Two-scale Tone Management for Photographic Look

Soonmin Bae   Sylvain Paris   Frédo Durand

Computer Science and Artificial Intelligence Laboratory

Massachusetts Institute of Technology

(a) input   (b) sample possible renditions: bright and sharp, gray and highly detailed, and contrasted, smooth and grainy
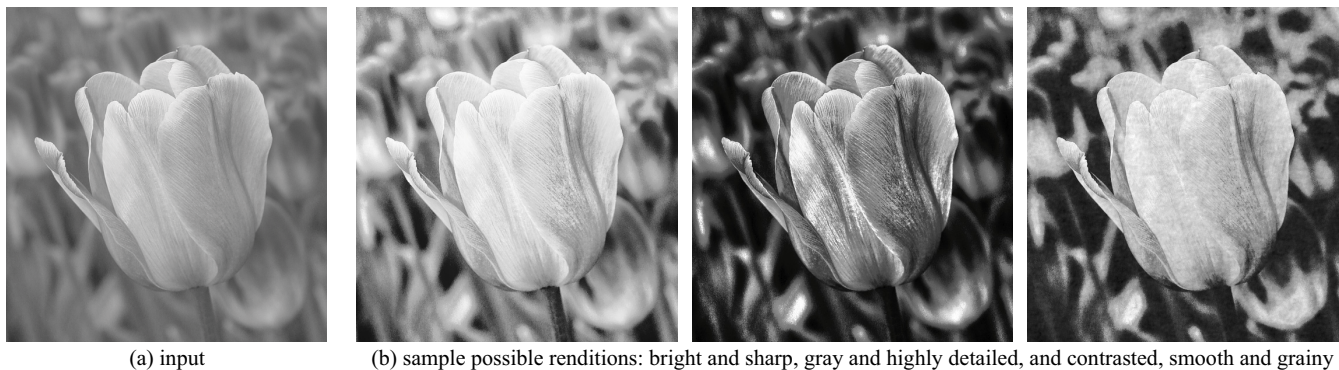
Figure 1: This paper describes a technique to enhance photographs. We equip the user with powerful filters that control several aspects of an image such as its tonal balance and its texture. We make it possible for anyone to explore various renditions of a scene in a few clicks. We provide an effective approach to æsthetic choices, easing the creation of compelling pictures.

Abstract

We introduce a new approach to tone management for photographs. Whereas traditional tone-mapping operators target a neutral and faithful rendition of the input image, we explore pictorial looks by controlling visual qualities such as the tonal balance and the amount of detail. Our method is based on a two-scale non-linear decomposition of an image. We modify the different layers based on their histograms and introduce a technique that controls the spatial variation of detail. We introduce a Poisson correction that prevents potential gradient reversal and preserves detail. In addition to directly controlling the parameters, the user can transfer the look of a model photograph to the picture being edited.

Keywords: Computational photography, high dynamic range, tone management, pictorial look, bilateral filter, image processing

1 Introduction

Much research has been dedicated to tone mapping for the display of high-dynamic-range images. These tools focus on contrast reduction, seeking a neutral reproduction, and are ideal when fidelity is needed. However, tone manipulation is also useful when the input has normal dynamic range, and many users seek to obtain a certain "look" for their pictures to convey a mood or an æsthetic. This is particularly significant for black-and-white photography where strikingly distinctive styles can be achieved. We present a new tone management approach that offers direct control over the "look" of an image for both high- and normal-dynamic-range inputs.

The "look" of images has been addressed in Non-Photorealistic Rendering and recent analogy approaches enable the imitation of texture or stylized images in a purely data-driven fashion, e.g. [Hertzmann et al. 2001]. However, to the best of our knowledge, no approach enables the imitation of a photographic "look" such as the ones achieved by master black-and-white photographers.

We argue that a large part of such a look deals with the management of tones, which advanced photographers perform through elaborate lighting, tedious work in the darkroom, or using photo editing software. Unfortunately, such painstaking work and advanced skill is out of reach of casual users. In addition, the issues of *workflow* and efficiency are becoming prevalent among professional users. The workflow describes the full process from image capture to printing and can include multiple software stages and manual retouching, all requiring much effort. Reducing the user work is critical to professionals, and many manuals and tools are dedicated to optimizing and automating all steps. For example, a wedding photographer takes hundreds of pictures and needs to give them a consistent look and to swiftly deliver them to clients. Individual retouching is hardly an option, and creative control over the look of each image is often sacrificed. Recently-introduced software such as Apple's Aperture and Adobe's Lightroom focuses on workflow optimization but offers little interactive editing capabilities.

To address these difficulties, we propose a tone-management technique dedicated to both casual and professional photographers. We focus on the tonal aspects of photos decoupled from their content. Issues such as framing and topic selection are out of the scope of our work. We nevertheless demonstrate the wide range of looks that our approach can produce. We provide simple controls and enable both global and local tone management. In addition to direct manipulation, users can transfer the look of a model picture,

(a) Clearing Winter Storm Ansel Adams,1942 or earlier
(reproduced with permission)

(b) Angkor #71, by Kenro Izu
(Original: 14"x20" film contact printed on Platinum/Palladium coated paper.)
Copyright 1994 Kenro Izu, reproduced with permission of the artist.

Figure 2: Typical model photographs that we use. Photo (a) exhibits strong contrast with rich blacks, and large textured areas. Photo (b) has mid-tones and vivid texture over the entire image.

thereby "showing" the desired look. This also allows professionals to apply the rendition of previous prints to new photographs.

This paper makes the following contributions.

Large-scale Tonal Balance Management: We control the large-scale spatial tonal variation over an image.

Spatial Detail Variation: We manipulate the amount of high-frequency detail or texture and its spatial variation. In particular, we introduce a computation of *textureness* that measures local high-frequency content while respecting strong edges.

Gradient Constraint: We employ a gradient reconstruction step to prevent gradient reversal and preserve detail.

Our exposition focuses on transfer between images because it demonstrates the relevance and robustness of the features we manipulate. However, direct control through the curve interface is equally powerful, though perhaps more suited to advanced users.

1.1 Related work

Tone Mapping Tone-mapping seeks the faithful reproduction of high-dynamic-range images on low-dynamic-range displays, while preserving visually important features [Reinhard et al. 2005]. Our work builds on local tone mapping where the mapping varies according to the neighborhood of a pixel [Pattanaik et al. 1998; Tumblin and Turk 1999; Reinhard et al. 2002; Durand and Dorsey 2002; Fattal et al. 2002; Li et al. 2005]. The precise characteristics of film have also been reproduced [Geigel and Musgrave 1997; Reinhard et al. 2002]. However, most techniques seek an objective rendering of the input, while we want to facilitate the exploration and transfer of particular pictorial looks.

Conversion to Grayscale Gooch et al. [2005] convert color images to grayscale while preserving salient features. They also seek fidelity to the original picture, whereas we explore stylistic variations. Their approach is nonetheless complementary to ours because it extracts compelling contrast from color images.

Gradient Image Processing A number of recent techniques have characterized images by their gradient and used Poisson reconstruction to perform tone mapping [Fattal et al. 2002] and montages [Pérez et al. 2003; Agarwala et al. 2004]. We also exploit the Poisson approach to ensure the quality of our result, because it naturally allows us to combat gradient reversal, a traditional plague of aggressive multi-scale manipulation.

Style Transfer and Stylization Style transfer has been explored for the textural aspects of non-photorealistic media, e.g. [Hertzmann et al. 2001; Drori et al. 2003], and DeCarlo et al. stylize photographs based on saliency [2002]. In contrast, we seek to retain photorealism and control large-scale effects such as tonal balance and the variation of local detail. In addition, our parametric approach leads to continuous changes supported by interactive feedback and enables interpolations and extrapolations of image look.

Visual Equalizer Our work is inspired by the ubiquitous visual equalizer of sound devices. Similarly, the modification of frequency bands can alter the "mood" or "style" of motion data [Bruderlin and Williams 1995]. The equivalent for images is challenging because of the halos that frequency decomposition can generate around edges. Our work can be seen as a two-band equalizer for images that uses non-linear signal processing to avoid halos and provides fine tonal and spatial control over each band.

1.2 Achieving a Photographic Look

The traditional darkroom offers remarkable global and local control over the brightness, contrast, and sharpness of images via a combination of chemical and optical processes [Rudman 1994; Adams 1995]. Black-and-white photographs vary in their tonal palette and how they deal with the dynamic range of a scene. A photographer like Adams (Fig. 2a) exhibits strong contrast with rich blacks, while an artist like Stieglitz (Fig. 15a) relies more on the mid-tones. This suggests the intensity histogram as a characterization of tonal look, but we show in this paper that the spatial distribution of tones must be taken into account because a histogram does not make the distinction between local and global contrast.

The amount of texture is crucial in photographs; some artists use vivid texture over the entire image (Fig. 2b), while other contrast large smooth areas with strong textures in other parts of the image (Fig. 2a). Furthermore, the human visual system is known to be more sensitive to local contrast than to low spatial frequencies.

Finally, a photograph is characterized by low-level aspects of the medium such as tone (e.g. sepia toning) and grain (controlled by the film and paper characteristics).

These observations drive our approach. We propose decompositions of an image that afford direct control over dynamic range, tonal distribution, texture and sharpness.
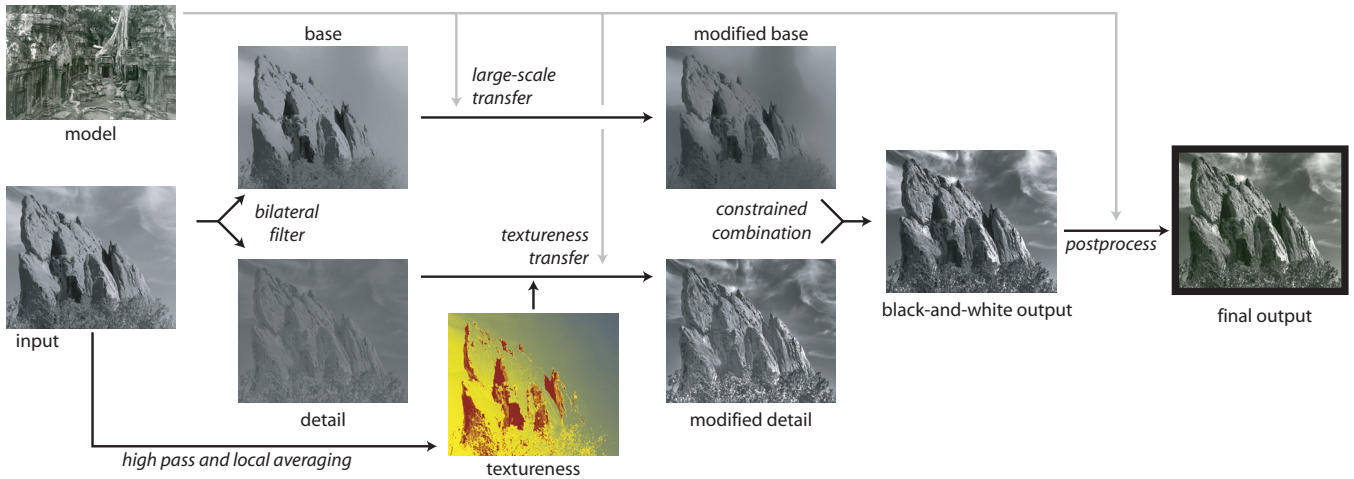
Figure 3: Overview of our pipeline. The input image is first split into base and detail layers using bilateral filtering. We use these layers to enforce statistics on low and high frequencies. To evaluate the texture degree of the image, we introduce the notion of *textureness*. The layers are then recombined and post-processed to produce the final output. The model is Kenro Izu's masterpiece shown in Figure 2b.

1.3 Overview

The previous discussion suggests that aspects such as the intensity distribution at different scales, spatial variations, and the amount and distribution of detail are critical to the look of a photograph. This inspires our use of a two-scale decomposition to control large-scale effects and the texture distribution. We quantify the look of an image using histograms over this decomposition, which affords both interactive control using a curve interface, and the ability to automatically transfer visual properties between images. In the latter, histograms of the components of a model image are forced upon a new input. Because we explore strong stylistic variations, we tend to perform larger modifications to the input than tone mapping. In particular, some looks require an increase in local contrast, which can produce halos if traditional techniques are used. We introduce a gradient constraint that prevents undesirable modifications. Finally, we post-process the image to achieve various effects such as soft focus, paper grain, and toning. Figure 3 summarizes this process.

2 Background

Before introducing our approach, we review important image-processing tools at the core of our technique.

Histogram Matching Matching histograms is the traditional solution to transferring an intensity distribution. Given an image I with histogram h_I and a reference histogram h_M, we seek a function $\ell_{I \rhd M}$ such that applying $\ell_{I \rhd M}$ to each pixel of I results in an image with histogram h_M. To build $\ell_{I \rhd M}$, we use the cumulative histograms c_M and c_I defined by $c(x) = \int_{-\infty}^{x} h$. It can be shown that $c_I(I)$ produces an image with a uniform histogram. Thus, we define:

$$\ell_{I \rhd M}(x) = c_M^{-1}[\, c_I(x) \,] \qquad (1)$$

and $\ell_{I \rhd M}(I)$ generates an image with the histogram h_M. More details can be found in image processing books, e.g. [Gonzales and Woods 2002] (p. 94). While histogram matching is a key tool in our approach, we observe that matching the pixel histogram is not sufficient to control the tonal look of an image (Fig. 11, 15 and 16).

Poisson Reconstruction Given a 2D field of 2D vectors $\mathbf{v}$, one can build an image I with a gradient ∇I as close as possible to $\mathbf{v}$, in the least square sense. This is achieved through a Poisson equation:

$$\partial I / \partial t = \Delta I - \operatorname{div}(\mathbf{v}) \qquad (2)$$

Pérez *et al.* [2003] have shown impressive image manipulations using this tool. We refer to their paper for detail.

Bilateral Filtering The bilateral filter [Tomasi and Manduchi 1998] smooths the input image while preserving its main edges. Each pixel is a weighted mean of its neighbors where the weights decrease with the distance in space and with the intensity difference. With $g_\sigma(x) = \exp(-x^2/\sigma^2)$, a Gaussian function, the bilateral filter of image I at pixel $\mathbf{p}$ is defined by:

$$bf(I)_{\mathbf{p}} = \frac{1}{k} \sum_{\mathbf{q} \in I} g_{\sigma_s}(\|\mathbf{p} - \mathbf{q}\|)\, g_{\sigma_r}\left(|I_{\mathbf{p}} - I_{\mathbf{q}}|\right)\, I_{\mathbf{q}} \qquad (3a)$$

$$\text{with:} \qquad k = \sum_{\mathbf{q} \in I} g_{\sigma_s}(\|\mathbf{p} - \mathbf{q}\|)\, g_{\sigma_r}\left(|I_{\mathbf{p}} - I_{\mathbf{q}}|\right) \qquad (3b)$$

where σ_s controls the spatial neighborhood, and σ_r the influence of the intensity difference, and k normalizes the weights. The bilateral filter is often used to create a two-scale decomposition where the output of the filter produces a large-scale layer (a.k.a. *base*) and the difference is called the *detail* layer [Durand and Dorsey 2002]. We use our fast version of the bilateral filter [Paris and Durand 2006].

3 Large-Scale Tonal Distribution

Our tone management relies on a two-scale decomposition based on the bilateral filter. We refine the standard usage of the bilateral filter in two ways: we introduce a gradient correction to prevent gradient reversals, and we apply histogram transformations instead of just scaling the large-scale component as in traditional tone mapping.

3.1 Bilateral Decomposition

We use a decomposition similar to that of Durand and Dorsey [2002]. Since contrast is a multiplicative effect, we work in the logarithmic domain. We define the base layer B and detail layer D from the input image I (where I, B and D have log values):

$$B = bf(I) \qquad \text{and} \qquad D = I - B \qquad (4)$$

The choice of σ_s and σ_r is crucial. σ_s specifies spatial scales and $\sigma_s = \min(\text{width}, \text{height})/16$ consistently produces good results. σ_r differentiates important edges from detail. We rely on the gradient

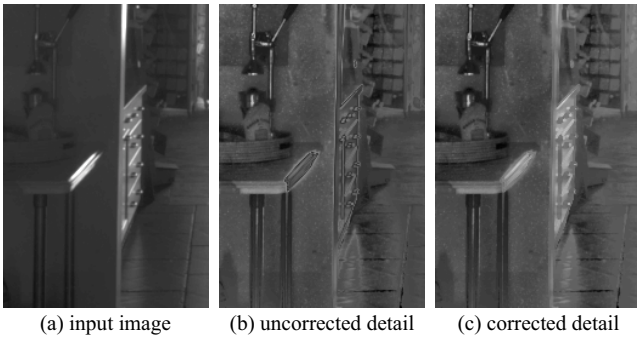(a) input image   (b) uncorrected detail   (c) corrected detail

Figure 4: The bilateral filter can cause gradient reversals in the detail layer near smooth edges. Note the problems in the highlights (b). We force the detail gradient to have the same orientation as the input (c). Contrast is increased in (b) and (c) for clarity.

norm to estimate the edge amplitude in the input. With p_{90} denoting the 90th percentile[1], $\sigma_{\mathrm{r}} = p_{90}(\|\nabla I\|)$ achieves consistently good results. These settings are robust to spatial and intensity scales.

Gradient Reversal Removal   Durand and Dorsey [2002] note that artifacts can occur when edges are not sharp. They introduce a "fix" that detects uncertain pixels and uses a smoothed base layer, but they highlight that this solution is not entirely satisfying. The problem is more acute in our case because we may *increase* the amount of detail (by a factor as high as 6 in some examples), which requires a reliable halo-free detail layer.

We address this by directly constraining the gradient of the decomposition to prevent reversal. We force the detail derivatives $\partial D/\partial x$ and $\partial D/\partial y$ to have the same sign as the input derivatives and an amplitude no greater than them. For this, we build a gradient field $\mathbf{v} = (x_{\mathbf{v}}, y_{\mathbf{v}})$:

$$x_{\mathbf{v}} = \begin{cases} 0 & \text{if} \quad \mathrm{sign}\,(\partial D/\partial x) \neq \mathrm{sign}\,(\partial I/\partial x) \\ \partial I/\partial x & \text{if} \quad |\partial D/\partial x| > |\partial I/\partial x| \\ \partial D/\partial x & \text{otherwise} \end{cases} \quad (5)$$

The y component $y_{\mathbf{v}}$ is defined similarly. The corrected detail layer is obtained by solving the corresponding Poisson equation (Eq. 2). We update the base layer accordingly: $B = I - D$. This approach results in a high-quality detail layer because it directly addresses gradient reversal and preserves other subtle variations (Fig. 4).

3.2 Tonal Balance

The base layer contains the large-scale spatial distribution of tones (Fig. 5). In contrast to tone mapping where the base layer is simply scaled down [Durand and Dorsey 2002], we want to enforce a large-scale distribution of tones that matches a model image. This is why we perform histogram matching and transfer the histogram of the model base B_M onto the new base B_I.
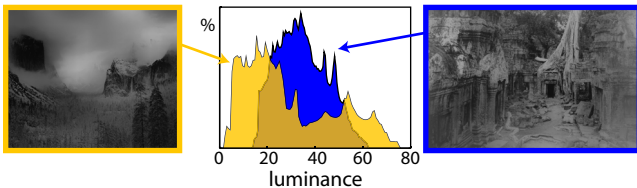

Figure 5: The luminance histogram of the base component is a good indicator of tonal balance. The photos are the same as in Figure 2.

---

4 Detail and Texture Management

The amount and spatial distribution of high-frequency texture is the natural complement of the large-scale tonal palette. The core contribution of our work is a technique that manipulates the amount of high-frequency content and its spatial variation. This contrasts with tone mapping approaches that usually do not modify the detail layer.

This step involves additional challenges compared to the base transform. First, we show that the detail layer does not capture all the high frequency content of the image. Second, we need to modify the spatial variation of detail without creating artifacts. In particular, we introduce a new technique to measure and modify local frequency content in an edge-preserving manner.

4.1 Detail Management based on Frequency Analysis

While the bilateral filter provides a decomposition that facilitates halo-free manipulation, the edge-preserving term $g_{\sigma_{\mathrm{r}}}$ results in substantial high-frequency content in the base layer (Fig. 6). While the choice of different parameters or more advanced filters [Choudhury and Tumblin 2003] can affect this issue, the very nature of such filter calls for high-frequency content in the base. In particular, the influence of the range Gaussian $g_{\sigma_{\mathrm{r}}}$ means that patterns that are high-frequency but high-contrast will mostly be in the base. While this is not an issue for tone mapping where the detail is unaffected, it is critical for our detail management. On the other hand, the manipulation of the detail layer is a safe operation that does not lead to the halo artifacts caused by linear image processing.

Our solution combines linear frequency *analysis* with the *manipulation* of the detail layer obtained from our nonlinear filter. We analyze the amount of texture (or high frequency) using a high pass filter applied to *both* the detail and the base layer. This ensures that all the frequency content is taken into account. We use this information to decide how the detail layer should be modified. In a nutshell, we get the best of the two approaches: reliable analysis of the high-pass filter, and the safe manipulation of the detail layer.

4.2 Textureness

We seek to characterize the local amount of high frequency content over the image to distinguish regions with high levels of detail from

(a) input   (b) high frequencies of input

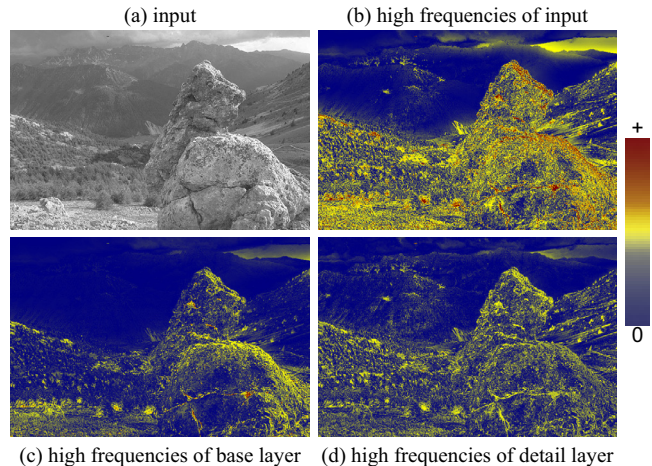(c) high frequencies of base layer   (d) high frequencies of detail layer

Figure 6: Because of the preserved edges, the high frequencies of an image (b) appear both in the base layer (c) and in the detail layer (d). This phenomenon has to be taken into account to achieve an appropriate analysis.
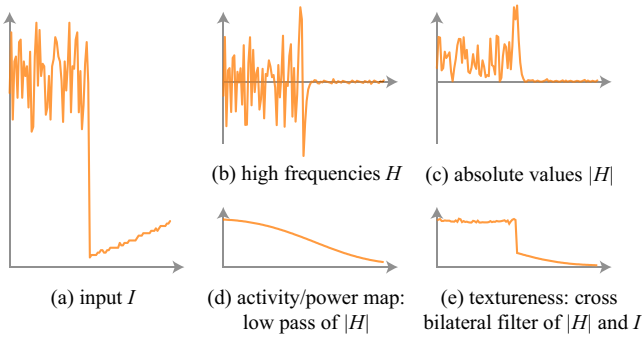
(b) high frequencies H (c) absolute values $|H|$

(a) input I (d) activity/power map: (e) textureness: cross
low pass of $|H|$ bilateral filter of $|H|$ and I

Figure 7: Textureness of a 1D signal. To estimate the textureness of the input (a), we compute the high frequencies (b) and their absolute values (c) . Finally, we locally average these amplitudes: Previous work based on low-pass filter (d) incurs halos (Fig. 8) whereas our cross bilateral filtering yields almost no halos (e) .

smooth regions. We build on the notion of power maps, e.g. [Su et al. 2005] and activity map [Li et al. 2005] where the local average of the amplitude of high frequencies is used. Figure 7 illustrates our computation of textureness for a 1D example where the left part has a high level of local contrast while the right part is smooth.

First, we compute a high-pass version H of the image using the same cutoff σ_s. Note that the local average of such a high-pass image is by definition zero: the low frequencies are removed. This is why we consider the *magnitude* (or absolute value) of H (Fig. 7c). Power maps or activity maps are then defined as the local average – obtained via low-pass filtering – of this magnitude (Fig. 7d). Such maps provide good characterization of highly-textured vs. smooth regions and the local level of detail can be altered by modifying the detail layer accordingly.

Unfortunately, such spatially-varying manipulation of detail can lead to artifacts at the boundary between highly detailed and smooth regions (Fig. 8). This is because the amount of detail on one side of the boundary influences the estimate on the other side, and the manipulation suffers from a halo effect similar to that observed in linear frequency decomposition of image intensity. This problem is the same as the one addressed by edge-preserving decomposition, except that we are dealing with a less spatially localized quantity, the magnitude of high frequency $|H|$. Strong edges are hard to characterize in $|H|$, which is why we define textureness using a *cross-bilateral filter* [Eisemann and Durand 2004; Petschnigg et al. 2004] where the intensity image defines the edge-preserving term to filter $|H|$. More precisely, our textureness is defined as

$$T(I)_{\mathbf{p}} = \frac{1}{k} \sum_{\mathbf{q} \in |H|} g_{\sigma_s}(\|\mathbf{p} - \mathbf{q}\|) \, g_{\sigma_r}\left(|I_{\mathbf{p}} - I_{\mathbf{q}}|\right) \, |H|_{\mathbf{q}} \quad (6a)$$

with: $\quad k = \sum_{\mathbf{q} \in I} g_{\sigma_s}(\|\mathbf{p} - \mathbf{q}\|) \, g_{\sigma_r}\left(|I_{\mathbf{p}} - I_{\mathbf{q}}|\right) \quad (6b)$

We set this cross filter with the same σ_r as for the base-detail computation, but with a larger σ_s (8 times larger in practice) to en-
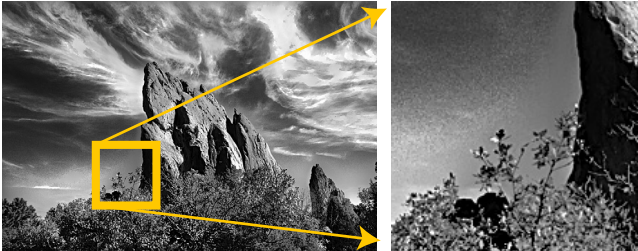

Figure 8: Using a Gaussian filter to locally average the high frequency amplitudes yields halos around strong edges. To prevent this defect, we use an edge-preserving filter.

sure smooth textureness variations on uniform regions (discontinuities can still happen at edges). Figure 9 shows how our textureness map captures the local amount of detail over the image.
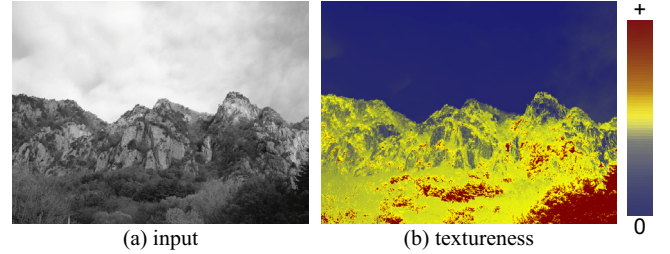

(a) input   (b) textureness

Figure 9: Our measure of textureness indicates the regions with the most contrasted texture.

Textureness Transfer   The input I and model M have textureness maps $T(I)$ and $T(M)$, respectively. Using histogram transfer, we enforce the histogram of $T(M)$ onto $T(I)$ to build the desired textureness map T'. To prevent halos, we modify only the detail layer D to approximate T'. We scale the values of D by a ratio ρ to match T' values while accounting for the textureness of the base B' modified by the tonal balance of the previous section:

$$\rho_{\mathbf{p}} = \max\left(0, \frac{T'_{\mathbf{p}} - T(B')_{\mathbf{p}}}{T(D)_{\mathbf{p}}}\right) \quad (7)$$

We do not apply negative ratios, thus preventing gradient reversals. Although this computation is done pixel-wise, we found that the textureness maps are smooth enough to ensure a smooth transformation. We linearly recombine the layers to produce the output: $O = B' + \rho D$.

4.3   Detail Preservation

As illustrated by Figure 10b, the previous result ($O = B' + \rho D$) may result in saturated highlights and shadows. These bright and dark regions are nevertheless of higher importance for photographers who aim for crisp details everywhere. We preserve these details in two steps.

First, we enforce the intensity histogram of the model M to the current output O, which brings back the values within the displayable range. Second, we modify the gradient field to ensure that no details are removed or overly emphasized. Similarly to our shock removal, we build a gradient field $\mathbf{v}$ that satisfies these constraints. We aim at preserving a portion α of the variations of the

(a) input   (b) unconstrained   (c) histogram matching   (d) gradient constraint
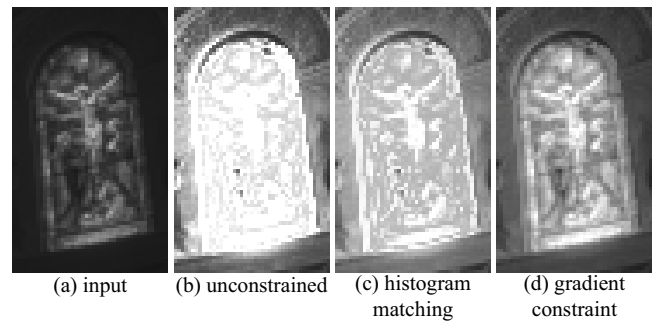
Figure 10: Without constraints, the result may lose valuable details (b) because the highlight are saturated. Enforcing the model histogram brings back the intensity values within the visible range (c). Finally, constraining the gradients to preserve some of the original variations (a) produces high quality details (d).
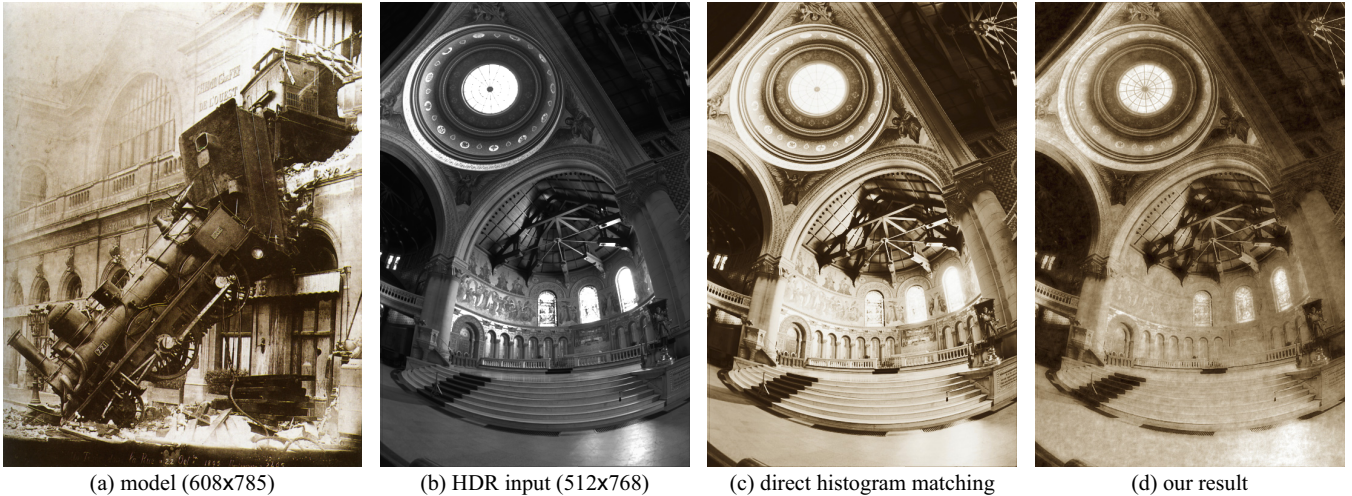
| (a) model (608x785) | (b) HDR input (512x768) | (c) direct histogram matching | (d) our result |

Figure 11: Our system can seamlessly handle HDR images . We can turn a sharp picture (b) into a soft grainy and toned photograph (d). We have toned the histogram-transferred version (c) to prevent biased comparison due to different color cast. The model (a) is Accident at the Gare Montparnasse from the Studio Lévy and Sons, 1895. The input (b) is courtesy of Paul Debevec, USC

input image, and we prevent the gradient being increased by a factor greater than β to avoid over-emphasizing noise. We define:

$$x_{\mathbf{v}} = \begin{cases} \alpha \, \partial I/\partial x & \text{if} \quad |\partial O/\partial x| < \alpha |\partial I/\partial x| \\ \beta \, \partial I/\partial x & \text{if} \quad |\partial O/\partial x| > \beta |\partial I/\partial x| \\ \partial O/\partial x & \text{otherwise} \end{cases} \quad (8)$$

The y component $y_{\mathbf{v}}$ is defined similarly, and the image is reconstructed with the Poisson technique. All that remains is to set α and β. We use percentiles to define $\phi = [p_{95}(O) - p_5(O)]/[p_{95}(I) - p_5(I)]$, which robustly estimates the contrast change induced by our processing. We then use a constant $\alpha = \phi/4$, and we make β depend on intensity in order to avoid increasing noise. We use a smooth-step function $v_\tau(x) = 0$ if $x < \tau$, 1 if $x > 2\tau$, and $1 - [1 - (x - \tau)^2/\tau^2]^2$ otherwise. Setting $\beta = 1 + 3v_\tau\phi$ performs consistently well with $\tau = 0.1$. As a result, we successfully preserve the richness of the input images as shown on Figure 10.

5 Additional Effects

While our focus is on the management of the tonal palette and the variation of detail, we have also developed simple filters to control low-level aspects of the look of a photograph.

Soft Focus and Sharpness The level of sharpness of a picture is a strong aspect of style as exemplified by soft-focus effects. To characterize sharpness, we use difference-of-Gaussian filters and analyze three octaves of the current output O. We set the parameters so that the highest band captures the wavelengths shorter than $\lambda_{\mathrm{h}} = \min(\text{width}, \text{height})/256$. For each band B_i^O, we evaluate the sharpness of the most contrasted edge with the 95th percentile $p_{95}(|B_i^O|)$. We divide this number by $p_{95}(O) - p_5(O)$ to make this measure invariant to intensity. The use of percentiles makes this estimation robust. To summarize, our sharpness estimator is a triplet of numbers $(\zeta_1, \zeta_2, \zeta_3)$ defined as $\zeta_i^O = p_{95}(|B_i^O|)/(p_{95}(O) - p_5(O))$. We compute the same measures for the model M and scale the bands B_i^O of the output by a factor ζ_i^M/ζ_i^O to transfer sharpness. See Figure 11, 15 and 17. In particular, in Figure 11, the intermediate frequencies are attenuated more than the highest frequencies, achieving a "soft-yet-sharp" rendition which is a convincing approximation of the effect produced by a soft-focus lens.

Film Grain and Paper Texture Some photographs exhibit a characteristic appearance due to the paper which they are printed on or because the film grain is visible. We reproduce this effect in two steps. First, since the grain is not part of the image content, we remove it from the model image with a bilateral filter on the luminance values, using $\sigma_{\mathrm{r}} = p_{75}(\|\nabla M\|)$. Then, we crop a sample from the residual (detail) of this bilateral filter in a uniform region. We generate a grain layer using texture synthesis [Heeger and Bergen 1995] (Fig. 1, 11, and 15).

Color and Toning To handle color images, we can use the original a and b channels in the CIE-LAB color space. a and b can be used directly, or they can be scaled by L_O/L_I where L_I and L_O are the luminance of the input and current output. The latter alters color saturation and is useful for HDR images because their chromaticity is often out of the displayable gamut [Fattal et al. 2002; Li et al. 2005]. Figures 13 and 17 show color renditions.

We produce toned pictures (e.g. sepia) using a one-dimensional color map. We use the *Lab* color space to build the functions $a(L)$ and $b(L)$ from the model by averaging a and b for the pixels with a given L. These functions are then applied to the L values of the current result (Fig. 11).

6 Results

We demonstrate our technique using models by different artists on a variety of inputs, including pictures by beginners using point-and-

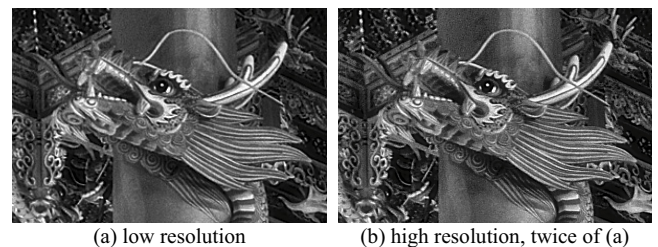| (a) low resolution | (b) high resolution, twice of (a) |

Figure 12: Results from lower resolution (a) provides quick previews and allow for interactive adjustments before rendering high resolution results (b). Limited differences are visible on the smallest details (*e.g.* in the background) because they are not well sampled in the low-resolution image.

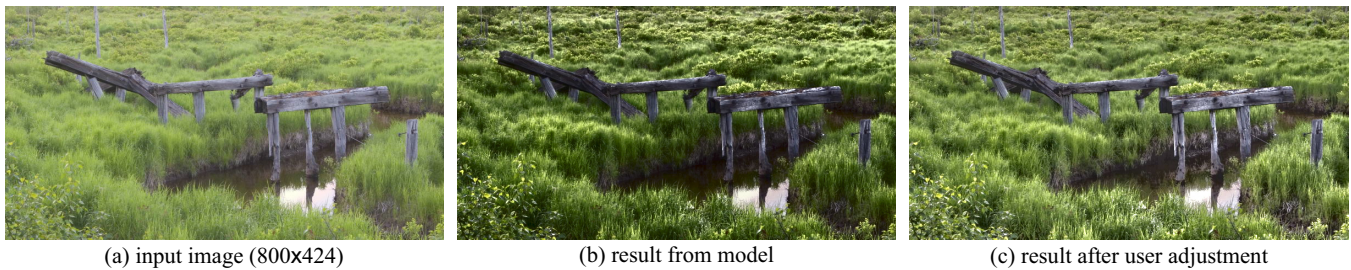| (a) input image (800x424) | (b) result from model | (c) result after user adjustment |

Figure 13: This rendition was obtained in two steps. We first used Kenro Izu's picture shown in Figure 2b as a model (b). Then, we manually increased the brightness and softened the texture to achieve the final rendition (c) that we felt is more suitable for the scene.

shoot cameras, photos by more advanced amateurs using SLRs, and high-dynamic-range images (Fig. 11).

Computation time varies roughly linearly with the number of pixels, thanks to our fast bilateral filter and a multigrid implementation of gradient reconstruction. For example, the full pipeline for a one megapixel image takes about six seconds on a 2.6GHz Opteron PC, and a four megapixel takes 23 seconds. However, note that we cache intermediate results such as the base, detail, and textureness map, which enables interactive feedback when using the user interface. In addition, results from downsampled images are faithful previews (Fig. 12) because our parameters are scale invariant, which enables fast interaction before a final computation at full resolution.

Our implementation enables interactive adjustment of the parameters through controls such as sliders for scalar parameters and, for the remapping function of the base layer, a spline interface inspired by the "curve" tool of photo-editing software. These adjustments can be saved and reused on subsequent inputs. We have also found that the interactive control is a great way to refine the result of an automatic transfer (Fig. 13).

Figures 15 and 16 shows a comparison of our results with a straightforward histogram matching from the model to the input. Histogram matching ignores the notion of texture and therefore overly increases or decreases the picture detail. In comparison, our technique yields results that are both more faithful to the model and higher quality, with rich shadows and detailed highlights.

Discussion The main cause of failure of our approach is poor input quality. In particular JPEG artifacts and noise noise can be amplified by our detail manipulation (Fig. 14). Apart from this, meaningful input/model couples (two landscapes, two trees, etc) consistently yield faithful transfers, close to our expectations. On more surprising pairs (*e.g.* a flower and a landscape), the process does not generate artifacts and the achieved mood is often pleasing, although one can always argue about the æsthetic quality of some results. Portraits are probably the most challenging type of input, and detail enhancement can lead to unflattering result because skin defects can be emphasized. It is then best to turn this feature off.
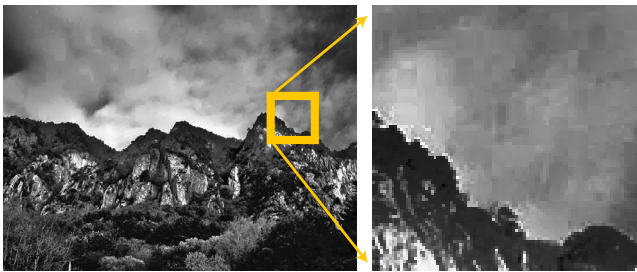

Figure 14: Our technique suffers from imperfections such as JPEG artifacts. In this example, the artifacts in the sky are not visible in the input image (Fig. 9a) but appear clearly after processing.

7 Conclusions

We have presented an approach to manipulate the tonal look of digital photographs. Using a combination of non-linear edge-preserving decomposition and linear analysis, we control both the large-scale tonal palette and the detail over an image. In particular, we manipulate the spatial variation of high-frequencies using a new textureness map that performs an edge-preserving analysis and manipulation of the high-frequency content. We have introduced a gradient constraint that preserves image content and prevents gradient reversal and halos.

Our method can be used to transfer the look of a model photograph or can be directly controlled using a simple interface. It allows for the exploration of a variety of styles and achieves high-quality results that are consistent from low-resolution previews to high-resolution prints.

This work opens several areas of future research. It should be combined with approaches to control the color components of pictorial style. While early experiments with videos have shown that our technique itself is stable, we have found that the biggest challenge is the fluctuation created by auto-exposure, autofocus and the variation of motion blur when the camera moves.

References

ADAMS, A. 1995. *The Print*. Little, Brown and Company.

AGARWALA, A., DONTCHEVA, M., AGRAWALA, M., DRUCKER, S., COLBURN, A., CURLESS, B., SALESIN, D. H., AND CO-HEN, M. F. 2004. Interactive digital photomontage. *ACM Trans. on Graphics 23*, 3. Proc. of ACM SIGGRAPH conf.

BRUDERLIN, A., AND WILLIAMS, L. 1995. Motion signal processing. In *Proc. of ACM SIGGRAPH conf.*

CHOUDHURY, P., AND TUMBLIN, J. E. 2003. The trilateral filter for high contrast images and meshes. In *Proc. of Eurographics Symp. on Rendering*.

DECARLO, D., AND SANTELLA, A. 2002. Stylization and abstraction of photographs. *ACM Trans. on Graphics 21*, 3. Proc. of ACM SIGGRAPH conf.

(a) model (811x1044)

(b) input (795x532)

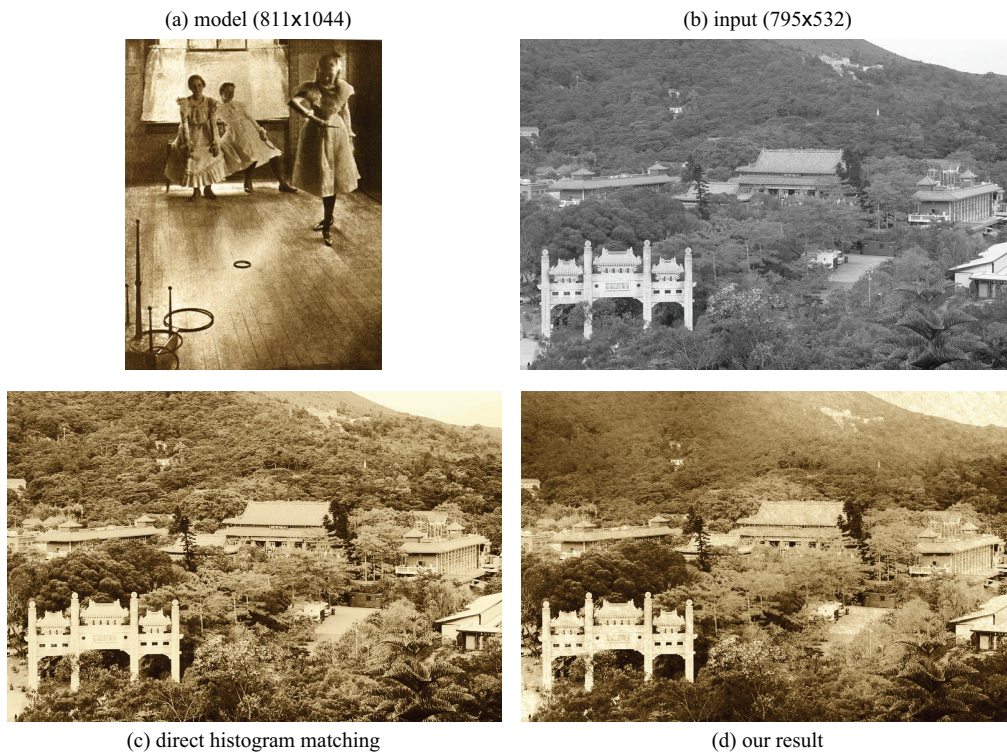(c) direct histogram matching

(d) our result

Figure 15: A simple histogram matching from the model (a) to the input (b) increases the texture level of the image (c) whereas the model has little texture. In comparison, we successfully reduce the texture and the sharpness to achieve large uniform gray regions similar to those in the model. The model is Ring Toss by Clarence H. White.

DRORI, I., COHEN-OR, D., AND YESHURUN, H. 2003. Example-based style synthesis. In *Proc. of IEEE conf. on Comp. Vision and Pattern Recognition*.

DURAND, F., AND DORSEY, J. 2002. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. on Graphics 21*, 3. Proc. of ACM SIGGRAPH conf.

EISEMANN, E., AND DURAND, F. 2004. Flash photography enhancement via intrinsic relighting. *ACM Trans. on Graphics 23*, 3. Proc. of ACM SIGGRAPH conf.

FATTAL, R., LISCHINSKI, D., AND WERMAN, M. 2002. Gradient domain high dynamic range compression. *ACM Trans. on Graphics 21*, 3. Proc. of ACM SIGGRAPH conf.

GEIGEL, J., AND MUSGRAVE, F. K. 1997. A model for simulating the photographic development process on digital images. In *Proc. of ACM SIGGRAPH conf.*

GONZALES, R. C., AND WOODS, R. E. 2002. *Digital Image Processing*. Prentice Hall.

GOOCH, A. A., OLSEN, S. C., TUMBLIN, J., AND GOOCH, B. 2005. Color2gray: Salience-preserving color removal. *ACM Trans. on Graphics 24*, 3. Proc. of ACM SIGGRAPH conf.

HEEGER, D. J., AND BERGEN, J. R. 1995. Pyramid-based texture analysis/synthesis. In *Proc. of ACM SIGGRAPH conf.*

HERTZMANN, A., JACOBS, C. E., OLIVER, N., CURLESS, B., AND SALESIN, D. H. 2001. Image analogies. In *Proc. of ACM SIGGRAPH conf.*

LI, Y., SHARAN, L., AND ADELSON, E. H. 2005. Compressing and companding high dynamic range images with subband architectures. *ACM Trans. on Graphics 24*, 3. Proc. of ACM SIGGRAPH conf.

PARIS, S., AND DURAND, F. 2006. A fast approximation of the bilateral filter using a signal processing approach. In *Proc. of Eur. Conf. on Comp. Vision*.

PATTANAIK, S., FERWERDA, J., FAIRCHILD, M., AND GREENBERG, D. 1998. A multiscale model of adaptation and spatial vision for realistic image display. In *Proc. of ACM SIGGRAPH conf.*

PETSCHNIGG, G., AGRAWALA, M., HOPPE, H., SZELISKI, R., COHEN, M. F., AND TOYAMA, K. 2004. Digital photography with flash and no-flash image pairs. *ACM Trans. on Graphics 23*, 3. Proc. of the ACM SIGGRAPH conf.

PÉREZ, P., GANGNET, M., AND BLAKE, A. 2003. Poisson image editing. *ACM Trans. on Graphics 22*, 3. Proc. of ACM SIGGRAPH conf.

REINHARD, E., STARK, M., SHIRLEY, P., AND FERWERDA, J. 2002. Photographic tone reproduction for digital images. *ACM Trans. on Graphics 21*, 3. Proc. of ACM SIGGRAPH conf.

REINHARD, E., WARD, G., PATTANAIK, S., AND DEBEVEC, P. 2005. *High Dynamic Range Imaging*. Morgan Kaufmann Publishers.

RUDMAN, T. 1994. *The Photographer's Master Printing Course*. Focal Press.

SU, S. L., DURAND, F., AND AGRAWALA, M. 2005. De-emphasis of distracting image regions using texture power maps. In *Proc. of Int. Workshop on Texture Analysis and Synthesis*.

TOMASI, C., AND MANDUCHI, R. 1998. Bilateral filtering for gray and color images. In *Proc. IEEE Int. Conf. on Comp. Vision*.

TUMBLIN, J., AND TURK, G. 1999. LCIS: a boundary hierarchy for detail-preserving contrast reduction. In *Proc. of ACM SIGGRAPH conf.*

(a) model (622x512)

(b) input with auto-levels (876x584)

(c) direct histogram transfer

(d) our result

Figure 16: Our approach is able to reproduce the level of texture observed in Adams' masterpiece (a) to achieve a compelling rendition (d). In comparison, Adobe® Photoshop® "auto-level" tool spans the image histogram on the whole intensity range. This reveals the small features of a picture but offers no control over the image look (b). And, a direct histogram transfer only adjusts the overall contrast and ignores the texture, thereby producing a dull rendition (c).

(a) input image (1200x900)
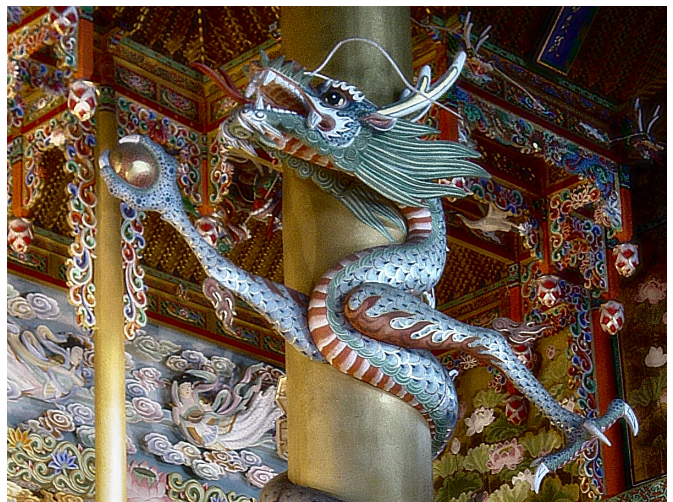
(b) our result

Figure 17: For color images, we process the luminance channel of the image and keep the original chrominance channels. In this example, the details are enhanced while the overall contrast and sharpness are increased. We used Adams' picture (Fig. 16a) as a model.