

Fast Bilateral Filtering for the Display of High-Dynamic-Range Images

Frédo Durand and Julie Dorsey
Laboratory for Computer Science, Massachusetts Institute of Technology

Abstract

We present a new technique for the display of high-dynamic-range images, which reduces the contrast while preserving detail. It is based on a two-scale decomposition of the image into a base layer, encoding large-scale variations, and a detail layer. Only the base layer has its contrast reduced, thereby preserving detail. The base layer is obtained using an edge-preserving filter called the *bilateral filter*. This is a non-linear filter, where the weight of each pixel is computed using a Gaussian in the spatial domain multiplied by an influence function in the intensity domain that decreases the weight of pixels with large intensity differences. We express bilateral filtering in the framework of robust statistics and show how it relates to anisotropic diffusion. We then accelerate bilateral filtering by using a piecewise-linear approximation in the intensity domain and appropriate subsampling. This results in a speed-up of two orders of magnitude. The method is fast and requires no parameter setting.

CR Categories: I.3.3 [Computer Graphics]: Picture/image generation—Display algorithms; I.4.1 [Image Processing and Computer Vision]: Enhancement—Digitization and image capture

Keywords: image processing, tone mapping, contrast reduction, edge-preserving filtering, weird maths

1 Introduction

As the availability of high-dynamic-range images grows due to advances in lighting simulation, e.g. [Ward 1994], multiple-exposure photography [Debevec and Malik 1997; Madden 1993] and new sensor technologies [Mitsunaga and Nayar 2000; Schechner and Nayar 2001; Yang et al. 1999], there is a growing demand to be able to display these images on low-dynamic-range media. Our visual system can cope with such high-contrast scenes because most of the adaptation mechanisms are *local* on the retina.

There is a tremendous need for contrast reduction in applications such as image-processing, medical imaging, realistic rendering, and digital photography. Consider photography for example. A major aspect of the art and craft concerns the management of contrast via e.g. exposure, lighting, printing, or local dodging and burning [Adams 1995; Rudman 2001]. In fact, poor management of light – under- or over-exposed areas, light behind the main character, etc. – is the single most-commonly-cited reason for rejecting



Figure 1: High-dynamic-range photography. No single global exposure can preserve both the colors of the sky and the details of the landscape, as shown on the rightmost images. In contrast, our spatially-varying display operator (large image) can bring out all details of the scene. Total clock time for this 700x480 image is 1.4 seconds on a 700Mhz PentiumIII. Radiance map courtesy of Paul Debevec, USC. [Debevec and Malik 1997]



Figure 2: Principle of our two-scale decomposition of the input intensity. Color is treated separately using simple ratios. Only the base scale has its contrast reduced.

photographs. This is why camera manufacturers have developed sophisticated exposure-metering systems. Unfortunately, exposure only operates via *global* contrast management – that is, it recenters the intensity window on the most relevant range. If the range of intensity is too large, the photo will contain under- and over-exposed areas (Fig. 1, rightmost part).

Our work is motivated by the idea that the use of high-dynamic-range cameras and relevant display operators can address these issues. Digital photography has inherited many of the strengths of film photography. However it also has the potential to overcome its limitations. Ideally, the photography process should be decomposed into a measurement phase (with a high-dynamic-range output), and a post-process phase that, among other things, manages the contrast. This post-process could be automatic or user-controlled, as part of the camera or on a computer, but it should take advantage of the wide range of available intensity to perform appropriate contrast reduction.

In this paper, we introduce a fast and robust operator that takes a high-dynamic-range image as input, and compresses the contrast while preserving the details of the original image, as introduced by Tumblin [1999]. Our operator is based on a two-scale decomposition of the image into a *base* layer (large-scale features) and a *detail*

layer (Fig. 2). Only the base layer has its contrast reduced, thereby preserving the detail. In order to perform a fast decomposition into these two layers, and to avoid halo artifacts, we present a fast and robust edge-preserving filter.

1.1 Overview

The primary focus of this paper is the development of a fast and robust edge-preserving filter – that is, a filter that blurs the small variations of a signal (noise or texture detail) but preserves the large discontinuities (edges). Our application is unusual however, in that the noise (detail) is the important information in the signal and must therefore be preserved.

We build on *bilateral filtering*, a non-linear filter introduced by Tomasi et al. [1998]. It derives from Gaussian blur, but it prevents blurring across edges by decreasing the weight of pixels when the intensity difference is too large. As it is a fast alternative to the use of anisotropic diffusion, which has proven to be a valuable tool in a variety of areas of computer graphics, e.g. [McCool 1999; Desbrun et al. 2000], the potential applications of this technique extend beyond the scope of contrast reduction.

This paper makes the following contributions:

Bilateral filtering and robust statistics: We recast bilateral filtering in the framework of robust statistics, which is concerned with estimators that are insensitive to outliers. Bilateral filtering is an estimator that considers values across edges to be outliers. This allows us to provide a wide theoretical context for bilateral filtering, and to relate it to anisotropic diffusion.

Fast bilateral filtering: We present two acceleration techniques: we linearize bilateral filtering, which allows us to use FFT and fast convolution, and we downsample the key operations.

Uncertainty: We compute the uncertainty of the output of the filter, which permits the correction of doubtful values.

Contrast reduction: We use bilateral filtering for the display of high-dynamic-range images. The method is fast, stable, and requires no setting of parameters.

2 Review of local tone mapping

Tone mapping operators can be classified into *global* and *local* techniques [Tumblin 1999; Ferwerda 1998; DiCarlo and Wandell 2000]. Because they use the same mapping function for all pixels, most *global* techniques do not directly address contrast reduction. A limited solution is proposed by Schlick [1994] and Tumblin et al. [1999], who use S-shaped functions inspired from photography, thus preserving some details in the highlights and shadows. Unfortunately, contrast is severely reduced in these areas. Some authors propose to interactively vary the mapping according to the region of interest attended by the user [Tumblin et al. 1999], potentially using graphics hardware [Cohen et al. 2001].

A notable exception is the global histogram adjustment by Ward-Larson et al. [1997]. They disregard the empty portions of the histogram, which results in efficient contrast reduction. However, the limitations due to the global nature of the technique become obvious when the input exhibits a uniform histogram (see e.g. the example by DiCarlo and Wandell [2000]).

In contrast, *local* operators use a mapping that varies spatially depending on the neighborhood of a pixel. This exploits the fact that human vision is sensitive mainly to local contrast.

Most local tone-mapping techniques use a decomposition of the image into different layers or scales (with the exception of Socolinsky, who uses a variational technique [2000]). The contrast is reduced differently for each scale, and the final image is a recomposition of the various scales after contrast reduction. The major pitfall of local methods is the presence of haloing artifacts. When dealing with high-dynamic-range images, haloing issues become

even more critical. In 8-bit images, the contrast at the edges is limited to roughly two orders of magnitude, which directly limits the strength of halos.

Chiu et al. vary a gain according to a low-pass version of the image [1993], which results in pronounced halos. Schlick had similar problems when he tried to vary his mapping spatially [1994]. Jobson et al. reduce halos by applying a similar technique at multiple scales [1997]. Pattanaik et al. use a multiscale decomposition of the image according to comprehensive psychophysically-derived filter banks [1998]. To date, this method seems to be the most faithful to human vision, however, it may still present halos.

DiCarlo et al. propose to use robust statistical estimators to improve current techniques [2000], although they do not provide a detailed description. Our method follows in the same spirit and focuses on the development of a fast and practical method.

Tumblin et al. [1999] propose an operator for synthetic images that takes advantage of the ability of the human visual system to decompose a scene into intrinsic “layers”, such as reflectance and illumination [Barrow and Tenenbaum 1978]. Because vision is sensitive mainly to the reflectance layers, they reduce contrast only in the illumination layer. This technique is unfortunately applicable only when the characteristics of the 3D scene are known. As we will see, our work can be seen as an extension to photographs. Our two-scale decomposition is very related to the texture-illumination decoupling technique by Oh et al. [2001].

Recently, Tumblin and Turk built on anisotropic diffusion to decompose an image using a new low-curvature image simplifier (LCIS) [Tumblin 1999; Tumblin and Turk 1999]. Their method can extract exquisite details from high-contrast images. Unfortunately, the solution of their partial differential equation is a slow iterative process. Moreover, the coefficients of their diffusion equation must be adapted to each image, which makes this method more difficult to use, and the extension to animated sequences unclear. We build upon a different edge-preserving filter that is easier to control and more amenable to acceleration. We will also deal with two problems mentioned by Tumblin et al.: the small remaining halos localized around the edges, and the need for a “leakage fixer” to completely stop diffusion at discontinuities.

3 Edge-preserving filtering

In this section, we review important edge-preserving-smoothing techniques, e.g. [Saint-Marc et al. 1991].

3.1 Anisotropic diffusion

Anisotropic diffusion [Perona and Malik 1990] is inspired by an interpretation of Gaussian blur as a heat conduction partial differential equation (PDE): $\frac{\partial I}{\partial t} = -\Delta I$. That is, the intensity I of each pixel is seen as heat and is propagated over time to its 4 neighbors according to the heat spatial variation.

Perona and Malik introduced an *edge-stopping* function g that varies the *conductance* according to the image gradient. This prevents heat flow across edges:

$$\frac{\partial I}{\partial t} = \text{div}[g(|\nabla I|) \nabla I]. \quad (1)$$

They propose two expressions for the edge-stopping function $g(x)$:

$$g_1(x) = \frac{1}{1 + \frac{x^2}{\sigma^2}} \quad \text{and} \quad g_2(x) = e^{-(x^2/\sigma^2)}, \quad (2)$$

where σ is a scale parameter in the intensity domain that specifies what gradient intensity should stop diffusion.

The discrete Perona-Malik diffusion equation governing the value I_s at pixel s is then

$$I_s^{t+1} = I_s^t + \frac{\lambda}{4} \sum_{p \in \text{neighb}_4(s)} g(I_p^t - I_s^t) (I_p^t - I_s^t), \quad (3)$$

where t describes discrete time steps, and $\text{neighb}_4(s)$ is the 4-neighborhood of pixel s . λ is a scalar that determines the rate of diffusion.

Although anisotropic diffusion is a popular tool for edge-preserving filtering, its discrete diffusion nature makes it a slow process. Moreover, the results depend on the stopping time, since the diffusion converges to a uniform image.

3.2 Robust anisotropic diffusion

Black et al. [1998] recast anisotropic diffusion in the framework of robust statistics. Our analysis of bilateral filtering is inspired by their work. The field of robust statistics develops estimators that are robust to outliers or deviation to the theoretical distribution [Huber 1981; Hampel et al. 1986].

Black et al. [1998] show that anisotropic diffusion can be seen as the estimate of a value I_s at each pixel s that is an estimate of its 4-neighbors, which minimizes an energy over the whole image:

$$\min \sum_{s \in \Omega} \sum_{p \in \text{neighb}_4(s)} \rho(I_p - I_s), \quad (4)$$

where Ω is the whole image, and ρ is an error norm (e.g. quadratic). Eq. 4 can be solved by gradient descent for each pixel:

$$I_s^{t+1} = I_s^t + \frac{\lambda}{4} \sum_{p \in \text{neighb}_4(s)} \psi(I_p - I_s), \quad (5)$$

where ψ is the derivative of ρ , and t is a discrete time variable. ψ is proportional to the so-called *influence function* that characterizes the influence of a sample on the estimate.

For example, a least-square estimate is obtained by using $\rho(x) = x^2$, and the corresponding influence function is linear, thus resulting in the mean estimator (Fig. 4, left). As a result, values far from the mean have a considerable influence on the estimate. In contrast, an influence function such as the *Lorentzian* error norm, given in Fig. 3 and plotted in Fig. 4, gives much less weight to outliers and is therefore more robust. In the plot of ψ , we see that the influence function is *redescending* [Black et al. 1998; Huber 1981]¹. Robust norms and influence functions depend on a parameter σ that provides the notion of scale in the intensity domain, and controls where the influence function becomes redescending, and thus which values are considered outliers.

Black et al. note that Eq. 5 is similar to Eq. 3 governing anisotropic diffusion, and that by defining $g(x) = \psi(x)/x$, anisotropic diffusion is reduced to a robust estimator. They also show that the g_1 function proposed by Perona et al. is equivalent to the Lorentzian error norm plotted in Fig. 4 and given in Fig. 3.

This analogy allows them to discuss desirable properties of edge-stopping functions. In particular, they show that *Tukey's biweight function* (Fig. 3) yields more robust results, because it completely stops diffusion across edges: The influence of outliers is null, as shown in Fig. 5, as opposed to the Lorentzian error norm that slowly goes to zero towards infinity. This also solves the termination problem, since diffusion then converges to a *piecewise-uniform* image.

¹Some authors reserve the term redescending for function that vanish after a certain value [Hampel et al. 1986].

Huber	Lorentz
$g_\sigma(x) = \begin{cases} \frac{1}{\sigma} & x \leq \sigma \\ \frac{1}{ x } & \text{otherwise} \end{cases}$	$g_\sigma(x) = \frac{2}{2 + \frac{x^2}{\sigma^2}}$
σ	$\sigma/\sqrt{2}$
Tukey	Gauss
$g_\sigma(x) = \begin{cases} \frac{1}{2} [1 - (x/\sigma)^2]^2 & x \leq \sigma \\ 0 & \text{otherwise} \end{cases}$	$g_\sigma(x) = e^{-\frac{x^2}{2\sigma^2}}$
$\sigma * \sqrt{5}$	σ

Figure 3: Robust edge-stopping functions. Note that ψ can be found by multiplying g by x , and ρ by integration of ψ . The value of σ has to be modified accordingly to use a consistent scale across estimators, as indicated below the Lorentz and Tukey functions.

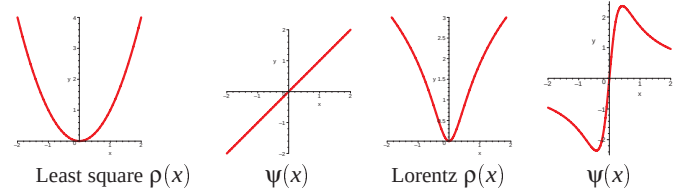


Figure 4: Least-square vs. Lorentzian error norm (after [Black et al. 1998]).

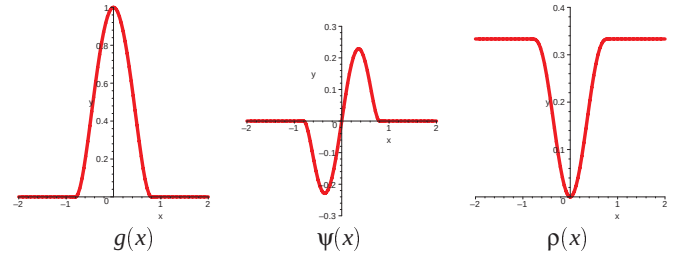


Figure 5: Tukey's biweight (after [Black et al. 1998]).

3.3 Bilateral filtering

Bilateral filtering was developed by Tomasi and Manduchi as an alternative to anisotropic diffusion [1998]. It is a non-linear filter where the output is a weighted average of the input. They start with standard Gaussian filtering with a spatial kernel f (Fig. 6). However, the weight of a pixel depends also on a function g in the intensity domain, which decreases the weight of pixels with large intensity differences. We note that g is an edge-stopping function similar to that of Perona et al. [1990]. The output of the bilateral filter for a pixel s is then:

$$J_s = \frac{1}{k(s)} \sum_{p \in \Omega} f(p-s) g(I_p - I_s) I_p, \quad (6)$$

where $k(s)$ is a normalization term:

$$k(s) = \sum_{p \in \Omega} f(p-s) g(I_p - I_s). \quad (7)$$

In practice, they use a Gaussian for f in the spatial domain, and a Gaussian for g in the intensity domain. Therefore, the value at a pixel s is influenced mainly by pixel that are close spatially and that have a similar intensity (Fig. 6). This is easy to extend to color images, and any metric g on pixels can be used (e.g. CIE-LAB).

Barash proposes a link between anisotropic diffusion and bilateral filtering [2001]. He uses an extended definition of intensity that includes spatial coordinates. This permits the extension of bilateral filtering to perform feature enhancement. Unfortunately,

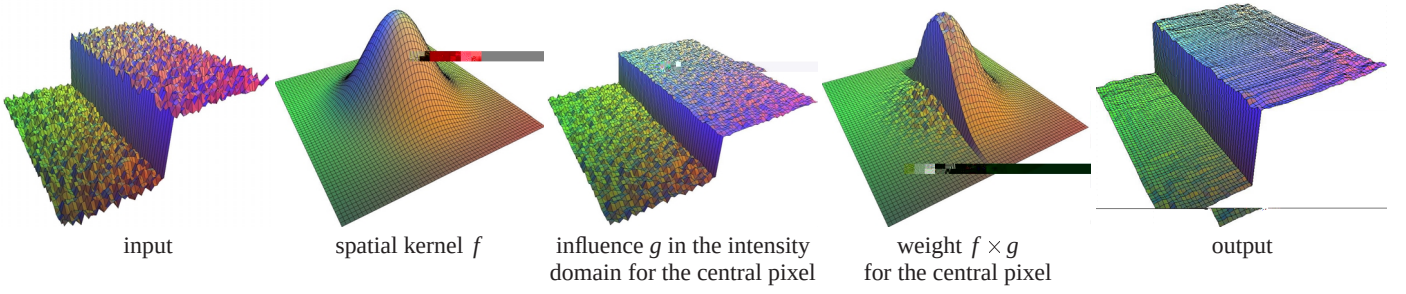


Figure 6: Bilateral filtering. Colors are used only to convey shape.

the extended definition of intensity is not quite natural. Elad also discusses the relation between bilateral filtering, anisotropic diffusion, and robust statistics, but he address the question from a linear-algebra point of view [to appear]. In this paper, we propose a different unified viewpoint based on robust statistics that extends the work by Black et al. [1998].

4 Edge-preserving smoothing as robust statistical estimation

In their paper, Tomasi et al. only outlined the principle of bilateral filters, and they then focused on the results obtained using two Gaussians. In this section, we provide a principled study of the properties of this family of filters. In particular, we show that bilateral filtering is a robust statistical estimator, which allows us to put empirical results into a wider theoretical context.

4.1 A unified viewpoint on bilateral filtering and 0-order anisotropic diffusion

In order to establish a link to bilateral filtering, we present a different interpretation of discrete anisotropic filtering. In Eq. 3, $I_p^t - I_s^t$ is used as the derivative of I^t in one direction. However, this can also be seen simply as the 0-order difference between the two pixel intensities. The edge-stopping function can thus be seen as preventing diffusion between pixels with large intensity differences. The two formulations are equivalent from a practical standpoint, but Black et al.'s variational interpretation [1998] is more faithful to Perona and Malik's diffusion analogy, while our 0-order interpretation is more natural in terms of robust statistics.

In particular, we can extend the 0-order anisotropic diffusion to a larger spatial support:

$$I_s^{t+1} = I_s^t + \lambda \sum_{p \in \Omega} f(p-s) g(I_p^t - I_s^t) (I_p^t - I_s^t), \quad (8)$$

where f is a spatial weighting function (typically a Gaussian), Ω is the whole image, and t is still a discrete time variable. The anisotropic diffusion of Perona et al., which we now call *local diffusion*, corresponds to an f that is zero except at the 4 neighbors. Eq. 8 defines a robust statistical estimator of the class of *M-estimators* (generalized maximum likelihood estimator) [Hampel et al. 1986; Huber 1981].

In the case where the conductance g is uniform (isotropic filtering) and where f is a Gaussian, Eq. 8 performs a Gaussian blur for each iteration, which is equivalent to several iterations of the heat-flow simulation. It can thus be seen as a way to trade the number of iterations for a larger spatial support. However, in the case of anisotropic diffusion, it has the additional property of propagating heat across ridges. Indeed, if the image is white with a black line in the middle, local anisotropic diffusion does not propagate energy

between the two connected components, while extended diffusion does. Depending on the application, this property will be either beneficial or deleterious. In the case of tone mapping, for example, the notion of connectedness is not important, as only spatial neighborhoods matter.

We now come to the robust statistical interpretation of bilateral filtering. Eq. 6 defines an estimator based on a weighted average of the data. It is therefore a *W-estimator* [Hampel et al. 1986]. The iterative formulation is an instance of *iteratively reweighted least squares*. This taxonomy is extremely important because it was shown that M-estimators and W-estimators are essentially equivalent and solve the same energy minimization problem [Hampel et al. 1986], p. 116:

$$\min \sum_{s \in \Omega} \sum_{p \in \Omega} \rho(I_s - I_p) \quad (9)$$

or for each pixel s :

$$\sum_{p \in \Omega} \psi(I_s - I_p) = 0, \quad (10)$$

where ψ is the derivative of ρ . As shown by Black et al. [1998] for anisotropic diffusion, and as is true also for bilateral filtering, it suffices to define $\psi(x) = g(x) * x$ to find the original formulations. In fact the second edge-stopping function g_2 in Eq. 2 defined by Perona et al. [1990] corresponds to the Gaussian influence function used for bilateral filtering [Tomasi and Manduchi 1998]. As a consequence of this unified viewpoint, all the studies on edge-stopping functions for anisotropic diffusion can be applied to bilateral filtering.

Eqs. 9 and 10 are not strictly equivalent because of local minima of the energy. Depending on the application, this can be desirable or undesirable. In the former case, the use of a very robust estimator, such as the median, to initialize an iterative process is recommended. In the case of tone mapping or texture-illumination decoupling, however, we want to find the local minimum closest to the initial pixel value.

It was noted by Tomasi et al. [1998] that bilateral filtering usually requires only one iteration. Hence it belongs to the class of *one-step W-estimators*, or *w-estimators*, which have been shown to be particularly efficient. The existence of local minima is however a very important issue, and the use of an initial median estimator is highly recommended. In contrast, Oh. et al. use a simple Gaussian blur [2001], which deserves further study.

Now that we have shown that 0-order anisotropic diffusion and bilateral filtering belong to the same family of estimators, we can compare them. They both respect *causality*: No maximum or minimum can be created, only removed. However, anisotropic diffusion is *adiabatic* (energy-preserving), while bilateral filtering is not. To see this, consider the energy exchange between two pixels p and s . In the diffusion case, the energy $\lambda f(p-s)g(I_p^t - I_s^t)(I_p^t - I_s^t)$ flowing from p to s is the opposite of the energy from s to p because the expression is symmetric (provided that g and f are symmetric). In contrast, in bilateral filtering, the normalization factor $1/k$

is different for the two pixels, resulting in an asymmetric energy flow. Energy preservation can be crucial for some applications, e.g. [Rushmeier and Ward 1994], but it is not for tone mapping or reflectance extraction.

In contrast to anisotropic diffusion, bilateral filtering does not rely on shock formation, so it is not prone to stairstepping artifacts. The output of bilateral filtering on a gradient input is smooth. This point is mostly due to the non-iterative nature of the filter and deserves further exploration.

4.2 Robust estimators

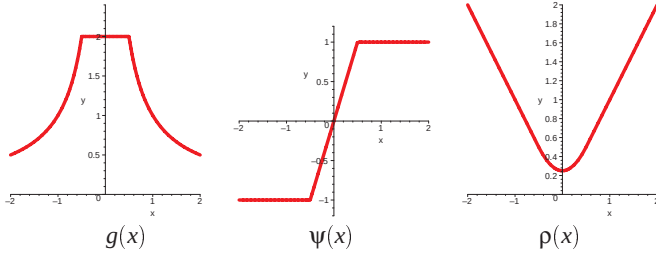


Figure 7: Huber's minimax (after [Black et al. 1998]).

Fig. 8 plots a variety of robust influence functions, and their Formulas are given in Fig. 3. When the influence function is monotonic, there is no local minimum problem, and estimators always converge to a global maximum. Most robust estimators have a shape as shown on the left: The function increases, then decreases, and potentially goes to zero if it has a finite rejection point.

These plots can be very helpful in understanding how an estimator deals with outliers. For example, we can see that the Huber minimax gives constant influence to outliers, and that the Lorentz estimator gives them more importance than, say, the Gaussian estimator. The Tukey biweight is the only purely *redescending* function we show. Outliers are thus completely ignored.

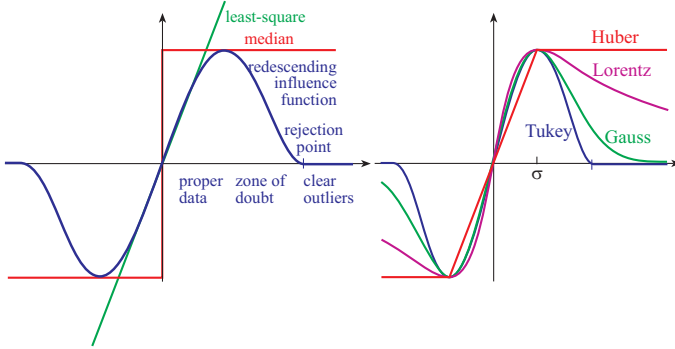


Figure 8: Comparison of influence functions.

We anticipate the results of our technique and show in Fig. 9 the output of a robust bilateral filter using these different ψ functions (or their g equivalent in Eq. 6). We can see that larger influences of outliers result in estimates that are more blurred and further from the input pixels. In what follows, we use the Gaussian or Tukey influence function, because they are more robust to outliers and better preserve edges.

5 Efficient Bilateral Filtering

Now that we have provided a theoretical framework for bilateral filtering, we will next deal with its speed. A direct implementation of

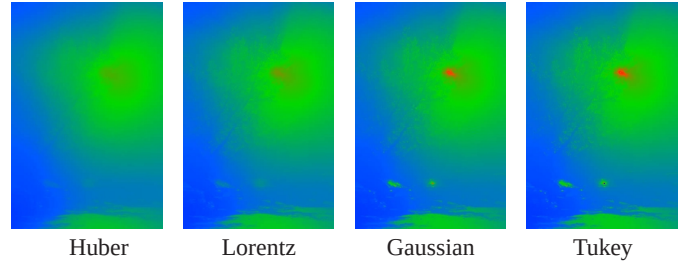


Figure 9: Comparison of the 4 estimators for the log of intensity of the foggy scene of Fig 15. The false-colored output is normalized to the log of the min and max of the input.

bilateral filtering might require $O(n^2)$ time, where n is the number of pixels in the image. In this section, we dramatically accelerate bilateral filtering using two strategies: a piecewise-linear approximation in the intensity domain, and a sub-sampling in the spatial domain. We then present a technique that detects and fixes pixels where the bilateral filter cannot obtain a good estimate due to lack of data.

5.1 Piecewise-linear bilateral filtering

A convolution such as Gaussian filtering can be greatly accelerated using Fast Fourier Transform. A $O(n^2)$ convolution in the primal becomes a $O(n)$ multiplication in the frequency domain. Since the discrete FFT and its inverse have cost $O(n \log n)$, there is a gain of one order of magnitude.

Unfortunately, this strategy cannot be applied directly to bilateral filtering, because it is not a convolution: The filter is signal-dependent because of the edge-stopping function $g(I_p - I_s)$. However consider Eq. 6 for a fixed pixel s . It is equivalent to the convolution of the function $H^{I_s} : p \rightarrow g(I_p - I_s)I_p$ by the kernel f . Similarly, the normalization factor k is the convolution of $G^{I_s} : p \rightarrow g(I_p - I_s)$ by f . That is, the only dependency on pixel s is the value I_s in g .

Our acceleration strategy is thus as follows: We discretize the set of possible signal intensities into NB_SEGMENT values $\{i^j\}$, and compute a linear filter for each such value:

$$\begin{aligned} J_s^j &= \frac{1}{k^j(s)} \sum_{p \in \Omega} f(p-s) g(I_p - i^j) I_p \\ &= \frac{1}{k^j(s)} \sum_{p \in \Omega} f(p-s) H_p^j \end{aligned} \quad (11)$$

and

$$\begin{aligned} k^j(s) &= \sum_{p \in \Omega} f(p-s) g(I_p - i^j) \\ &= \sum_{p \in \Omega} f(p-s) G^j(p). \end{aligned} \quad (12)$$

The final output of the filter for a pixel s is then a linear interpolation between the output J_s^j of the two closest values i^j of I_s . This corresponds to a piecewise-linear approximation of the original bilateral filter (note however that it is a linearization of the whole functional, not of the influence function). The pseudocode is given in Fig. 10.

Fig. 11 shows the speed-up we obtain depending on the size of the spatial kernel. Quickly, the piecewise-linear version outperforms the brute-force implementation, due to the use of FFT convolution. The formal analysis of error remains to be performed, but no artifact was noticeable for segments up to the size of the scale σ_r .

This could be further accelerated when the distribution of intensities is not uniform spatially. We can subdivide the image into sub-images, and if the difference between the max and min of the

```

PiecewiseBilateral
  (Image I, spatial kernel  $f_{\sigma_s}$ , intensity influence  $g_{\sigma_r}$ )
  J=0 /* set the output to zero */
  for j=0..NB_SEGMENTS
     $i^j = \min I + j \times (\max(I) - \min(I)) / \text{NB\_SEGMENTS}$ 
     $G^j = g_{\sigma_r}(I - i^j)$  /* evaluate  $g_{\sigma_r}$  at each pixel */
     $K^j = G^j \otimes f_{\sigma_s}$  /* normalization factor */
     $H^j = G^j \times I$  /* compute H for each pixel */
     $H^{*j} = H^j \otimes f_{\sigma_s}$ 
     $J^j = H^{*j} / K^j$  /* normalize */
  J = J + J^j  $\times$  InterpolationWeight(I,  $i^j$ )

```

Figure 10: Pseudo code of the piecewise-linear acceleration of bilateral filtering. Operations with upper cases such as $G^j = g_{\sigma_r}(I, i^j)$ denote computation on all pixels of the image. $\otimes$ denotes the convolution, while $\times$ is simply the per-pixel multiplication. InterpolationWeight is the “hat” interpolation weight for linear interpolation. In practice, we use $\text{NB_SEGMENT} = (\max(I) - \min(I)) / \sigma_r$.

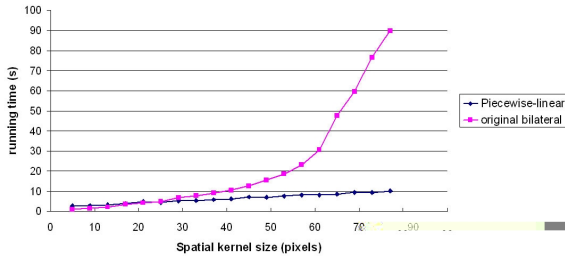


Figure 11: Speed-up of the piecewise-linear acceleration for 17 segments and a 576x768 image.

```

FastBilateral
  (Image I, spatial kernel  $f_{\sigma_s}$ , intensity influence  $g_{\sigma_r}$ ,
   downsampling factor z)
  J=0 /* set the full-scale output to zero */
   $I' = \text{downsample}(I, z)$ 
   $f'_{\sigma_s/z} = \text{downsample}(f_{\sigma_s}, z)$ 
  for j=0..NB_SEGMENTS
     $i^j = \min I + j \times (\max(I) - \min(I)) / \text{NB\_SEGMENTS}$ 
     $G^j = g_{\sigma_r}(I' - i^j)$  /* evaluate  $g_{\sigma_r}$  at each pixel */
     $K^j = G^j \otimes f'_{\sigma_s/z}$  /* normalization factor */
     $H^j = G^j \times I'$  /* compute H for each pixel */
     $H^{*j} = H^j \otimes f'_{\sigma_s/z}$ 
     $J^j = H^{*j} / K^j$  /* normalize */
   $J^j = \text{upsample}(J^j, z)$ 
  J = J + J^j  $\times$  InterpolationWeight(I,  $i^j$ )

```

Figure 12: Pseudo code of the downsampled piecewise-linear acceleration of bilateral filtering. Parts at the full resolution are in green, while downsampled operations are in blue, and downsampled images are denoted with a prime.

intensity is more reduced in the sub-images than in the whole image, fewer segments can be used. This solution has however not been implemented yet.

5.2 Subsampling

To further accelerate bilateral filtering, we note that all operations in Fig. 10 except the final interpolation aim at low-pass filtering. We can thus safely use a downsampled version of the image with little quality loss. However, the final interpolation must be performed using the full-scale image, otherwise edges would not be respected, resulting in visible artifacts. Fig. 12 shows the new algorithm.

We use nearest-neighbor downsampling, because it does not modify the histogram. The acceleration we obtain is plotted in Fig. 13 for an example. While a formal study of error/acceleration remains to be done, we did not notice any visible artifact up to downsampling factor of 10 to 25. At this resolution, the cost of the upsampling and linear interpolation outweighs the filtering operations, and no further acceleration is gained by more aggressive downsampling.

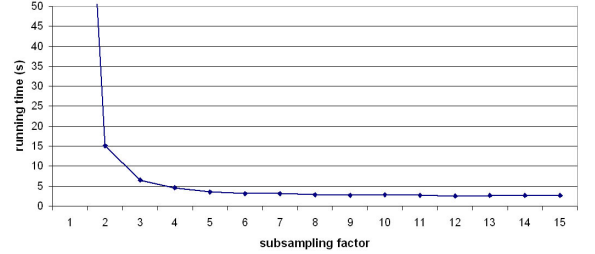


Figure 13: Speed-up due to downsampling for 17 segments and a 576x768 image. The value for the full-scale filtering is 173 sec.

5.3 Uncertainty

As noted by Tumblin et al. [Tumblin 1999; Tumblin and Turk 1999], edge-preserving contrast reduction can still encounter small halo artifacts for antialiased edges or due to flare around high-contrast edges. We noticed similar problems on some synthetic as well as real images. We propose an explanation in terms of signal/noise ratio. These small halos correspond to pixels where there is not enough information in the neighborhood to decouple the large-scale and the small-scale features. Indeed, the values at the edges span the whole range between the upper and the lower values, and there are very few pixels in the zone of proper data of the influence function. We thus compute a statistical estimator with very little data, and the variance is quite high.

Fortunately, bilateral filtering provides a direct measure of this uncertainty: The normalization factor k in Eq. 6 is the sum of the influence of each pixel. We can therefore use it to detect dubious pixels that need to be fixed. In practice, we use the log of this value because it better extracts uncertain pixels.

The fixing strategy we use is then simple. We compute a low-pass version $\tilde{J}$ of the output J of the bilateral filter, using a small Gaussian kernel (2 pixels in practice), and we assign to a pixel the value of a linear interpolation between J and $\tilde{J}$ depending on the log of the uncertainty k .

6 Contrast reduction

We now describe how bilateral filtering can be used for contrast reduction. We note that our method is not strictly a tone reproduction operator, in the sense of Tumblin and Rushmeier’s [1993], since it does not attempt to imitate human vision.

Building on previous approaches, our contrast reduction is based on a multiscale decomposition e.g. [Jobson et al. 1997; Pattanaik et al. 1998; Tumblin and Turk 1999]. However, we only use a two-scale decomposition, where the “base” image is computed using bilateral filtering, and the detail layer is the division of the input intensity by the base layer. Fig. 2 illustrates the general approach. The base layer has its contrast reduced, while the magnitude of the detail layer is unchanged, thus preserving detail.

Following Tumblin et al. [Tumblin 1999; Tumblin and Turk 1999], we compress the range of the base layer using a scale factor in the log domain. We compute this scale factor such that the whole

range of the base layer is compressed to a user-controllable *base contrast*. In practice, a base contrast of 5 worked well for all our examples, but in some situations where lights sources are visible, one might want to vary this setting.

Our treatment of color is simple. We perform contrast reduction on the intensity of pixels and recombine color after contrast reduction [Schlick 1994; Tumblin 1999; Tumblin and Turk 1999]. We perform our calculations on the logs of pixel intensities, because pixel differences then correspond directly to contrast, and because it yields a more uniform treatment of the whole range.

Our approach is faithful to the original idea by Chiu et al. [1993], albeit using a robust filter instead of their low-pass filter. It can also be viewed as the decomposition of the image into intrinsic layers of reflectance and illuminance [Oh et al. 2001], followed by an appropriate contrast reduction of the illuminance (or base) layer [Tumblin et al. 1999].

For the filtering phase, we experimented with the various influence functions discussed in Section 4.2. As expected, the Huber minimax estimator decreases the strength of halos compared to standard Gaussian blur, but does not eliminate them. Moreover, the results vary with the size of the spatial kernel. The Lorentz function performed better, but only the Gaussian and Tukey’s biweight were able to accurately decompose the image. With both functions, the scale σ_s of the spatial kernel had little influence on the result. This is important since it allows us to keep σ_s constant to a value of 2% of the image size.

The value $\sigma_r = 0.4$ performed consistently well for all our experiments. Again, this property is quite important because the user does not have to set a complex parameter. The significance of this value might come from two complementary origins, which are still areas of future research. First, it might be due to characteristics of the local sensitivity of the human visual system. Perhaps beyond this value, we notice no difference. Second, it might be related to the physical range of possible reflectance values, between a perfect reflector and a black material.

As a conclusion, the only user-controlled parameters of our method are the overall brightness and the base contrast. While the automatic values perform very well, we found it useful to provide these intuitive degrees of freedom to allow the user a control over the “look” of the image. The base contrast provides a very intuitive alternative to the contrast/brightness setting of image-editing software. It controls the overall appearance of the image, while still preserving the fine details.

6.1 Implementation and results

We have implemented our technique using a floating point representation of images, and the Intel image processing library for the convolutions. We have tested it on a variety of synthetic and real images, as shown in the color plates. All the examples reproduced in the paper use the Gaussian influence function, but the results with Tukey’s biweight are not different. The technique is extremely fast, as can be seen in Fig. 14. We have tested it on an upsampled 10Mpixel image with contrast of more than 1:100,000, and the computation took only 6s on a 2GHz Pentium 4. In particular, due to our acceleration techniques, the running time grows sub-linearly. This is a dramatic speed-up compared to previous methods.

Our technique can address some of the most challenging photographic situations, such as interior lighting or sunset photos, and produces very compelling images. In our experiments, Tumblin and Turk’s operator [1999] appears to better preserve fine details, while our technique better preserves the overall photorealistic appearance (Figs. 21 and 22).

Image	resolution	# segments	z	timing (s)
Grove D	710 * 480	15	4	0.33
Memorial	512 * 768	11	4	0.31
Hotel room	750 * 487	13	4	0.31
Vine	710 * 480	10	4	0.23
Fog	1130 * 751	12	8	0.45
Grove C	709 * 480	14	4	0.30
Window	2K*1.3K	10	16	2.73
Interior	2K*1.3K	19	16	2.19
Interior*2	2.6K * 4K	19	24	6.03

Figure 14: Results of our new technique. Timings on a 2GHz P4.

7 Discussion

This paper opens several avenues of future research related to edge-preserving filtering and contrast reduction. The unified viewpoint on bilateral filtering and anisotropic diffusion offers some interesting possibilities. The robust statistical framework we have introduced suggests the application of bilateral filtering to a variety of graphics areas where energy preservation is not a major concern.

The treatment of uncertainty deserves more attention. The correction scheme based on a Gaussian blur by a small kernel works well in the cases we have tested, but a more formal analysis is needed. Other approaches might involve the use of a different range scale σ_r .

In terms of contrast reduction, future work includes the development of a more principled fixing method for uncertain values, and the use of a more elaborate compression function for the base layer, e.g. [Tumblin et al. 1999; Larson et al. 1997]. White balance is an important issue for indoor scenes that also exhibit outdoor portions, as can be seen in Fig. 23. A strategy similar to Pattanaik et al.’s operator [Pattanaik et al. 1998] should be developed. The inclusion of perceptual aspects is a logical step. The main difficulty stems from the complex interaction between local adaptation and gaze movements. The extension to animated sequences is an exciting topic. Initial experiments are very encouraging.

Finally, contrast reduction is only one example of *pictorial techniques* to cope with the *limitations of the medium* [Durand 2002]. We believe that these techniques are crucial aspects of the digital photography and video revolution, and will facilitate the creation of effective and compelling pictures.

Acknowledgments

We would like to thank Byong Mok Oh for his help with the radiance maps and the bibliography; he and Ray Jones also provided crucial proofreading. Thanks to Paul Debevec and Jack Tumblin for allowing us to use their radiance maps. Thanks to the reviewers for their careful comments. This research was supported by NSF grants CCR-0072690 and EIA-9802220, and by a gift from Pixar Animation Studios.

References

- ADAMS, A. 1995. *The Camera+The Negative+The Print*. Little Brown and Co.
- BARASH, D. 2001. A fundamental relationship between bilateral filtering, adaptive smoothing and the nonlinear diffusion equation. *IEEE PAMI*. in press.
- BARROW, H., AND TENENBAUM, J. 1978. Recovering intrinsic scene characteristics from images. In *Computer Vision Systems*. Academic Press, New York, 3–26.
- BLACK, M., SAPIRO, G., MARIMONT, D., AND HEEGER, D. 1998. Robust anisotropic diffusion. *IEEE Trans. Image Processing* 7, 3 (Mar.), 421–432.
- CHIU, K., HERE, M., SHIRLEY, P., SWAMY, S., WANG, C., AND ZIMMERMAN, K. 1993. Spatially nonuniform scaling functions for high contrast images. In *Proc. Graphics Interface*, 245–253.

- COHEN, J., TCHOU, C., HAWKINS, T., AND DEBEVEC, P. 2001. Real-time high-dynamic range texture mapping. In *Rendering Techniques 2001: 12th Eurographics Workshop on Rendering*, Eurographics, 313–320.
- DEBEVEC, P. E., AND MALIK, J. 1997. Recovering high dynamic range radiance maps from photographs. In *Proceedings of SIGGRAPH 97*, ACM SIGGRAPH / Addison Wesley, Los Angeles, California, Computer Graphics Proceedings, Annual Conference Series, 369–378.
- DESBRUN, M., MEYER, M., SCHRÖDER, P., AND BARR, A. H. 2000. Anisotropic feature-preserving denoising of height fields and bivariate data. In *Graphics Interface*, 145–152.
- DICARLO, J., AND WANDELL, B. 2000. Rendering high dynamic range images. *Proceedings of the SPIE: Image Sensors* 3965, 392–401.
- DURAND, F. 2002. An invitation to discuss computer depiction. In *Proc. NPAR'02*.
- ELAD, M. to appear. On the bilateral filter and ways to improve it. *IEEE Trans. on Image Processing*.
- FERWERDA, J. 1998. Fundamentals of spatial vision. In *Applications of visual perception in computer graphics*. Siggraph '98 Course Notes.
- HAMPEL, F. R., RONCHETTI, E. M., ROUSSEUW, P. J., AND STAHEL, W. A. 1986. *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.
- HUBER, P. J. 1981. *Robust Statistics*. John Wiley and Sons, New York.
- JOBSON, RAHMAN, AND WOODSELL. 1997. A multi-scale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. on Image Processing: Special Issue on Color Processing* 6 (July), 965–976.
- LARSON, G. W., RUSHMEIER, H., AND PIATKO, C. 1997. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics* 3, 4 (October - December), 291–306.
- MADDEN, B. 1993. Extended intensity range imaging. Tech. rep., U. of Pennsylvania, GRASP Laboratory.
- MCCOOL, M. 1999. Anisotropic diffusion for monte carlo noise reduction. *ACM Trans. on Graphics* 18, 2, 171–194.
- MITSUNAGA, T., AND NAYAR, S. K. 2000. High dynamic range imaging: Spatially varying pixel exposures. In *IEEE CVPR*, 472–479.
- OH, B. M., CHEN, M., DORSEY, J., AND DURAND, F. 2001. Image-based modeling and photo editing. In *Proceedings of ACM SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, 433–442.
- PATTANAIK, S. N., FERWERDA, J. A., FAIRCHILD, M. D., AND GREENBERG, D. P. 1998. A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings of SIGGRAPH 98*, ACM SIGGRAPH / Addison Wesley, Orlando, Florida, Computer Graphics Proceedings, Annual Conference Series, 287–298.
- PERONA, P., AND MALIK, J. 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE PAMI* 12, 7, 629–639.
- RUDMAN, T. 2001. *The Photographer's Master Printing Course*. Focal Press.
- RUSHMEIER, H. E., AND WARD, G. J. 1994. Energy preserving non-linear filters. In *Proceedings of SIGGRAPH 94*, ACM SIGGRAPH / ACM Press, Orlando, Florida, Computer Graphics Proceedings, Annual Conference Series, 131–138.
- SAINT-MARC, P., CHEN, J., AND MEDIONI, G., 1991. Adaptive smoothing: a general tool for early vision.
- SCHECHNER, Y. Y., AND NAYAR, S. K. 2001. Generalized mosaicing. In *Proc. IEEE CVPR*, 17–24.
- SCHLICK, C. 1994. Quantization techniques for visualization of high dynamic range pictures. *5th Eurographics Workshop on Rendering*, 7–20.
- SOCOLINSKY, D. 2000. Dynamic range constraints in image fusion and visualization. In *Proc. Signal and Image Processing*.
- TOMASI, C., AND MANDUCHI, R. 1998. Bilateral filtering for gray and color images. In *Proc. IEEE Int. Conf. on Computer Vision*, 836–846.
- TUMBLIN, J., AND RUSHMEIER, H. 1993. Tone reproduction for realistic images. *IEEE Comp. Graphics & Applications* 13, 6, 42–48.
- TUMBLIN, J., AND TURK, G. 1999. Lcis: A boundary hierarchy for detail-preserving contrast reduction. In *Proceedings of SIGGRAPH 99*, ACM SIGGRAPH / Addison Wesley Longman, Los Angeles, California, Computer Graphics Proceedings, Annual Conference Series, 83–90.
- TUMBLIN, J., HODGINS, J., AND GUENTER, B. 1999. Two methods for display of high contrast images. *ACM Trans. on Graphics* 18, 1, 56–94.
- TUMBLIN, J. 1999. *Three methods of detail-preserving contrast reduction for displayed images*. PhD thesis, College of Computing Georgia Inst. of Technology.
- WARD, G. J. 1994. The radiance lighting simulation and rendering system. In *Proceedings of SIGGRAPH 94*, ACM SIGGRAPH / ACM Press, Orlando, Florida, Computer Graphics Proceedings, Annual Conference Series, 459–472.
- YANG, D., GAMAL, A. E., FOWLER, B., AND TIAN, H. 1999. A 640x512 cmos image sensor with ultrawide dynamic range floating-point pixel-level adc. *IEEE Journal of Solid State Circuits* 34, 12 (Dec.), 1821–1834.



Figure 15: Foggy scene. Radiance map courtesy of Jack Tumblin, Northwestern University [Tumblin and Turk 1999].

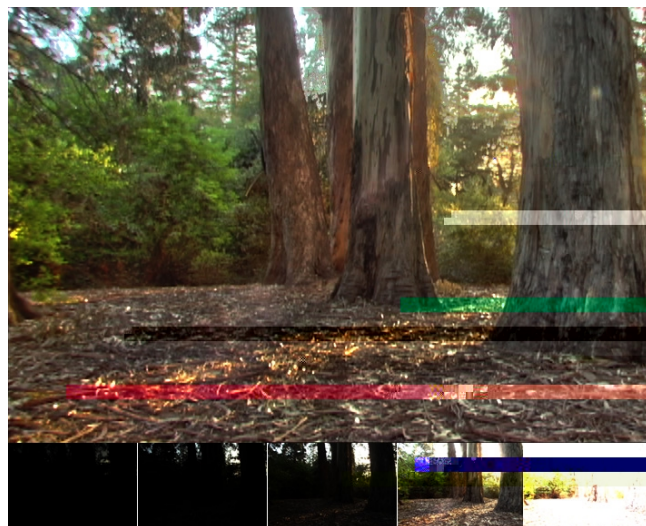


Figure 16: Grove scene. Radiance map courtesy of Paul Debevec, USC [Debevec and Malik 1997].

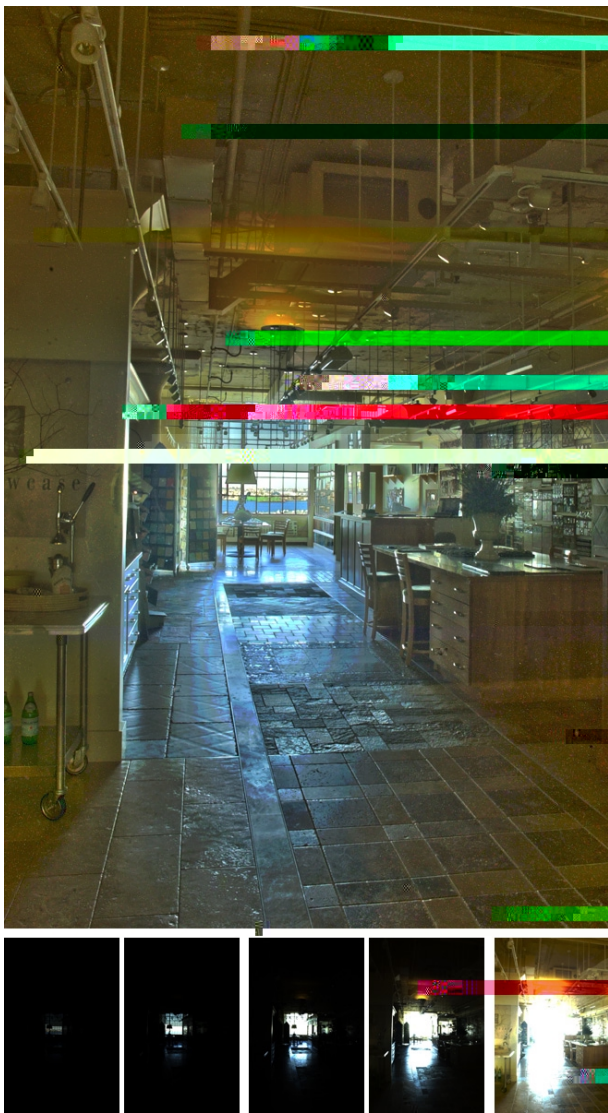


Figure 17: Interior scene.

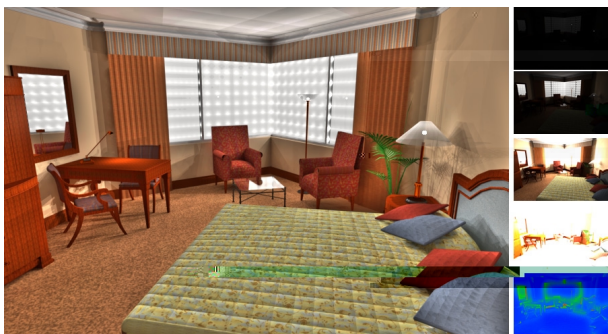
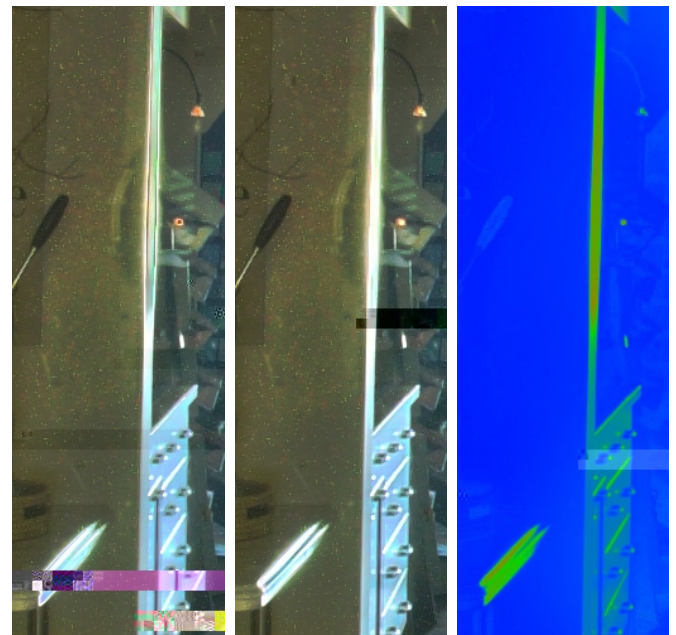


Figure 18: Hotel room. The rightmost image shows the uncertainty. Designed and rendered by Simon Crone using RADIANCE [Ward 1994]. Source image: Proposed Burswood Hotel Suite Refurbishment (1995). Interior Design - The Marsh Partnership, Perth, Australia. Computer simulation - Lighting Images, Perth, Australia. Copyright (c) 1995 Simon Crone.



without with uncertainty fix uncertainty

Figure 19: Zoom of Fig. 17. The haloing artifacts in the vertical highlight and in the lamp are dramatically reduced. The noise is due to the sensor.

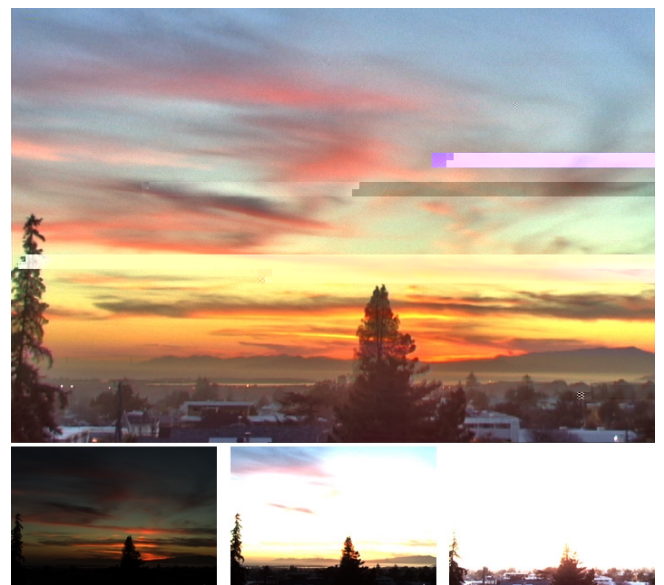
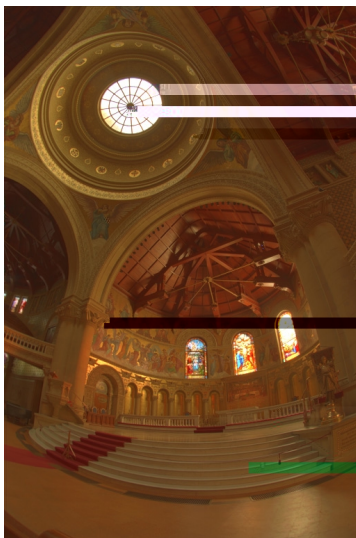
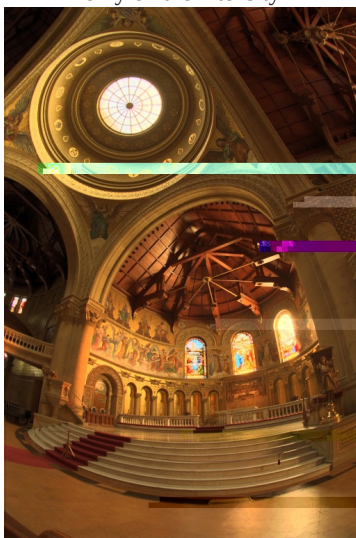


Figure 20: Vine scene. Radiance map courtesy of Paul Debevec, USC [Debevec and Malik 1997].



User-optimized gamma correction only on the intensity



Histogram adjustment [Larson et al. 1997]



LCIS. Image reprinted by permission, copyright © 1999 Jack Tumblin [Tumblin and Turk 1999]

Figure 21: Stanford Memorial Church, displayed with different methods.



Figure 22: Stanford Memorial Church displayed using bilateral filtering. The rightmost frame is the color-coded base layer. Radiance map courtesy of Paul Debevec, USC [Debevec and Malik 1997].

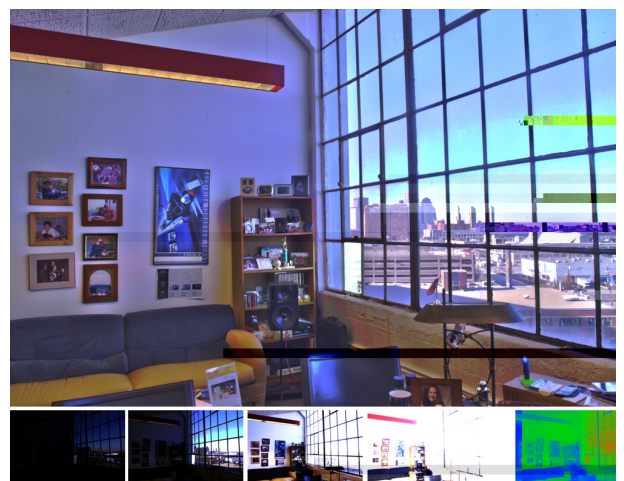


Figure 23: Window scene. The rightmost image shows the color-coded base layer.