Hardware-Efficient Virtual High Dynamic Range Image Reproduction

Wei-Ming Ke, Tsun-Hsien Wang, and Ching-Te Chiu
Department of Computer Science, National Tsing Hua University, Taiwan

ABSTRACT

High dynamic range (HDR) images keep dynamic range of luminance from 10^5 to 10^8 and preserve more details than low dynamic range (LDR) images. Conventional acquisition of HDR images requires several images with different exposure settings of one scene, so multiple cameras or static scene are necessary. Besides, in order to transform HDR images onto normal LDR display, tone-mapping algorithms are required which need intensive computations. In this paper, we propose a hardware-efficient virtual HDR image synthesizer that includes virtual photography and local contrast enhancement. Only one LDR image is enough to generate HDR-like images, with fine details and uniformly-distributed intensity. A real-time hardware display system suitable for image or video contrast enhancement is also implemented. Under UMC90nm technology, we can process video sequences with NTSC 720×480 resolution at 60 frames per second (FPS), running at 100MHz and consume 0.3mm^2 silicon area.

Index Terms**—**High dynamic range (HDR), virtual photography, contrast enhancement

1. INTRODUCTION

The dynamic range of natural luminance intensity can reach 10^8:1 while traditional acquisition devices only show images in dynamic range of 100~1000:1. It is due to the dynamic range of the capturing devices and the limitation of the bit resolution in the image format. That discrepancy makes the accurate display of real world scenes difficult. For example, some objects of the scenes are unidentifiable because the background is too dark or too bright. Therefore, how to generate images with high dynamic range and preserve details on display devices becomes an interesting topic.

Debevec and Malik [1]; Nayar and Mitsunaga [2], sequentially capture multiple images of the same scene using different exposures to synthesize HDR images. They are only suitable for static scenes. Ikeda [3]; Street [4] proposed HDR generating algorithms by multiple sensors or detectors. However, these approaches require sophisticated devices to capture the scene simultaneously, and the cost is too expensive. On the other hand, due to the limitation of display, the performance evaluation should focus on the final displayable result.

To reveal image details in real-time systems, surveillance systems for example, local contrast enhancement has been researched. Dah-Chung Chang et al. [5] observe that the gain to local contrast enhancement is in inverse proportion to the local standard deviation (LSD). The hardware implementation is complicated. Schutte [6] and Sascha [7] propose multi-scale approach to enhance local contrast, but the LSD computations and divisions are also involved. In addition, for properly enhancing the pixel according to its neighboring region, the number of windows is decided by experience, and the maximum window size is usually large.

We propose a hardware-efficient virtual HDR synthesizer that includes two parts, the virtual photography and local contrast enhancement. The effect of mapping the virtual real-scene image into LDR display is considered and a simplified HDR-like mapping curve is proposed. Processed by only one LDR image, the intensity distribution becomes more uniform, and the local contrast is enhanced. The result achieves the same quality as using conventional high dynamic range (HDR) image generation and tone-mapping with respect to detail revelation. The proposed schemed is implemented in hardware using UMC 90nm technology.

The rest of this paper is organized as follows. Section 2 proposed our virtual photography and improved contrast enhancement algorithms. Section 3 illustrates the corresponding hardware implementation. Section 4 provides experimental results. Finally, we draw our conclusion in Section 5.

2. PROPOSED ALGORITHM

Our algorithm is divided into two parts: *Virtual Photograph(VP)* and *Contrast Enhancement*. As in Fig. 1, the original LDR image is first processed by virtual photography. The virtual photography can be achieved in two ways. The first approach is to generate multiple virtual real-scene images, fuse them into a single virtual HDR one, then tone mapping for LDR display. For hardware design consideration, we propose an alternative way using polynomial rotation to realize the function of traditional virtual photography. Afterwards, we make use of local-adaptive contrast enhancement to enhance the image details.

2.1 Virtual real-scene image generation and fusion

There are several challenges for generating a virtual real-scene image with high dynamic range from a single LDR one. First, the information contained in a single LDR image is limited. Second, once data has been captured and encoded, its quality is determined by the capture device and state-of-the-art of image coding technology. Hence the existing data must be enhanced for further processing. Moreover, the camera response function

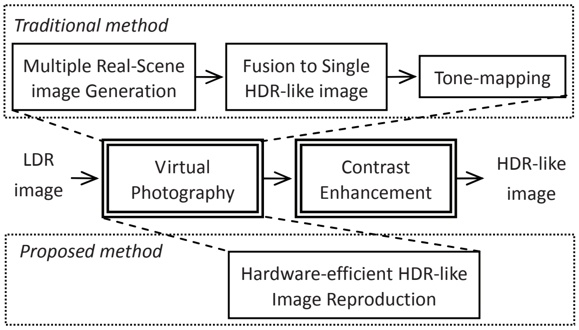

Fig. 1. Processing flow of high-contrast, virtual HDR image generation.

is non-linear, which is complicated to recover or simulate.

We emulate a virtual camera response function as a reversible S curve [8]. The human eye response of luminance range is considered in the range of 10^{-3}~10^{5}. As in eq. (1), a reference LDR photograph L_d is used to generate a virtual real-scene luminance L_w. Besides, the different brightness photographs should be created by different exposure values in the camera. Based on this concept, we design an adaptive function for L_a, as in eq. (2), where μ is a constant and *EV* indicates the exposure value. Combined with eq. (1) and (2), we can generate multiple virtual real-scene images with different exposure values. The less exposure value represents the less average luminance value of the whole image. Furthermore, we design a different adaptation factor *P* for different exposure values such that the range between the brightest image and the darkest image can be enhanced.

According to the original image and above response functions, several virtual real scene images with different exposure values can be generated. These images and the original one are combined into one intensity-balanced image using a weighted-average scheme, as shown in eq. (3)(4), which can theoretically increase the dynamic range. *N* is the total number of differently-exposed images, L_k is the k_{th} image intensity, and w_k is the weighting function of L_k [9].

$$L_w = \frac{10^{-P} \cdot L_d}{L_a\left(1 - L_d\right)} \quad (1)$$

$$L_a = \frac{1}{(1 + e^{-\mu EV})} \quad (2)$$

$$L_{VP} = \frac{\sum_{k=1}^{N} w_k \cdot L_k}{\sum_{k=1}^{N} w_k} \quad (3)$$

$$w_k = \exp\left(-4 \cdot \frac{(L_k - 127.5)^2}{127.5^2}\right) \quad (4)$$

2.2 Hardware-efficient HDR-like Image Reproduction

In our system, however, contrast enhancement will perform right after virtual photography, so the consideration of details could be removed from the weighing function. Moreover, division and exponential operations are needed in eq. (3)(4). They would be obstacles to the hardware implementation. Therefore, a more instinctive approach is proposed to perform virtual photography. First, we adopt a parabolic weighting function as in eq. (5).

$$w_k = -A(L_k - p)^2 + 1 \quad (5)$$

The highest weighting value occurs when a pixel value equals to *p*, and *A* is the degradation amplitude. Because this weighing function for a certain pixel in L_k only depends the pixel itself, the overall virtual photography turns into a 1-to-1 mapping function providing *N* and the different applied exposure values are fixed. In our case, we assume that the input LDR intensity is with the middle exposure value (L_3), and we apply other four fixed exposure values to produce L_k, *k=1, 2, 4, 5*. Nevertheless, the division operation in eq. (3) is still needed, and therefore we propose a new polynomial L_{VP_app} to approximate the L_{VP} with eq. (5), as shown in eq. (6).

$$L_{VP\_app} = L + A'L(L-1)(L-p) \quad (6)$$

A' is the amplitude of the "darkening" and "brightening" effect. The comparison of virtual photography methods using different weighting function and approximated L_{VP_app} is shown in Fig. 2, where *A=3*, *A'=2*, and *p=0.5*. Note that the root mean square errors (RMSE) of L_{VP_app} and L_{VP} with eq. (5) are mostly less than 0.04. The L_{VP_app} enhances the detail around middle luminance than L_{VP} with eq. (4). After our virtual photography scheme, the dark pixels would be brighter and vice versa. As in Fig. 3, the intensity-histogram becomes smoother and closer to the middle value.

2.3 Local Contrast Enhancement

Multi-scale contrast enhancement is a popular way to enhance image details [6][7]. However, for hardware realization, these approaches need more memory buffer to store the sliding windows with different sizes, which would increase the power consumption and chip area. On the other hand, local standard deviation and division operations increase the computing complexity tremendously. After virtual photography, image intensity is redistributed and some details in dark/bright area are already revealed. Using multi-scale approaches would cause a lot of noise artifact. By experiments, we use a filter with 5×5 sliding window to perform contrast enhancement, as in eq. (7).

$$\begin{aligned} m(i,j) &= \frac{1}{25}\sum_{p=-2}^{2}\left(\sum_{q=-2}^{2} L_{VP}(i+p, j+q)\right) \\ L_{CE}(i,j) &= L_{VP}(i,j) + (L_{VP}(i,j) - m(i,j)) \end{aligned} \quad (7)$$

To evaluate the resulting image quality, we proposed two terms for quantification. First, the details should be enhanced as much as possible. Just-noticeable-difference (JND) is a well-known property to determine the perception limit of the human visual system (HVS). For 2-D images, the background intensity of a target pixel is applied to calculate the JND threshold [10], which is done by taking the average intensity of a pixel's surrounding 5×5 block. Then we can decide the JND ratio, which is defined as the ratio of the amount of JND-satisfied pixels to the image size, to assess the degree of observable details in the image. However, the original noise distortion should not be emphasized during contrast enhancement process. Accordingly, assuming that *B* is the set of all 5×5 sub-blocks of L_{CE}, and *std(b)* is the standard deviation of sub-block *b*, we can define the peak-to-noise ratio for smooth area (PSNR-SA) as eq. (8).

$$\begin{aligned} S &= \{(p,q) \mid (p,q) \in b, b \in B, std(b) \le 1\} \\ MSE\text{-}SA &= \frac{\sum_i \sum_j (L_{VP}(i,j) - L_{CE}(i,j))^2}{element\ count\ in\ S}, (i,j) \in S \\ PSNR\text{-}SA &= 10\log_{10}\left(\frac{255^2}{MSE\text{-}SA}\right) \end{aligned} \quad (8)$$

The software simulation result is shown in Fig. 4. Schutte [6] introduces obvious noise due to original JPEG artifacts. Sascha [7] enhances not so much because it needs larger window sizes and more scales to gather regional information. According to Fig. 4(d) and Table 1, we can see that the details are apparently revealed with acceptable noise because we compromise the trade-off between details enhancement and the extent of noise distortion. In terms of design complexity, Schutte and Sascha's approaches need one square root and one division for each scale operation, while none of those operations are required in ours.

3. HARDWARE IMPLEMENTATION

The hardware architecture of our system is shown in Fig. 5. Input 8-bit image intensity is first processed by approximated virtual photography module, which is implemented without dividers. The internal data precision is 16-bit, but the output L_{VP} is rounding to 8-bit. Afterwards, L_{VP} is feeding into a memory bank with 5 line-buffers, which providing 5 parallel

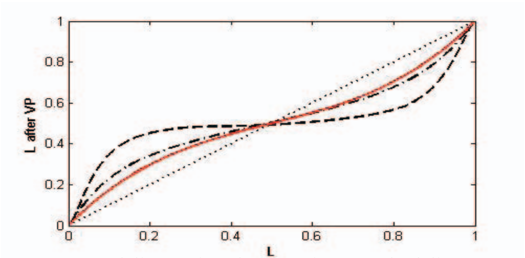

Fig. 2. Dotted line: Directly mapping. Dashed line: L_{VP} with eq. (4), Dashed dotted line: L_{VP} with eq. (5). Solid line: Approximated result L_{VP_app}.

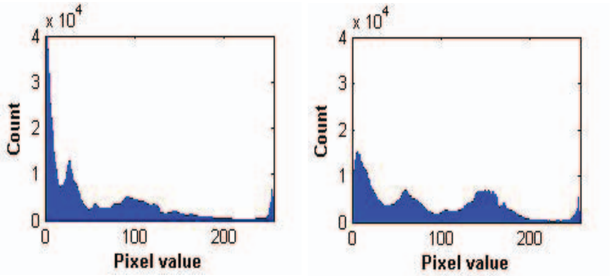

Fig. 3. Normalized intensity-histogram of (a)original chateau image, (b)chateau image after virtual photography.


(a) (b) (c) (d)

Fig. 4. (a) Original foyer image. (b) Schutte [6] with three scales. (c) Sascha [7] with three scales. (d) Our work.

Table 1. Quality assesement for various contrast enhancement algorithms. The target image is Fig. 4(a).

	Original	After VP	Schutte[6]		Sascha[7]		**Our work**
			scale=1	scale=3	scale=1	scale=3	
JND ratio	20.6%	28.0%	35.4%	46.8%	32.9%	34.0%	**40.7%**
PSNR-SA(dB)	N/A	N/A	63.05	56.08	64.34	56.90	**60.61**

pixel-outputs of one sub-block. The main computing block accepts these pixels, calculate the row sum in this sub-block, and select the target pixel. After a 5-level mean filter of the row sum, we can derive the average of the 5×5 sub-block, and combine the target pixel, then the final enhanced result is produced.

On the other hand, the row/column counter keeps counting as the input pixel is valid. And according to the counter values, the control unit sends control signals to the memory bank for read/write operations in the correct line buffer, and the main computing block for managing boundary special cases.

4. EXPERIMENTAL RESULTS

Compared to traditional HDR tone-mapping algorithms, the precision of input and output pixel in our system is only 8-bit and the computing complexity is relatively lower. So we can realize the hardware implementation with no distortion caused by the hardware bit-precision limitation. Fig. 6 presents the comparison of the results processed by complex tone-mapping algorithm [11] and our system. We use HDRSoft [12] to synthesize HDR radiance map. In our work, the intensity is uniformly distributed and the capability to enhance details is even much higher than the HDR tone-mapping approach. After function verification, the RTL code is transformed into real chip using UMC90nm technology. The performance and hardware area cost is excellent compared to other tone-mapping hardware, as shown in Table 2. Note that for tone-mapping algorithms in Table 2, the overhead of additional HDR-generating hardware is not taken into account yet.

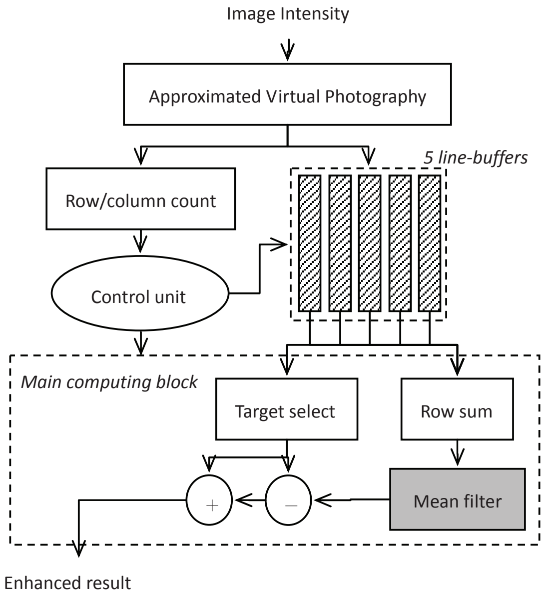

Fig. 5. Hardware architecture of virtual photography and contrast enhancement.

(a) (b) (c)

Fig. 6. (a) Original chateau image with middle exposure. (b) Photographic tone-mapping result [11] using synthesized HDR image. (c) Our work by taking (a) as the input image.

Table 2. Performance and cost comparison to tone-mapping hardware.

	Graphics HW [13]	FPGA [14]	Photographic Tone-mapping [15]	Gradient Tone-mapping [16]	Integrated Tone-mapping [17]	**Our work**
Frame size	512×512	1024×768	720×480	720×480	1024×480	**720×480**
FPS	20	60	30	30	60	**60**
Area(mm^2)	200	800	4.18	12	13.8	**0.3**

5. CONCLUSION

We propose an efficient algorithm to generate high-contrast, virtual HDR image using only one LDR image. Different from algorithms for generating HDR radiance maps, we focus on the details of the final displayable LDR image. Therefore we implement our hardware system with much lower complexity compared to traditional HDR generation and tone-mapping algorithm. Using UMC90nm technology, our system can achieve real-time for 720×480 video frames at 60FPS. Moreover, because we don't need differently-exposed images of one scene or expensive HDR camera, the proposed hardware can easily applied to current image capture devices, and suitable for both static and dynamic scenes.

6. REFERENCE

[1] Paul E. Debevec and J. Malik, "Recovering High Dynamic Range Radiance Maps from Photographs". In SIGGRAPH 97, pp.369-378, August 1997.
[2] S. K. Nayar and T. Mitsunaga, "High dynamic range imaging: spatially varying pixel exposures," in Proc. Computer Vision and Pattern Recognition (CVPR), pp.473-479, 2000.
[3] E. Ikeda. "Image data processing apparatus for processing combined image signals in order to extend dynamic range", *U.S. Patent 5801773*, Sep. 1998.
[4] R. A. Street. "High dynamic range segmented pixel sensor array", *U.S. Patent 5789737*, Aug. 1998.
[5] D.C. Chang and W.R. Wu, "Image contrast enhancement based on a histogram transformation of local standard deviation," IEEE Trans. MI, vol. 17, no. 4, pp. 518-531, Aug. 1998.
[6] K. Schutte, "Multi-Scale Adaptive Gain Control of IR images," Infrared Technology and Applications XXIII, Proceedings of SPIE, Vol. 3061, pp. 906-914 (1997)
[7] Sascha D. Cvetkovic, Johan Schirris, and Peter H.N. de With, "Locally-Adaptive Image Contrast Enhancement without Noise and Ringing Artifacts," IEEE International Conference on Image Processing, vol. 3, pp. 551-560, 2007
[8] http://support-tw.canon-asia.com/contents/TW/EN/8100096300.html
[9] M.A, Robertson, S. Borman and R.L. Stevenson, "Dynamic Range Improvement through Multiple Exposures", Image Processing, 1999. ICIP 99. Proceedings. 1999 International conference on Volume 3, pp.159–163 , Oct. 1999.
[10] Chun-Hsien Chou and Yun-Chin Li, "A perceptually tuned subband image coder based on themeasure of just-noticeable-distortion profile," IEEE Trans. Circuits Syst. Video Technol., vol. 5, no. 6, pp. 467–476, Dec.. 1995
[11] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic Tone Reproduction for Digital Images," ACM SIGGRAP, pp. 267-276, 2002
[12] HDRSoft, "http://www.hdrsoft.com"
[13] N. Goodnight, R. Wang, C. Wooleey, and G. Humphreys, "Interactive time-dependent tone mapping using programmable graphics hardware," Rendering Techniques, 14th Eurographics Symposium on Rendering, pp. 26-37, 2003.
[14] F. Hassan, and J. E. Carletta, "A real-time FPGA-based architecture for a Reinhard-like tone mapping operator," Processing of 22nd ACM SIGGRAPH/EUROGRAPHICS, pp. 65-71, 2007.
[15] Tsun-Hsien Wang, Wei-Su,Wong, Fan-Chu Chen and Ching-Te Chiu,"Design and Implementation of a real-time global Tone Mapping Processor for High Dynamic Range Video," IEEE ICIP2007, vol. VI pp. 209-212, 2007.
[16] Tsun-Hsien Wang, Wei-Ming Ke, Ding-Chuang Zwao, Fan-Chu Chen and Ching-Te Chiu, "Block-Based Gradient Domain High Dynamic Range Compression Design For Real-Time Application," IEEE ICIP2007, vol. III, pp. 561~ 564, 2007.
[17] Ching-Te Chiu, Tsun-Hsien Wang, Wei-Ming Ke, Chen-Yu Chuang, Jhih-Rong Chen, Rong Yang and Ren-Song Tsay, " Design Optimization of a Global/Local Tone Mapping Processor on ARM SoC Platform for Real-time High Dynamic Range Video", IEEE International Conference on Image Processing 2007