sets, their similarities and differences. The bottom plot shows the output edge, comparing between the ideal edge (solid), the first set of parameters (circles) and second set (stars). We note that the center slope is identical with all three curves. The first set of parameters $(s = 3)$ has a slight overshoot in values, which is due to the wider lobe of the generated L_{-1} curve, while the second set of parameters $(s = 5 + \text{BP})$ has a set of slight undershoots and overshoots before settling to final edge value.

Fig. 9 shows the use of the parameters sets in more natural environments, real image settings. Two images are shown, a natural outdoor scene, and a high-frequency building image. In both examples, the original image is shown top left, with a blurred version of it, top right. We assume that we are presented with the blurred image as the input image. Enhancement results, with both parameter settings are presented on the bottom. The overall perceptual enhancement effect is evident in both cases, and with quite similar results. Careful perceptual testing by a set of independent human viewers reveal that most perceive the second set of parameters (with the additional BP filtering operation) to have the slight effect of preserving higher frequencies of the original image content.

We have validated the use of both parameter sets, in both synthetic as well as real-world scenarios. In the following examples presented, we choose to use the enhancement scheme in its entirety, using the second set of parameters. We demonstrate zoom-in applications and provide some additional insight into the frequency augmentation of the enhancement procedure.

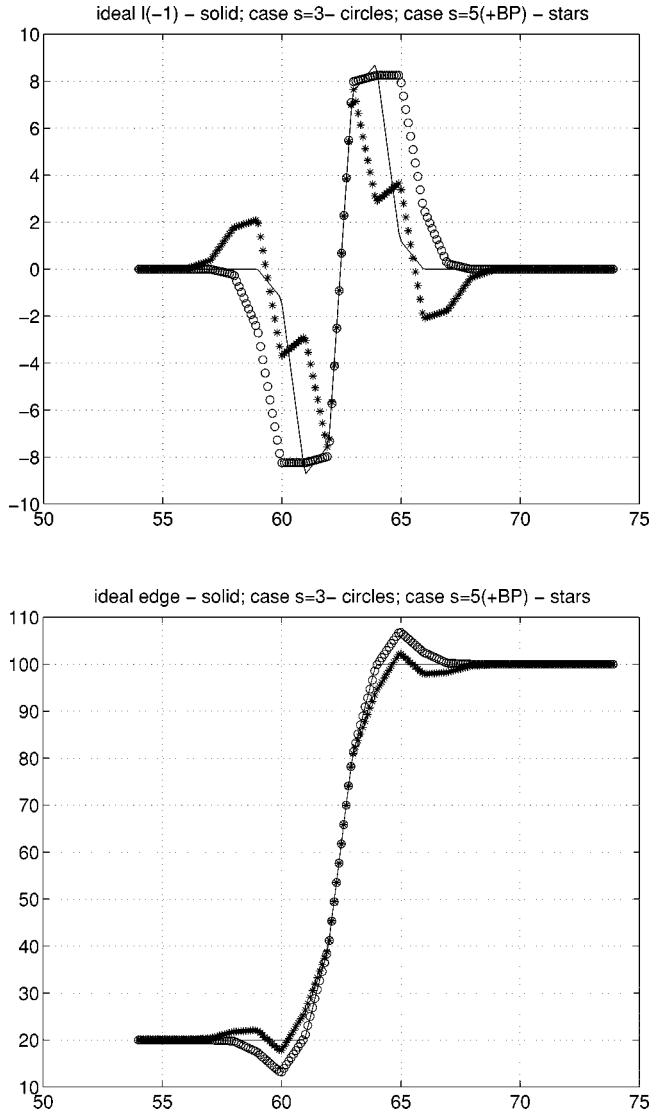

Fig. 8. Top plot: comparison between two parameter sets, with respect to the desired Laplacian curve (solid). Bottom plot: comparison between the two parameter sets as related to the ideal edge output (solid).

Fig. 10 exemplifies a zoom-in application. Here, an image is expanded by a $2:1$ ratio using the expand operation described in Appendix A. The absence of the high-spatial frequencies makes the image appear soft or blurry. We wish to see if the system can enhance the image sufficiently, thus saving in the required bandwidth for the transmission of a full resolution image. Fig. 10 presents the given and enhanced images, left to right, respectively, together with their corresponding power spectrum characteristics (bottom). In the enhanced image, Fig. 10 (top right), we can see much more detail in the hair region and perceive the texture in the nose region and the important eye region much more clearly. The enhancement is overall evident. Looking at the power spectrum characteristics (bottom) it is evident that the input power spectra is augmented. The enhancement process actually extrapolates to higher frequencies, thus producing the satisfying enhanced result.

In the following example, we wish to exemplify the difference between high-frequency enhancement techniques available in the literature (sometimes referred to as the unsharp masking method), and the nonlinear analysis scheme presented in this paper. Fig. 11 top exhibits the full monkey image. A blurred input is presented at the top-left corner. The image was blurred using 5-tap Gaussian filter. Enhancing the high-frequency components present in the given image results in an enhanced image, as shown top-right. Here we are using the unsharp masking algorithm in which a scaled version of the high-frequency components of the input image, is added back to the input image. The scale parameter was chosen to give the best perceptual results possible without generating too many distracting ringing side-effects and overshoots. The result of applying the algorithm presented in this paper is depicted in the bottom of the figure. We get an overall enhancement perception. The differences are evident in the hair, whiskers and eye regions. Fig. 12 displays the corresponding power-spectra characteristics. The power spectra at the bottom of the figure has its higher frequencies augmented.

Fig. 13 presents a third example in which we examine the ability of the algorithm to enhance images which are blurred in a $4:1$ fashion. The blurred input to the system is presented left with the enhanced result depicted on the right. On the top is the case of zooming-in by two, i.e., we start with G_1 level of the pyramid (left) and estimate level G_0 (right). On the bottom is an extension to additional blurring—the case of zooming-in by four, or starting with level G_2 of the pyramid (left) and estimating level G_1 (right). We note the strong resemblance between the bottom right image and the top left image. The increase in resolution (bottom to top) is evident. In both cases, we can see much more detail in the tree and bush stems as well as perceive the texture of the scene much more clearly. The enhancement is overall evident. For the extension to color analysis see Appendix B.

Our final example is a rockscene image displayed in Fig. 14. The top figure presents the enhancement results. The bottom figure displays the corresponding power characteristics. The blurred input, which can be the result of cutting-off high-frequencies due to bandwidth considerations or a zoom-in application, is presented at the top-left corner. The original image, which we are assuming is not available to the system and which we wish to reproduce, is presented on the top right. The result of applying the algorithm presented to the blurred input is depicted in the bottom of each figure. We get an overall enhancement perception. The enhanced image very closely matches the original one and the power spectra of the enhanced image is very close to the original power spectra.

VI. Computational Cost

Specific consideration was given for simplicity of computations and ease of implementation. In the results presented, a 5×5 filter was used in the extraction of the L_0 edge map. A separable LPF was used of the form: [1/16, 1/4, 3/8, 1/4, 1/16]. The BPF was defined next as (1 − LPF). This initial BPF is part of any enhancement algorithm. If fewer multiplications are required a 3×3 filter can be used.[3] The nonlinearity stage of the proposed algorithm involves bounding the L_0 map followed by scalar multiplication of the resultant image. The scalar multiplication can be incorporated as a filter gain, or as part of the look-up table. It is therefore the look-up table which is the core

[3]Pyramid chips are currently available at the David Sarnoff Laboratories [12].

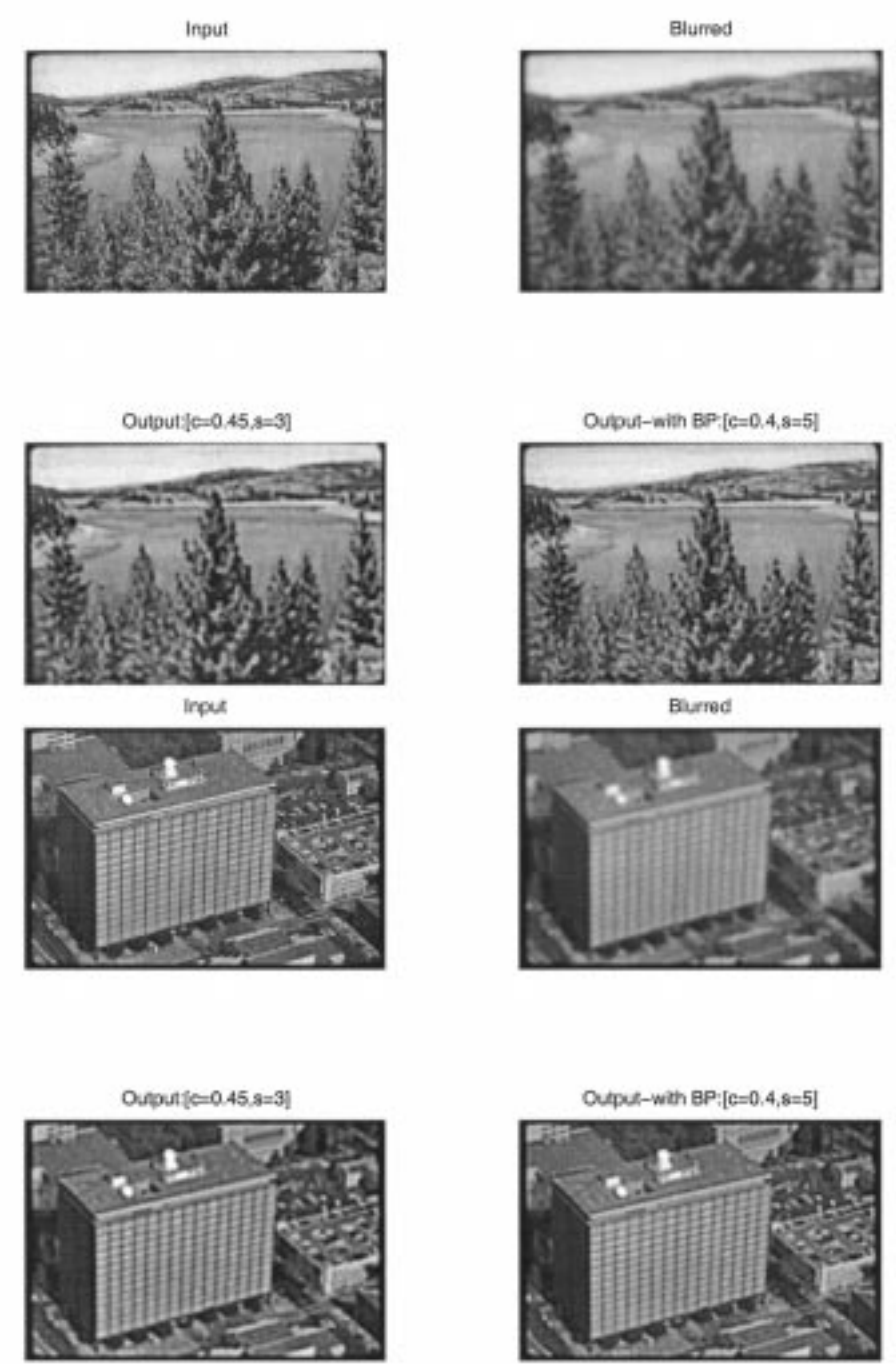

Fig. 9. Enhancement results on two natural scenery, gray level images. In each case, top left image is the original, high-resolution image ("ideal"), top right is the input blurred image, bottom left is the enhanced result using the first parameter set, and bottom right is the enhanced result using the second parameter set.

of the nonlinearity operation. If resources allow, a second filtering stage can be added at this time in order to remove any low-frequencies present in the resultant L_{-1} map, thus adding only the high-frequency response to the given input G_0. Experiments have shown that the second filtering operation is not critical for achieving good enhancement results and can therefore be ignored for real-time application domains.

VII. Comparison with Other Work

In our enhancement scheme the focus is on high-frequency augmentation with phase-coherent characteristics. We are not aware of any other work in the literature that follows similar objectives.

Common techniques for contrast enhancement generally fall into one of two categories [6]: in the first category, techniques such as histogram equalization, modify the brightness of each pixel from the statistical information of an image. In the second category are the techniques which first separate the high- and/or low-frequency components of images, manipulating them separately and them recombining them together. The unsharp masking method, as well as the method presented in this paper, belong to this category.

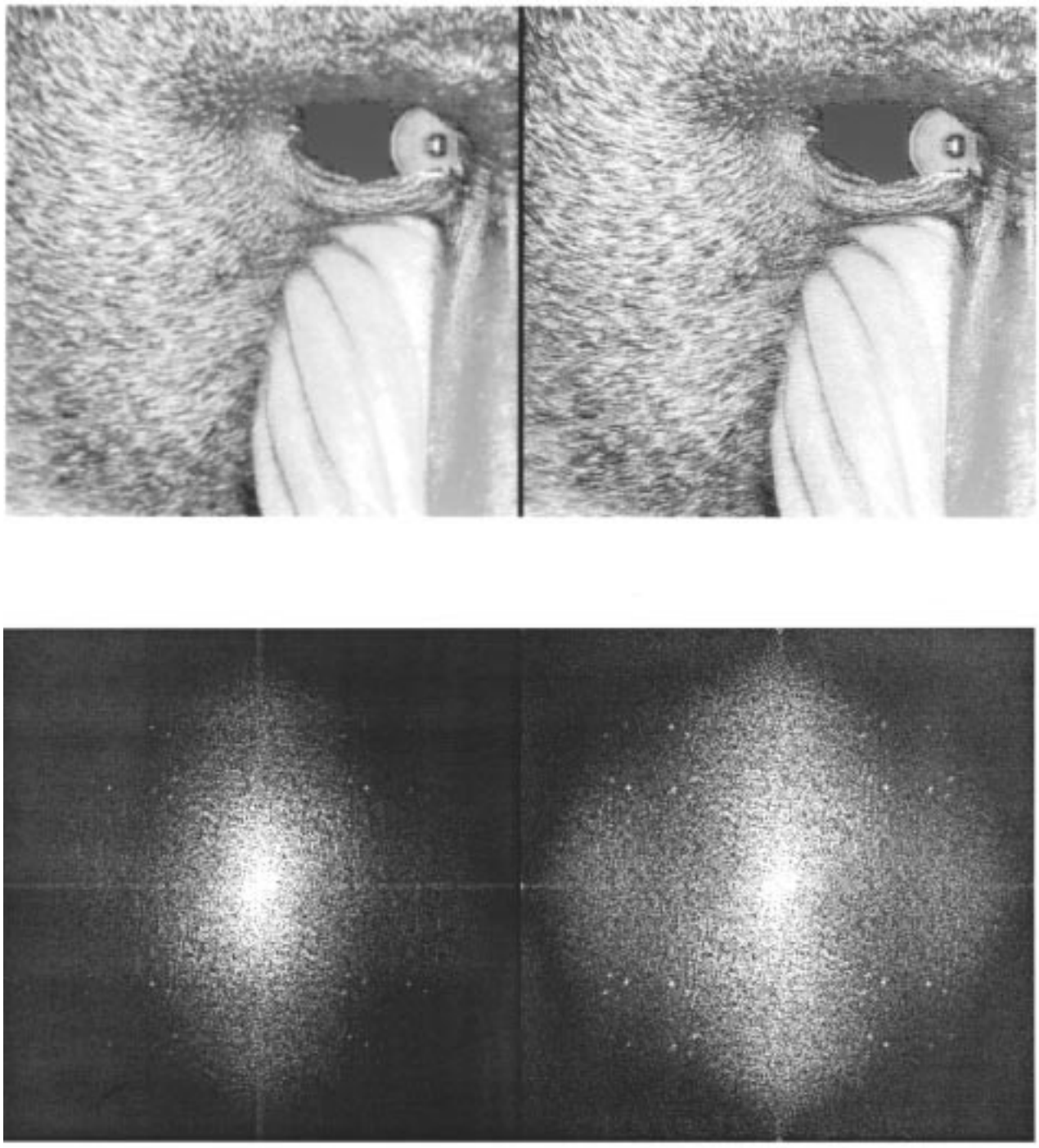

Fig. 10. Given and enhanced images, left to right, respectively, together with their corresponding power spectrum characteristics (bottom). It is evident that the input power spectra is augmented. The enhancement process actually extrapolates to higher frequencies, thus producing the satisfying enhanced result.

A different nonlinear filter for image enhancement has been proposed in [7]. The filter behaves like a local-mean-weighted highpass filter. The basic diagram of the image enhancement algorithm is the same (see Fig. 2). Still, the motivating background is very different resulting in interesting differences. In our algorithm a homogeneous filter is used for the extraction of the high-frequency L_0 map. This is part of the pyramid generation. In [7], the filtering is biased either in the vertical and horizontal directions or along the diagonal directions—with the latter chosen as giving better performance. A visual-based comparison between the two nonlinear filter performances, was presented in [4] demonstrating slight differences, with less aliasing effects evident using the isotropic Laplacian filter.

In [11], a detailed analysis is pursued into the Volterra family of filters in order to define the most isotropic filter in this class. The Volterra filter belongs to a general class of mean-weighted high-pass filters. The nonlinearity of the enhancement process is introduced by the multiplication of the high-pass filter by the local mean. In [10] and [11], Volterra filters are used for edge-enhancement and image-zooming, respectively. The

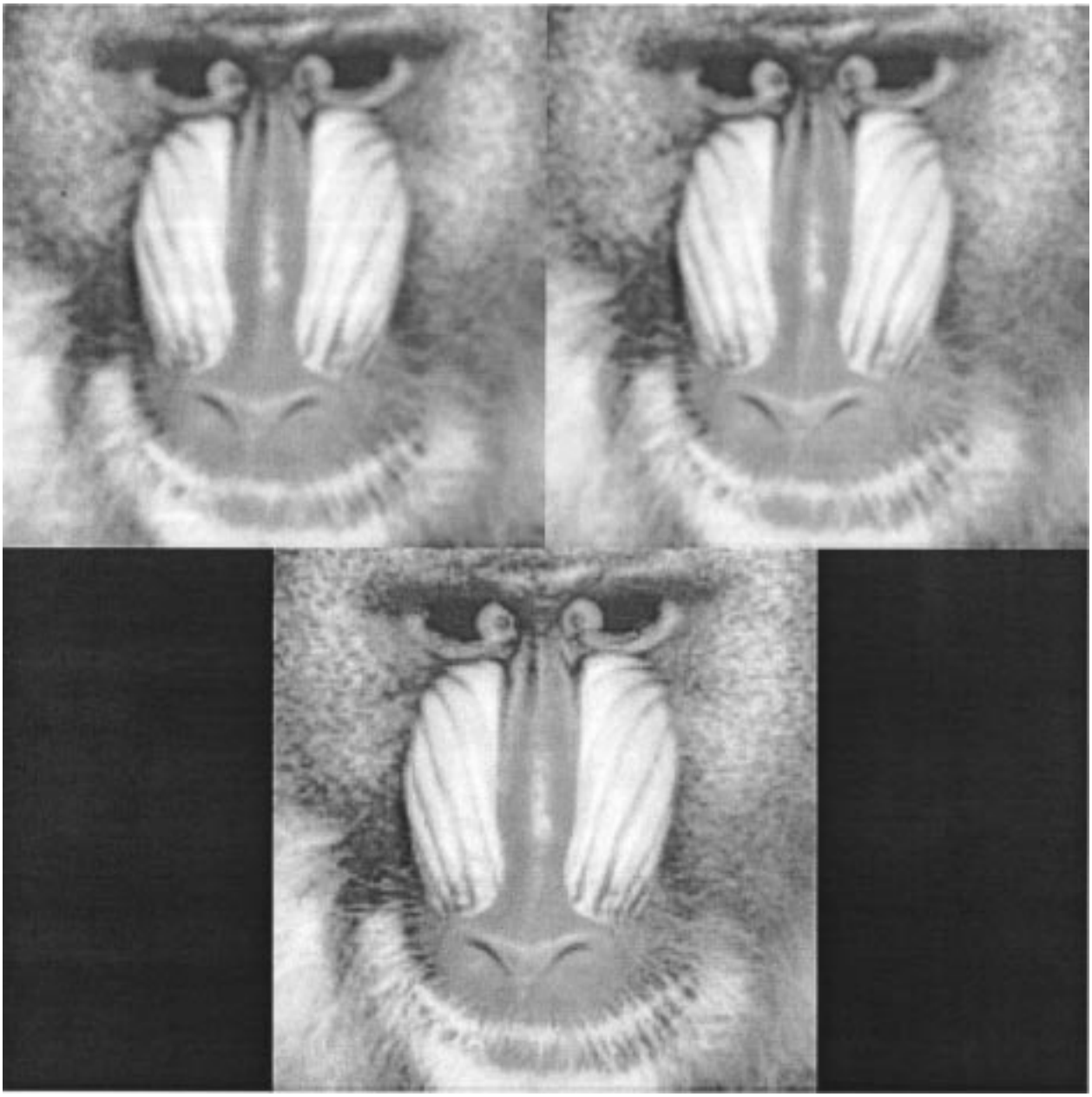

Fig. 11. Top: Enhancement algorithm results, monkey image. A blurred input is presented at the top-left corner. Enhancing the high-frequency components present in the given image results in an enhanced image, as shown top-right. The result of applying the algorithm presented in this chapter is depicted in the bottom of the figure.

introduction of the nonlinearity is motivated by the need to handle noise in the enhancement procedure. Unsharp masking, although quite effective for enhancing low-contrast images, does not discriminate between actual image information and noise. Thus, noise is enhanced as well. The Volterra filter has the effect of adding less of the high-frequency components to the dark regions and more to the brighter ones, and can be desirable for a smoother perception of the enhanced result.

In our approach, the nonlinearity is introduced specifically with the goal of augmenting frequency content. The nonlinearity which we chose includes a bounding function. This also has the effect of introducing a stronger high-frequency component to brighter areas than to darker ones. A major difference between the two nonlinearities is that the phase of the edges is preserved in our algorithm, following the scale-space formalism. The procedure outlined in [7], however, has the effect of shifting the phase toward the brighter region. Computationally, the main difference between the approaches is in the chosen nonlinearities. Here, the core of our enhancement algorithm is the look-up table. This competes with the multiplication of the highpass filter output with the local mean.

Our focus is on global processing which allows for computational efficiency. Localized processing can definitely be useful in noise reduction; region-based interpolation schemes for the image-zooming application, as presented in [10], allow for better tuning of the edge-sharpening process, at an increased computational cost. Some of our own investigation into localization is presented in Section VIII.

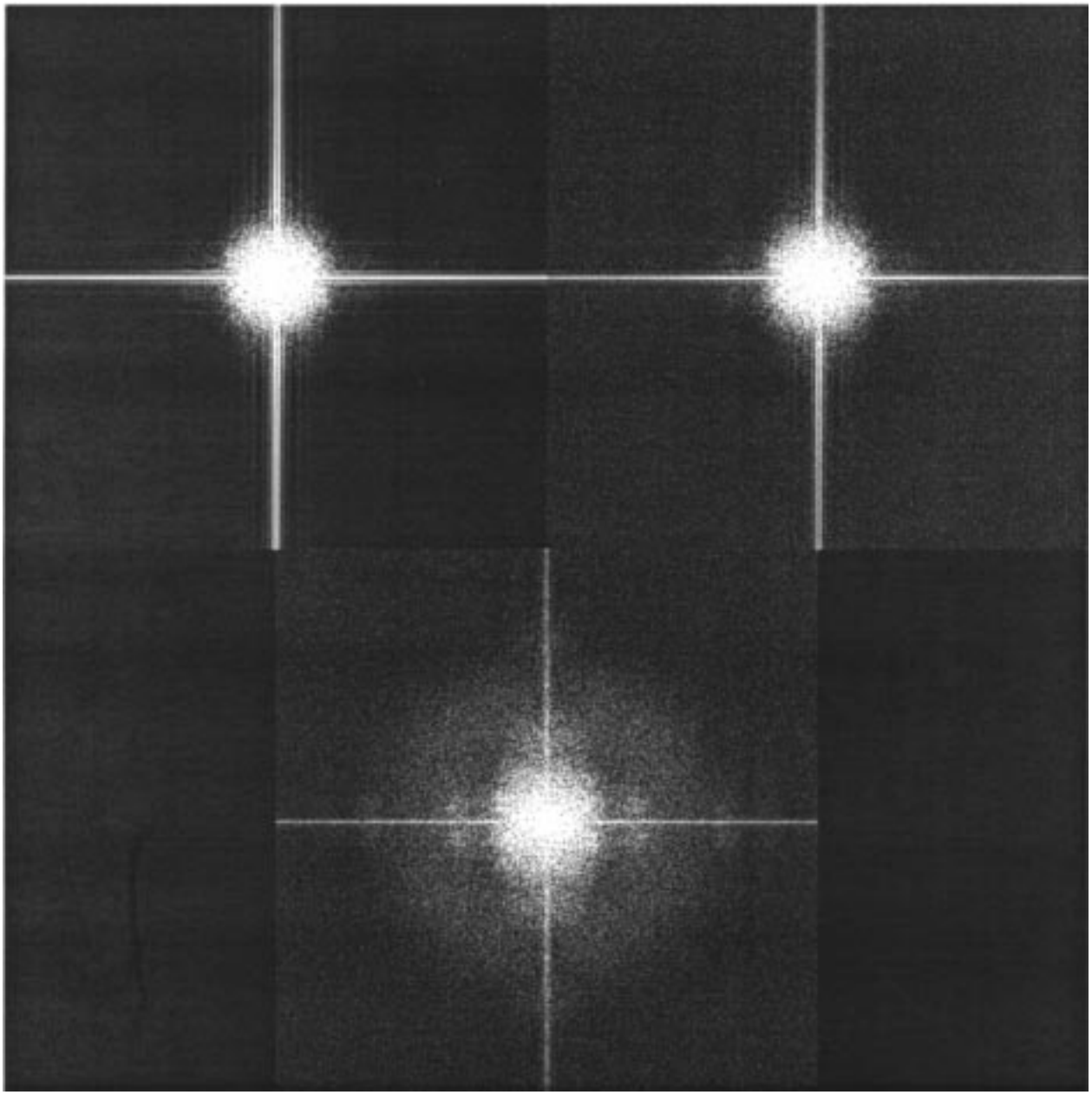

Fig. 12. Power spectra characteristics of the monkey image.

VIII. Summary and Discussion

We have described an enhancement scheme that could very well address the most important features for producing visually pleasing enhanced versions of existing images. The simplicity of the computations involved and ease of implementation enable it to be incorporated in real-time applications such as high-definition television (HDTV). For the combination of the enhancement scheme with pyramid compression see [5].

1) Value of the Proposed Method: One may argue that for the case of step edges, the classical signal-processing approach to achieving the goal of nonlinear interpolation would consist of the following steps.

1) Linear (sinc) interpolation from the samples into a pseudo-analog representation to construct G_0U.
2) Nonlinear local maximum and minimum filtering. Using a filter radius of the order of twice the original sampling distance, replace each pixel by the local maxima or minimum value, whichever is closer, to reconstruct U.
3) Linear filtering with G_{-1} to construct the bandlimited $G_{-1}U$.
4) Resampling at half the original sampling distance.

The above scheme provides for a more sophisticated extrapolation in the spatial frequency domain. Yet, it has the following disadvantages:

- step 1) is costly in time and memory;
- step 2) is nonseparable if isotropic;
- step 3) is again costly in time and memory.

The algorithm presented in this paper avoids these disadvantages, while creating higher harmonics (with the correct phase)

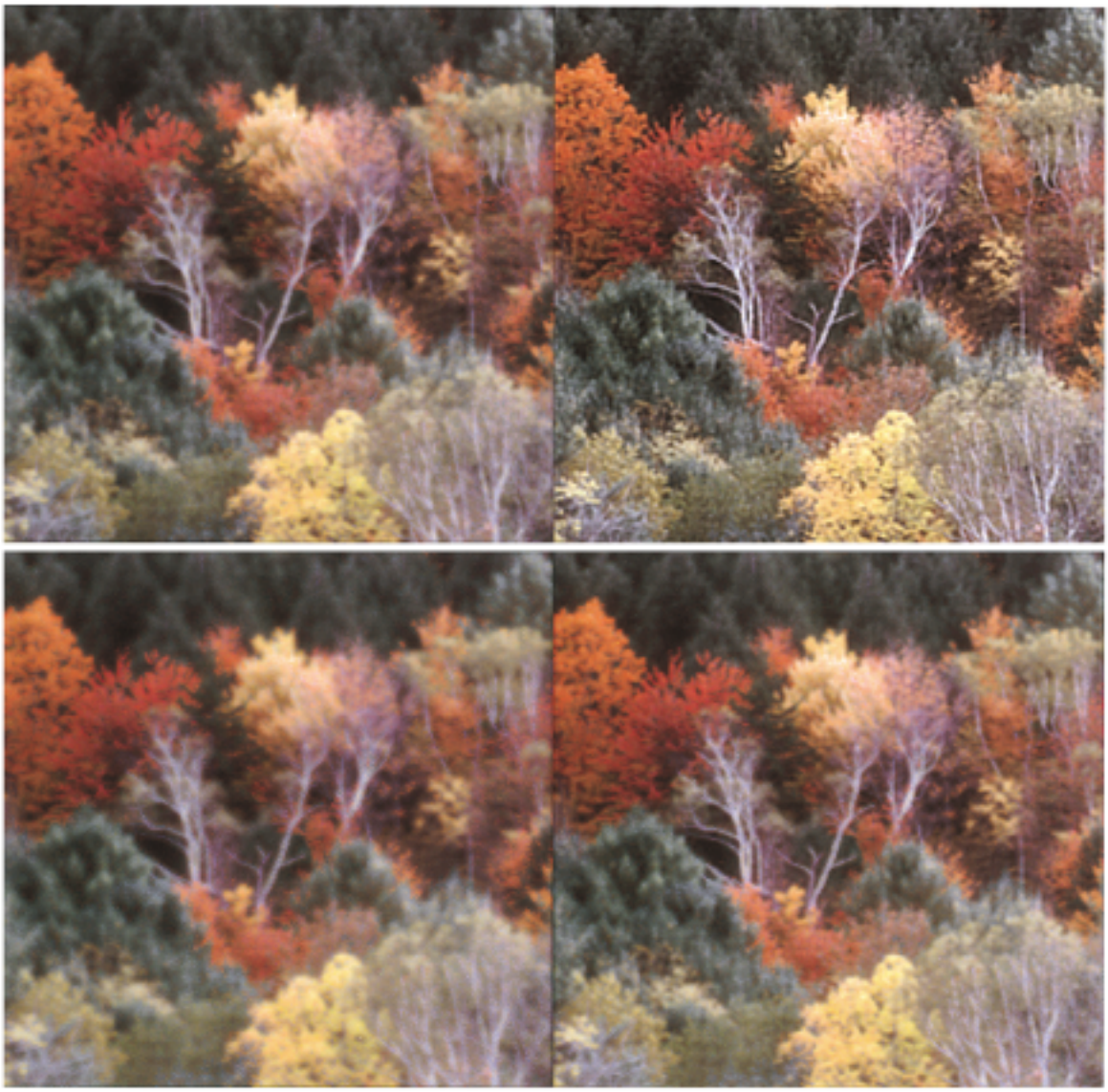

Fig. 13. Global enhancement: Blurred (left) and enhanced (right). The tree scene is blurred with 5-tap (top) and 9-tap (bottom) Gaussian filters.

in edge transitions. The algorithm thus achieves the perceptual effect of image enhancement, with both time and storage savings.

2) Arbitrary Objects: The presented enhancement technique is only a first step toward what can be accomplished by extrapolation across scale. Edges are only one major scale invariant feature. Lines and dots, for example, require additional analysis.

In the current implementation, we do not address the issue of other objects in the image. In particular, points are most likely converted to sharpened circles with sharp edges and lines are converted into ribbons with sharp edges. It is our conclusion from the presented results, that the images look overall sharpened, and are visually agreeable to the user's perception. This is most probably the case as edges are the dominant features where most of the high frequencies exist, whereas points and lines occur less frequently. It could also be that the presence of high frequencies created around points and lines increases the perception of sharpness. We leave the theoretical investigation into additional image primitives to future study.

3) Global versus Local Processing: Global enhancement techniques work very well for many images. However, there are often more complex situations where different parts of the image require different amounts of sharpening (or scaling). An example is presented in the synthetic image of Fig. 15. An original sharp image is presented top left, and the blurred version is presented top right. In the case of already sharp regions, the global enhancement parameters may result in ringing which is visibly disturbing to the eye (center row, left). When blurring high-frequency regions the magnitude of the input signal reduces substantially. In this case the global enhancement parameters do not scale enough (center row, right). The sensitivity of the human perception to ringing side-effects in enhanced images, is strongly dependent on local characteristics such as local contrast and local frequency. We follow [9] in adapting the scaling function to these characteristics of the

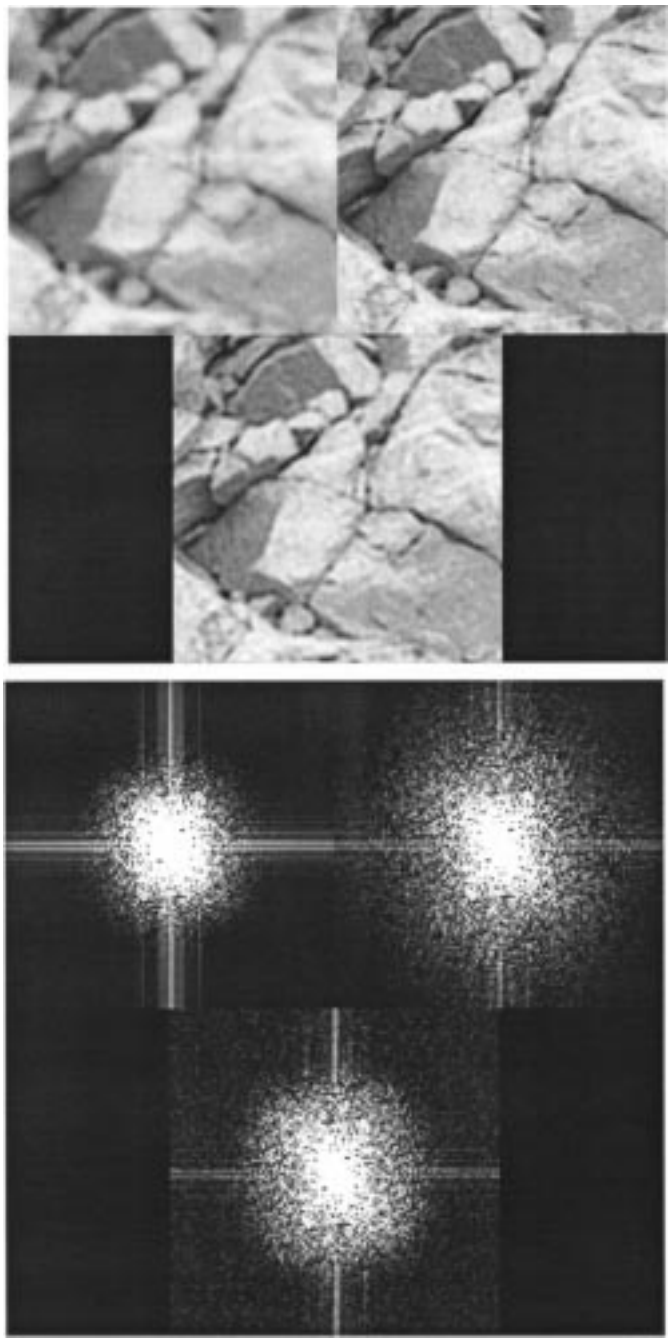

Fig. 14. (Top) Enhancement results. (Bottom) Corresponding power-spectrum characteristics. In each of the above figures the blurred input and original image are presented (top left and top right, respectively) followed by the enhanced output (bottom). Both visual perception enhancement and power-spectrum augmentation are evident.

human visual system. An initial investigation into localizing the enhancement scheme is the following: We divide the image into 50% overlapping 8×8 blocks. In each block we define the local contrast, c_L, and the local spatial frequency, f_L. (Many schemes can be used; we follow [9]). The scaling constant per block, $s(i,j)$, is set relative to these two parameters: For a contrast above a predetermined threshold, $s(i,j) = 1$. Otherwise, $s(i,j)$ increments with increasing frequencies, f_L. In order to determine the effect of f_L on $s(i,j)$, synthetic grids of different frequencies were blurred, and human subjects were

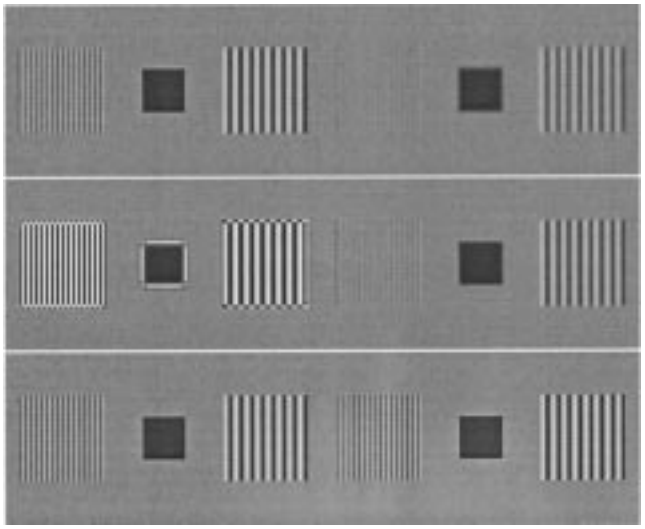

Fig. 15. Global enhancement versus local enhancement. The original input and a blurred version are presented top row, left and right, respectively. Results of the global enhancement scheme are presented in the center row. The results of adaptive enhancement are shown on the bottom.

asked to name the largest scale factor just prior to perceiving ringing effects. The chosen scales were recorded. Once the scaling matrix is obtained, it is interpolated to the original image size and applied to the clipped L_0, as in (3).

The result of incorporating local considerations can be seen in the bottom row of Fig. 15. The local contrast check provides for the resultant left-hand side, while the adaptive scaling based on local frequencies allows for the sharpened result of the right-hand side.

4) Noise Considerations: As is always the case, a tradeoff exists between high-frequency enhancement and noise generation. The enhancement scheme will work best on input images which have been previously processed for noise. One possibility is to use reconstruction schemes which remove noise while unavoidably blurring the image, and then enhancing the resultant image with the algorithm described in this work.

APPENDIX I
ZOOM-IN APPLICATION

The algorithm for a "zoom in" application includes the following five steps.

1) Extract the high- frequency components of an image $L_0 = F_{bp} \times G_0$. We have used $F_{bp} = 1 - W$, as utilized in the formation of a Laplacian pyramid, or the more complex reduce, expand filter technique utilized in the formation of a Burt Laplacian pyramid.
2) Create a double sampled (along both dimensions) version, L_0^e, of L_0 through an interpolation procedure. The standard Pyramid expand technique is to insert 0's at alternate pixels and lines, smooth the result with the lowpass filter, and then multiply the result by a factor of 4, which can be combined in an efficient subroutine.
3) Clip L_0^e, which amounts to setting the magnitude of the signal to a predetermined level if it exceeds that value [(3)].

4) Create L_{-1} by bandpass filtering the clipped output to reshape the new transients so they have the desired spatial frequency components. This was done using the same bandpass filter as in step 1, but other variants are possible. It has been found that this step can be eliminated and still produce pleasing results.
5) Add a scaled version of L_{-1} to an expanded version of G_0 created in the same fashion as L_0^e in step 2 [(4)].

APPENDIX II
COLOR ANALYSIS

Enhancing color images is achieved by performing the edge enhancement algorithm on the illuminance (intensity) band alone. The following procedure is used: Eight-bit RGB components (0-1 range) are first linearly transformed to an essentially uncorrelated YUV representation. The intensity component, Y, is enhanced, to get Y′. The original U and V components are then combined with Y′, and the (Y′UV) set is transformed back to obtain R′G′B′. Like the original RGB components, these new components must also be eight-bit in the (0-1) range. Any values exceeding the range are thresholded. It is interesting to note that in the YUV representation, the U and V color bands can be substantially compressed, with only a very low-resolution version needed to accompany the Y (or Y′) band. In Fig. 13 the U and V components are compressed by using the G_2 level of their respective Gaussian pyramids.

REFERENCES

[1] C. H. Anderson, "A filter-subtract-decimate hierarchical pyramid signal analyzing and synthesizing technique," U.S. Patent 718 104, 1987.
[2] P. Burt and R. Kolczynski, "Enhanced image capture through fusion," in *Proc. ICCV*, Berlin, Germany, 1993, pp. 173–182.
[3] P. J. Burt and E. A. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, pp. 532–540, 1983.
[4] H. Greenspan and C. H. Anderson, "Image enhancement by nonlinear extrapolation in frequency space," in *Proc. SPIE, Image Video Processing II*, vol. 2182, 1994, pp. 2–13.
[5] H. Greenspan and M.-C. Lee, "Combining image processing and image compression scheme," Jet Propulsion Lab., Feb. 15, 1995.
[6] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[7] S. K. Mitra, H. Li, I. Lin, and T. Yu, "A new class of nonlinear filters for image enhancement," in *Proc. ICASSP M5.1*, 1991, pp. 2525–2528.
[8] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1978.
[9] M. K. Sundareshan, T. L. Ji, and H. Roehrig, "Adaptive image contrast enhancement based on human visual properties," *IEEE Trans. Med. Imag.*, vol. 13, Dec. 1994.
[10] S. Thurnohofer and S. Mitra, "Edge-enhanced image zooming," *Opt. Eng.*, vol. 35, no. 7, pp. 1862–1870, 1996.
[11] ——, "A general framework for quadratic volterra filters for edge enhancement," *IEEE Trans. Image Processing*, vol. 5, pp. 950–963, June 1996.
[12] G. S. VanderWal, "The Sarnoff pyramid chip," in *Proc. Computer Architecture Machine Perception*, Dec. 1991.
[13] A. Witkin, "Scale-space filtering," in *Proc. IJCAI*, Karlsruhe, Germany, 1983, pp. 1019–1021.
[14] A. L. Yuille and T. Poggio, "Fingerprint theorems for zero crossings," *Opt. Soc. Amer. A*, pp. 683–692, 1985.
[15] ——, "Scaling theorems for zero-crossings," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, pp. 15–25, Jan. 1988.

Hayit Greenspan received the B.S. and M.S. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, in 1986 and 1989, respectively, and the Ph.D. degree in electrical engineering from the California Institute of Technology, Pasadena, in 1994.

She was a Postdoc with the Computer Science Division, University of California, Berkeley, from 1995 to 1997. In 1997, she joined the BioMedical Engineering Department, Faculty of Engineering, Tel-Aviv University, Tel-Aviv, Israel. Her research interests include content-based image search and retrieval, medical image processing, texture recognition, image enhancement, and pattern recognition.

Charles H. Anderson received the B.S. degree in physics from the California Institute of Technology, Pasadena, in 1957, and the Ph.D. degree in physics from Harvard University, Cambridge, MA, in 1962.

He did research in many areas for 24 years while with David Sarnoff Laboratories, Princeton, NJ, then moved to the Jet Propulsion Lab, Pasadena, in 1987, where he worked on image analysis. In 1992, he moved to his current position of Research Professor of neuroscience, Washington University School of Medicine, St. Louis, MO. He has published papers in experimental physics, visual neurophysiology, computational vision, and computational neuroscience, and holds numerous patents.

Dr. Anderson is a Fellow of the American Physical Society and a member of the AAAS and Society for Neuroscience.

Sofia Akber received the B.S. degree in electrical engineering from the California Institute of Technology, Pasadena, in 1994. She is currently a graduate student in the department of geophysics at the University of California, Berkeley, where she does research in the physics of materials under high pressure.