LOCALLY ADAPTIVE MULTISCALE CONTRAST OPTIMIZATION

Nicolas Bonnier and Eero P. Simoncelli

Center for Neural Science, and Courant Institute for Mathematical Sciences,
New York University, New York, NY 10003
Email: nbonnier@cns.nyu.edu, eero.simoncelli@nyu.edu

ABSTRACT

We describe a method for automatically and adaptively boosting the visibility of local features in an image. A log intensity image is first decomposed into a set of subbands at multiple scales and orientations. Operating successively from coarse frequency bands to fine, the coefficients of each subband are adjusted so as to move their locally averaged amplitudes toward a target value using a gamma operation. Target values are chosen to fall linearly over scale, consistent with a scale-invariant spectral model. To avoid enlarging the range of image intensity values, in those locations where the local mean is near the minimal or maximal values of the image and the local contrast is being boosted significantly, the local mean is moved toward the global mean. Finally, a spatial mask is applied in the pixel domain to ensure that the enhancements are applied only in the vicinity of image features. The resulting image appears to be both sharper and of higher contrast.

Introduction

Essentially all devices used for capture or reproduction of visual images are incapable of representing the full range of intensities found in the visual world. Engineered devices often handle this problem by compressing or truncating the intensity range. For example, the process of film exposure, development and printing reproduces light intensities according to a sigmoidal function that compresses the contrast of low and high intensity values. Other such "tone-mapping" solutions include clipping, linear re-mapping, gamma (exponential) corrections, and histogram equalization. During any such process, regions of the scene that are contrast-reduced (i.e., where the slope of the mapping function is low) can become difficult or impossible to see. Professional photographers have learned to compensate for these problems in the darkroom, by selectively "dodging" or "burning" local image regions where the loss of contrast would be detrimental to the desired appearance.

The human visual system also uses sensors with limited response range. It does perform global adjustments to adjust the intensity range (for example, by adjusting the size of the iris). Perhaps more importantly, it uses spatially adaptive processing such as local gain control in order to "see" details in all locations within the image. Although these adaptive biological processes are not yet fully understood, it is clear that digital image processing offers the flexibility to implement such solutions and a variety of methods have begun to take advantage of such principles [e.g., 1–5].

In this paper, we describe a methodology for adaptively adjusting contrast within a digital image, without introducing visible artifacts or expanding the overall image intensity range. The image is decomposed into multiple frequency bands, and the coefficients in each band are modified using a nonlinear "gamma" operation that moves their local average magnitude toward a target value. Target values for each band are chosen to fall with with a slope of $-2/\mathrm{octave}$, consistent with a a $1/f^2$ spectral model. The method can be applied to a conventional digital image, in order to enhance the visibility of features that might otherwise be lost when displayed. It is also relevant for processing of high-dynamic range (HDR) images, in order to render them more visibly on a low-dynamic range display.

Preprocessing

We start by preprocessing the image pixel values so that they represent log light intensities. This kind of processing roughly mimics the transformation achieved by the retina, and has been studied by a number of authors (see [6] for review). The test images used to demonstrate the method are shown in Fig. 1.

A number of authors have advocated the use of multiscale representations for contrast adjustment [e.g., 2,3,7,8]. We decompose our images using the steerable pyramid, a multiscale subband representation whose basis functions are derivatives of a radial blurring function [9]. For this paper, we use the complex-valued version of this decomposition, as described in [10], with two orientation bands (vertical and horizontal).

The enhancement method is implemented in a coarse-to-fine iterative fashion. For each step, we operate on a subband, as well as the lowpass residual that is obtained by reconstructing all subbands at lower frequencies. The coefficients of this lowpass residual band represent the local mean of the image, and the coefficients of the subband represent the variations around this mean. The enhancement procedure is a combination of three basic operations, which are described in the following sections.

I: Equalization of local contrast

Perhaps the simplest means of enhancing contrast is to linearly boost high frequencies (known as "unsharp masking" in the photographic literature). Within a multi-scale pyramid, this can be accomplished by multiplying the coefficients in each subband by a scalar whose value is larger for higher-frequency bands. Although appealing for its simplicity, this solution is not satisfactory because high contrast and low contrast features are boosted equally. In general, contrast varies widely across a typical image, and the primary goal of our method is to reduce this variation by boosting contrast in those regions where it is low or moderate, while leaving it unchanged in regions where it is high.

To appear in *Proc. ICIP2005, September 11-14, 2005, Genoa, Italy*
© IEEE 2005

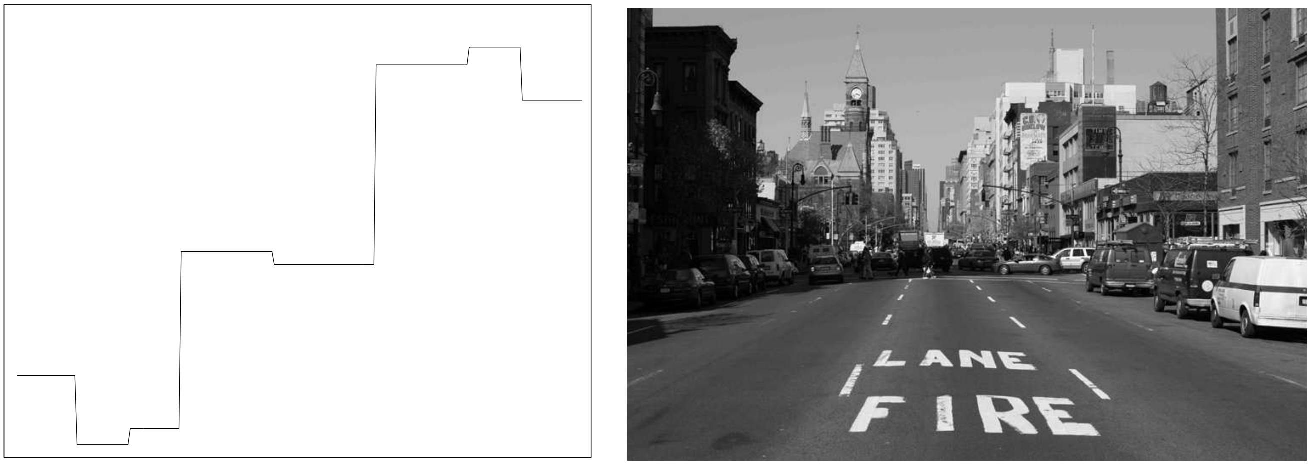

Fig. 1. Test images. Left: horizontal horizontal slice of a test image consisting of vertical step edges. Right: photographic image, taken from a 12-bit digital Canon 10D camera.

We use a nonlinear operation to boost contrast selectively. For each subband, a local contrast measure is extracted, based on the average local magnitude of the subband coefficients:

$$c(x, y) = \sqrt{\{g * |b|^2\}(x, y)}, \tag{1}$$

where g is a blurring filter (Gaussian, with standard deviation of five samples), and b represents the complex subband coefficients. Note that perceptually, contrast is usually defined as ratio of signal variation to signal mean [4, 7, 11]. Here we use only the signal strength, because the initial log-domain representation has already implicitly taken the mean into account.

To reduce the variation in contrast across the image, each coefficient is boosted according to the strength of the local contrast signal:

$$b'(x, y) = m(x, y)b(x, y),$$

where $b(x, y)$ is the original coefficient, $b'(x, y)$ the updated coefficient, and

$$m(x, y) = \left[(1 - \epsilon)\frac{c(x, y)}{t_c} + \epsilon\right]^{(\gamma-1)} \tag{2}$$

The parameter $\gamma \in [0, 1)$ determines the strength of the effect (small gamma produces a large effect, and γ=1 produces no effect), and the parameter ϵ (set to a value of 0.01 in our experiments) prevents amplification of noise in low-signal areas. The contrast ***target***, t_c, represents the contrast level toward which c is moved, and is described below.

This type of "gamma" adjustment is widely used in the intensity domain to compensate for the nonlinearities of devices such as cathode ray tubes. The particular version used here will push all contrasts toward the target contrast, producing a proportionately larger change in those values that are far from the target than in those that are near. Rewritten in the log domain, this adjustment corresponds to a weighted average of the original contrast and the target contrast, with the weight determined by γ. A related gamma adjustment was developed in [3] for rendering HDR images.

A simple choice of target contrast t_c is the global maximum of the contrast of the subband. Alternatively, one can simultaneously choose t_c across all bands of the pyramid, so as to achieve a particular spectral shape. Since the Fourier spectra of natural images have been shown by many authors to follow a power law (see [6] for review), with an exponent of roughly -2, we choose a set of target contrasts that fall at this rate with scale.

Figure 2 shows the results of this enhancement procedure, applied to a test image containing step edges, as well as a 12-bit linearized digital camera image. Note the substantial increase in apparent contrast at edges and in detail regions, as well as the appearance of edge ("ringing" and "halo") artifacts.

II: Compensatory adjustment of local mean

In regions of very low or high intensity, the amplification of subband coefficients can lead to an expansion in the total pixel intensity range. Those extremal values then need to be clipped, thus partly eliminating the effect of the contrast enhancement. Clipping can be avoided by globally adjusting (tone-mapping) the pixel values, but tends to lower the global contrast.

Our solution for this problem is again adaptive. For those locations undergoing substantial boosting and having very low (or high) local mean, we adjust the lowpass signal, moving it toward the global mean:

$$\begin{aligned} l'(x, y) &= l(x, y) \\ &+ \quad h * \lceil \max\nolimits_{x,y} [l(x, y) + c(x, y)] - (l(x, y) + c'(x, y)) \rceil \\ &+ \quad h * \lfloor \min\nolimits_{x,y} [l(x, y) - c(x, y)] - (l(x, y) - c'(x, y)) \rfloor \end{aligned}$$

where $c'(x, y)$ is the contrast of the modified coefficients (computed by applying Eq. (1) to $b'(x, y)$), $\lfloor \cdot \rfloor$ indicates the positive part and $\lceil \cdot \rceil$ indicates the negative part, and h is is a blurring filter (Gaussian, with standard deviation of two samples). The result of this operation is shown in Fig. 3. Note the increased contrast of details in the shadow region on the left side of the photographic image.

It is interesting to consider this adjustment in the case when $c(x, y)$ is constant. Under these homogeneous contrast conditions, $m(x, y)$ is constant, and the lowpass adjustment depends

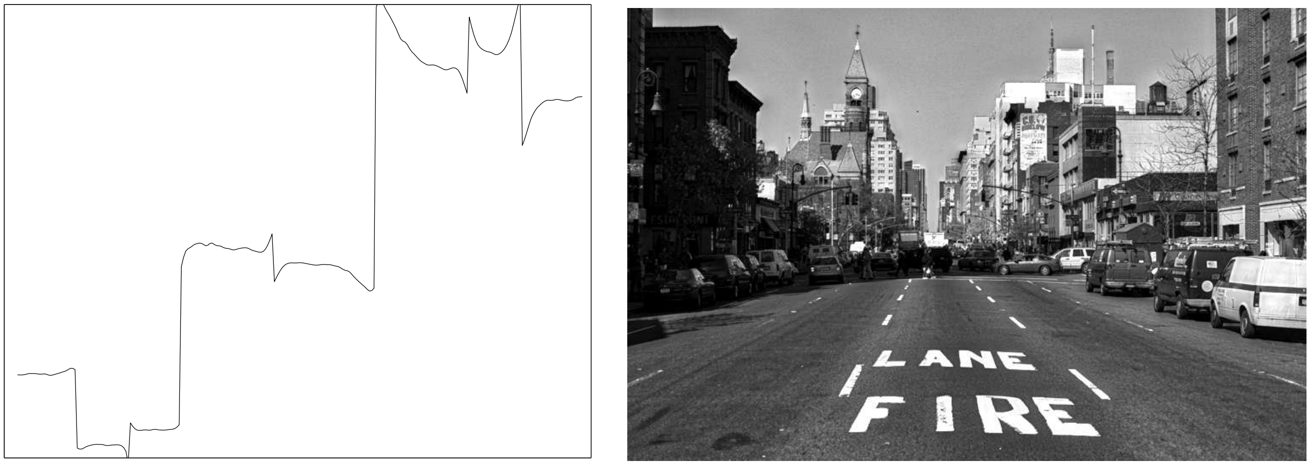

Fig. 2. Enhancement results computed by applying a "gamma" adjustment ($\gamma = 0.5$) to the contrast of each subband of a two-orientation steerable pyramid. Original test images are shown in Fig. 1.

only on the values of the lowpass coefficients themselves. The resulting function is approximately a sigmoidal nonlinearity, as is commonly used to compress overall dynamic range in film photography.

III: Spatial masking of features

The two concepts described above generate a desirable increase in apparent local contrast in the image. We find, however, that an equal modification of energy on two coefficients with identical values in different parts of the image is not perceived as equal if the surrounding of these coefficient is different. This is a "masking" effect and it suggests that we should adapt the modification of a given coefficient according to its spatial surroundings. In addition, we also find that the method produces ringing or halo artifacts near strong edges, especially if they are adjacent to flat regions (see Figs. 2 and 3). This is due to the extent of the spatial filters used in the pyramid decomposition, and to the fact that each of the coefficients that contribute to the representation of these edges are being boosted differently. Recent work on display of high dynamic range images eliminates such artifacts using robust nonlinear filters to generate lowpass bands [8, 12]. Here, we prefer to develop a solution that operates on the linear pyramid representation.

Both the masking and halo problems can be overcome by spatially masking the enhancements so that they are applied primarily in the immediate vicinity of image features. Specifically, we compute a "feature mask" by taking the mean of the log contrast across all pyramid bands at each spatial location.

$$f(x, y) = \sum_k log_2 \left(c_k(x, y) \right).$$

This mask is normalized to have a maximum value of one. Finally, the result image is computed by taking an average of the original image and the enhanced pyramid image, weighted by this feature mask:

$$r(x, y) = f(x, y)I'(x, y) + [1 - f(x, y)]I(x, y),$$

where $I'(x, y)$ is the enhanced image that is derived from the reconstructed pyramid, and $I(x, y)$ is the original image. The result of the full algorithm is shown on the two example in figure 4.

Conclusion

We have described a simple multiscale algorithm to enhance the visibility of local features in an image. The method is based on a gamma-like correction to the amplitudes of coefficients in a multiscale decomposition, similar to that proposed by several other authors [3, 4]. In addition, we adjust the local mean (lowpass residual) in those locations where it is extremal and the changes to the subband coefficients would lead to pixel values exceeding the original range. Finally, we apply the changes only in regions associated with significant local contrast. We've demonstrated the behavior of the method on two example images, and although it appears promising, a much more extensive set of tests on a wide variety of images is needed for proper validation.

We envision a number of improvements and extensions to this approach. The development of the algorithm in terms of three distinct operations is conceptually convenient, but it is difficult to guarantee that these operations will behave compatibly across all images. We believe it should be possible to combine the adjustments into a single unified operation. It would also be desirable to set the parameters (e.g., γ) automatically, based on the input image, in such a way that the algorithm becomes idempotent (i.e., an image that has already been enhanced is unaffected if the algorithm is applied again). Finally, we see contrast enhancement as a portion of a more general framework for automatic improvement of image quality, with a full solution potentially handling sharpening, denoising, and color balance.

1. REFERENCES

[1] J. P. Antoine and P. Vandergheynst, "Contrast enhancement in images using the two-dimensional wavelet transform," in *Proc. 3rd Int'l Wkshp Sig. and Im. Proc.* (B. Mertzios and P. Liatsis, eds.), pp. 65–68, Elsevier, Amsterdam, 1996.

[2] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg, "A multiscale model of adaptation and spatial vi-

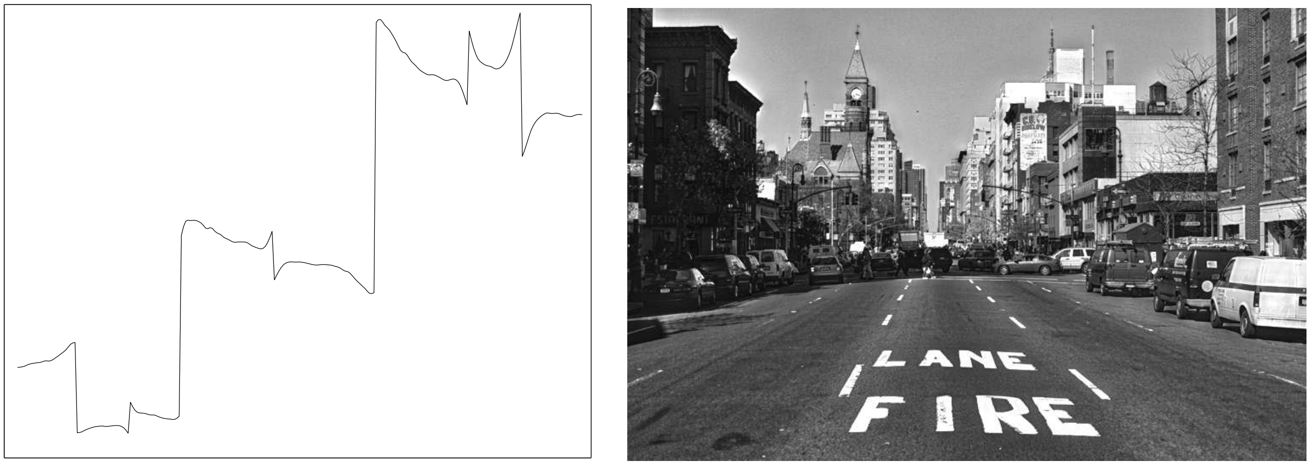

Fig. 3. Enhancement results computed by applying a "gamma" adjustment to the contrast of each subband, and a compensatory adjustment of the lowpass band. Original test images are shown in Fig. 1.

Fig. 4. Enhancement results computed from the full algorithm, which includes a "gamma" adjustment to the contrast of each subband, a compensatory adjustment of the lowpass band, and a feature mask in the pixel domain.

sion for realistic image display," in *Proc. 25th Conf. Computer Graphics*, pp. 287–298, 1998.

[3] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," in *Proc. ACM SigGraph*, Jul 2002.

[4] J. L. Starck, F. Murtagh, E. Candes, and D. L. Donoho, "Gray and color image contrast enhancement by the curvelet transform," *IEEE Trans. Im. Proc.*, vol. 12, pp. 706–717, 2003.

[5] J. Tang, E. Peli, and S. Acton, "Image enhancement using a contrast measure in the compression domain," *IEEE Signal Processing Letters*, vol. 10, no. 10, pp. 289–292, 2003.

[6] D. L. Ruderman, "The statistics of natural images," *Network: Computation in Neural Systems*, vol. 5, pp. 517–548, 1996.

[7] S. Winkler and P. Vandergheynst, "Computing isotropic local contrast from oriented pyramid decompositions," in *Proc. Int'l Conf on Image Proc*, vol. 4, pp. 420–424, Oct 1999.

[8] J. DiCarlo and B. Wandell, "Rendering high dynamic range images," in *Proc. SPIE Electronic Imaging*, vol. 3965, (San Jose, CA), pp. 392–401, 2000.

[9] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multi-scale transforms," *IEEE Trans. Info. Theory*, vol. 38, pp. 587–607, Mar 1992.

[10] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *Int'l J. Computer Vision*, vol. 40, pp. 49–71, Dec 2000.

[11] E. Peli, "Contrast in complex images," *J Opt Soc Am A*, vol. 7, no. 10, pp. 2032–2040, 1990.

[12] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *Proc. ACM SigGraph*, ACM Press, 2002.