https://doi.org/10.2352/ISSN.2470-1173.2017.17.COIMG-440

Localized High Dynamic Range Plenoptic Image Compression

Chuan-Chung, Chang[1], Hsin-Hsiang, Lo[1], Han-Hsuan, Lin[1], Zhi-Rong, Fan[2], Shao-Hsuan, Cheng[1], Chih-Hung, Lu[1], Fu-Ming, Chuang[1], Jiun-In, Guo[2]

[1]Coretronic Corporation, Science Park, Hsinchu, Taiwan

[2]Department of Electronics Engineering, National Chiao-Tung University, Hsinchu, Taiwan

Abstract

Based on multi-view characteristics of plenoptic imaging, a new method for generating localized high dynamic range (LHDR) compression for plenoptic image is proposed, which determine HDR compression region, and maintains image quality for different plenoptic-views. The proposed method is realized by physically segmented sub-images through plenoptic optics. And plenoptic image with wider dynamic range is processed by calculated local contrast of combined sub-images via block based harsh environment detection, and then followed by HDR process or dynamic local contrast enhancement (DLCE) for each block. Experiment results for proposed LHDR and reference HDR are also reported, which using self-designed plenoptic camera with low distorted sub-image, 60° field of view, 1920x1080 pixels, and 90 fps 2/3" CMOS imager. The results show that the proposed method is capable of wider dynamic range compression, while preserving fine details, fewer artifacts. Extended dynamic range from 84.6 dB to 108.9 dB is realized. And processing time of proposed method is 85% compared to reference HDR processing.

Introduction

In recent years, plenoptic camera has become one of most important topics in computational imaging, which providing new approaches and possibilities: digital refocusing [1], selective or all-in-focusing [2], depth imaging [3]. Several plenoptic cameras can be purchased in market already, like LYTRO ILLUM with constant F-number, 8X optical zoom (9.6^0 to 54^0 viewing angle) [4] which aiming for consumer market; RAYTRIX's plenoptic module for machine vision, microscopy, fluid imaging or some of others with changeable field of view by different primary lens (4.2^0 to 38.4^0 for R8, 4.5^0 to 41.7^0 for R42) [5]. For plenoptic camera with wider viewing angle, high dynamic range (HDR) will be necessary because of undesired or non-controlled lighting condition may occur. Plenoptic camera with HDR, which using special made micro lens array (MLA) with additional optical prosperities (e.g.: different aperture size, transmittance, polarization, spectrum, or others) has been proposed, which can be used for one-shot HDR, and HDR image was processed by built-in HDR function in Photoshop [6].

Relative Work

HDR imaging which is a popular technique in the past decade, it can recover real world luminance of scenes to radiance map based on fusing from multiple exposures images with limited dynamic range (LDR). When HDR image is generated, tone mapping operator can further compress HDR image to suitable for dynamic range for displaying.

In this section, brief review of previous approaches of tone mapping algorithms is discussed. It can be classified into two groups: (1) global method and (2) local method. Global mapping algorithms include gamma correction and curve mapping. Local method has many research, it can be classified into (i) detail manipulation, (ii) gradient domain method, and (iii) local gamma correction. In gradient domain method, R. Fattal [7] proposed an effective method that manipulates the gradient field of the luminance image by attenuating the magnitudes of large gradients, and a new, low dynamic range image was obtained by solving a Poisson equation on the modified gradient field. Their results achieved drastic dynamic range compression with fine details and avoiding common artifacts in bright and dark regions. In local gamma correction, F. Durand [8] used a bilateral filter to reduce the overall contrast while preserving local details in an image. It is based on a two-scale decomposition of the image into a base layer, encoding large-scale variations, and a detail layer. The base layer is obtained using the bilateral filter to reduce its contrast preserving detail. And E. Reinhard [9] proposed another method which used automatic dodging-and-burning and makes possible local contrast adjustment.

Optical Design for Plenoptic Camera

It is well known that non-focused plenoptic optics suffers from huge loss in spatial resolution [10]. So focused plenoptic which proposed by A. Lumsdaine and T. Georgiev [11] was considered in study. For plenoptic camera with wider field of view, it can be found that paraxial design will be quite limited. More consideration for optical aberrations, especially for pupil aberration of primary lens should be well-controlled. In order to provide a wide angle plenoptic camera, some considerations are shown in Figure 1 for example.

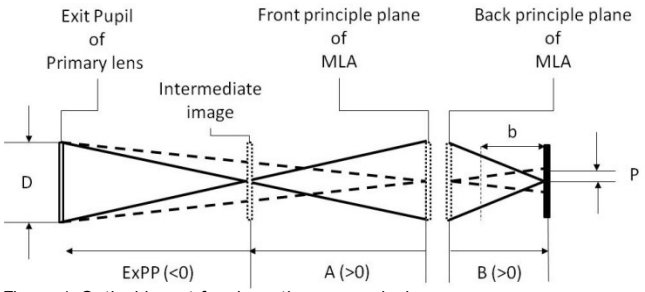

Figure 1. Optical layout for plenoptic camera design

For relationship of parameters shown in Figure 1, design rule for wide angle plenoptic camera are listed as follows with limited magnification ratio (B/A) with absolute range 0.5 to 0.005:

$$B = (1 - M) \times f_{MLA} \geq b \tag{1}$$

$$p = D \times B / (-ExPP + A) \tag{2}$$

where D is exit pupil diameter of primary lens, ExPP is distance between the exit pupil to intermediate image plane, p is pitch of micro lens array (MLA).

According equations listed above, it can be found that pitch of sub-lens in MLA will be function of D, ExPP and magnification ratio of MLA. For primary lens with wide or ultra wide field of view, size and position of exit pupil will be changed with field angle which called pupil aberration and pupil shifting. Although it can be found that MLA with variable size of sub-lens can be used

to compensate pupil aberration in theory, but it is not a practical way because of complexity of manufacturing and cost limitations.

If considering equation (2) for more detail, it can be notices that p will be close to a fix value when ExPP large enough, so it provide a chance to design a plenoptic camera with wide or ultra wide field of view.

Localized High Dynamic Range Imaging

The concept flowchart of proposed LHDR is shown in Figure 2. Firstly, multiple plenoptic raw images are taken by histogram based exposure time selection [12] for capturing suitable source images from current scene, then dividing them into several blocks which combine a group of sub images. Secondly, analyzing and determining whether block is in harsh environments by local contrast. If the selected blocks are in harsh environments, it will be processed by the multi-camera HDR; otherwise, it will be processed by DLCE. Finally, adoption neighboring blocks to do interpolation for eliminating block effects.

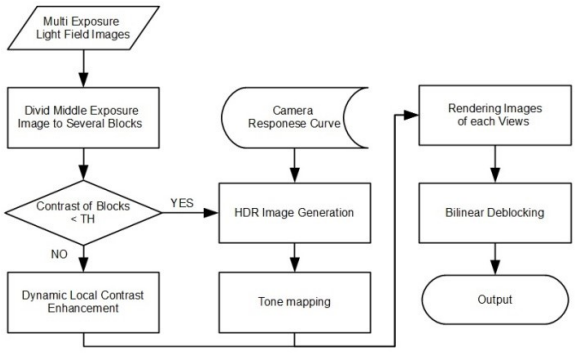

Figure 2. The proposed localized HDR method

Harsh Environment Detection

In general, images in harsh environments have lower contrast that can be defined by sum of intensity difference between pixel under evaluation and average value of given search window. It calculates contrast to detect whether the images are in harsh environments or not. Figure 3 shows the method for harsh environments detection. The contrast of selected block is calculated by Eq. (3), where $P(x,y)$ is current pixel value, $A(x,y)$ is average value of search window, $D(x,y)$ is absolute difference and $M.N$ are width and height of the selected block. The sum of difference is defined as contrast of the selected block.

$$Contrast = \frac{\sum_{x=0}^{M}\sum_{y=0}^{N} D(x,y)}{M \times N} \tag{3}$$

$$D(x,y) = \left| P(x,y) - A(x,y) \right| \tag{4}$$

Dynamic Local Contrast Enhancement

Processing for DLCE is composed by following steps:

1. Clipping histogram of each block, where H_{avg} is average of histogram, α is a constant (2.56 using in the paper).

$$clip = (1-\beta) \times H_{avg} + \alpha \times H_{avg} \tag{5}$$

$$\beta = contrast \ / 255 \tag{6}$$

2. Clipping histogram of each block, where H_{avg} is average of histogram, α is a constant (2.56 used in the paper).

3. Adjusting mapping function based on the contrast dynamically, where CDF(t) is cumulative distribution function.

$$\mathrm{Re}(t) = (H'_{\max} - H'_{\min}) \times CDF(t) + H'_{\min} \tag{7}$$

$$H'_{\max} = H_{\max} - \beta \times (H_{\max} - H_{\min}) \tag{8}$$

$$H'_{\min} = H_{\min} + \beta \times (H_{\max} - H_{\min}) \tag{9}$$

4. Generating new pixel values by interpolation.

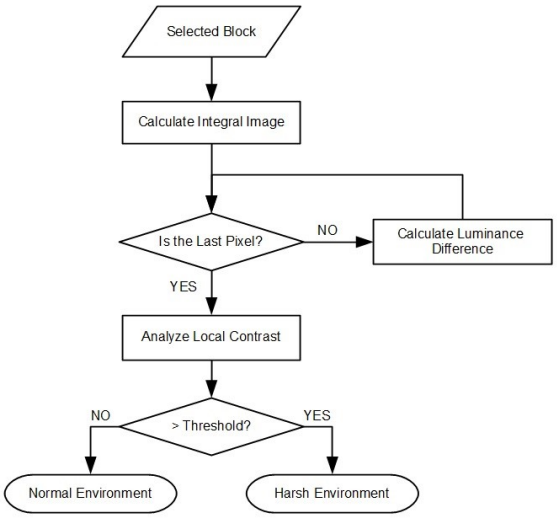

Figure 3. Harsh environment detection method

High Dynamic Range through Multi-images

Conventional cameras can capture scene with limited dynamic range (e.g.:70 dB). In comparison, human eye has great ability to adapt to high or low levels of light, the dynamic range of human eyes can percept reaches almost 10000:1. So far cameras cannot perform as good as human eyes do, and neither can the common monitor to display. The process between capturing scene radiance and generating final digital values consists of many non-linear mapping functions. In order to find the relationship between radiance and pixel values, Paul E. Debevec proposed a camera response function to construct this non-linear mapping [13].

Camera Response Function

When the pixel value and exposure time are known, these pixel values are then brought into quadratic objective function as shown in Eq. (10):

$$o = \sum_{i=1}^{N}\sum_{j=1}^{P} \{w(Z_{ij})[g(Z_{ij}) - \ln E_i - \ln t_j]\}^2 + \lambda \sum_{z=Z_{\min}+1}^{Z_{\max}-1} [w(z)g''(z)]^2 \tag{10}$$

where g is camera response function, N is the pixel location, P is the number of image, Z_{max} and Z_{min} are the largest and smallest values of pixels, w(z)and g"(z) are shown in the following:

$$w(z)=\begin{cases} z, z<30 \\ 30+2\times(z-30)\ ,30\le z<113 \\ 194+(z-113)\ ,113\le z<143 \\ 223-(z-143)\ ,143\le z<173 \\ 194-3\times(z-173)\ ,173\le z<220 \\ 50-(z-220)\ ,220\le z<256 \end{cases} \tag{11}$$

$$g''(z)=g(z-1)-2g(z)+g(z+1) \tag{12}$$

where w(z) is a new weighting function to enhance HDR image quality as shown in Figure 4, the luminance which is located in high of center is given a strong weight and in low area is given more less weight to suppress more noises.

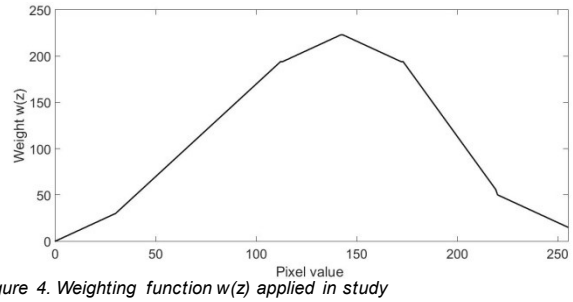

Figure 4. Weighting function w(z) applied in study

Tone Mapping

After applying the above mentioned methods, a high dynamic range image is generated. However, most of monitors can only display LDR image. In order to achieve high dynamic range compression, the initial luminance mapping is mentioned [9] as follows:

$$L_w^{'}=\exp\left(\frac{1}{N}\sum\nolimits_{x,y}\log(\delta+L_w(x,y))\right) \tag{13}$$

where N is total number of pixels in the image, L_w is the world luminance for the position (x,y), and δ is a small constant to avoid equation error.

$$L(x,y)=\frac{\alpha}{L_w^{'}}L_w(x,y) \tag{14}$$

where L(x,y) is the scaled luminance, and α is a key value that dominates the brightness of the image scene. According to [9], typically vary α from 0.18 up to 0.36 and 0.72 and vary it down to 0.09, and 0.045. A tone mapping operator that allow high luminance is given by:

$$L_d(x,y)=\frac{L(x,y)\left(1+\frac{L(x,y)}{L_{white}^2}\right)}{1+L(x,y)} \tag{15}$$

where L_d is the value which has been adjusted, L_{white} is the maximum of L(x,y). Finally, the proposed tone mapping method majorly uses photographic compression and image blending to keep more comprehensive information.

$$I_{result}(x,y)=(1-\alpha)\times I_{photograph}(x,y)+\alpha\times I_{source}^{'}(x,y) \tag{14}$$

$$I_{source}^{'}(x,y)=(1-\beta)\times I_{source1}(x,y)+\beta\times I_{source2}(x,y) \tag{15}$$

$$\alpha=\gamma\times\exp\left(-4\frac{(I_{photograph}(x,y)-255)^2}{(255-I_{threshold})^2}\right) \tag{16}$$

$$\beta=u+v\times\exp\left(-4\frac{(I_{source2}(x,y)-255)^2}{255^2}\right) \tag{17}$$

where $I_{photograph}(x,y)$ is the pixel value after photographic compression. $I_{source1}(x,y)$, and $I_{source2}(x,y)$ is the pixel value of the lowest exposure image and second low exposure image. $I_{threshold}$ is 0.7 times the maximum luminance, γ is scaling parameter which ranges from 0 to 1 but cannot equal to zero for image continuity, u is a constant value which dominates the weighting to the two image's pixel values and v is a scaling parameter. In experiment, the parameters sets are γ = 0.4, u = 0.3, and v = 0.3. By using the proposed method, the result image will preserve details in both bright region and dark region.

Plenoptic rendering

In the paper, full resolution rendering [14] is applied directly, where assume plenoptic raw data is composed by $N_x\times N_y$ sub-images with size $N_x\times N_y$ pixels. The full resolution rendering algorithm is performed by piecing patches in each sub-image together. For a given patch size P, rendering image with P x $N_x\times P\times N_y$ pixels can be achieved. Following equations show the relationship function between the coordinate of the raw data and rendered image:

$$i=floor\left[(s/p)M_x+0.5f\right]+(s\%P)+shift\_X+dx \tag{18}$$

$$j=floor\left[(s/P)M_y+0.5f\right]+(t\%P)+shift\_Y+dy \tag{19}$$

where (i, j) is coordinate of raw data; (s,t) is the relative one of rendered image; (shiftX, shiftY) is cutting boundary and (dx, dy) is coordinate of the patch.

De-blocking

By using the proposed method that combines HDR tone mapping and DLCE processing, the result image will preserve details in both bright region and dark region with block artifacts which based on block size. More details can be enhanced by using smaller block size, but with more block artifacts, it is shown in up column of Figure 6 left. The proposed de-blocking method using reference based bilinear interpolation to adjust the luminance of each blocks. The de-blocking result is shown in down column of Figure 6.

Assuming a similar point in reference image can be matched, and using them to adjust the luminance by bilinear interpolation. As Figure 5, there are 4 blocks a, b, c, d have block artifacts on each edge, where f(a) is the current pixel value which need to be adjust, f(b), f(c), and f(d) are the highly related pixel value of other three blocks which are searched in reference image. α is the distance between point f(a) and the left edge of block, ß is the distance between point f(a) and the top edge of block. Bilinear interpolation is shown in the following:

$$f_1(p)=f(a)\times(1-\alpha)+f(b)\times\alpha \tag{20}$$

$$f_2(p)=f(c)\times(1-\alpha)+f(d)\times\alpha \tag{21}$$

$$f_3(p)=f_1(p)\times(1-\beta)+f_2(p)\times\beta \tag{22}$$

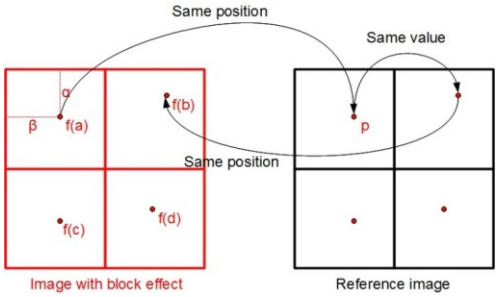

Figure 5. Bilinear interpolation de-blocking

Figure 6. LHDR image without (up) & with de-blocking (down) process

Experiment and Results

An off-the-shelf lens (Edmund Optics 68-215) with F/3.0, image circle for 2/3 inch image sensor (ON-Semi VITA2000) was chosen for primary lens. MLA with F/2.5, 0.75 mm focal length, 0.3 mm pitch was designed according Eq. (1) & (2) and Figure 7 shows simulated sub-image distribution of the plenoptic camera. It shows that low distorted sub-images and well control for F-number matching has been achieved. The designed MLA was further manufactured by diamond turning process and then integrated with the primary lens and image sensor. The finished plenoptic camera was used for following experiment.

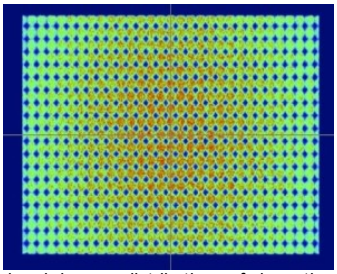

Figure 7. Simulated sub-images distributions of plenoptic camera with 60^0 field of view

Raw plenoptic images with full resolution 1920x1080 pixels and exposure time 0.0032, 0.008 and 0.0128 sec were shown in top column in Figure 8, and rendered result with each one are shown in down column in Figure 8. The captured mages were further processed by a reference HDR [12] and proposed LHDR algorithms on PC which hardware specifications are listed in Table 1.

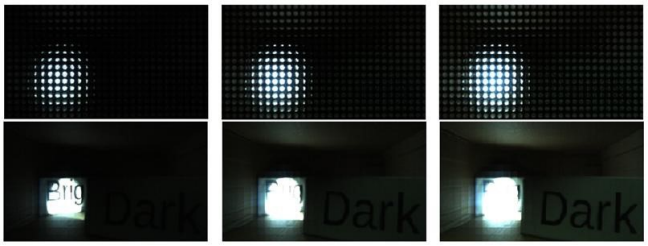

Figure 8. Up: captured plenoptic images by self-designed plenoptic camera; Down: central view with plenoptic rendering. (from left to right, exposure time increasing respectively)

Table 1: PC Specification

Processor	Intel dual core I5-5200U (2.2Ghz, L1:128Kb, L2:512Kb)
Memory	8 GB
Input image size	1920 × 1080
Output image size	320 × 188
Operation System	Windows 7

In our experiment, a basic plenoptic rendering [14] was used and multi-view output image with low resolution (320×188 pixels). The result of reference HDR [12] and proposed LHDR algorithm are shown in Figure 9, it can be observed that LHDR keep higher contrast and rich details in both bright and dark region, i.e. "Dark" in (b), "VS229" in (f) have stronger contrast than in (a), (e) and "HDMI" is cannot be seen in (c). It can be seen clearly in (d); where in bright region, i.e. "Brig" in (b), "hp f300" in (d) have more details than (a), (c), LHDR can avoid graying out effect in bright region which is shown (c) and (d).

Figure 9. (a), (c), (e): the result of center view by reference HDR [12]; (b), (d), (f): the result of center view by proposed LHDR.

Dynamic range is further measured based on modified ISO-14524 chart with 12 different gray level patches and an extra light source which putting on center of the test chart. Multiple dynamic range scenes can be created by adjusting brightness of environment

and light source. During experiment, scene with dynamic range from 45 to 111 dB can be generated, which measured and confirmed by luminance meter (Topcon BM-7). Results of Auto-exposure, processed by reference HDR and proposed method are shown in Figure 10, and contrast between each nearby patches are calculated by [15] to sum the numbers of different gray levels called "score", more calculated results are shown in Figure 11. Comparing the "score" and its dynamic range of LHDR, Auto-exposure, and HDR, it can be found that LHDR can deal with 108.9 dB in dynamic range which is higher than HDR's 84.6 dB and Auto-exposure's 72.1 dB.

It is also noticed that some blue block artifacts both shown in bright region, which coming from non-well-controlled stray light of MLA.

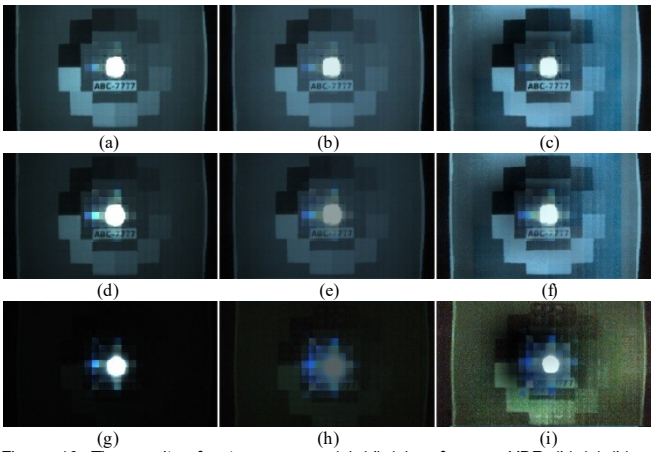

Figure 10. The results of auto-exposure (a) (d) (g), reference HDR (b) (e) (h), and proposed LHDR (c) (f) (i) in different dynamic scene, from top to bottom are 57.6 dB, 72.1 dB, 105.7dB.

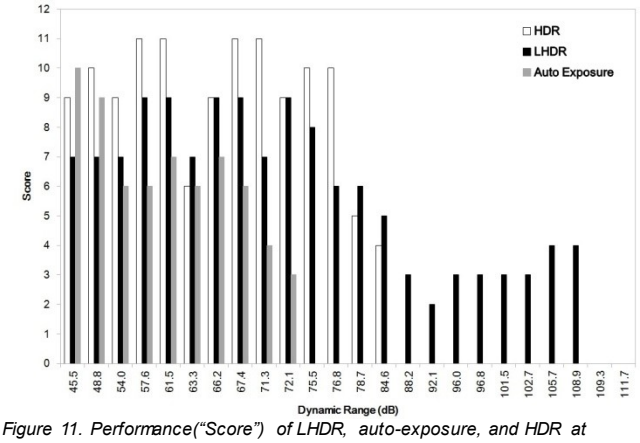

Figure 11. Performance("Score") of LHDR, auto-exposure, and HDR at different dynamic.

Computational time of reference HDR and DLCE are listed in Table 2, which shows that HDR spends 0.24 sec and DLCE spends 0.06 sec for image with 1920 ×1080 pixels. The reference HDR is 12 times longer than DLCE because of HDR process includes HDR generation and tone mapping. Comparison of computational time for the three tested scenes of reference HDR [12] and proposed LHDR are shown in Table 4, the results show that processing time of LHDR is 75% ~ 92% compared to reference HDR, which depends on scene content, more blocks are selected processing by DLCE, more time can be saved, the proportion of two block processing type foe each scene is shown in Table 3. The average time reduction of proposed LHDR is 85.12% compare to reference HDR [12].

Table 2: Computational time of HDR and DLCE

	HDR (sec)	DLCE (sec)
HDR generating	0.120	N/A
Tone mapping	0.060	0.015
Total	0.180	0.015

Table 3: The proportion of two block processing type foe each tested scene

	HDR	DLCE
Scene 1	66.7 %	33.3 %
Scene 2	79.2 %	20.8 %
Scene 3	87.5 %	12.5 %

Table 4: Comparison of total processing time of reference HDR [12] and proposed LHDR

	HDR (sec)	LHDR (sec)	Time comparison
Scene1	0.240	0.180	75 %
Scene2	0.250	0.220	88 %
Scene3	0.250	0.230	92 %
Average	0.2467	0.210	85 %

Conclusion

LHDR, which basing on multi-view plenoptic LDR images is proposed in the paper. The results show that the proposed LHDR algorithm achieved higher contrast and more detail in both of bright and dark region, dynamic range can be extended from 84.6 dB to 108.9 dB, and average 85% processing time compared to reference HDR.

Acknowledgements

This work was supported by Project for Industry-Academia Collaboration on Innovative R&D at Hsinchu Science Park, grant no. 104A01.

References

[1] A. Isaksen, L. McMillan and S.J. Gortler, "Dynamically Reparameterized Light Fields," Signal Recovery and Synthesis, Proceedings of the 27th annual conference on Computer graphics and interactive techniques (SIGGRAPH 2000), pp. 297-306, 2000.

[2] A. Kubota, K. Takahashi, K. Aizawa, and T. Chen, "All-focused light field rendering," Proceedings of Eurographics Symposium on Rendering (EGSR2004), pp. 235–242, 2004.

[3] T. E. Bishop and P. Favaro, "Plenoptic depth estimation from multiple aliased views," Computer Vision Workshops (ICCV Workshops), IEEE 12th International Conference on, pp. 1622-1629, 2009.

[4] https://www.lytro.com/illum/specs

[5] http://www.raytrix.de/produkte/Light Field Camera R42.PDF

[6] T. Georgiev, A. Lumsdaine, and S. Goma, "High Dynamic Range Image Capture with Plenoptic 2.0 Camera," Signal Recovery and Synthesis, OSA Optics & Photonics Technical Digest, paper SWA7P, 2009.

[7] R. Fattal, D. Lischinski, and M. Werman, "Gradient Domain High Dynamic Range Compression," ACM Transactions on Graphics (TOG), Vol. 21.3, pp. 249-256, July, 2002.

[8] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range image," ACM Transactions on Graphics (TOG), Vol. 21.3, 2002, pp. 257-266.

[9] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images." ACM Transactions on Graphics (TOG) 21.3 (2002): 267-276.

[10] R. Ng, M. Levoy, M. Bŕedif, G. Duval, M. Horowitz, and P. Hanrahan, "Light Field Photography with a Hand-held Plenoptic Camera," Stanford Tech Report CTSR, 2005.

[11] A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," Computational Photography (ICCP), IEEE International Conference on, 2009.

[12] P.-H. Huang, Y.-H. Maio, and J.-I. Guo , "High dynamic range imaging technology for micro camera array", APSIPA, pp.1 4, 9-12 Dec. 2014

[13] P. E. Debevec and J. Malik "Recovering High Dynamic Range Radiance Maps from Photographs," SIGGRAPH97, Proceedings of the 24th annual conference on Computer graphics and interactive techniques, pp 369-378, 1997.

[14] T. Georgiev and A. Lumsdaine, "Focused Plenoptic Camera and Rendering," Journal of Electronic Imaging, vol. 19, p. 021106, 2010.

[15] A. Gebejes, and R. R. Huertas, "Texture Characterization based on Grey-Level Co-occurrence Matrix", Information and Communication Technologies - International Conference (ICTIC), pages 375 – 378, 2013.

Author Biography

Chuan-Chung, Chang received his M.S. in optics and photonics (2000) and Ph.D. in optics and photonics (2009) both form National Central University, Taiwan. After receiving MS degree, he joint Industrial Technology Research Institute for more than thirteen years and now in Coretronic. He is interesting for optical testing, optical system and integrated design & solution for reference optics and computational imaging.

Hsin-Hsiang, Lo received his M.S. in optics and photonics form National Central University, Taiwan (2001). Since then he join Industrial Technology Research Institute in Taiwan for twelve years and now in Coretronic. He is interesting for lens design and optical system developing.

Han-Hsuan, Lin received her M.S. in communication engineering form National Central University, Taiwan (2014). Since then she join Coretronic. She is interesting for video decoding and image compression.

Zhi-Rong, Fan received his B.S. in electronic engineering form National Chung-Hsing University, Taiwan (2014). He is graduate student of department of Electronics Engineering in National Chiao Tung University.

Shao-Hsuan, Cheng received his M.S. in astronomy form National Central University, Taiwan (2012). Since then he join National Space Organization and working for image enhancement for Satellite imagery. He joins Coretronic since 2015.

Chih-Hung, Lu received his M.S. in communication engineering form National Central University, Taiwan (2012). Since then he join National Chung-Shan Institute of Science and Technology and now in Coretronic. He is interesting for video decoding, image compression and object reorganization.

Fu-Ming, Chuang received his Ph.D. in optics and photonics form National Central University, Taiwan (1996). He is interested for lens design and has more than thirty years in imaging lens, projection system, and micro-optics. He has been CTO of Coretronic Crop. since 2007.

Jiun-In Guo received B.S. and Ph.D. in electronics engineering from National Chiao Tung University, Taiwan, in 1989 and 1993, respectively. He is currently a full Professor of the Dept. of Electronics Engineering, National Chiao Tung University. His research interests include DSP, VLSI design, SOC design, and intelligent vision processing. He is the author of over 200 technical papers on the research areas of low-power algorithm, architecture, and system design for DSP/Multimedia/Vision processing applications.