Enhancing Photographs with Near Infrared Images

Xiaopeng Zhang, Terence Sim, Xiaoping Miao
School of Computing
National University of Singapore
{zhangxi7,tsim,miaoxiao}@comp.nus.edu.sg

Abstract

Near Infra-Red (NIR) images of natural scenes usually have better contrast and contain rich texture details that may not be perceived in visible light photographs (VIS). In this paper, we propose a novel method to enhance a photograph by using the contrast and texture information of its corresponding NIR image. More precisely, we first decompose the NIR/VIS pair into average and detail wavelet subbands. We then transfer the contrast in the average subband and transfer texture in the detail subbands. We built a special camera mount that optically aligns two consumer-grade digital cameras, one of which was modified to capture NIR. Our results exhibit higher visual quality than tone-mapped HDR images, showing that NIR imaging is useful for computational photography.

1. Introduction

The radiance from natural scenes usually spans a very wide dynamic range, far exceeding what a digital camera can capture. For instance, in a sunny outdoor environment, the dynamic range could reach as high as 10^9. In contrast, a professional-grade digital camera that uses 14 bits per channel can capture a range of only 10^4. Consumer-grade cameras are even worse. One common technique around this problem is to first compute an high dynamic range (HDR) image, usually from multiple shots of varying exposures, and then to map this into a lower dynamic range (LDR) image suitable for display devices. However, such a tone-mapping procedure does not usually produce a perceptually pleasing result. Usually, pixels end up becoming too bright or too dark, and rich scene information such as color and texture are almost completely lost. Fig. 1(a) shows a typical photo taken under an HDR environment, where the footpath is very bright but the region inside the building can barely be seen.

Recently, many methods have been proposed to recover HDR radiance map and convert it to LDR images. Debevec and Malik presented a way to recover HDR radiance from multiply exposed photographs [4]. Mapping HDR to LDR, also known as tone mapping, can be roughly divided into two categories: spatially uniform mapping and spatially varying mapping. For instance, Reinhard *et al.* used a spatially uniform operator to compress dynamic range globally and then manipulated contrast locally based on luminance values [13]. A review of tone mapping techniques can be found in [5]. From HDR photos, it is possible to retrieve all details and color correctly. But obtaining an HDR map requires multiple images captured with different exposures. This in turn requires the scene to be static, which greatly limits its applicability.

Another class of techniques involves modifying the imaging sensor. For example, Nayar and Branzoi [10] use an LCD panel to modulate the light falling onto the sensor, while Nayar *et al.* [11] use a digital micromirror array for the same purpose. The goal is to adaptively control the exposure of small groups of pixel according to the radiance falling on them.

Another possible solution, widely used by professional photographers, is to take photos in RAW format and manually adjust contrast region by region. Usually RAW pictures use 12 or 14 bits per channel to record scene radiance, thus resulting in a higher dynamic range than normal JPEG photos. Such manual adjustment is tedious and requires experience, and the dynamic range of RAW format is still quite limited.

In contrast, our method uses Near Infrared (NIR) light. This lies between visible red light and Long Infra-Red (LIR) light in the electromagnetic spectrum. NIR light has wavelength in the range $750 - 1400$ nm, which is longer than visible light ($380 - 750$ nm). Human eyes can not see NIR light but most digital cameras can sense it very well. For example, some models of SONY digital cameras or camcorders have a *Night Shot* mode which increases cameras visual range by letting the sensor acquire more NIR light. However, most manufacturers insert an IR cutoff filter over the camera sensor to filter out NIR light, to avoid some unwanted artifacts. In fact, NIR images usually have better brightness contrast and provide rich texture details, as seen

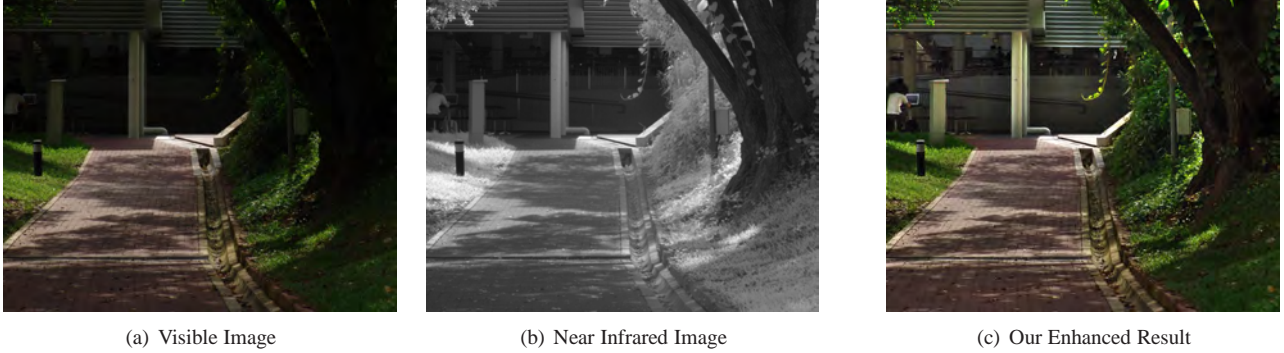| (a) Visible Image | (b) Near Infrared Image | (c) Our Enhanced Result |

Figure 1. We proposed a novel image enhancement method by transferring contrast and texture from near infrared image to visible image. Fig. 1(a) and Fig. 1(b) are an improper exposed photo taken under high dynamic range environment and its corresponding near infrared photo. With only these two input images, our approach can adaptively and automatically adjust contrast and enrich visible details in over- or under-exposed areas, as shown in Fig. 1(c).

in Fig. 1(a) and 1(b). The details of trees and leaves are barely seen in the visible image, but look clear and sharp in the NIR image. We exploit this fact in our work.

Inspired by the camera's ability to record NIR light and by recent works on tonal transfer [2, 12], we propose a novel method that can adjust a photograph's contrast adaptively and enrich texture details fully automatically with just one shot (*i.e.* one VIS/NIR image pair). NIR photography is not new; it is commonly appreciated for its artistic value [8], but has not been fully exploited in computational photography. Morris *et al*. observed that the wavelet coefficients of long infrared (LIR) natural images closely follows the Laplacian distribution [9]. Encouraged by their work, we build a dual-camera system that can capture visible photo and NIR photo of the same scene simultaneously, and find that NIR images also have similar statistical properties. Moreover, we notice that NIR images of natural scenes usually exhibit lower dynamic range and contain rich texture details. In terms of transfer techniques, the Neumann brothers showed how to transfer color style from a source image to an arbitrary target image by applying histogram matching [12]. Similarly, Bae *et al*. also presented a method to transfer tonal quality from one image to another, using histogram matching and bilateral filter [2]. In the light of their work, we propose an original way of using NIR information to enhance visible photographs. Given as input one VIS image and its corresponding NIR image, our approach can adaptively detect unsatisfactory pixels in the VIS image, and transfer contrast information and high frequency texture from its NIR counterpart. We use histogram matching in the gradient domain to transfer contrast and use wavelet coefficients to transfer texture information. We were able to achieve very pleasing results (see Fig.1).

The highlights of our work include:

- A novel method that uses NIR information to adjust the contrast of photographs adaptively, and to enrich

texture details automatically with one single shot.

- A threshold-free method to detect regions that require enhancement.

- A study of the statistical properties of NIR images of natural scenes.

As far as we can tell, we are the first to use NIR for photograph enhancement.

2. Near Infrared Imaging

2.1. Dual-camera system

NIR light lies adjacent to visible red light in the electromagnetic spectrum, and has longer wavelength than visible light. NIR is not visible to human eyes, but can be recorded by CCD or CMOS sensors. However, most manufacturers of digital cameras install an IR cutoff filter over the sensor to suppress infrared light and avoid unwanted artifacts.

To capture both visible and NIR pictures for the same scene simultaneously, we built a dual-camera system which comprises two Sony F828 digital cameras and one hot mirror. A hot mirror is a specialized dielectric mirror which can reflect NIR light when incident light arrives at a certain angle. We used a $45°$ hot mirror, meaning it can reflect NIR light with angle of incidence of $45°$ but does not block visible light. Fig. 2(a) illustrates how our system works. Although the Sony F828 has built-in *Night Shot* mode which can temporarily move the IR cutoff filter away to allow NIR imaging, Sony has intentionally limited such NIR imaging to only allow long exposure times. Our modified camera does not suffer from this limitation. We also modified the remote control of the camera so that it can trigger two cameras at the same time. We have carefully setup two cameras to ensure that they are optically aligned. They also share the same camera settings, such as focal length and aperture

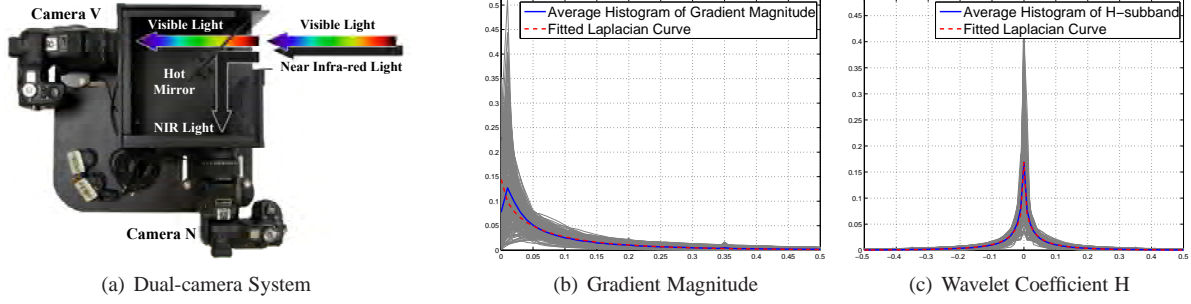| (a) Dual-camera System | (b) Gradient Magnitude | (c) Wavelet Coefficient H |

Figure 2. (a): Our VIS-NIR dual-camera prototype. Camera V and N are optically aligned and connected to the same remote control, allowing a VIS/NIR image pair of the same scene to be captured with a single shot. (b-c): Statistical properties of NIR images. (b) shows distribution of gradient magnitude, similar to statistics of visible images [7]. (c) shows distribution of H wavelet subband of Haar transform, similar to statistics of IR images [9]. Subbands V and D have similar distribution.

size, to guarantee the geometric alignment of the image pair. Currently, we do not force the two cameras to use the same shutter speed, because digital cameras are designed to be less sensitive to NIR thus requiring a slightly longer exposure.

The NIR picture captured in this way is actually an RGB color image and looks reddish since NIR light is just adjacent to red light. However, because of the filters we use, the NIR light we capture is almost monochromatic and should not contain any color information. So we use only intensity information by converting to *HSV* color space and using *V* channel. Fig.1(a) and Fig. 1(b) show an example image pair captured by our dual-camera system. Our prototype hardware may look bulky, but this can be miniaturized. Our goal is to show the usefulness of NIR images.

2.2. Statistics of NIR images

Huang and Mumford [7] have shown that the gradient histograms of natural images follow a generalized Laplace distribution which can be expressed as Eq. 1:

$$P(x) = k \cdot e^{-|x/s|^{\alpha}}. \tag{1}$$

Recently, Morris *et al.* [9] found wavelet coefficients of LIR (wavelength lies in $4000 - 120000$ nm) images of natural scenes can also be well fitted with a Laplacian curve. In this paper, we show that NIR natural images share similar statistical properties, as illustrated in Fig. 2. We collect a total of 220 NIR photos for statistical analysis. Some of them are collected from web and others are captured by ourselves, mostly covering subjects of natural scene and people. Similar to [7, 9], we use gradient magnitude and the Haar wavelet coefficients. We calculate the histograms of all images for both gradient magnitude and wavelet coefficients in horizontal (H), vertical (V) and diagonal (D) directions. All these histograms are calculated on logarithm of the actual values and normalized based on image pixels. In Fig. 2, all gray lines denote the actual histograms, the blue

lines show the average histogram distribution, and the red dash lines show the fitted Laplacian curve (Eq. 1). We can see that the fit is good, meaning that NIR images have similar statistical properties as visible and LIR images (Please refer details to [7, 9]). In Sec.3.2.2 we will show how we can use these statistical properties to guide the enhancement process.

3. Visible Image Enhancement

The workflow of our approach is illustrated in Fig. 3. There are three main steps: computing the weighted region mask, transferring contrast and transferring texture. The details are explained in Sec. 3.1, Sec. 3.2 and Sec. 3.3 respectively. Note that all inputs are the logarithm of original image values as we mentioned in the previous section.

3.1. Computing the weighted region mask

Intuitively, regions that suffer a loss of details are typically too bright or too dark, and have low saturation. From this observation, a weighted mask can be calculated based on saturation and brightness value. Let W_s and W_v denote weighted mask of saturation and brightness, W denote the final weighted region mask indicating areas to be enhanced. Then W can be obtained using following equations:

$$
\begin{aligned}
W_s &= 1 - e^{-p_s|s-1|}, \quad p_s \in [0,1], \ s \in [0,1] \\
W_v &= 1 - e^{-p_v|v-0.5|}, \quad p_v \in [0,1], \ v \in [0,1] \\
W &= W_s \cdot W_v,
\end{aligned} \tag{2}
$$

where s and v are the saturation and brightness intensity, and p_s and p_v denote the probability that s and v appear in visible image respectively. p_s and p_v can be easily obtained from the normalized histograms of channels *S* and *V*. The meaning of p_s and p_v is that the pixels to be enhanced should distribute over large areas, rather than in small regions. Enhancing large areas while ignoring small regions usually achieves better perceptual quality.
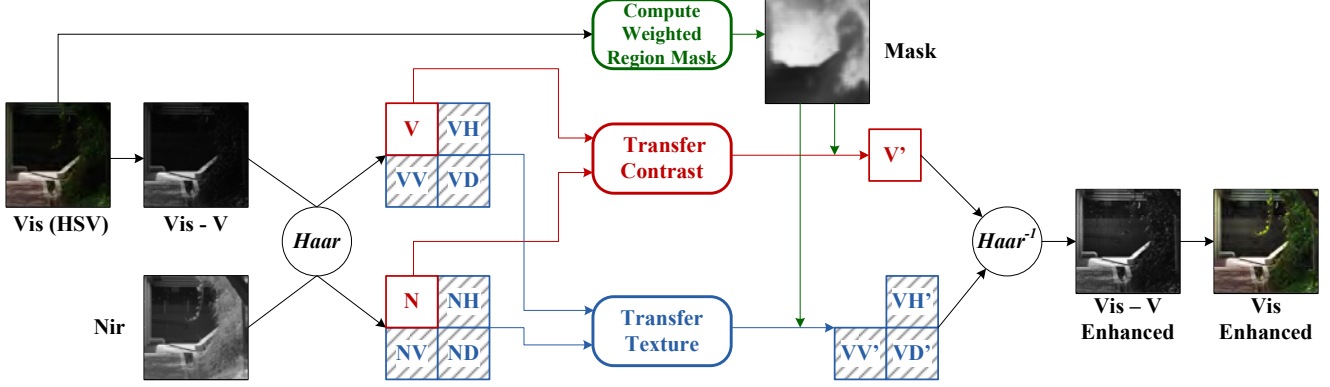
Figure 3. The workflow of our approach. The enhancement process uses the Haar wavelet decomposition, and comprises three major steps: computing the weighted region mask, transferring contrast, and transferring texture. See text for details.
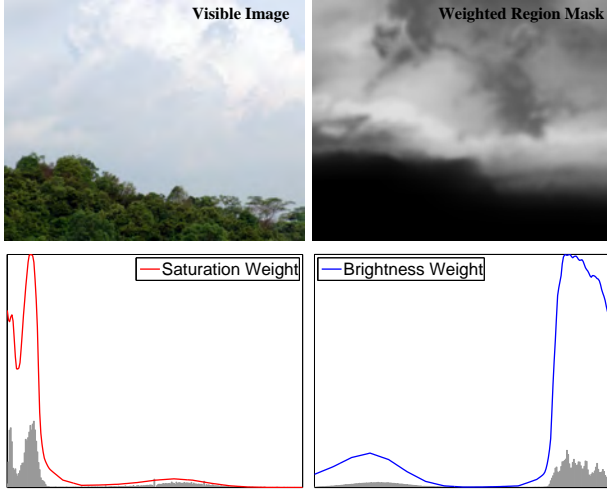

Figure 4. Weighted region mask computation. The second row shows the histograms of saturation and brightness channel (gray regions). The red and blue lines illustrate the weights computed according to Eq. 2. The curves are smoothed for reducing noise. The sky and cloud have high brightness and relatively low saturation, and thus have higher weights; vice versa for the tree tops.

Note that W is calculated adaptively and fully automatically, not requiring any thresholds. The weighted region mask is used as a mask for brightness and texture transfer later. Fig. 4 shows what our weighted region mask looks like and how it is generated. A higher value in W means more information will be transferred from NIR image, and vice versa. To reduce noise, a Gaussian blurring is first applied on W.

3.2. Transferring contrast

Our contrast transfer method is based on histogram matching and gradient techniques. Bae *et al*. also applied histogram matching to transfer photograph tonality [2]. Instead of matching histogram in the intensity domain as they

did, we show that histogram matching in the gradient magnitude can achieve better and reliable results.

3.2.1 Histogram matching

The histogram matching problem can be simply defined as: given an image I and a target histogram (pdf) $h(z)$, the problem is to find a new image J by transforming I, so as to make histogram of J be as same as h. The problem can be solved by using the cumulative distribution function (CDF), f. Define $f(x) = \int_o^x h(z)dz$, where x is image intensity. Let I_{ij} and J_{ij} denote each pixel intensity in I and J. Then the desired image J can be obtained using Eq. 3, and the detailed proof can be found in [6].

$$J_{ij} = f_J^{-1}(f_I(I_{ij})) \tag{3}$$

3.2.2 Large-Scale contrast transfer

The brightness contrast of a visible image is affected by environment illumination, as well as object shape and texture in the scene. Therefore, the brightness map of an image should change smoothly while preserving major features such as strong edges. To achieve a smooth brightness map of visible image V and NIR image N (V and N are actually the average subbands in the Haar decomposition, as shown in Fig. 3), we apply bilateral filtering [14] to decompose images to large-scale layer and detail layer, and use the larger-scale layer as brightness map, as in Eq. 4:

$$\begin{aligned} V_L &= bf(V), \quad V_D = V - V_L \\ N_L &= bf(N), \quad N_D = N - N_L. \end{aligned} \tag{4}$$

V_L and N_L are large-scale layers, and V_D, N_D are corresponding detail layers (after taking the logarithm). We use a similar definition for the bilateral filter function bf and parameter selection as in Bae *et al*.'s work.

We implement three different methods to transfer contrast from the NIR image to the VIS image. A comparison of their results can found in Fig. 5 and 6.

Method 1: Histogram Matching Inspired by Bae *et al.*'s method [2], we can simply match intensity histogram of V_L with N_L to transfer intensity distribution. This method is easy and efficient, but histogram matching blindly alters pixel values and thus very possibly destroy illumination consistency. From Fig. 5, we see that histogram matching does improve the contrast significantly. However, we also see that pixels in the tree bark are over brightened and inconsistent with the illumination in the original image. After applying the gradient constraint, the result looks more natural.

Method 2: Histogram Matching with Gradient Constraint To maintain illumination consistency, we can check the gradient direction of the altered brightness map pixel by pixel. Once we find the gradient direction that is reversed or changed too much from the original brightness map, we force them to be zero. After applying the gradient constraint, the enhanced result looks more natural compared with method 1 (see Fig. 5). But in some cases, where gradients change abruptly along their original directions due to the histogram matching step, this constraint will fail, as shown in Fig. 6. The gradient constraint cannot remove the banding-effect on the pillar and wall, because the gradients in those areas are not actually reversed.

Method 3: Gradient Magnitude Matching To strictly maintain smoothness of the transferred brightness map, we match the histogram of brightness gradient magnitude instead of brightness intensity. We define V_G and N_G as the gradient magnitude of V_L and N_L:

$$V_G = \sqrt{V_{G_x}^2 + V_{G_y}^2} = \sqrt{(\frac{\partial V_L}{\partial x})^2 + (\frac{\partial V_L}{\partial y})^2}$$
$$N_G = \sqrt{N_{G_x}^2 + N_{G_y}^2} = \sqrt{(\frac{\partial N_L}{\partial x})^2 + (\frac{\partial N_L}{\partial y})^2}. \qquad (5)$$

In Sec. 2.2 we have shown that gradient magnitude histogram of NIR image can be well fitted with a generalized Laplacian curve. Because N_L is a smoothed version of the NIR image, its gradient magnitude N_G also has same statistical property. Let l denote the Laplacian curve that can fit histogram of N_G. Instead of matching histogram of V_G with histogram of N_G directly, we use l as the target histogram to produce a smoother and noise-free distribution transfer. In this case, the functions f_I and f_J in Eq. 3 are the CDFs of l. Let $V_{G'}$ denote the histogram matching result, we can easily compute new gradients by scaling V_{G_x} and V_{G_y} along their original directions respectively:

$$V_{G'_x} = \frac{V_{G'}}{V_G} \cdot V_{G_x}$$
$$V_{G'_y} = \frac{V_{G'}}{V_G} \cdot V_{G_y}. \qquad (6)$$

(a) Visbile Image (b) NIR Image
(c) Histogram Matching (Method 1) (d) HM + Gradient Constraint (Method 2)
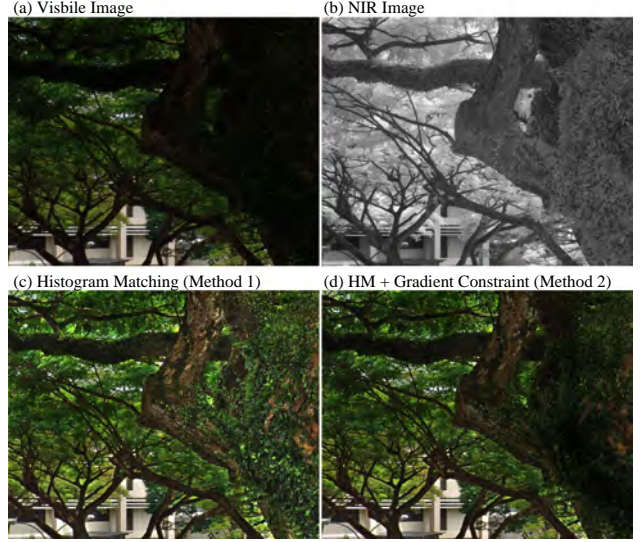
Figure 5. Comparison of histogram matching (method 1) with histogram matching with gradient constraint (method 2). The result of histogram matching (Fig. 5(c)) looks artificial because it breaks the consistency of overall brightness distribution. After applying gradient constraints, Fig. 5(d) looks more natural.

From $V_{G'_x}$ and $V_{G'_y}$, we reconstruct new large-scale brightness map $V_{L'}$ by using Agrawal *et al.*'s improved Poisson solver [1]. The final contrast transferred V' is obtained by blending enhanced brightness map and its original version V together using alpha-blending

$$V' = W \cdot (V_{L'} + V_D) + (1 - W) \cdot V, \qquad (7)$$

where the weighted map W is used as the alpha channel and $|\cdot|$ denotes pixel-wise multiplication. This method naturally maintains illumination consistency and achieves the best result among these three methods. See Fig.6 for comparison. Note that the banding-effects of methods 1 and 2 are completely suppressed, while overall contrast has been improved.

3.3. Transferring texture

As we state in our workflow (Fig.3), after applying Haar wavelet transformation, the wavelet subbands in horizontal, vertical, and diagonal directions actually contain rich texture information. To transfer those details, we use alpha blending again to combine corresponding subbands together:

$$VH' = W \cdot NH + (1 - W) \cdot VH. \qquad (8)$$

VV' and VD' are obtained similarly. The new subbands VH', VV', and VD' not only inherit texture details from the VIS image, but are also enhanced by rich high frequency details from NIR image. Fig. 7(g) show the result with high frequency details transferred. The textures on the roof in the original image are almost lost completely. By transferring

(a) Visible Image

(b) Histogram Matching (Method 1)

(c) HM + Gradient Constraint (Method 2)

(d) HM of Gradient Magnitude (Method 3)

Figure 6. Comparison of Methods 1, 2, and 3. Note that the band-effects (regions in red box) due to blind histogram matching have been successfully suppressed by our guided histogram matching of gradient magnitude. The result of Method 3 achieves the least artifacts and best perceptual quality.

high frequency details from NIR to visible image, those lost textures are successfully recovered, and those weak textures are also reinforced greatly. Finally, we apply inverse Haar wavelet transform to enhance the V channel.

4. Experiments and Results

A common HDR scene is the natural outdoor environment under bright sunlight. To demonstrate the strength of our techniques, we test our approach with pictures taken in such HDR situations. All input visible and NIR image pairs have been geometrically aligned. In outdoor daylight, tree leaves and some objects, such as cloth and skin, reflect NIR light strongly, so they look bright and have much details even in shaded areas (see Fig. 1, 8, and 9). Such features in NIR images are useful for enhancing visible images.

We also find that contrast transfer and texture transfer are both equally important for enhancement. As shown in Fig. 7: Fig. 7(e) is the enhanced result with only texture

(a) Original Visible Image   (b) Weighted Region Mask

(c) Cropped Roof Visible Image

(d) Cropped Roof NIR Image

(e) Texture Transfer Only

(f) Contrast Transfer Only
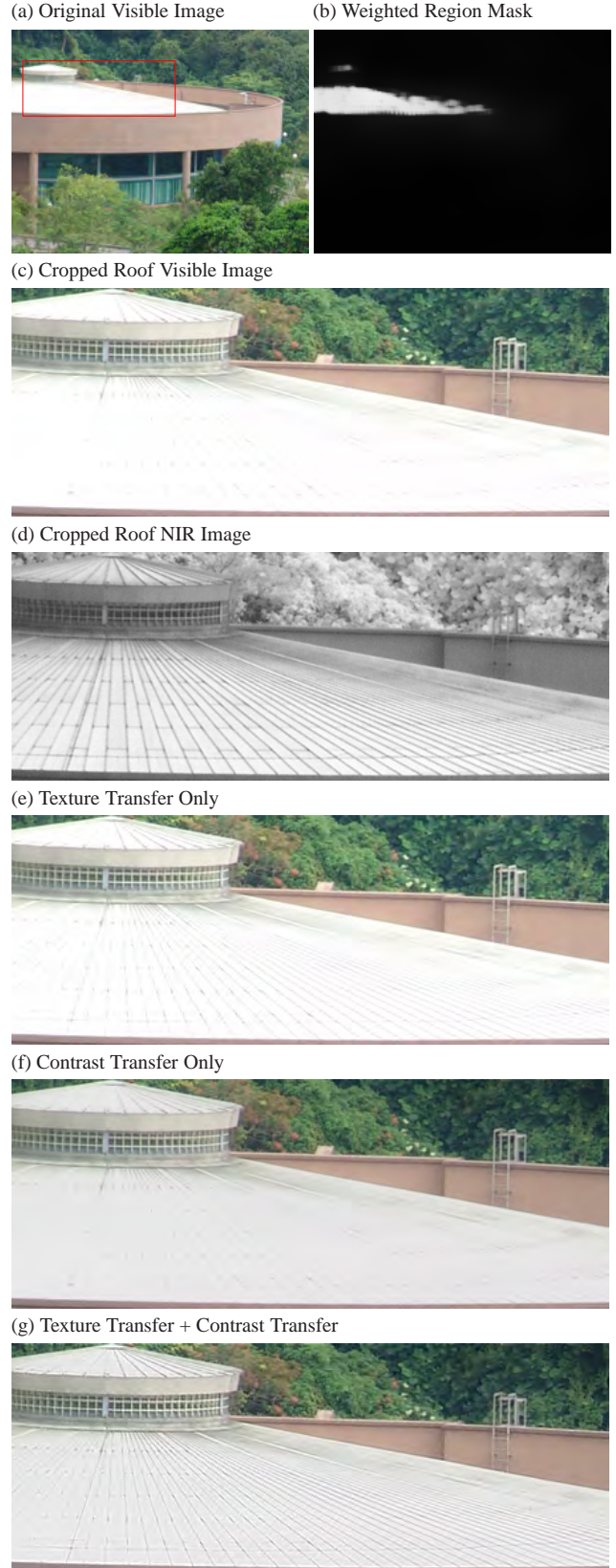
(g) Texture Transfer + Contrast Transfer

Figure 7. Comparison of results with either contrast transfer or texture transfer. Fig. 7(c-g) show the zoomed-in roof details.

transferred, where most of roof details are successfully recovered but the picture still looks over-exposed; Fig. 7(f) is the result with only contrast transferred, which has lower contrast but roof details are still lost. Obviously, after transferring the contrast and texture, Fig. 7(g) exhibits better visual quality.

To show that histogram matching of gradient magnitude (Method 3) can preserve overall illumination map and achieve higher perceptual quality, we compare of our results with a naively blended output (see Fig. 8(d) and Fig. 9(d)). This trivial result is obtained using alpha-blending of V and N based on the weighted region mask W, *i.e.* using each pixel value in W as alpha value.

We also compare our results with tone-mapped HDR images of same scenes, as shown in Fig. 8(h) and Fig. 9(e). To get the HDR image, we take multiple images with different exposure (usually 5-7 pictures with exposure difference of 1 stop), and assemble them using the HDR Shop software developed by Debevec [3]. We recovered the camera response curve of our camera to generate HDR image more precisely, and we applied Reinhard *et al.*'s algorithm for tone mapping [13]. Tone-mapped HDR images are supposed to produce a range-compressed image with rich details and high visible quality. However, tone mapping algorithms usually have a strict assumption that the scene must be static, *i.e.* no moving objects throughout the whole image sequence. Such an assumption is easily broken in outdoor scenes, as shown in Fig. 8(h). The walking pedestrian and leaves waving in the wind cause serious "ghosting effect" (shown in red boxes) in the tone mapped results. Because the inputs of our approach are captured in a single shot, our results are free of such artifacts. Besides, our approach preserves consistency of overall illumination distribution while recovering scene details, therefore our results gain better perceptual quality on brightness contrast than tone-mapped results, as shown in Fig. 8(h) and Fig. 9(e).

5. Conclusions and Discussions

In this paper, we presented an approach of enhancing visible photograph using NIR information based on a dual-camera prototype. Without manual segmentation or interaction, our method can calculate the enhancing weight for each pixel automatically and transfer brightness and texture details from the NIR image to the visible image. We show that our histogram matching of gradient magnitude can well maintain large-scale illumination and achieve better perceptual quality. Also, by combining the wavelet H, V, and D subbands of NIR and visible image high frequency details are effectively enhanced.

Since common digital cameras are capable of recording NIR light, ours is a practical method to enhance LDR images. Compared with tone mapping an HDR image, our method requires only one shot (*i.e.* one VIS/NIR image pair). Benefiting from the better contrast in the NIR images and a spatially varying weighted mask, our method can produce results that are aesthetically more pleasing than tone-mapped methods.

As far as we can tell from searching the published literature, we are the first to make use of NIR information for photograph enhancement. We believe that the unique property of NIR light has great potential for computational photography. Photography in low light condition may also benefit from NIR imaging, because in low visible light situation there could still be sufficient NIR light and objects that strongly reflect NIR can be captured. In the near future, we hope to investigate how NIR images can be used to guide HDR tone-mapping.

6. Acknowledgments

References

[1] A. Agrawal, R. Raskar, and R. Chellappa. What is the range of surface reconstructions from a gradient field? In *ECCV*, 2006. 5

[2] S. Bae, S. Paris, and F. Durand. Two-scale tone management for photographic look. In *SIGGRAPH*, 2006. 2, 4, 5

[3] P. Debevec. HDR Shop. http://www.hdrshop.com/. 7

[4] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH*, 1997. 1

[5] K. Devlin. A review of tone reproduction techniques. Technical report, University of Bristol, 2002. 1

[6] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice Hall, 2002. 4

[7] J. Huang and D. Mumford. Statistics of natural images and models. In *CVPR*, 1999. 3

[8] C. Maher and L. Berman. Fine art infrared photography. http://www.infrareddreams.com/. 2

[9] N. J. W. Morris, S. Avidan, W. Matusik, and H. Pfister. Statistics of infrared images. In *CVPR*, 2007. 2, 3

[10] S. Nayar and V. Branzoi. Adaptive Dynamic Range Imaging: Optical Control of Pixel Exposures over Space and Time. In *ICCV*, 2003. 1

[11] S. Nayar, V. Branzoi, and T. Boult. Programmable Imaging using a Digital Micromirror Array. In *CVPR*, 2004. 1

[12] L. Neumann and A. Neumann. Color style transfer techniques using hue, lightness and saturation histogram matching. In *Computational Aesthetics in Graphics, Visualization and Imaging*, 2005. 2

[13] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. In *SIGGRAPH*, 2002. 1, 7

[14] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV*, 1998. 4

(a) Visible Image   (b) NIR Image   (c) Weighted Region Map   (d) Alpha-blending

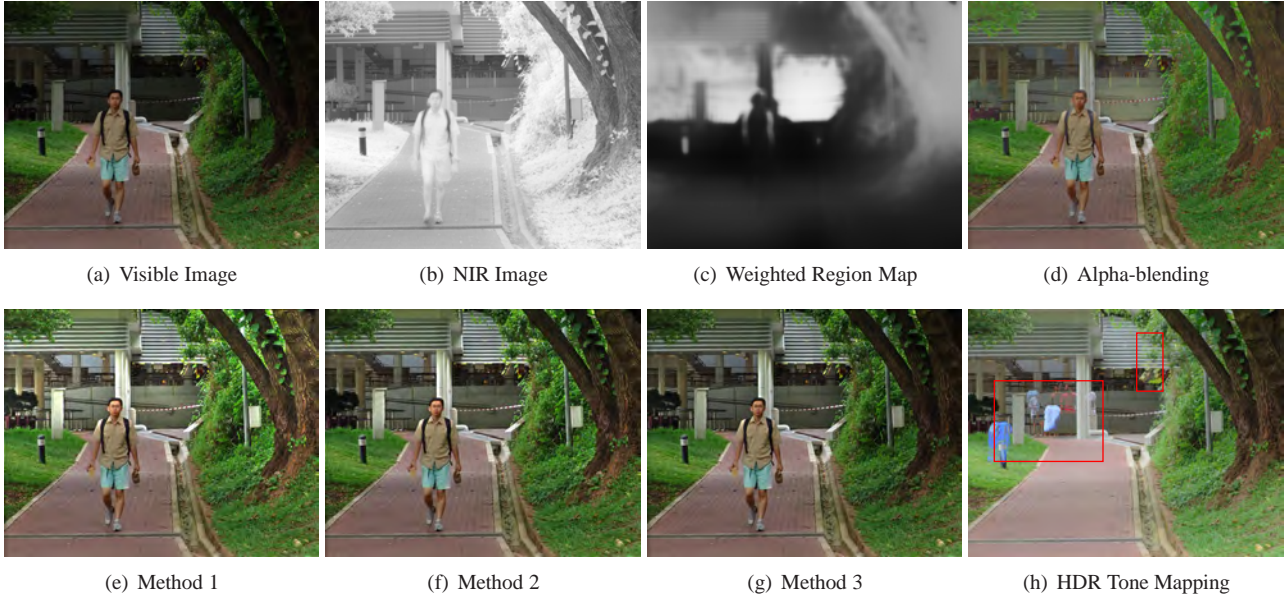(e) Method 1   (f) Method 2   (g) Method 3   (h) HDR Tone Mapping

Figure 8. Comparison of our approach with alpha-blending and HDR tone mapping. The naive alpha-blending result appears bad, since simple pixel-wise blending cannot transfer overall contrast. As for HDR tone mapping, the result exhibits "ghosting effect" (shown in red boxes) because objects have moved during the capture of multiple exposures.

(a) Visible Image   (b) NIR Image   (c) Weighted Region Map

(d) Alpha-blending   (e) HDR Tone Mapping   (f) Our Result (Method 3)

Figure 9. Another comparison of our approach with alpha-blending and HDR tone mapping. Our method successfully enhances brightness and texture by using NIR information.