# Image enhancement by non-linear extrapolation in frequency space

H. Greenspan

Department of Electrical Engineering, 116-81
California Institute of Technology
Pasadena, CA 91125

C. H. Anderson

Dept. of Anatomy and Neurobiology
Washington University School of Medicine
St. Louis, MO 63110

## ABSTRACT

A procedure for creating images with higher resolution than the sampling rate would allow is described. The enhancement algorithm augments the frequency content of the image using shape-invariant properties of edges across scale by using a non-linearity that generates phase-coherent higher harmonics. The procedure utilizes 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).

## 1. INTRODUCTION

We present a procedure for creating images with higher resolution than the sampling rate would allow. The procedure outlined here is applicable in several domains. In one case we assume that a given input image is blurred and no degradation model is known. If such a model exists, restoration techniques can be applied, together with other frequency enhancement techniques (unsharp masking) present in the literature[1]. 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 the pyramid image representation[2,3] and can be described using the scale-space formalism[4,5]. Image representation across scale is described in section 2. In section 3 we formalize the enhancement procedure. This procedure includes a simple extrapolation across scale representations using a non-linearity that generates phase coherent higher harmonics. The enhancement algorithm is schematically summarized in Fig. 1. It shares the basic structure of other high-frequency enhancement methods, except that the linear filter is replaced by a nonlinear filter operation. Experimental results depicting the power-spectra augmentation and the visual enhancement of several images are presented in section 4. Here we show possible improvements over the existing reconstruction techniques. Some similarities exist with Mitra's recent work on nonlinear filtering[6]. A brief discussion of the similarities and differences is presented in section 5. Section 6 concludes this work with suggested future directions.

## 2. IMAGE REPRESENTATION ACROSS SCALE

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)[7]. 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. We concentrate on the edge representation of an image across different image resolutions.


0-8194-1477-8/94/$6.00

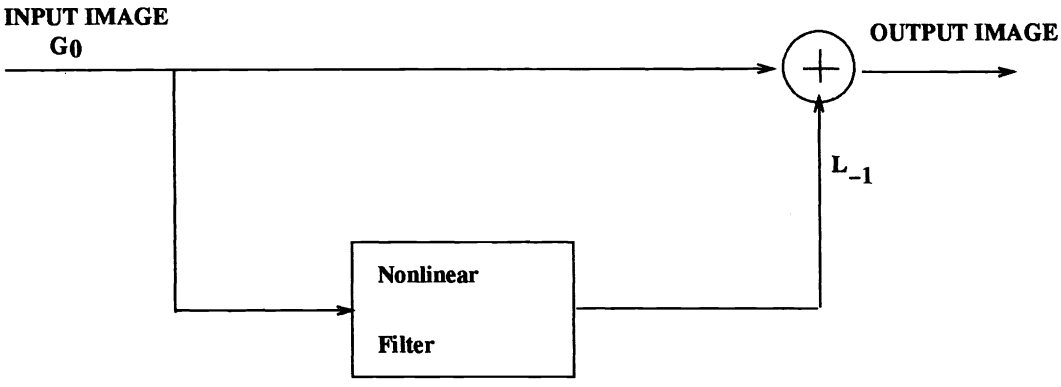


Figure 1: Basic diagram of the image enhancement algorithm

2.1 The pyramid representation

We are interested in viewing an image in a multi-resolution framework for characterizing edges across scale. Gaussian and Laplacian pyramids, as described by Burt[2] or Anderson[3], are utilized. 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. The Burt and Anderson Laplacian pyramids differ in the details of when the subsampling step is applied and have slightly different bandpass characteristics. 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:

$$G^0_{n+1} = W * G_n \ ; \ L_n = G_n - G^0_{n+1} \ ; \ G_{n+1} = \text{Subsampled } G^0_{n+1} \tag{1}$$

Generally, the weighting function W is Gaussian in shape and normalized to have the sum of its coefficients equal to 1. The values used for the LPF, which is a 5-sample separable filter, are (1/16, 1/4, 3/8, 1/4, 1/16). Fig. 2 presents an example of a Laplacian pyramid representation.

It has been shown[2] that the Laplacian pyramid forms an overcomplete representation of the image, thus enabling full reconstruction. The reconstruction process entails adding to a given LPF version of the image $G_N$ the bandpass images $L_n(n = (N-1)..0)$, thus reconstructing the Gaussian pyramid, level by level, up to the original input image $G_0$. This is a recursive process as in equation (2):

$$G_n = L_n + \hat{G}_{n+1}; n = (N-1)..0 \tag{2}$$

where $\hat{G}_{n+1}$ is the interpolated version of $G_{n+1}$.

The Laplacian pyramid is a special case of the wavelet representation[8] and as such it preserves the shape and phase of the edge maps across scale (for example see Fig. 2). The application of the Laplacian transform to an ideal edge transition results in a series of self-similar transient structures as illustrated in Fig. 3 (a). An edge of finite resolution would produce a decrease in amplitude of these transients with increasing spatial frequency, with it going to 0 at frequencies above the Nyquist limit (see Fig. 3 (b) ). An edge of finite resolution can be created by starting with a low resolution Gaussian image and then adding on all the bandpass transient structures shown in Fig. 3 (b). 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 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 scaled spatial frequency response is similar. The simple procedure described in this paper creates localized transients for $L_{-1}$ that satisfy all these constraints except

Figure 2: Multi-scale sequence of edge maps. Presented from left to right are the Laplacian pyramid components: $L_0$, $L_1$ and $L_2$ respectively.

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.

The pyramid representation can be viewed as a discrete version of the scale-space description of ZC which has been introduced in the literature.[4,5] The scale-space formalism gives the position of the ZC across a continuum of scales. One of the main theorems 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 based on the pyramid representation and in agreement with the consistency characteristic of the scale-space formalism.

## 3. THE ENHANCEMENT SCHEME

Our objective is to form the next higher harmonic of the given signal while maintaining phase. Fig. 4 illustrates a 1-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; i.e., it is not possible 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 non-linear operator. The waveform as in (5) is the result of bounding 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. It was found that clipping $L_0$ with a threshold of 0.04 times the maximum signal's amplitude (i.e., 10 out of 256), and then multiplying by a factor of 6 gives the closest resemblance to the ideal-edge output without much ringing side-effect. The enhanced edge output is presented in (6).

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

$$L_{-1} = const(BOUND(L_0)) \tag{3}$$

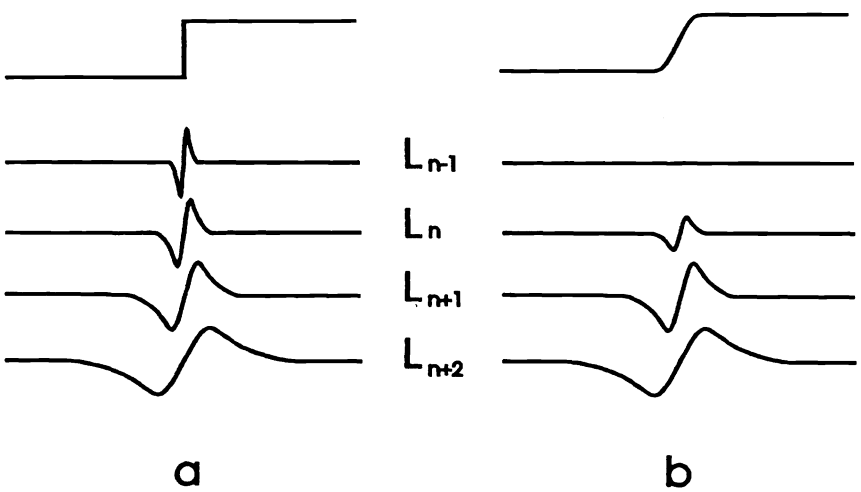


Figure 3: Laplacian transform on an edge transition

where $BOUND(S)$ is the following function:

$$BOUND(S) = \begin{cases} T & \text{if } S > T \\ S & \text{if } -T \leq S \leq T \\ -T & \text{if } S < -T \end{cases}$$

Here, T = $0.04(G_0)_{max}$.

Generating the new output image entails taking the $L_{-1}$ image as the high-frequency component of the pyramid representation. Based on the reconstruction capability of the pyramid representation (equation (2)), the new output is generated next as the sum of the given input $G_0$ and $L_{-1}$, as in equation (4).

$$OutputImage = L_{-1} + G_0 \tag{4}$$

The algorithm for a "zoom in" application includes the following 5 steps:

1. Extract the high frequency components of an image, $L_0 = F_{bp} * G_0$. 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].
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 (equation 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 (equation 4).

The approach as presented above, together with the non-linear operator characteristics which were found for the high-contrast ideal edge scenario, are used in the examples provided later in the paper.

### 3.1 Computational cost

Specific consideration is given for simplicity of computations and ease of implementation. In the results presented below, a 5*5 filter is used in the extraction of the $L_0$ edge map. A separable LPF is used of the form: [1/16, 1/4, 3/8, 1/4, 1/16]. The BPF is 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 (pyramid chips are currently available at the David Sarnoff Laboratories[9]). The non-linearity 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 of the non-linearity 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. We therefore choose to ignore it for real-time application domains.

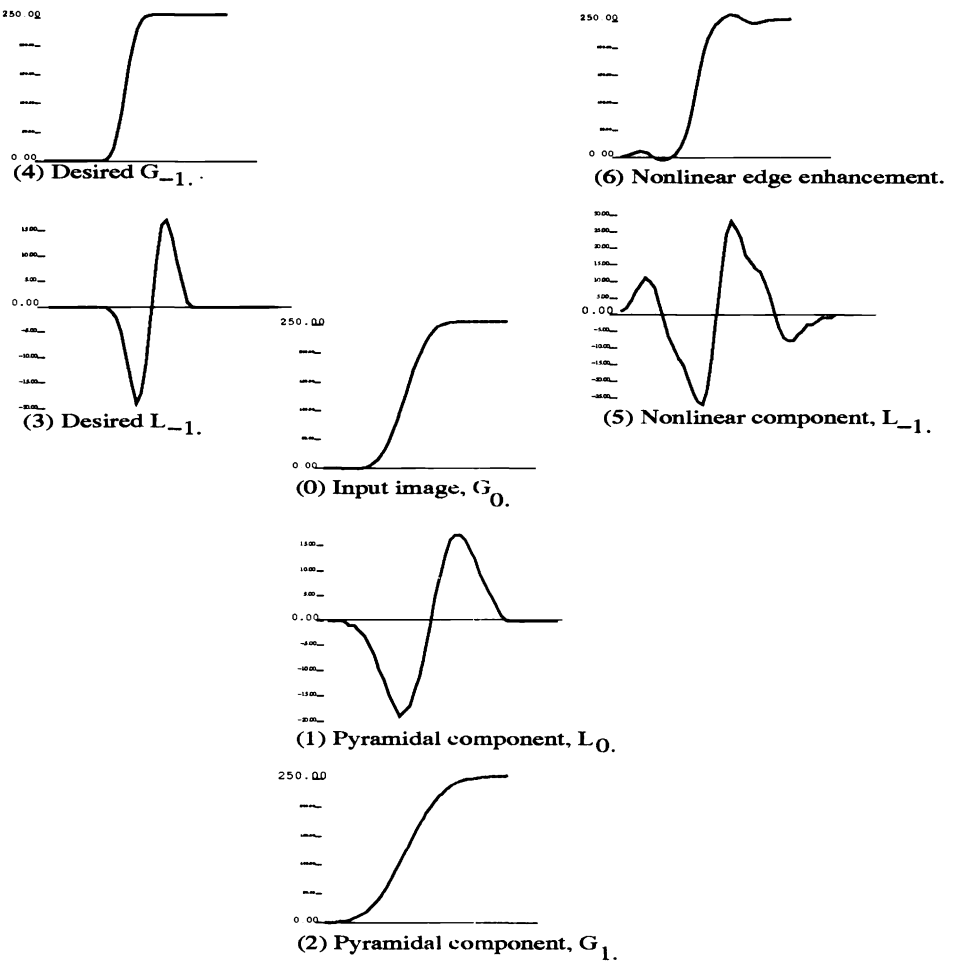


Figure 4: One-dimensional ideal-edge scenario

## 4. 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.

The first result exemplifies a zoom-in application. Here, an image is zoomed-up by a 2 to 1 ratio using the expand operation described above (section 3). 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. 5 presents the given and enhanced images, left to right respectively, together with their corresponding power spectrum characteristics (bottom). The zoomed-in input is shown top left. This is part of a monkey's head. The output of the proposed enhancement technique is presented 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 bottom of the figure, 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. 6 exhibits the monkey image example (part of this image was used in Fig. 5). 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 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. 7 displays the corresponding power-spectra characteristics. The power spectra at the bottom of the figure has its higher frequencies augmented.

Our final example is a rockscene image displayed in Fig. 8. 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.

## 5. 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. A different nonlinear filter for image enhancement has been proposed recently by Mitra[6]. The filter behaves like a local-mean-weighted highpass filter. The basic diagram of the image enhancement algorithm is the same (see Fig. 1). Still, the motivating background is very different resulting in interesting differences. In our algorithm a circularly symmetric filter is used for the extraction of the high-frequency $L_0$ map (as part of the pyramid generation). This allows for a uniform handling of edges at all possible orientations. In Mitra's work, the filtering is biased either in the vertical and horizontal directions or along the diagonal directions - with the latter chosen as giving better performance. The difference between the 2 filter performances can be detected in the example of Fig. 9. The input image is presented top left. The enhanced output of Mitra's algorithm is shown top right. The output of our enhancement algorithm is shown on the bottom. Both outputs are nice enhancement results. The differences can mainly be detected in the pants region, where our algorithm has less aliasing effects, and in the scarf (ribbon area) - where an undesired zig-zag effect is detected in Fig. 9 top right. These, we believe, are directly related to the filtering characteristics.

The non-linearity of the enhancement process is introduced in Mitra's work via a multiplication of the highpass filter by the local mean. This 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 non-linearity is introduced via 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 non-linearities is that the phase

of the edges is preserved in our algorithm, following the scale-space formalism. The procedure outlined in Mitra's algorithm, however, has the effect of shifting the phase towards the brighter region. It remains to be investigated if this might cause any undesired effects.

Computationally, the main difference between the approaches is in the chosen non-linearities. 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 in Mitra's algorithm.

## 6. SUMMARY

The presented enhancement technique is only a first step towards what can be accomplished by extrapolation across scale. Edges are only one major scale invariant features. Lines and dots for example require additional analysis. It may be desirable to use an adaptive threshold rather than the constant value used, although this has a side-effect of introducing phase shifts whose effects on the perceived sharpness are unknown at the present. 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.

In conclusion, we have described an enhancement scheme that could very well address the most important features required in producing visually pleasing enhanced resolution 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).

## 7. ACKNOWLEDGMENTS

C. H. Anderson was with the Jet Propulsion Laboratory, Pasadena CA, at the time of the research. This research was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

## 8. REFERENCES

[1] W. K. Pratt, *Digital Image Processing*, John Wiley and Sons, Inc. 1978.
[2] P. J. Burt and E. A. Adelson, "The Laplacian Pyramid as a compact image code," *IEEE Trans. Commun.*, Vol. COM-31, pp. 532-540, 1983.
[3] C. H. Anderson, "A Filter-Subtract-Decimate Hierarchical Pyramid Signal Analyzing and Synthesizing Technique," *United States Patent* 4,718,104, 1987.
[4] A. Witkin, "Scale-space filtering," in *Proc. IJCAI*, Karlsruhe, West Germany, pp. 1019-1021, 1983.
[5] A. L. Yuille and T. Poggio, "Scaling theorems for zero-crossings," Massachusetts Inst. Technol., Cambridge, A. I. Memo 722, 1983.
[6] S. K. Mitra, H. Li, I. Lin and T. Yu, "A new class of Nonlinear Filters for Image Enhancement," *ICASSP 91*, M5.1, pp. 2525-2528, 1991.
[7] D. C. Marr and E. C. Hildreth, "Theory of edge detection," in *Proc. Roy. Soc. London B*, vol. 207, pp. 187-217, 1980.
[8] J. W. Woods, *Subband Image Coding*, Kluwer Academic Publishers, 1991.
[9] G. S. van der Wal, "The Sarnoff Pyramid Chip," *Proc. Computer Architecture for Machine Perception*, CAMP-91, Paris, Dec. 16, 1991.

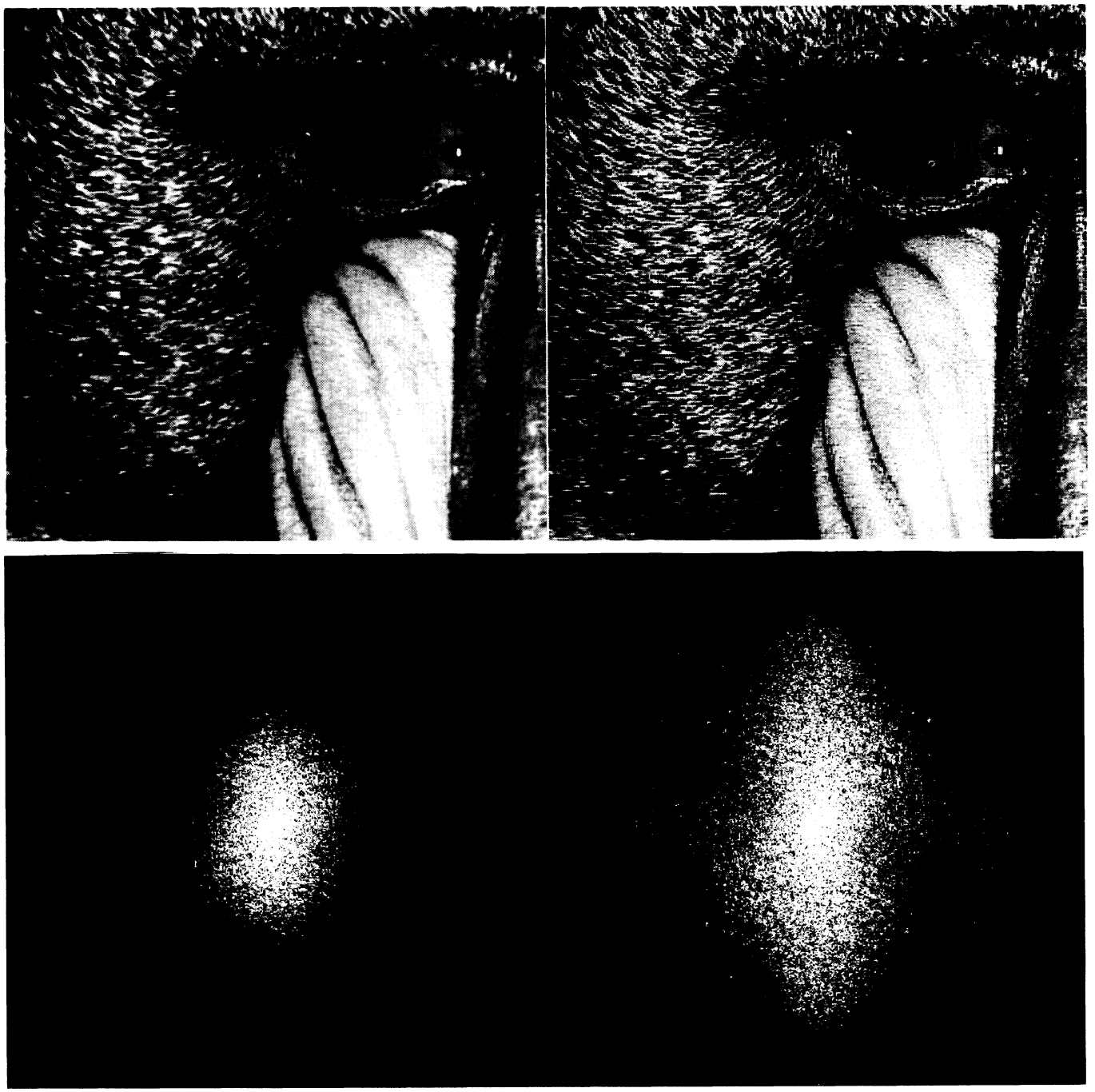

Figure 5: Given and enhanced images, left to right respectively, together with their corresponding power spectrum characteristics (bottom). It is evident that the input power spectra is augmented. The enhancement process extrapolates to higher frequencies, thus producing the satisfying enhanced result.

Figure 6: 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. Extrapolating to higher frequencies via the proposed enhancement algorithm results in the bottom image.

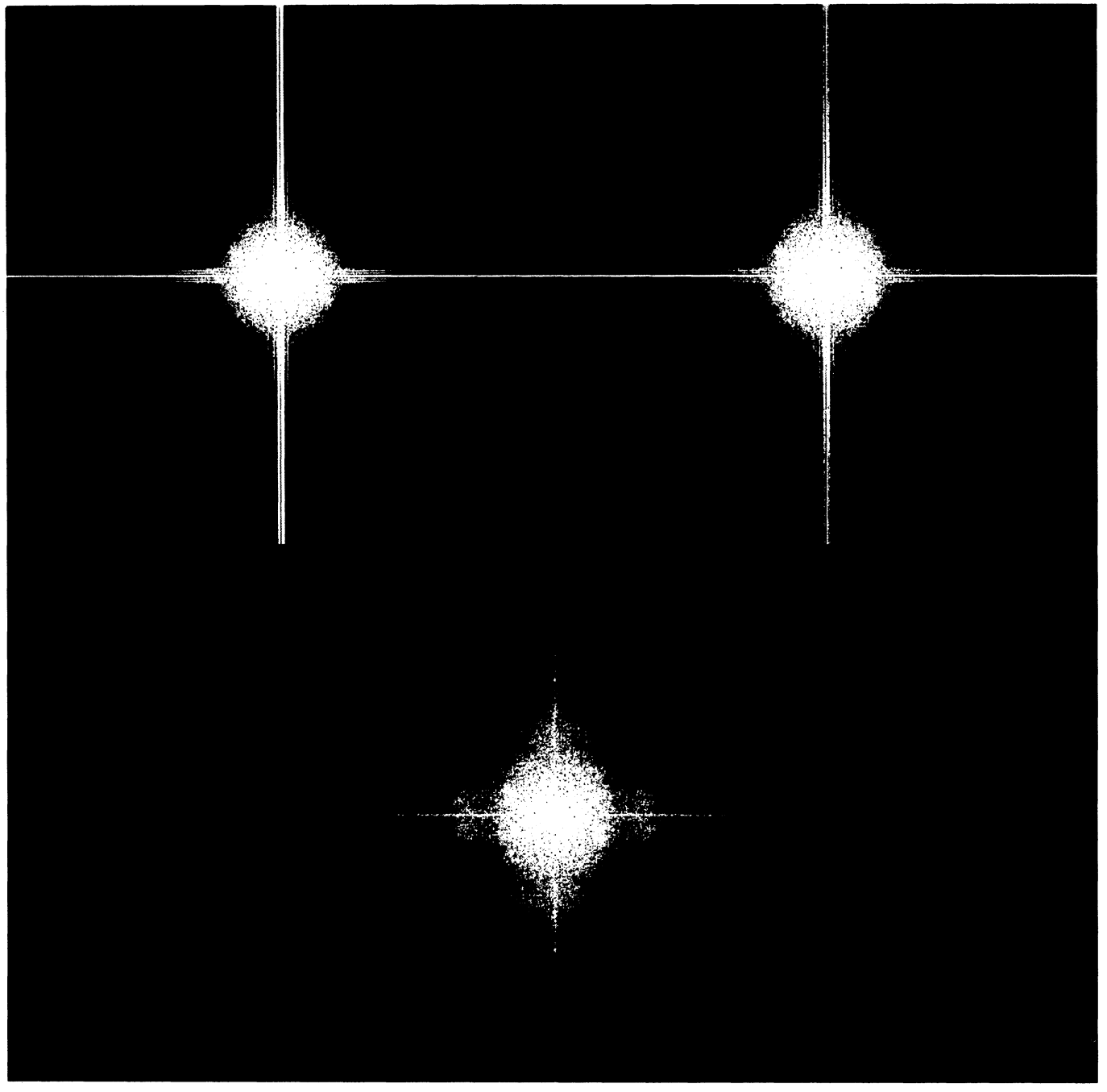

Figure 7: Power spectra characteristics of the monkey image. The blurred input power spectrum, the high-frequency enhanced image power spectrum and the spectrum of the image enhanced with the proposed scheme are presented top left, top right and bottom, respectively.

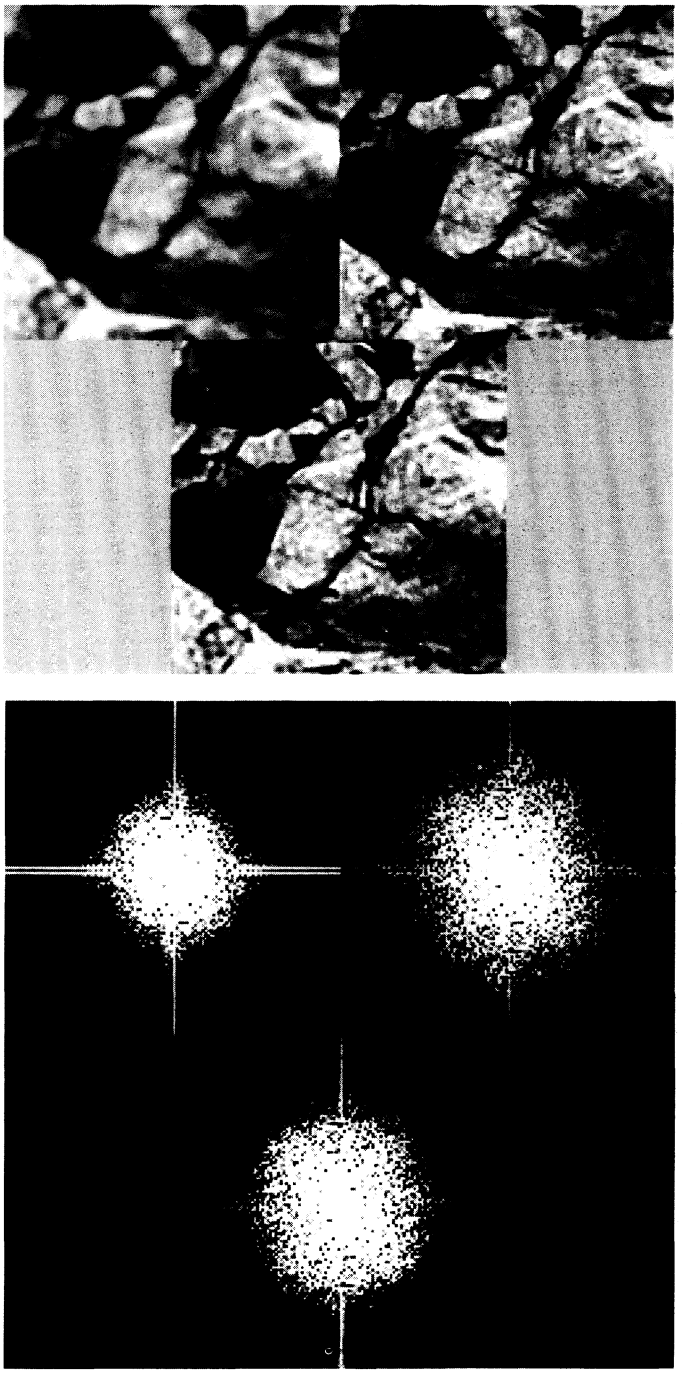

Figure 8: Enhancement results and corresponding power-spectrum characteristics. In each of the above figures the blurred input and original image are presented (top left and top right, respectively) followed by the enhanced output (bottom). Both visual perception enhancement and power-spectrum augmentation are evident.

Figure 9: Comparison to Mitra's algorithm. Blurred input is presented top left. Mitra's algorithm result presented top right. The proposed enhancement algorithm produces the result presented at the bottom of the figure.