Available online at www.sciencedirect.com
ScienceDirect
J. Vis. Commun. Image R. 18 (2007) 366–376
JOURNAL OF VISUAL Communication & IMAGE Representation
www.elsevier.com/locate/jvci

Noise reduction in high dynamic range imaging

Ahmet Oğuz Akyüz [a,b,*], Erik Reinhard [a,c]

[a] University of Central Florida, 4000 Central Florida Blvd. Orlando, FL 32816, USA
[b] Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, Tuebingen 72076, Germany
[c] University of Bristol, Department of Computer Science, Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, UK

Received 9 November 2006; accepted 3 April 2007
Available online 13 April 2007

Abstract

A common method to create high dynamic range (HDR) images is to combine several different exposures of the same scene. In this approach, the use of higher ISO settings will reduce exposure times, and thereby the total capture time. This is advantageous in certain environments where it may help minimize ghosting artifacts. However, exposures taken at high sensitivity settings tend to be noisy, which is further amplified by the HDR creation algorithm. We present a robust and efficient technique to significantly reduce noise in an HDR image even when its constituent exposures are taken at very high ISO settings. The method does not introduce blur or other artifacts, and leverages the wealth of information available in a sequence of aligned exposures.

Keywords: Noise; Noise reduction; High dynamic range (HDR) imaging; Image enhancement

1. Introduction

High dynamic range imaging allows the capture of faithful representations of real world scenes. Therefore it is rapidly becoming more widely used in film, photography, and computer graphics. A common technique to create HDR images is to merge multiple low dynamic range (LDR) frames, each taken with a different exposure time [1–3]. Digital cameras that allow the user to change the exposure time per frame are suitable for this purpose. However, a weakness of current digital cameras is that they tend to be noisy under low light conditions or high sensitivity settings. The process of HDR creation will amplify noise present in the shorter exposures, which may render the result unusable.

In this paper, we present a method to effectively reduce the noise in an HDR image by processing its constituent frames. The resultant HDR image will have a significantly higher signal-to-noise ratio as shown in Fig. 1. The ability to remove noise allows photographers to capture at higher speeds (i.e. higher ISO settings) thus mitigating the inherent limitations of the multiple exposure technique such as the occurrence of ghosting.[1] It would also allow HDR videos to be captured at higher frame rates [4].

We first briefly review noise and its characteristics in Section 2, and explain the HDR assembly process in Section 3. We present our novel HDR noise reduction algorithm in Section 4. Results are given in Section 5, followed by a comparison with the current state-of-the-art in Section 6. Conclusions are drawn in Section 7.

2. Sources of noise

Noise can be defined as undesired random degradations in images which may occur during capture, transmission, and processing. There are three primary sources of noise in digital cameras, namely photon shot noise, dark current noise, and read noise. Photons arrive at random intervals

[1] Doubling the ISO value requires halving the exposure time to record the same image. Thus, a camera can capture the same scene 128 times faster at ISO 6400 than at ISO 50. This gain could be crucial in eliminating ghosting artifacts.

* Corresponding author. Address: University of Central Florida, USA.
E-mail address: oguz@cs.ucf.edu (A.O. Akyüz).

1047-3203/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2007.04.001

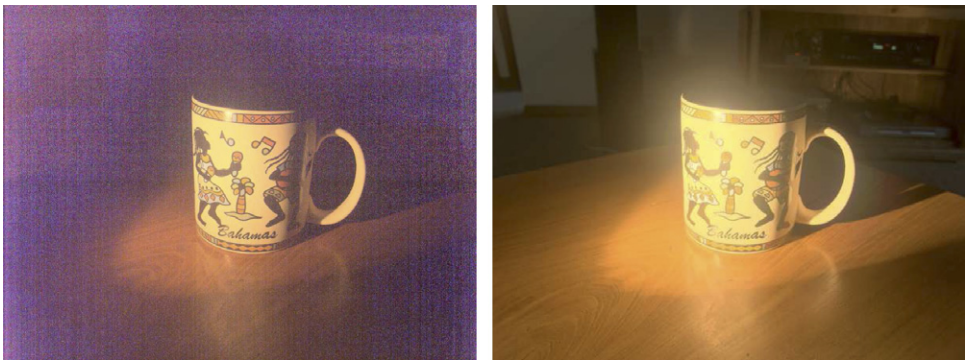

Fig. 1. A standard multiple exposure technique was applied to create the left image, whereas the right image shows a result obtained with techniques described in this paper.

at any detector, and impose a hard limit on the noise performance of digital cameras. The time between arrival of photons may be modeled with Poisson statistics. Thus, the uncertainty of the number of photons collected over a given period of time is linear in the square root of the signal amplitude. This type of noise is called *photon shot noise*, and cannot be reduced via better camera design.

Dark current noise, which is also governed by a Poisson distribution, is caused by the statistical variation of the number of thermally generated electrons due to the heating of the camera sensors. It may be reduced by cooling the sensors. Other types of noise that may occur during the amplification and quantization of signals are collectively referred to as *read noise*, which may be minimized by careful design of camera circuitry.

Noise removal methods are usually classified as linear and non-linear methods. Linear methods, such as Gaussian or (weighted) mean filters, are based on blurring the image and thus trade noise for visible artifacts [5].

Non-linear methods preserve the details of the image better although artifacts may still occur. These methods include order statistic filters, such as the median filter [6], morphological filters [7], the bilateral filter [8], anisotropic diffusion based techniques [9], and wavelet based noise reduction techniques [10].

If multiple images of the same scene are available, then these may be "frame averaged" [5]. Frame averaging is a blur free method if the frames are registered and there is no object movement in the captured scene. These requirements may make frame averaging difficult to apply in standard image processing, so linear or non-linear methods that require only one photograph of the captured scene may be preferred. However these requirements should already be met to create an HDR image, thus making frame averaging a sensible choice to be used in HDR imagery. In this paper, we show how to adapt frame averaging to remove noise before it amplifies during HDR creation.

As standard frame averaging techniques are not applicable when the exposure is varied between frames, the key contribution of our paper is to show how frame averaging can be extended for use with multiple exposure techniques.

3. High dynamic range image creation

Dependent on the scene, a digital photograph may have image areas that are over- and under-exposed, as well as properly exposed areas. If we make the exposure time longer, previously under-exposed regions may become properly exposed. Similarly, by making the exposure time shorter, previously over-exposed regions may become well exposed. Thus, with a sequence of differently exposed images of the same scene, all image areas are properly exposed in at least one image. We can create a high dynamic range image by combining these exposures, thereby exploiting the fact that each image area is well exposed in one exposure or the other.

To create an HDR image one would ideally bring the low dynamic range images into the same domain by dividing each image with its exposure time (i.e. normalization) and then summing the corresponding pixels of normalized images. These steps, however, cannot be directly performed because most cameras apply non-linear processing to the incident light as it passes through camera circuitry. The net effect of this non-linear processing is called the *camera response* and it should be inverted before creating an HDR image.

The non-linearity of a camera can be inverted if its response function is known. Response functions are usually proprietary and not disclosed by camera manufacturers. However, in recent years algorithms were developed to recover the response function of a camera by using only a set of images of the same scene taken with different exposure times. [1–3,11]. Once the response function is recovered from an image sequence, it can be used to linearize other images taken with the same camera.

After linearization, corresponding pixels from the image sequence are summed up to compute final irradiances. However in each exposure some pixels will be under or over-exposed, which contain no useful information. Therefore these pixels should be excluded from the summation. Also not all the pixels are equally reliable due to the non-linear processing of cameras. These issues can be addressed by applying a weight function to the pixels in each individual frame during the summation.

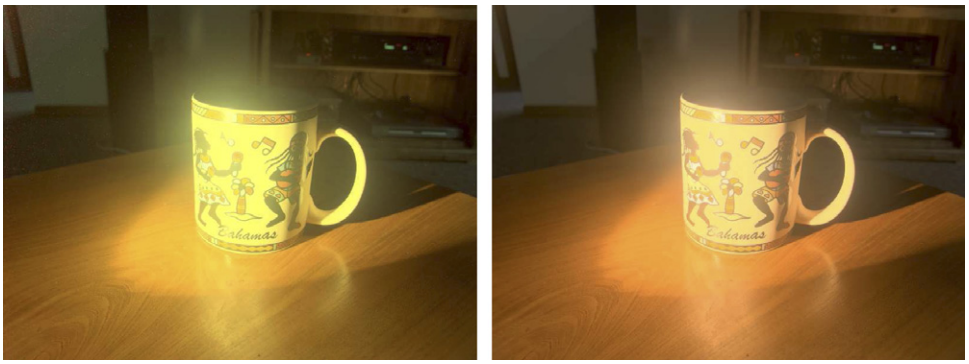

Fig. 2. The left image is created by using pixel values as input to the broad hat function. For the right image, the luminance of each pixel is used instead, which prevents the green color cast and reduces noise. Both of these images have relatively low noise compared to the left image in Fig. 1 since they are created from exposures captured at the lowest ISO setting of the camera whereas the image in Fig. 1 was created from exposures captured at the highest ISO setting (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Several methods have been proposed to select a proper weight function. Mann and Picard [1] and Robertson et al. [11] propose to use the derivative of the camera response function as their weighting function by arguing that the reliability of pixel values is correlated with the camera's sensitivity to light changes.

Debevec and Malik [2] use a simple hat shaped function by assuming that the pixels that are in the middle of the range are more reliable. Mitsunaga and Nayar [3] multiply Mann and Picard's weighting function by the linearized camera output since signal-to-noise ratio (SNR) increases with signal intensity. Ward suggests to multiply Mitsunaga and Nayar's weighting function with a broad hat filter to exclude dubious pixels near extremes [12].[2]

Although the weighting function proposed by Ward eliminates unreliable pixels near extremes it is vulnerable to fluctuations in pixel values. For instance it is not uncommon to have a pixel attain a value of 255 in a shorter exposure and drop to 254 or 253 in a longer exposure. Moreover this fluctuation may happen in each color channel independently causing a color cast in the final HDR image. For this reason we suggest to use the luminance of the pixel as the input to the broad hat function. In Fig. 2 we compare the results in case of using these two weighting functions. The weighting functions discussed above are given in Table 1 and plotted in Fig. 3.

All the methods described above alter the weighting function to find a good compromise between the noisy, under-/over-exposed, and usable pixels. A different approach is suggested by Grossberg and Nayar based on computing the optimum exposures before starting the capture process [13]. To this end, the authors minimize an objective function which includes a separate term to account for camera noise. Therefore, in a sense, the authors suggest a recipe to exclude noisy exposures upfront, albeit at the cost of solving a potentially lengthy optimization problem.[3]

Table 1
Several weighting functions used in HDR image generation

Mann and Picard $w(x) = f'(x)$	Debevec and Malik $w(x) = \begin{cases} x & x \leqslant 127 \\ 255 - x & x > 127 \end{cases}$
Mitsunaga and Nayar $w(x) = f^{-1}(x)f'(x)$	Ward $w(x) = f^{-1}(x)f'(x)[1 - (\frac{x}{127.5} - 1)^{12}]$
Our weighting function $w(x, L) = f^{-1}(x)f'(x)[1 - (\frac{L}{127.5} - 1)^{12}]$	

Here, $f(x)$ is the recovered camera response function and L is luminance.

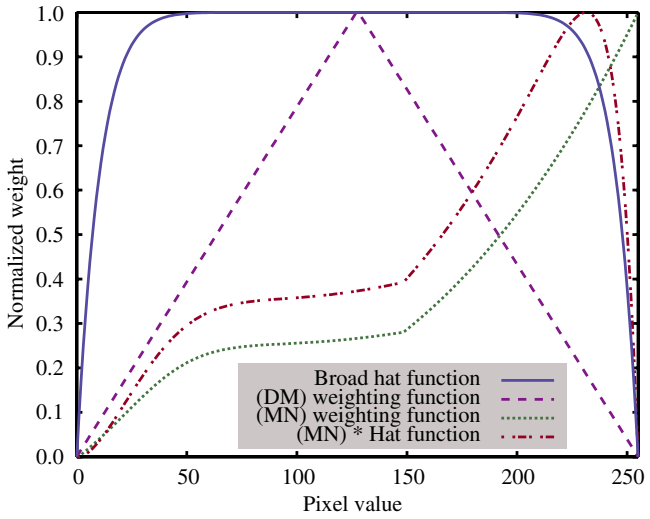

Fig. 3. Several weighting functions used in HDR creation. Debevec and Malik's (DM) function is independent of the camera response whereas Mitsunaga and Nayar's (MN), and (MN) times broad hat function are derived from the camera response. These weighting functions are derived from the red channel response of a Nikon D2H camera.

Having computed the camera response f and the weight function w, the final (scaled) irradiance value[4] I_p of pixel p

[2] The broad hat function used in this book is given by $1 - (\frac{x}{127.5} - 1)^{12}$.

[3] To minimize this cost, the authors derive exposure tables which could be embedded into the camera firmware.

[4] We call this quantity scaled irradiance because for absolute irradiances the camera should be calibrated in SI units.

Fig. 4. The top row shows a selection of exposures used to create the left image in Fig. 1. The bottom row shows the corrected versions of these images with our technique, which yielded the right image in Fig. 1.

is then computed as a weighted average of the corresponding pixel values in its N constituent frames

$$I_p = \sum_{a=1}^{N} \frac{f^{-1}(p_a)w(p_a)}{t_a} \Big/ \sum_{a=1}^{N} w(p_a) \tag{1}$$

where f^{-1} is the inverse camera response, p_a is the value of pixel p in image a and t_a is the exposure time of image a. For color images the red, the green, and the blue components of a pixel should be linearized by using their respective response curves.

This HDR assembly process may produce noisy images if the exposures used to create an HDR image are noisy (see Fig. 4). Although a weighted average is applied to the pixels during HDR assembly, in many cases this is not sufficient and the output images are rendered unusable (Figs. 1, 13, and 14).

At low light conditions (including standard room lighting) and high sensitivity settings, short exposures are more vulnerable to fluctuations in the number of photons that impinge on the camera sensors than long exposures. Thus, their signal-to-noise ratio is lower. During HDR assembly when a short exposure is divided by its exposure time its noise gets amplified. For instance, if the darkest frame is captured in 1/8000th of a second, division by the exposure time amplifies its noise 8000 times.

Combining such a noisy frame with longer (less noisy) exposures will result in noise in the final HDR. This is not sufficiently mitigated by the noise averaging that naturally occurs by combining exposures.

On the other hand, removing such exposures from consideration would result in a less noisy HDR image. However, such short exposures tend to carry useful information in a small number of pixels, usually depicting highlights or light sources. Ignoring these exposures will therefore reduce the overall dynamic range, and result in a loss of detail in highlights and light sources. To reduce noise while maintaining a high dynamic range, we outline our noise removal algorithm in the following section.

4. Algorithm

We preprocess the constituent frames of an HDR image prior to HDR assembly to reduce noise and prevent it from being further amplified.

To correct a noisy frame, we first order the frames by exposure time and linearize all the frames using the inverse response function f^{-1}. We then bring each frame into the same domain by dividing by its exposure time t_j. Each frame is then corrected by applying a weighted average to it and several successive longer exposures in the same sequence. Finally each frame is converted back to its original domain by multiplying it with its exposure time and unlinearizing using the response function f so that the output of our algorithm can be used in the subsequent HDR creation algorithm without any modifications. This yields the following equations:

$$g(p_j, a) = \begin{cases} v_j & j = a \\ \tau(p_j)v_j, & j \neq a \end{cases} \tag{2}$$

$$c_a = \sum_{j=a}^{a+s} \frac{f^{-1}(p_j)g(p_j, a)}{t_j} \Big/ \sum_{j=a}^{a+s} g(p_j, a) \tag{3}$$

$$p'_a = f(t_a c_a) \tag{4}$$

where g is a weighting function explained below and s is the number of subsequent exposures used during averaging; which we call the *cluster size*. For instance, if s is 4, the first frame is averaged with the second, the third, and the fourth frames. The second frame is averaged with the third, the fourth, and the fifth frames and so on. This process is depicted in Fig. 5 and a formal analysis is given in Appendix A.

Fig. 5. Cluster separation in the case of using 4 frames per cluster. The first frame is updated with the average of the first cluster, the second frame is updated with the average of the second cluster and so on.

The function g, shown in Fig. 6, is the product of the weights v given to each exposure and another function τ used to exclude over-exposed and unreliable pixels from averaging. We first explain how τ is designed and then proceed with setting the weights v and the cluster size s.

In digital images fluctuations may occur for pixels close to the extremes due to noise and compression artifacts. Such fluctuations may have a negative effect on our algorithm mainly because they will be emphasized by the weights v used during averaging. In addition, averaging with over-exposed pixels may cause a loss in the dynamic range giving a "washed out" appearance to the final HDR image. To avoid such scenarios we design τ such that pixels larger than 249 are excluded from the averaging. The averaging amount is smoothly increased for pixels between 249 and 200 using Hermite interpolation. All the pixels smaller that 200 are fully averaged. In principle these constants can be made user parameters. They serve as a gauge between the amount of noise reduction desired and how acceptable it is to use over-exposed and uncertain pixels in averaging. Hence, τ is defined by the following equation and its influence on the final weight is shown in Fig. 6.

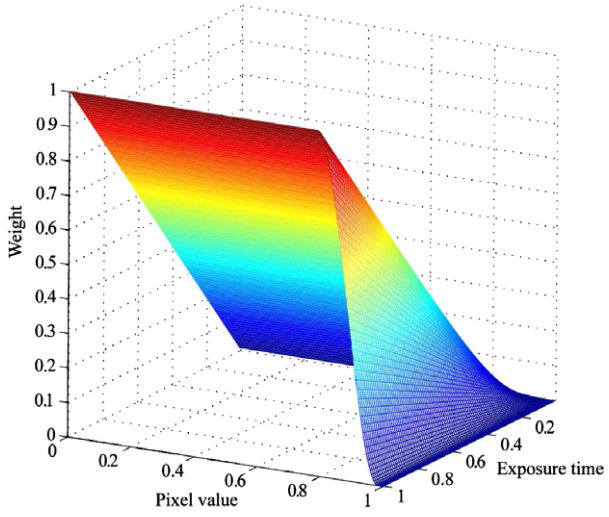

Fig. 6. Our weighting function g for noise reduction. Note that g depends on both exposure time and pixel value: we give more weight to pixels captured with longer exposure times, but at the same time prevent averaging with over-exposed (i.e. unreliable) pixels. The pixel values and exposure times are normalized.

$$h(x) = 1 - \frac{250 - x}{50} \tag{5}$$

$$\tau(x) = \begin{cases} 1 & 0 \leqslant x < 200 \\ 1 - 3h(x)^2 + 2h(x)^3 & 200 \leqslant x < 250 \\ 0 & 250 \leqslant x \leqslant 255 \end{cases} \tag{6}$$

When averaging within a cluster, it is possible that some pixels will be larger than 249 in all the images in the cluster. If τ is applied to all the images in that cluster, those pixels will be black in the final rendering. To address this problem τ is applied to all the images other than the image being corrected. For example, if we are correcting the 1st image using the 2nd and the 3rd images (i.e. the cluster size is 3), τ prevents leaking of over-exposed and unreliable pixels from the 2nd and the 3rd images to the 1st image. This condition is encoded in Eq. (2).

4.1. Weights and the cluster size

v_j is the weight given to exposure j during the averaging. To set v appropriately, we should consider the effect of the exposure time on noise, which is the only varying factor across exposures.

Theoretically, a change in the exposure time may have both negative and positive effects on the amount of noise in a photograph. For instance, dark current noise which is due to heating of camera sensors is expected to occur more in longer exposures than for shorter ones. In contrast, photon shot noise which is due to randomness in the behavior of photons tends to occur more in shorter exposures than longer ones.

We analyzed numerous exposure sequences and observed that short exposures have smaller signal-to-noise ratio than long exposures; which indicates that in our image sequences photon shot noise is more dominant than dark current noise. This may be due to using the cameras at high ISO values. At high ISO values fluctuations in pixel values are further emphasized while shorter capture times mitigate heating problems. The image sequence in Fig. 4 is in accord with this observation; the shorter exposures have a smaller SNR than longer ones.

Photon shot noise is modeled by the Poisson distribution for which the *mean* is equal to the *variance*. Assume that an area of the image receives N photons per unit time. Then a frame with a unit exposure will catch N photons on average with a variance of also N. The second frame exposed for t units of time (where $t > 1$) will catch tN photons on the average with a variance of tN. Dividing by the exposure time, the expected value of photons in the second image will be $tN/t = N$. However, its variance will be $tN/t^2 = N/t$ [5]. Thus the frame exposed t times longer will have t times less variance. We exploit this fact by setting the weighting coefficients equal to the exposure times, $v_j = t_j$.

Instead of using all of the lighter exposures to correct a noise frame, we separate the frames into clusters of size s

and average each cluster separately. Using bigger cluster sizes provides better noise reduction. However, after a certain cluster size using more images does not give better noise reduction due to thresholding with τ, and may unnecessarily increase the duration of averaging. For the same reason bigger clusters do not cause any artifacts (such as decreasing the dynamic range, banding or haloing). For example, the noise of the image in Fig. 1 is reduced by using a cluster size of 6 (out of 16 frames). Using a bigger cluster size for this image produces the same result. Smaller cluster sizes produce noisier images as shown in the following section.

Eq. (2) is repeated for all the pixels in each exposure and all the exposures that will be used in HDR creation. The only exceptions to this rule are the exposures at the light end of the sequence. The cluster size is automatically decreased to fit to the number of remaining frames. For instance, if we are correcting for the 14th image of a 16 image sequence, only the last 3 images are used for averaging even if the cluster size is bigger. Averaging with fewer images at the end of the sequence theoretically affects noise reduction for those images. However, in practice it does not compromise noise reduction since noise is mostly due to short exposures.

Once all the frames are corrected, the resulting frames can be used to create an HDR image. However some of the operations in HDR generation, such as linearization and division by exposure time, can be eliminated since these are already performed during noise reduction. Consequently the following equations may be used for noise removal and HDR creation, respectively.

$$c_a = \sum_{j=a}^{a+s} \frac{f^{-1}(p_j) g(p_j, a)}{t_j} \Big/ \sum_{j=a}^{a+s} g(p_j, a) \tag{7}$$

$$I_p = \sum_{a=1}^{N} c_a w(f(t_a c_a), L_{c_a}) \Big/ \sum_{a=1}^{N} w(f(t_a c_a), L_{c_a}) \tag{8}$$

Note that the weighting function w used during HDR generation is the one defined in Table 1. Its two parameters are the pixel value and the luminance of the pixel, respectively.

5. Results

We conducted our experiments using two different cameras. The first camera, a Minolta DiMAGE A1, is a prosumer camera which allows to change the shutter speed, aperture size, and ISO settings separately. The second is a professional Nikon D2H which allows its ISO value to be increased up to 6400, which is eight times more than the maximum sensitivity of the Minolta DiMAGE A1. We used a tripod to minimize camera movement during image capture. The response functions of both cameras (Figs. 7 and 8) are created from the image sequence shown in Fig. 10 by using Mitsunaga and Nayar's algorithm [3]. Note that these response functions are recovered only once for each camera and are used for the other images taken with the same camera.

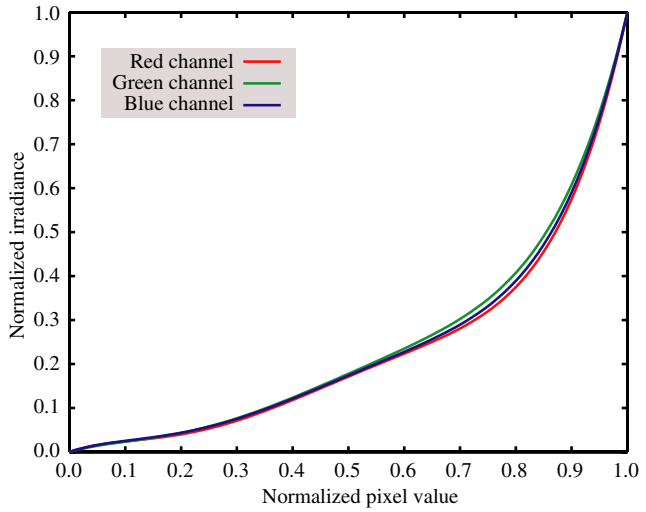

Fig. 7. Response curves of the Minolta DiMAGE A1 camera created by Mitsunaga and Nayar's algorithm from the image sequence shown in Fig. 10.

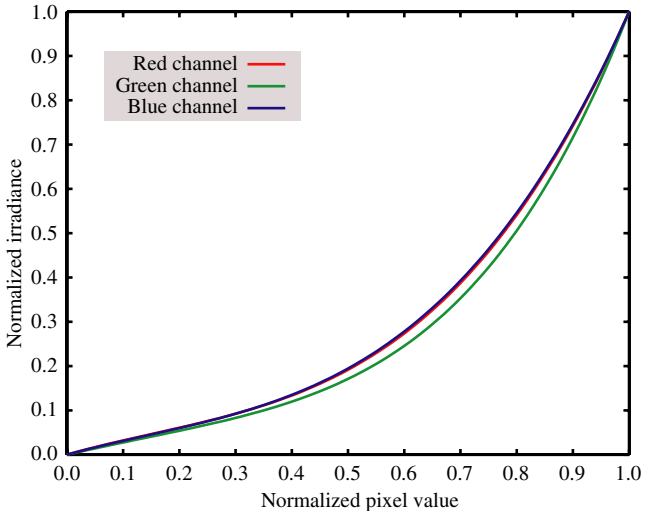

Fig. 8. Response curves of the Nikon D2H camera created by Mitsunaga and Nayar's algorithm from the image sequence shown in Fig. 10.

We compare the response curves of our cameras with a linear response and the sRGB gamma curve in Fig. 9. As seen from the figure, response curves of both cameras deviate from the standard sRGB curve. However, the response curve of the professional Nikon camera is closer to the sRGB curve.

The image in Fig. 1 is created from 16 exposures using the Minolta camera with an ISO setting of 800. As expected the resultant HDR image on the left is more noisy than any of its constituent frames. The right image in the same figure is corrected with our algorithm with a cluster size of $s = 6$ frames. To display the HDR images we reduced the dynamic range of both with the photographic tone reproduction operator using identical parameters [14]. Note that the right image is essentially noise free and it also has a significantly higher dynamic range than the left image.

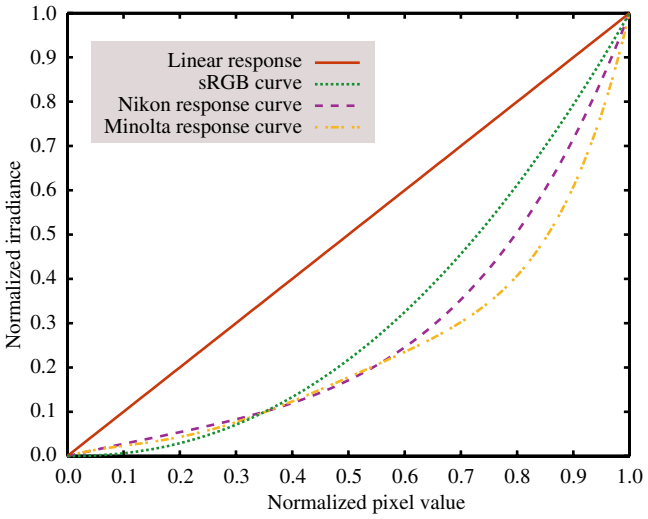

Fig. 9. Green channel response curves of the Minolta and the Nikon cameras in comparison with a linear and the sRGB response curve (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Since noise mainly occurs in the darker exposures, one might consider excluding some of the frames from the dark end of the image sequence. However, as Figs. 11 and 12 demonstrate, excluding frames from the dark end of the sequence causes a severe loss of dynamic range and the noise reduction achieved is substantially less than the reduction achieved by our method. For instance as the close-up views shown in Fig. 11, excluding frames does not reduce noise as effectively as our algorithm does and it causes loss of details in the highlights.

In addition to being visually distracting, noise may also cause tone reproduction operators to yield unexpected results. For instance, the average luminance of HDR images might be affected, and this may result in an over or under compression of the irradiance values. Fig. 13 shows this effect on three different tone reproduction operators. The top left image is tone mapped with the Tumblin and Rushmeier operator [15] without noise reduction. The bottom left image is tone mapped with the same operator and the same parameters after reducing the noise with our algorithm. The reduction of the noise prevents over compression of the radiance values. The images in the middle are tone mapped using a Bilateral filter [16] which produces a brighter scene for the noisy image than for the corrected image. The right most images are tone mapped with the photographic tone reproduction operator [14], which produces a similar overall appearance for both the noisy image in the top right corner and the corrected image below it.

The images at the top row in Fig. 14 are created from 20 frames captured by the Minolta camera at an ISO value of 800. The top right image is corrected with our algorithm by using a cluster size of 6 images. At the bottom row of the same figure images of a bull statue are created from 9 frames captured by the Nikon camera at an ISO value of 6400. Due to the quality of the camera and better lighting conditions this image is considerably less noisy than the previous one. However, as the close-up shows it still contains a fair amount of noise, which is successfully reduced by averaging with a cluster size of 5 frames.

Our algorithm does not degrade an HDR image if we run it on an initially noise free sequence (a formal reasoning is given in Appendix A). Fig. 15 demonstrates two images created from a noise free sequence of 9 images captured by the Nikon camera at ISO value 200. The left image is created without application of our algorithm and the right image is created after applying our algorithm with a cluster size of 5. Both images are tone mapped with the photographic tone mapping operator.

6. Comparison

Robertson et al. [11] proposed an algorithm for both recovering the response curve of a camera and creating

Fig. 10. Image sequence used to create the response curve of the Minolta camera. The same scene is captured also with the Nikon camera to recover its response. The sequence is captured by using a tripod to minimize alignment problems. The exposure times from darkest to lightest exposure vary from $\frac{1}{16{,}000}$ to $\frac{1}{4}$ s, with doubling at each exposure.

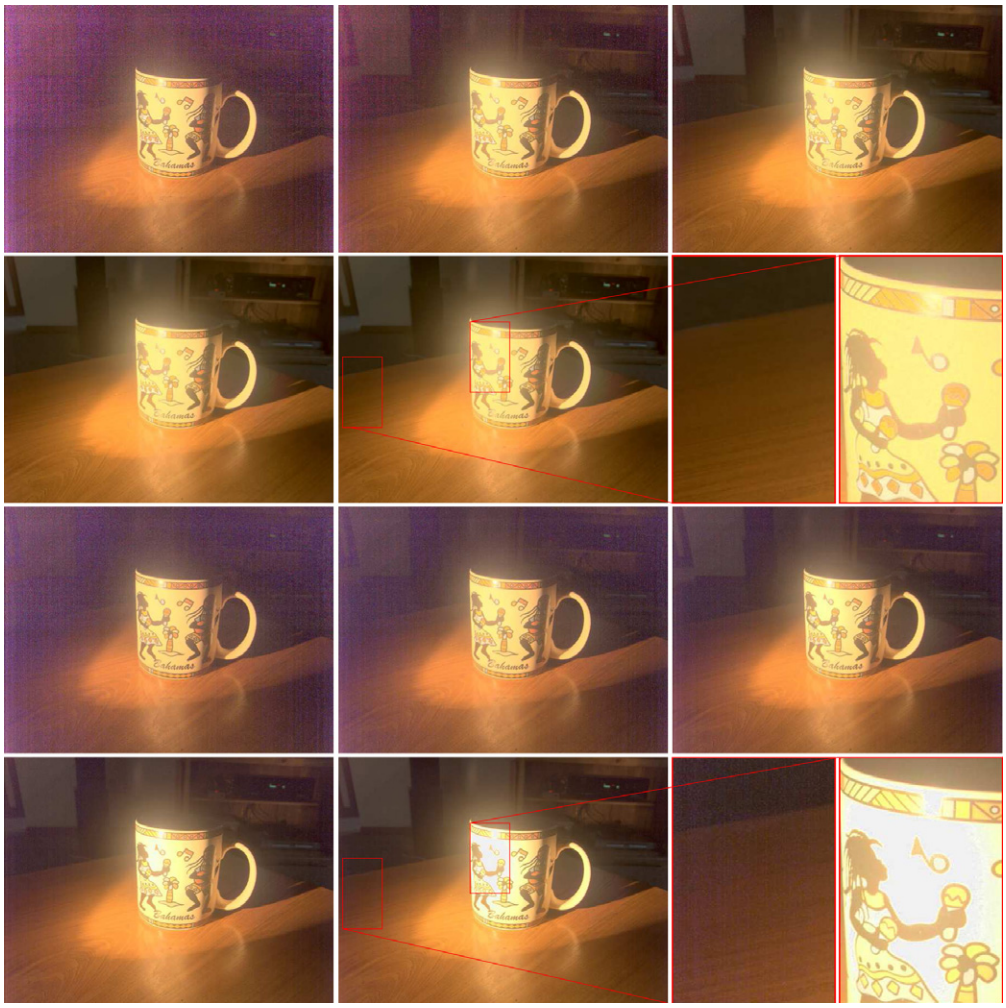

Fig. 11. Comparison of our algorithm with excluding frames from the dark end of an image sequence. Our results are shown in the first two rows. The cluster size from left to right and top to bottom is 2, 3, 4, 5, and 6. The last image of the second row shows a close-up view of the marked regions. For the last two rows, results in case of excluding 2, 3, 4, 5, and 6 images from the dark end of the sequence are shown. The dynamic ranges of all the images are reduced with the photographic tone reproduction operator [14]. The graph in Fig. 12 depicts the change in the dynamic range for each approach.

HDR images with reduced noise. To our knowledge Robertson et al.'s algorithm is the current state-of-the-art for noise reduction in HDR imagery. In this section, we briefly explain this algorithm, and discuss its differences from our algorithm.

For HDR creation, Robertson et al. propose to weigh pixels coming from longer exposures more than pixels from shorter ones since longer exposures have a higher signal-to-noise ratio. The authors define HDR assembly as follows:

$$I_p = \sum_{a=1}^{N} \frac{f^{-1}(p_a)}{t_a} w(p_a) t_a^2 \Big/ \sum_{a=1}^{N} w(p_a) t_a^2 \tag{9}$$

$$= \sum_{a=1}^{N} f^{-1}(p_a) w(p_a) t_a \Big/ \sum_{a=1}^{N} w(p_a) t_a^2 \tag{10}$$

Thus each pixel is weighted with the squared exposure time of its image.

Robertson et al. also propose a new method to recover the camera response for use in the previous equation. To recover the camera response they minimize the following objective function O:

$$O = \sum_{i,j} \bar{w}(p_{ij}) \left(f^{-1}(p_{ij}) - t_i I_j\right)^2 \tag{11}$$

$$\bar{w}(x) = \exp\left(-4 \frac{(x - 127.5)^2}{127.5^2}\right) \tag{12}$$

where i and j are indices over images and pixels, respectively, and $\bar{w}$ is a Gaussian shaped weight function which is used only during the recovery of the camera response. Once the camera response is recovered a cubic spline is fitted to it and the derivative of the spline is used as the weight function for HDR recovery (in Eq. (10)).

The cubic spline is generated such that its derivative is 0 at both ends of the pixel range. This feature is important in Robertson's algorithm, because the derivative, which is later used as a weight function, then tends to zero at either end of the scale. Giving zero weight to under- and over-exposed pixels is necessary to obtain correct results.

However with this approach the camera response is forced to have an S-shape even though the actual shape of the camera response may be different. In fact most

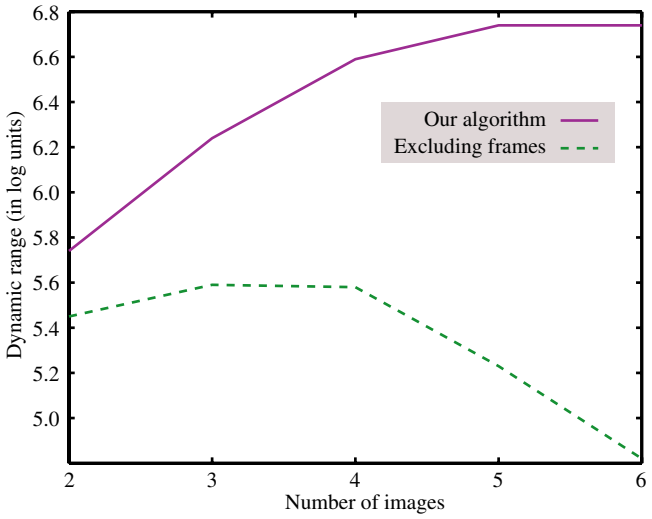

Fig. 12. The dynamic range in $\log_{10}$ units of the final HDR images in Fig. 11 in the case of applying our algorithm and excluding frames from the dark end of the sequence.

cameras do not have an S-shaped response function [17]. The mean response of a collection of 201 different film, digital, and video cameras is shown in Fig. 16, together with the sRGB gamma curve and the response curves of the Minolta camera created by Robertson et al.'s and Mitsunaga and Nayar's algorithms. We see that Mitsunaga and Nayar's response curve is qualitatively similar to the average response curve. However, Robertson's algorithm yields a response curve that is fundamentally different from the actual response of the camera.

In Fig. 17, we present a side-by-side comparison of our best result with the best result we obtained with Robertson et al.'s algorithm. Note that our algorithm preserves the highlight on the mug correctly and does not cause banding artifacts.

7. Conclusions

In this paper, we outline an efficient, effective, and simple noise reduction technique for high dynamic range

Fig. 13. Effect of noise on different tone mapping operators. The images on the top row are tone mapped without application of our algorithm. The images on the bottom row are corrected with our algorithm prior to the tone mapping. For both rows, the operators used for tone mapping from left to right are Tumblin and Rushmeier operator [15], the bilateral filter [16], and the photographic tone mapping operator [14], respectively. The constituent frames of this HDR image are captured with Nikon D2H with an ISO setting of 6400.

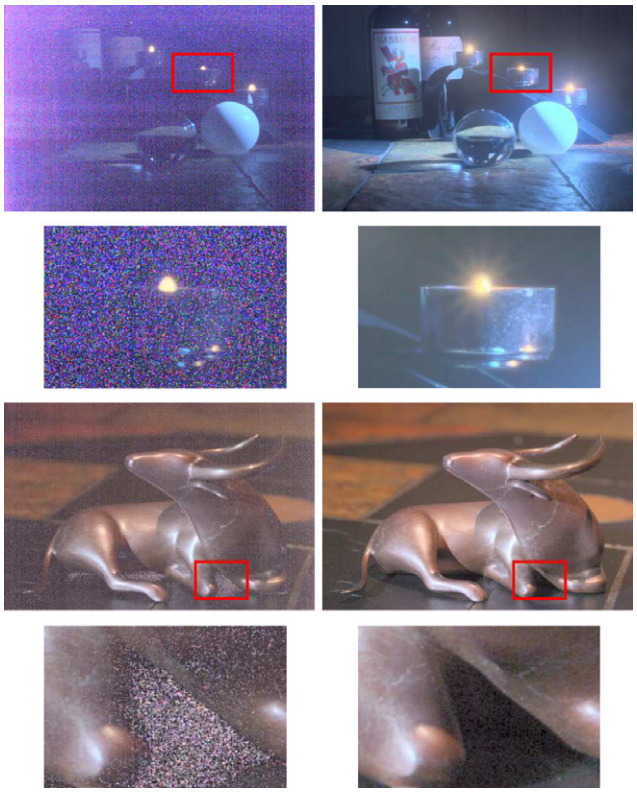

Fig. 14. The images on the first and the second rows are created from the exposures taken by Minolta DiMAGE A1, and Nikon D2H, respectively. The images on the right are corrected with our algorithm. The close-ups show the amount of noise reduction achieved in both of the HDR images. The HDR images are tone mapped with the photographic tone reproduction operator [14].

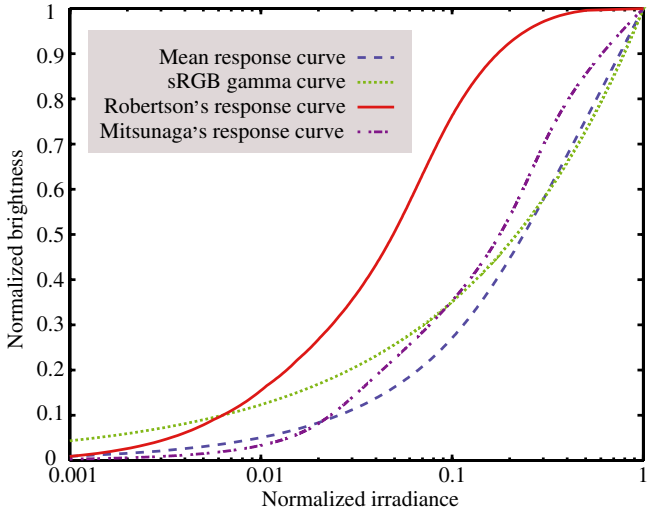

Fig. 16. The mean response of the 201 cameras created by Grossberg et al. [17], the sRGB gamma curve, and the response curve of the Minolta camera created by Robertson et al.'s algorithm.

images. Our algorithm requires only the response curve of the camera to be known or recovered from an image sequence. Response curves may be recovered with any existing recovery algorithm.

Using our technique photographers can increase the camera sensitivity without introducing noise problems, thus capturing the same dynamic range in a shorter amount of time. This especially holds true if auto-bracketing is used. Modern high-end digital cameras are able to take seven (Canon) or nine (Nikon) auto-bracketed exposures, each spaced 1 EV apart. Data acquisition is therefore directly related to the choice of exposure time. Shorter exposure times can be achieved by increasing the ISO setting on the camera. This normally yields noisier images, which may be corrected with our algorithm. The advantage of this approach is that important problems such as ghosting and motion blur occur to a significantly lesser extent.

We believe that high dynamic range video applications may also benefit from this work. By increasing the sensitivity of the video camera higher frame rates may be achieved. The noise can then be removed by using our algorithm as a post process on the recorded frames.

Appendix A

Here, we provide a more formal analysis of our noise reduction algorithm. Assume that an image region has irradiance I. In each image capture j, the camera records the irradiance together with an additive noise term n_j associated with the capture process. If the photon shot noise is dominant, its variance decreases with increasing capture time, thus $n_j \sim \sigma_j/t_j$ where σ_j denotes the variance for unit

Fig. 15. Application of our algorithm to an originally noise free sequence. The left image is created without applying our algorithm and the right image is created after applying our algorithm. Note that the application of our algorithm to a noise free sequence does not cause any side effects.

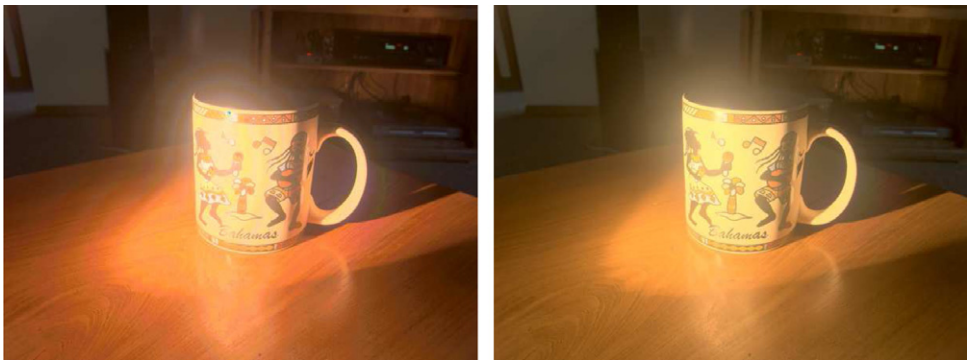

Fig. 17. Both HDR images are created from the sequence shown in Fig. 4; whereas the left image is created by using the algorithm described in Robertson et al. [11]. The right image is created by using our algorithm. We used photographic tone reproduction to display both images [14].

time. Therefore, the irradiance recorded in exposure j is equal to:

$$I_j = (I + \sigma_j/t_j)t_j \tag{13}$$

In our algorithm, we first normalize these terms by the exposure time and then calculate their weighted average,

$$I'_j = \frac{\sum(I + \sigma_j/t_j)w}{\sum w} \tag{14}$$

where the weight term is shown by w for brevity. We can rearrange the above equation as

$$I'_j = I + \frac{\sum(\sigma_j/t_j)w}{\sum w} \tag{15}$$

Setting $w = t_j$ as we do in our algorithm we obtain

$$I'_j = I + \frac{\sum \sigma_j}{\sum t_j} \tag{16}$$

which approaches to I as t_j goes to infinity. Although in practice we cannot use infinite exposure times, this analysis shows that the processed irradiance approximates the true irradiance as the number of averaged images grows. Furthermore it also demonstrates that, if applied on an originally noise free sequence (i.e. $\sigma_j = 0$), the process does not introduce any side effects.

References

[1] S. Mann, R. Picard, Being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures, Proceedings of IS&T 46th annual conference, 1995, pp. 422–428.

[2] P.E. Debevec, J. Malik, Recovering high dynamic range radiance maps from photographs, in: SIGGRAPH 97 Conference Proceedings, Annual Conference Series, August 1997, pp. 369–378.

[3] T. Mitsunaga, S.K. Nayar, Radiometric self calibration, in: Proceedings of CVPR, vol. 2, June 1999, pp. 374–380.

[4] S.B. Kang, M. Uyttendaele, S. Winder, R. Szeliski, High dynamic range video, ACM Trans. Graph. 22 (3) (2003) 319–325.

[5] A. Bovik (Ed.), Handbook of Image and Video Processing, Academic Press, New York, 2000.

[6] N.C. Gallagher, G.L. Wise, A theoretical analysis of the properties of median filters, IEEE Trans. Acoust. 29 (6) (1981) 1136–1141.

[7] N.D. Sidiropoulos, J.S. Baras, C.A. Berenstein, Optimal filtering of digital binary images corrupted by union/intersection noise, IEEE Trans. Image Process. 78 (1994) 382–403.

[8] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: Proceedings of the 1998 IEEE International Conference on Computer Vision, IEEE, 1998, pp. 839–846.

[9] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. Pattern Anal. Mach. Intell. 12 (7) (1990) 629–639.

[10] D.L. Donoho, I.M. Johnstone, Ideal spatial adaptation via wavelet shrinkage, Biometrika 81 (1994) 425–455.

[11] M.A. Robertson, S. Borman, R.L. Stevenson, Estimation-theoretic approach to dynamic range enhancement using multiple exposures, J. Electron. Imaging 12 (2) (2003) 219–228.

[12] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting, Morgan Kaufmann Publishers, San Francisco, 2005.

[13] M.D. Grossberg, S.K. Nayar, High dynamic range from multiple images: which exposures to combine? in: ICCV Workshop on Color and Photometric Methods in Computer Vision (CPMCV), October 2003.

[14] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda, Photographic tone reproduction for digital images, ACM Trans. Graph. 21 (3) (2002) 267–276.

[15] Jack Tumblin, Holly Rushmeier, Tone reproduction for computer generated images, IEEE Comput. Graph. Appl. 13 (6) (1993) 42–48.

[16] F. Durand, J. Dorsey, Fast bilateral filtering for the display of high-dynamic-range images, ACM Trans. Graph. 21 (3) (2002) 257–266.

[17] M.D. Grossberg, S.K. Nayar, What is the space of camera response functions? in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2003, pp. 602–609.