Perceptually Motivated Automatic Sharpness Enhancement Using Hierarchy of Non-Local Means

Anustup Choudhury and Gérard Medioni
University of Southern California
Department of Computer Science
Los Angeles, California 90089. USA
{achoudhu,medioni}@usc.edu

Abstract

We address the problem of sharpness enhancement of images. Existing hierarchical techniques that decompose an image into a smooth image and high frequency components based on Gaussian filter and bilateral filter suffer from halo effects, whereas techniques based on weighted least squares extract low contrast features as detail. Other techniques require multiple images and are not tolerant to noise.

We use a single image to enhance sharpness based on a hierarchical framework using a modified Laplacian pyramid. In order to ensure robustness, we remove noise by using an extra level in the hierarchical framework. We use an edge-preserving Non-Local Means filter and modify it to remove potential halo effects. However, these effects are only reduced but not removed completely after similar modifications are made to bilateral filter.

We compare our results with existing techniques, including commercial packages and show better decomposition and enhancement using our method. Based on validation by human observers, we introduce a new measure to quantify sharpness quality, which allows us to automatically set parameters in order to achieve preferred sharpness enhancement. This causes blurry images to be sharpened more and sufficiently sharp images to not be sharpened.

1. Introduction

The human visual system can perceive a scene with precise representation of detail and color. However, while trying to capture the scene and display it, due to limitations of devices such as camera, monitors etc., the details of the scene in an image may not be clearly visible. Also, for old people with vision problems, the image of any scene may appear blurred. In such situations, we need to enhance the sharpness of the image. Examples of our sharpness enhancement can be seen in Figure 1.

Some of the existing techniques to achieve sharpness enhancement use a hierarchical framework and decompose the image into different levels - a smooth image and several fine/detail components. To decompose the image, different types of filters are used. However, due to the inability of these filters to preserve edges, these techniques suffer from halo effects or produce loss of structure in the image. Other techniques use multiple images taken under different lighting conditions, which may not always be practical.

In this paper, we present a robust sharpness enhancement technique using a single image, based on the Non Local means (NL-means) algorithm [1] using a hierarchical framework to decompose the image into one smooth level and several fine levels. We modify the Laplacian pyramid framework as follows - 1) We do not sub-sample the image and 2) While decomposing the image, we smooth the original image by a higher degree. We found NL-means filter to be well-suited for progressive smoothing of images resulting in better extraction of fine levels. We modify this filter using segmentation information to remove halo effects. The robustness of our method lies in using an extra level of decomposition by smoothing the image using a very low value of the filtering parameter. Once the different levels of the image are obtained, we can enhance those levels individually and combine them to get an enhanced image.

We introduce a new method to measure the sharpness quality of images based on image gradients and derive our measure by conducting experiments on human observers. The advantages of using this measure are that 1) it correlates well with perceived sharpness and 2) it can be used to predict the sharpness of an image which then allows automatic estimation of the “best” enhancement parameters.

The rest of the paper is organized as follows: In Sec. 2, we review previous work. In Sec. 3, we describe the proposed method. In Sec. 4, we show results and compare our results with other techniques. We also present results from human evaluation. Finally, we conclude our paper in Sec. 5.
2. Previous Work

Sharpness enhancement can be achieved by many methods. Peaking [13] increases the amplitude of the high frequency component by adding the 2nd derivative at the edge. LTI/CTI improves the perceived sharpness [16] but it may result in Moire effect on video sequences.

Other methods use a hierarchical framework as mentioned earlier. In order to decompose the image, various edge-preserving filters can be used. Greenspan et al. [12] use a Gaussian filter. Recently, Farbman et al. [9] have proposed a weighted-least squares (WLS) approach to perform edge-preserving smoothing. Bilateral filter [22] and its variations are used by Durand and Dorsey [8] for edge-preserving smoothing. Chen et al. [4] construct a bilateral pyramid using bilateral filter by increasing the width of the kernels to do smoothing for progressive abstraction of videos. Fattal et al. [11] recursively applies bilateral filter and combines the images and their high frequency components from multiple light sources.

Farbman et al. [9] have shown how the WLS approach results in an edge preserving smoothing by showing absence of blur across edges and smoothing in flat regions of the smooth image. However, in order to show preservability of edges by smoothing techniques, it is important to look at method noise [1], which is the difference between the images before and after smoothing. The "best" method produces noise uncorrelated with the input image. As shown in [1], a Gaussian filter blurs the edge during smoothing while removing noise and this effect can be seen in the method noise. A bilateral filter with high value of range filter results in loss of structure causing potential halo effects.

Figure 2 compares the results of smoothing using WLS and NL-means filter. In case of WLS, it "seems" that the edges are preserved but if we visualize the method noise, we can see that edge information is lost. As shown in [1] and Figure 2(c), NL-means filter does the best edge-preservation. There is no structure present in the method noise and the smooth image has no blur across edges.

Tumblin and Turk in [23] have used a variant of anisotropic diffusion that works well for preserving edges in an image. However, this technique tends to oversharpen edges and may result in artifacts. Zhang and Allebach [25] modify the range filter of the bilateral filter to perform both sharpness enhancement and noise removal. However the performance is constrained by the choice of training dataset. Subr et al. [20] acquire information about oscillations from the local extrema of an image at multiple scales, and enhance details in an image. Recently, Fattal [10] proposed a multi-resolution analysis framework based on wavelets.

3. Our Method

Our method is summarized in the flowchart shown in Figure 3. It consists of 4 key steps: (1) Noise Removal (2) Decomposition using hierarchical framework (3) Segmentation (4) Enhancement of the decomposed levels. We highlight why we prefer NL-means filter over existing approaches. The image is converted from RGB color space to CIELAB color space and only the lightness channel is modified, in order to preserve the color properties.

We use the NL-means filter [1] to remove noise and to decompose the image. The hypothesis behind using this filter [1] is that for any image, the most similar pixels to a given pixel need not be close to it. They could lie anywhere in the image. For comparing how similar the pixels are, the difference between the neighborhood of the pixel is considered. This technique uses self-similarity in an image to
reduce the noise. The formulation of the NL-means filter is:

\[ NLu(x) = \frac{1}{N(x)} \int e^{-(G_p * |u(x+) - u(y+)|^2)/2} u(y)dy, \]

where \( u(x) \) is the observed intensity at pixel \( x \), \( G_p \) is a Gaussian kernel with standard deviation \( \rho \), \( h \) is the filtering parameter that controls the amount of smoothing and \( N(x) \) is the normalizing factor. Equation 1 means that every image pixel \( u(x) \) is replaced by the weighted average of other pixels in the image \( u(y) \) and this weight is given by the similarity between the Gaussian neighborhood of pixel \( x \) and pixel \( y \). Ideally, for a given neighborhood of pixel, \( W \), we should search the entire image to find a similar neighborhood. But for efficient computation, we consider a smaller local search area, \( S \). The values of the search area and neighborhood are \( S = 7 \times 7 \) and \( W = 5 \times 5 \). Enhancement is achieved by increasing the intensity of each level and then adding back the enhanced levels.

### 3.1. Noise removal to ensure robustness

All existing enhancement techniques make the assumption that the input image is noise-free. This need not be the case, and most existing techniques fail in the presence of noise. This is because, after smoothing, noise is still present in the fine level and enhancement of the fine level will result in enhancement of noise, thus spoiling the visual quality of images.

In order to remove noise, as an initialization, we filter the image using NL-means filter with a low value of \( h \). If noise is present in the image, it will be removed before enhancement, thus preventing the degradation of image quality. If noise is not present in image, smoothing using low values of \( h \) causes minimal loss of structure. Removal of this detail from image does not result in much difference when visualizing the image. We can represent our system as shown in Equation 2.

\[ J = I + Residue, \]

where \( J \) is the input image and \( I = NL(J) \). If \( J \) is noisy, \( Residue \) is noise, else \( Residue \) is negligible. \( I \) can be decomposed as shown in Equation 3 and enhanced as shown in Equation 5 whereas \( Residue \) is added after enhancement to preserve the original properties of the image. From Section 2 and [1], we find that NL-means is better at both edge-preservation and noise removal than existing techniques.

### 3.2. Hierarchical Decomposition with NL-means

The hierarchical framework is inspired from the Laplacian pyramid [2] and we build it using NL-means filter with filtering parameter, \( h_i, i \in [1, k] \) where \( k \) is level of hierarchy. A similar framework is used by Liu et al. [17] to exploit self-similarity at different hierarchical levels for denoising. In our case, \( h_i < h_{i+1} \). The filtering parameter, \( h_i \) can be used for abstractions of the different levels of hierarchy to obtain images at different spatial scales. The effects of \( h_i \) on enhancement can be seen later in Section 4.1. Unlike Laplacian pyramid, we do not sub-sample the smooth images because the different levels are obtained by smoothing using an edge-preserving filter. As a result, the non-noisy input image is decomposed into several fine levels \( f_i \) and the smoothest level \( s \). This can be expressed as

\[ I = s + \sum_{i=1}^{k} f_i, \]

where \( s = NL_{h_0}(n_0) \). \( NL_{h_i}(x) \) is the smooth image obtained by applying the NL-means filter with filtering parameter, \( h_i \), on image \( x \) and \( n_0 \) is the input non-noisy image \( I \).

The smooth image is obtained by applying the filter on the original image with a high value of filtering parameter, \( h_k \). The fine image at any level is calculated as the difference between the original image and its smooth image. This can be expressed as shown in Equation 4.

\[ f_i = n_{i-1} - NL_{h_i}(n_0), \]

where, \( n_i = NL_{h_i}(n_0) \).

### 3.3. Segmentation

As mentioned earlier in Section 3.2, to get abstractions at different spatial scales of the image, the value of \( h \) is progressively increased at every level. As we can see in the first row of Figure 4(a), a high value of \( h \) results in blurring of edges and consequently enhancements of the detail layer(bottom row of Figure 4(a)) will cause halo artifacts.
where we decompose using bilateral filter with adaptive 
other example of these effects can be seen in Section 4.2 
(served, the smoothing of the image is not done effectively 
We observe that even though the edges of the image are pre- 
varied \( \sigma \) noticeable effect only at higher levels. 
image as shown in Figure 4(b). Though the segmented im-
aries of the image and thus removes the halo effect from the 
hierarchy is chosen. This preserves high contrast bound-
image can be used at all levels of the hierarchy, it will have 
noticeable effect only at higher levels. 
In order to remove blurring of edges, we first segment the 
image using Mean-shift segmentation algorithm [7] and find 
out high contrast edges in the image. Then, while smooth-
ing the image using NL-means filter, we adaptively modify 
the smoothing parameter such that if an edge occurs in the 
pre-segmented image, a lower value of \( h \) is chosen, oth-
wise the default value of \( h \) for that particular level of the 
hierarchy is chosen. This preserves high contrast bound-
daries of the image and thus removes the halo effect from the 
image as shown in Figure 4(b). Though the segmented im-
can be used at all levels of the hierarchy, it will have 
noticeable effect only at higher levels. 
In Figure 4(c), we used bilateral filter for smoothing and 
varied \( \sigma \) depending on the presence or absence of edges. 
We observe that even though the edges of the image are pre-
erved, the smoothing of the image is not done effectively 
(noise is still present in the image). This will only help to 
reduce halo effects but will not remove it completely. An-
other example of these effects can be seen in Section 4.2 
where we decompose using bilateral filter with adaptive \( \sigma \).

3.4. Sharpness Enhancement

Once the different levels are obtained, sharpness en-
hancement can be achieved as shown in Equation 5.

\[
\text{EnhancedImage} = l_0 * s + \sum_{i=1}^{k} l_i * f_i, \tag{5}
\]

where \( k \) denotes the number of decomposition levels, \( l_i \)
is the enhancement factor for every fine level and \( l_0 \) is 
the enhancement factor for smooth level. The parameters 
\( l_i, i \in [0, k] \) can be user-defined to modulate sharpness. The 
default parameters in this paper are \( k = 2, l_0 = 1, l_1 = 5 \)
and \( l_2 = 1 \). Modifying \( l_1 \) creates most noticeable effect 
and its value can be automatically set to achieve preferred 
sharpness enhancement depending on the value of Sharp-
ness Preference Index (SPI) as defined in Section 4.3. The 
default values of the filtering parameters \( h_i \)’s are \( h_1 = 0.1 \)
and \( h_2 = 1 \) for uniform regions. In the presence of an edge, 
the value of the filtering parameter is \( h_0 = 0.01 \).

4. Results and Discussion

In this section, we show results of our sharpness en-
hancement and decomposition and compare it with existing 
techniques. We also introduce a new method to measure 
sharpness enhancement that facilitates automatic enhance-
ment and discuss the computational costs of the algorithm.

4.1. Sharpness Enhancement

The results of our enhancement is shown in Figure 1. 
The increase in enhancement is only due to increase in the 
value of \( l_1 \) from 5 to 10. \( l_1 \) corresponds to fine level \( f_1 \) 
which contains the minor details of the image and therefore 
increasing \( l_1 \) boosts the minor details of the image. Increas-
ing \( l_0 \) increases the intensity of the smooth level, \( s \) resulting 
in a brighter enhanced image.

![Figure 5. Abstractions at different spatial scales by changing filtering parameter \( h \). (a) is original image. (b) uses \( h = 0.01 \) at level 1 and (c) uses \( h = 0.1 \) at level 1. \( l_1 = 5 \) for both enhancements](image)

Changing the value of \( h \) causes abstractions at different 
spatial scales. As can be seen in Figure 5(b), using very 
low values of \( h \) results in the abstraction of very fine details 
of the image (sand particles) that are not preferred. Using 
higher values of \( h \) \( (h = 0.1) \) results in the abstraction of 
preferred and more perceptible details such as the patterns 
in sand and the rock formations. 

We have compared our results with the results obtained 
by Fattal et al. [11] as shown in Figure 6. We can see that 
our technique results in better delineation of the veins of the 
leaf. Note that our method uses only one input image 
whereas Fattal et al. use 3 input images under different light 
source directions.

We compare our result with the result of applying Photo-
shop’s unsharp mask on Figure 1(a) as shown in Figure 7. 
Comparing this with our enhancement as shown in Fig-
ure 1(c), we can see that although sharpness is enhanced 
using an unsharp mask, there is a very distinct undesirable 
halo effect along the perimeter of the flower that is 
not present in our method. The recent technique by Subr 
et al. [20] also suffer from subtle halo effects as can be seen

![Figure 4. Filtering using edge information. The top row is the 
smooth image and the bottom row is the method noise. The red 
rectangle is zoomed for clarity. (a) uses NL-means filter with high 
value of \( h = 0.5 \) and no edge information (b) uses edge and vary-
ing \( h \) (c) uses bilateral filter along with edge and varying \( \sigma \)](image)
Figure 6. Comparison of our enhancement (bottom) with Fattal et al. [11](top). Details are more clearly visible in the bottom image by the white colored regions that is present at certain parts along the boundary of the flower and the leaves.

Figure 7. Halo effect (white) along the boundary of flower and leaves using existing methods and the lack of it using our method (Figure 1(c)). Left image is enhanced with Photoshops’s unsharp mask and right image is from [20]

4.2. Comparison of image decomposition

In this section, we compare the decomposition at multiple scales of our technique with existing approaches. Given an image, it is hard to quantify the spatial scale of details of image and there are no ground truths available for either natural or synthetic images. Consequently, we cannot use any metric such as PSNR or SSIM for quantitative evaluation. RMSE is known to not correlate well with perceived image quality [25]. We use visual assessment to evaluate the performance of different techniques. At the risk of visual assessment being subjective, we point out the differences during decomposition of the different approaches.

As shown in Figure 8, in the bilateral pyramid by Chen et al. [4], strong edges of the image are not retained and visible ringing can be seen in the fine levels. The decompositions by LCIS [23] and by Trilateral filter [6] also results in visible ringing in the fine levels. Enhancement of the levels that suffer from ringing artifacts result in halo effects. The decomposition technique used by Fattal et al. [11] results in gradient reversal. WLS extracts low contrast features such as clouds present at the top of the image in the fine level.

Segmentation is an important part of our method. As shown in Figure 8(h), ringing can be observed in the fine level when we use NL-means filter without segmentation information. Our technique does not suffer from such artifacts. If decomposition is performed with bilateral filter using the edge information, then the ringing artifacts are reduced but not removed completely.

4.3. Quantitative Results and Human Evaluation

To measure the sharpness of images, we use modified version of the Tenengrad criterion [15, 21], that can be represented as shown in Equation 6.

\[\text{Tenengrad Criterion} = \frac{\sum_{x,y} \sqrt{\text{grad}_x^2 + \text{grad}_y^2}}{n}, \quad (6)\]

where \(n\) is the number of pixels in the image, \(\text{grad}_x\) and \(\text{grad}_y\) are the morphological gradients along the horizontal and vertical directions and \(\text{gradient}(image) = \text{dilate}(image) - \text{erode}(image)\). Existing enhancement techniques [5, 19, 3] consider the sharper image to be better which need not always be true. Since perception of sharpness is subjective, we conduct experiments to analyze the...
response of human observers regarding the sharpness of an image. Our experimental setup consists of 50 images that were obtained from [9] and Flickr. We split this dataset into a training set of 28 images and a test set of 22 images. All the images were enhanced offline by increasing the value of $l_1$. Each subject was seated at a distance of 20" from the 24"(diagonal) screen. Our psychophysical experiments can be divided into 3 parts - A) Check if Tenengrad criterion correlates with perceived sharpness B) Introduce a new metric for sharpness to aid automatic sharpness enhancement using the training set, and C) Validate the new metric for sharpness using the test set. 10 observers were used for both the correlation experiment and the training procedure and 15 observers were used for the test procedure.

4.3.1 Experiment A: Correlation Experiment
In the first experiment, we find out how well the Tenengrad Criterion correlates with perceived sharpness. From the training set, we select 4 distinct images and enhance those images by increasing $l_1$ from $l_1 \in [1 \ldots 8]$. For each set of images, we randomly present those 8 images independently to 10 observers and ask them to rank the images in increasing order of sharpness and note their responses. We did not use higher values of $l_1$ because the enhancements become indistinguishable and the Tenengrad criterion also converges. We compare the ranks of the sharpness that is assigned by the Tenengrad criterion with the ranks that are provided by the human observers.

We use the Spearman’s rank correlation coefficient ($\rho$) for each image to give a quantitative measure of the correlation of ranks given by the metric and those given by the observers. The value of $\rho$ ranges from $[-1, +1]$. A value of +1 implies that all pairs of ranks are equal. The value of $\rho$ can be computed as shown in Equation 7.

$$\rho = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)},$$  \hspace{1cm} (7)

where $n = 8$ is the number of levels of image and $d_i$ is the difference between the $i^{th}$ pair or ranks given by the metric and the observer. The median value of $\rho$ is 0.9762 which implies that the Tenengrad criterion correlates well with perceived sharpness and this correlation is statistically significant ($p < 0.03$, Median of $\rho = 0.0003$).

4.3.2 Experiment B: Sharpness Metric
As motivated earlier, existing sharpness metrics do not correspond to human perception. We conduct experiments to introduce a new metric for sharpness measure to perform automatic sharpness enhancement. We use the training set of 28 images and increase the value of $l_1$ from 1 to 25 at equal intervals, thus increasing the sharpness of images.

![Figure 10. Tenengrad values on 9 images. Responses are marked by circles (Preferred images) and squares (Transition to “Too-detailed”). The size corresponds to the number of responses.](image)

Our experiments demonstrate that the preferred image need not be the sharpest image. This can be seen by the position of the circles in Figure 10, which is not present at the highest value of $l_1$. In fact, after a certain value of $l_1$, as the image becomes sharper the human observers report the details being too much, which is demonstrated by the position of the squares. We also observe that the Tenengrad criterion converges for large values of $l_1$, typically 500.

![Figure 9. Correlation between the ranks assigned using Tenengrad criterion and the ranks assigned by human observers for (left) image with the best average $\rho (\rho = 0.9714)$ and (right) image with the worst average $\rho (\rho = 0.9524)$ across all observers. The identity line depicts perfect agreement between the different ranks.](image)
amount of detail present in an image.

4.3.3 Experiment C: Automatic Enhancement

Given an image, since all computations can be performed off-line, we can compute its SPI and predict the sharpness quality of image. We use the test set of 22 images and enhance those images automatically with the lowest value of $l_1$ such that its SPI is in the range $[0.23, 0.36]$. For every image, we present the original and the preferred images simultaneously and randomly (either left or the right side) on the screen. When the original image is selected as the preferred image, we present the next higher level of enhanced image as the original image. Then each of the 15 observers were asked independently to rate the image on the “right” side as “Better”, “Same” or “Worse” relative to the image on the “left” side of the screen. Scores were thus assigned as 1 for Better, 0 for Same and -1 for Worse to the selected image. These responses were used to evaluate the “preference” for the image selected by our automatic mechanism.

Figure 12 shows the preference ratings for the automatically selected images. We can see that the observers shows a strong preference for the automatically selected images. We use the Wilcoxon signed-rank test [24] to show that the preference for automatically selected images are statistically significantly different from that for the original images. The empirical value of $A_z$ for the different observers are shown in Figure 13. For all the observers on an average, ($A_z = 0.69 \pm 0.19$), the selected images are deemed to have better quality than the original images. Using the Wilcoxon signed-rank test [24], our automatically selected images have statistically significantly better quality than original images ($p = 0.0074$).

Although preference for our automatic selection is statistically significant, in some cases (5 images) as shown in Figure 12 the observers prefer the original image. Apart from personal preferences, this is due to the characteristics of the image. The observers prefer less enhancement of good quality natural images. However, they show preference for enhancements in images having an object in the foreground. Categorization of images and tuning the parameters accordingly may help in improving the efficiency of our implementation. We can use prior information about the image such as its blur content, saliency map and texture analysis and accordingly enhance the images.

4.4. Computational Cost

Since using NL-means filter is computationally expensive, we use a GPU implementation of the filter [14] and for a 800 X 533 image, every level of the hierarchy takes 0.07 seconds on a Windows PC with Xeon processor having GeForce GTX 480 GPU and Visual C++ environment.

5. Conclusion and Future Work

We have proposed a sharpness enhancement technique from a single image based on a hierarchical framework us-
ing an adaptive Non-Local means filter. Our technique does not suffer from the limitations of existing approaches - halo artifacts, extraction of low contrast features as details, multiple images, noise enhancement etc. We have compared our results with that from existing techniques and we use visual evaluation to show that our technique results in better enhancement of the images. On the basis of human perception, we introduced a new measure for sharpness enhancement using which we automatically sharpen images.

As future work, we would like to explore how this system scales from images to videos. We would like to use prior information about image category and saliency and test SPI on a larger dataset with more people. Our method improves the perceptual sharpness of the image for people with normal vision. We would like to explore whether it helps people with age-related Macular Degeneration.

Figure 14. Automatic sharpness enhancement of images. Left are the original images and right are the enhanced images. The box represents the preferred images according to our metric. For the first image, the enhanced image ($l_1 = 3$) is preferred whereas for the $2^{nd}$ image, the original image ($l_1 = 1$) is preferred. The enhancements are better visible at the original high resolution.

6. Acknowledgement
This research was supported by the National Institutes of Health Grant EY016093.

References