https://doi.org/10.2352/ISSN.2470-1173.2018.05.PMII-269

Light Field Perception Enhancement for Integral Displays

Basel Salahieh, Yi Wu, Oscar Nestares; Intel Corporation, 2200 Mission College Blvd, Santa Clara, CA 95054, USA

Abstract

Light fields can be captured by plenoptic cameras and observed on integral displays supporting auto-stereoscopic viewing and full parallax. However due to the limited aperture of plenoptic cameras and the resampling process needed to overcome the resolution mismatch between the capturing and displaying devices, the perceived parallax of the resampled light fields is considerably reduced, limiting the viewing experience. We propose a light field retargeting technique that enhances the perceived parallax by first slicing the captured light fields according to their disparity, and then translating these slices with proper magnitudes and directions prior to the resampling stage. The resampled light field conserves enough parallax to be perceived in the target display. The developed technique gives users control over the depth of field and the axial location of the rendered objects as seen through the integral display.

Introduction

Light fields (LFs) [1, 2] are a collection of rays emanating from a real-world scene at various directions which enables naturally three dimensional (3D) perception when these rays are projected back to the viewers. One way to record and project light fields is by inserting a lenslet array in front of a sensor in the capture device known as plenoptic camera [3, 4] and in front of digital screen in the display device known as integral display [5, 6]. However, light fields captured by plenoptic cameras have little parallax due to the limited aperture of the main lens. In addition, the spatial and angular resolution delivered by the integral displays are constrained by the screen resolution, the displayed depth of field, and the viewing zone [6]. These factors make the display resolution different from that in the capture highlighting the need to resample the captured LF to fit on the display.

To match the display resolution given the captured LFs, simple resampling is traditionally implemented using 2D interpolation [7] which locally averages the spatial pixels within a window to produce an undersampled version of the original LF views. In the case where neighboring pixels of the LF view belong to different depth planes in 3D world coordinates, this undersampling can result in relatively shallow depth perception (e.g., parallax) on the integral display, producing unsatisfactory results. Alternatively, a 4D interpolation [8] of the LF, based on joint spatio-angular local averaging, may be employed. This approach, however, may only provide slight enhancement in the synthesized views, but at greatly increased computational cost. Due to the limited spatial resolution of current LF displays, 2D bilinear and 4D quad linear interpolation methods will reduce the disparity between the different LF views.

Birklbauer and Bimber [9] proposed a retargeting algorithm based on z-stack seam carving to rescale light fields to different resolutions while maintaining the angular consistency. Yet, this method doesn't tackle the limited parallax content of LF capturing devices. A stereoscopic warping method utilizing non-linear disparity mapping [10] is proposed to control the perceived depth. This helped retargeting stereoscopic 3D video to a different disparity range (i.e. compressed or stretched), based on visual importance of scene elements. However, it works for stereoscopic 3D display, not LF display. A similar idea in [11] produces multi-perspective stereoscopy from light fields by allowing per-pixel control over disparity and graph-cut computations, however, the runtime for a VGA resolution LF input takes a few minutes.

This paper proposes an efficient lightweight LF retargeting approach to boost the parallax content of the LF prior to the resampling step using estimated disparity maps, which significantly enhances the perceived depth and improves the motion parallax seen on the integral display.

Light Field Retargeting

The developed technique is a comprehensive end-to-end LF retargeting pipeline between the LF capture stage and the multi-view display stage, delivering an enhanced viewing experience to the users. Figure 1 shows a block diagram illustrating the stages of the method.

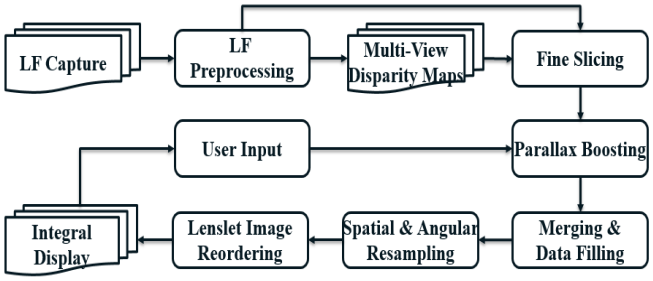

Figure 1. Block diagram of light field retargeting.

Light Field Capture and Preprocessing

The retargeting procedure begins capturing a sample of the LF which can be accomplished by different methods: a plenoptic camera, where parallax content is limited; a focal stack, which also suffers from limited parallax [12]; and a multi-camera array with controllable baseline. While our LF retargeting is of higher value in cases where the small aperture limits the amount of captured parallax (e.g., plenoptic and focal stack), it still provides interactive rendering control for all acquired LFs. Once captured, initial preprocessing operations are performed such as extraction of raw view images in rectangular grid format, denoising, color correction, and rectification.

Multi-View Disparity Estimate

We estimate sub-pixel disparity for all views. For computational efficiency, we only calculate disparities for a few preselected reference views using two-step approach: integer disparity is calculated first at a coarse level, and then is refined to subpixel precision within a small range. Since the search range is in a small range instead of whole search space, we can afford to have an accurate subpixel disparity algorithm to generate disparity in a continuous depth space with 1/20 subpixel accuracy. This two-step approach can significantly reduce runtime without sacrificing precision compared to running an expensive sub-pixel disparity algorithm directly. The estimated results are then propagated to

other views through projection models to form a complete set of multi-view disparity maps [13]. The total runtime for estimating disparity for all views (24 reference views + 172 propagated ones) is 11 second on an Intel i7 3GHz 8-cores machine retargeting LFs at resolutions specified in the experimental section.

Fine Slicing

At the fine slicing step, the estimated disparity maps are uniformly quantized into a pre-set number of levels and the LF data are sliced accordingly such that pixels close together in depths (i.e. belonging to the same quantization level of the associated disparity map) are brought together on same slice. Having more slices will result in a better parallax boosting and data filling but this is upper bounded by the subpixel accuracy of the estimated disparity maps and will require more computation.

Parallax Boosting

The resulting slices then undergo relative shifts proportional to the angular view and the distance of their disparity plane to a reference plane selected as the zero disparity plane (ZDP), according to these equations:

$$Shift_{x_{i,k}} = \lfloor Ang_{x_i} * (QuantD_k - ZDP) * DIncrement \rceil$$

$$Shift_{y_{j,k}} = \lfloor Ang_{y_j} * (QuantD_k - ZDP) * DIncrement \rceil$$

Where Ang_x and Ang_y are the normalized angular coordinates [-0.5, 0.5] indexed by i and j = 1, …, number of views in one dimension, $QuantD$ is the normalized quantized disparity map [0,1] indexed by k = 1, …, number of slices, ZDP is in normalized coordinates where $ZDP = 0$ for pop-up mode (shift with respect to background), $ZDP = 0.5$ for halved mode (shift with respect to center), and $ZDP = 1$ for virtual mode (shift with respect to front layer), $DIncrement$ is the maximum shift, and $\lfloor . \rceil$ is the rounding operator to the nearest integer for best filling results.

To better understand the shift equations, let us consider a case where a scene is sliced into 5 slices and the viewing angle is changing in 1D for simplicity. At the central view, Ang_x and $Ang_y = 0$ hence there will be no shifts in any slices regardless of the selection of ZDP, see upper-left sketch in Fig. 2. As the viewing angle is changed, the selection of ZDP will be determining the direction and amplitude of shifts imposed on layers. The layers in front the selected ZDP plane will be shifted in the opposite direction of the viewing angle while those behind ZDP will be moved in the same direction. The ZDP will experience no shift and the shift amplitude will be linearly increasing (in disparity domain) as the layers are further from ZDP. The rest of the sketches in Fig. 2 illustrate the direction and amplitude of shifts for a left viewing angle and various ZDP selections.

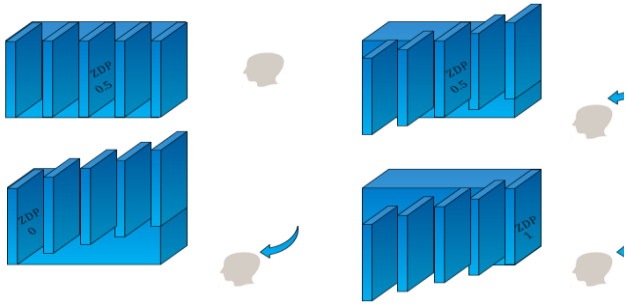

Figure 2. Parallax boosting for a scene sliced into 5 layers at two different viewing angles and ZDP modes: central view and ZDP = 0.5 at upper-left, left view and ZDP = 0.5 at upper-right, left view and ZDP = 0 at bottom-left, and left view and ZDP = 1 at bottom-right.

Merging and Data Filling

When boosting the parallax, we are actually synthesizing novel views of larger baseline and with new occlusion relations. The shifted slices per view are then merged together in ascending depth order so that the front layers can overwrite the back ones in the occluded regions to support depth cues. Note that the slicing and shifting operations result in holes (black regions) due to the new occlusion relations within the synthesized views, hence a filling operation is required.

We constrained the slices' shifts to integer values so that intensity values at sliced boundaries may not be spread over neighboring pixels. This is critical to get good filling results in the black regions since the boundary pixels will be utilized to interpolate these missing data. A nearest-neighbor interpolation is used for the data filling so the sharpness is maintained in the filled regions. The filling process is then followed by a median filtering within a small window (3×3) to impose consistency in the filled regions.

Resampling, Reordering, and Displaying

After merging and filling, the synthesized views are spatially and angularly resampled using bilinear interpolation to match the target integral display's resolution. Finally, the resampled multi-view images are reordered so pixels of same lateral indices are stacked together to form a lenslet image. This reordered image is to be displayed beneath the microlenses to steer the retargeted LF views back to viewers in the proper directions.

User Input

Two key parameters can be controlled by the user interactively through either gesturing or keystrokes to experience different rendering modes: $DIncrement$ and ZDP. The user controls the perceived depth of field (DOF) by adjusting $DIncrement$ to experience stretching and compressing effects in depth. Also, they can assign which disparity slice to be rendered at display surface (usually the lenslet plane) by adjusting ZDP. This determines the slices to be floating on top of the display as opposed to those to be virtually created inside it. Once adjusted, the program is looped back to the parallax boosting stage to synthesize the new LF content accordingly. Note that the interactive loop including parallax boosting, merging, filling, resampling, and reordering takes less than 50 millisecond to execute on Intel i7 3GHz 8-cores machine retargeting LFs at resolutions specified in the following section.

A flow diagram summarizing the main steps is shown in Fig 3. Note that parallax boosting is the novel step leading to an enhanced LF perception on the integral displays.

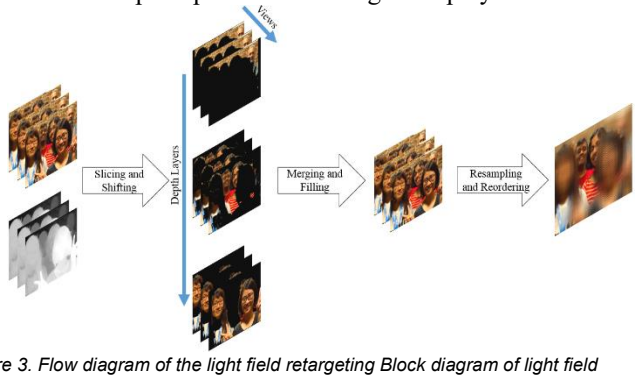

Figure 3. Flow diagram of the light field retargeting Block diagram of light field retargeting technique shown for 3 horizontal views and 3 disparity layers with central ZDP.

Experimental Results

In this section we illustrate and analyze our LF retargeting results as compared with other approaches. Unless stated differently, the retargeting parameters used in the study were set as follows: 100 slices, $ZDP = 0.5$, $DIncrement = 100$, and all results are shown at the display resolution (after resampling). Lytro Illum [14], an example of commercial plenoptic camera, was used for capturing LFs at 14×14×375×541 angular-spatial resolution. The data is then retargeted for an integral display composed of a Sony Xperia Z5 4K screen and a lenslet array of 0.5 mm lens pitch separated by its focal length from the screen plane to have a homogenous resolution over wide DOF in a focused display mode [15]. The resulting integral display has 16×16×135×240 angular-spatial resolution. Further simulation results evaluating the parallax boosting at different $DIncrement$ and ZDP values can be found in [16]. All the results were verified visually on the actual integral display, shown in Fig. 4. We have intentionally left part of the Sony screen uncovered in Fig. 4–left so readers can see how the scene looks like with and without the lenslet array. We also show samples of two different views in Fig. 4–right as seen on the integral display.

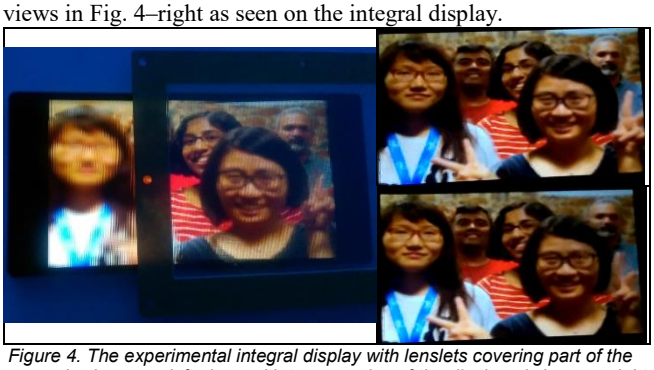

Figure 4. The experimental integral display with lenslets covering part of the screen is shown on left along with two samples of the displayed views on right

Comparison of Resampling Techniques

Figure 5 shows an example using 2D bilinear, 4D quad linear interpolation and our method. To help illustrate the differences between these methods, we have selected two extreme views, the top-left view and the right-bottom view, that are displayed in Fig. 5-1st and 2nd columns. To compare the differences between the two corner views, an overlaid image is generated such that pixels are shown in gray values where intensities are matched between the two views and in different color bands (i.e. green and magenta) where they are differed. Note that more color regions means better parallax. These color regions are significant in the case our LF retargeting approach is used as opposed to 2D and 4D interpolation, see overlaid views in Fig. 5-3rd column. Thus, the people in the scene can be uniquely observed from different perspectives as observers are moving from top-left to bottom-right within the eyebox (viewing zone) of an integral display.

Figure 5. Comparison of 2D, 4D interpolation and ours.

Comparison with Lytro View Synthesis Results

We compare our parallax-enhanced LF retargeting results with those generated by the view synthesis tool available in Lytro power tools (LPT) [17]. The tool can be used to synthesize views beyond the physical aperture such that they will have larger parallax but this comes with artifacts. It is unclear how the tool is implemented, but we share the comparison results here for completeness. For a fair comparison, we synthesize views of similar parallax content by both techniques, as shown in Fig. 6. This is done by extracting views over 16× the aperture size in LPT tool while imposing $DIncrement = 150$ in our LF retargeting method.

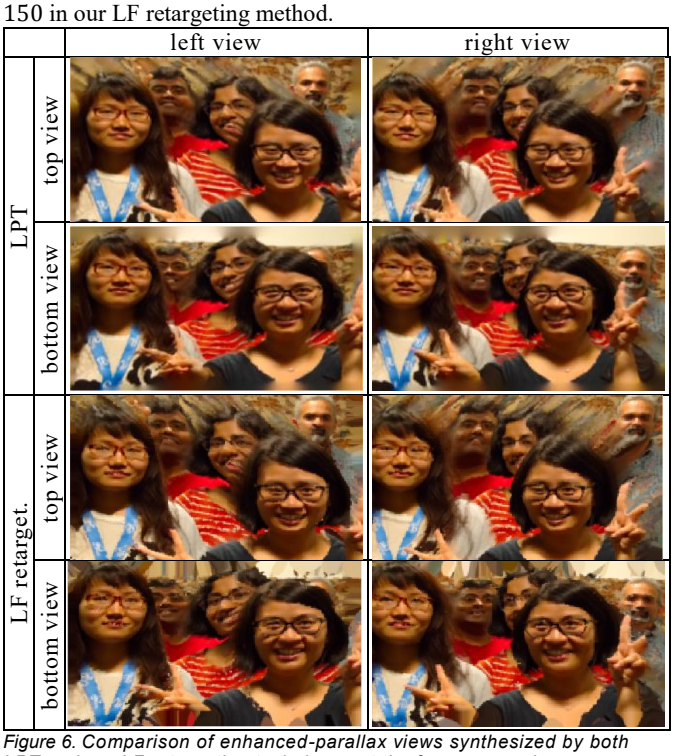

Figure 6. Comparison of enhanced-parallax views synthesized by both LPT and our LF retargeting techniques at the four corner views.

A closer look at certain regions in the synthesized bottom-right view, shown in Fig. 7, reveals some differences: The hand at the left images shows aliasing artifacts and there is noticeable blur in regions extrapolated by the LPT tool, while our LF retargeting result is sharper and doesn't show noticeable aliasing. The hair separating the two faces at the central images seems to be unrealistically thinned in LPT while ours preserves its natural appearance. Finally, the index finger in right images seems to have a wavy shape in the LPT result while it is straighter in ours. Admittedly the middle finger in ours seems to be cut, this is due to being at the boundary region where part of it got shifted outside the image frame during the parallax boosting step. This can be solved by padding the images prior to shifts while extracting the original frames after the filling.

Figure 7. Zoomed bottom-right views to illustrate visual differences between LPT and LF retargeting.

Validating the parallax enhancement on other LFs

We further validate our LF retargeting technique on additional datasets generated by different cameras and show the improvement after the parallax boosting as opposed to a simple 2D resampling of original LFs; the first dataset is a group of people (Fig. 8) captured by Lytro Illum (Plenoptic 1.0) camera [14], the second is an electronic board (Fig. 9) captured by Raytrix R29 (Plenoptic 2.0) camera [18], the last is a synthesized LF for Mona scene (Fig. 10) from Heidleberg database [19].

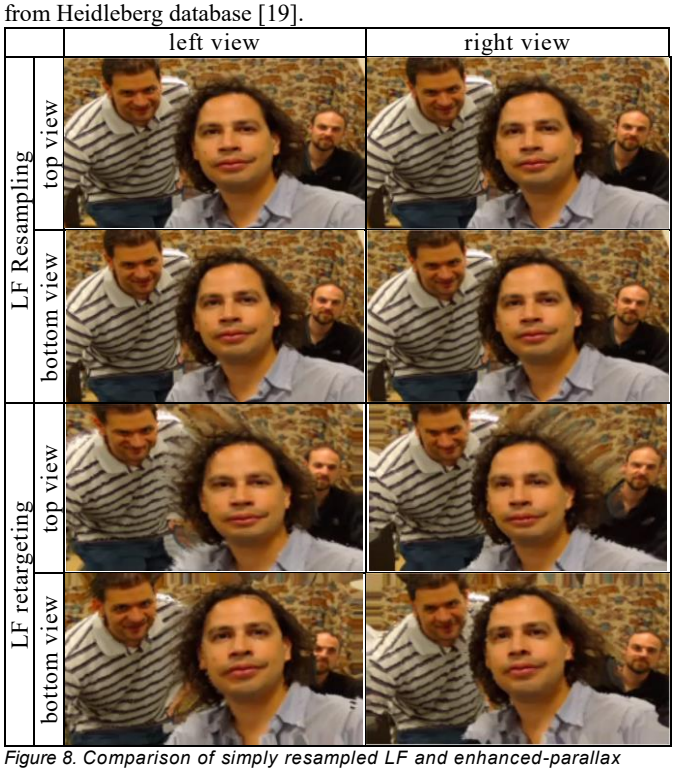

Figure 8. Comparison of simply resampled LF and enhanced-parallax retargeted LF captured by Lytro Illum camera at the four corner views.

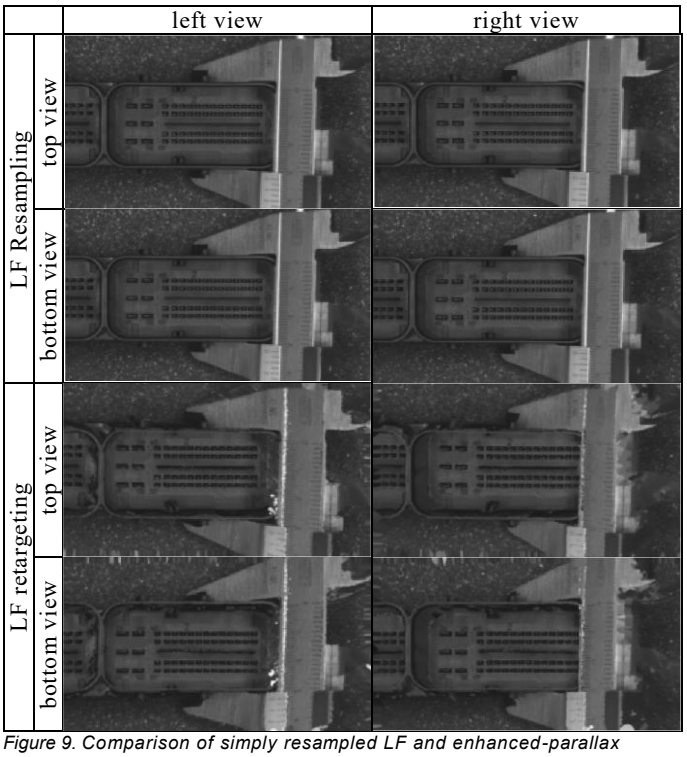

Figure 9. Comparison of simply resampled LF and enhanced-parallax retargeted LF captured by Raytrix R29 camera at the four corner views.

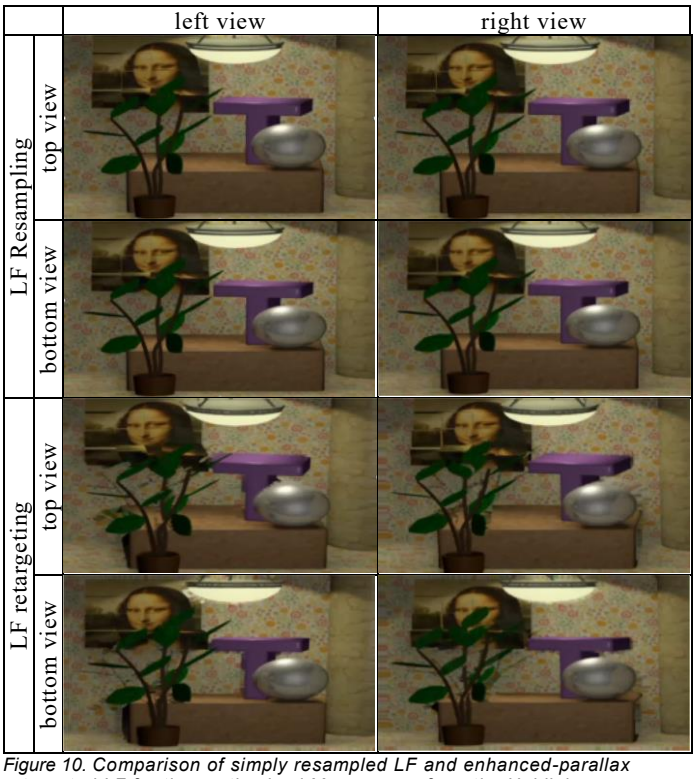

Figure 10. Comparison of simply resampled LF and enhanced-parallax retargeted LF for the synthesized Mona scene from the Heidleberg database.

Conclusions

Light field capturing systems, especially single aperture ones, may restrict the amount of parallax in the captured light field. In addition, capture-display mismatch resolution requires resampling operations which in turn results in shallower depth content as seen on the integral displays. Our light field retargeting technique address these problems by synthesizing views with enhanced parallax as if they were captured by cameras with a larger virtual baseline. The developed technique provides better rendering experiences by enabling the user selecting different depth of field and axial translation effects through the display surface.

Future work will include enhancing the quality of the filled regions by incorporating advanced light field inpainting algorithms, including training models, semantic analysis, and saliency constraints. We will further research relaxing the need for multi-view disparity estimate stage by better understanding the scene and use this knowledge to do efficient slicing and parallax boosting.

References

[1] E. H. Adelson and J. R. Bergen, "The Plenoptic Function and the Elements of Early Vision," In Computation Models of Visual Processing, MIT Press, Cambridge, pp. 3–20, 1991.

[2] M. Levoy and P. Hanrahan, "Light Field Rendering," in Proceedings of the 23rd annual conference on Computer Graphics and Interactive Techniques, SIGGRAPH, 31-42, 1996.

[3] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light Field Photography with a Hand-Held Plenoptic Camera," Computer Science Technical Report CSTR 2 (11): 1-11, 2005.

[4] A. Lumsdaine and T. Georgiev, "The Focused Plenoptic Camera," International Conference on Computational Photography (ICCP), IEEE, 1-8, 2009.

[5] D. E. Roberts, "History of Lenticular and Related Autostereoscopic Methods," White Paper, Leap Technologies LLC., 2003.

[6] S. W. Min, J. Kim, and B. Lee, "New Characteristic Equation of Three-Dimensional Integral Imaging System and Its Applications," Jpn. J. Appl. Phys. 44, L71–L74, 2005.

[7] F. B. Atalay and D. M. Mount, "Interpolation over Light Fields with Applications in Computer Graphics," Proc. of the 5th Workshop on Algorithm Engineering and Experiments (ALENEX'03), SIAM, 56-68, 2003.

[8] C. Zhang and T. Chen, Light Field Sampling, Morgan & Claypool Publishers, 2006.

[9] C. Birklbauer and O. Bimber, "Light-Field Retargeting," Computer graphics forum, 2012.

[10] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, "Nonlinear Disparity Mapping for Stereoscopic 3D," in ACM transactions on graphics. Proceedings of SIGGRAPH, vol. 29(4), 2010.

[11] C. Kim, A. Hornung, S. Heinzle, W. Matusik, and M. Gross, "Multi-Perspective Stereoscopy from Light Fields," ACM Transactions on Graphics (TOG), 30.6, 2011.

[12] A. Levin and F. Durand, "Linear View Synthesis Using a Dimensionality Gap Light Field Prior," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 1831-1838, 2010.

[13] B. Salahieh1, S. Hunter, Y. Wu, O. Nestares, "Light Field Retargeting for Multi-Panel Displays," (Submitted for Publication).

[14] Lytro Illum Camera, https://www.lytro.com/imaging, (Accessed on Aug., 2017).

[15] J. Hong, Y. Kim, H.-J. Choi, J. Hahn, J.-H. Park, H. Kim, S.-W. Min, N. Chen, and B. Lee, "Three-Dimensional Display Technologies of Recent Interest: Principles, Status, and Issues," Appl. Opt. 50, H87-H115, 2011.

[16] B. Salahieh, Y. Wu, and O. Nestares, "Light Field Retargeting from Plenoptic Camera to Integral Display," 3D Image Acquisition and Display, OSA Imaging and Applied Optics, 2017.

[17] Lytro Power Tools, https://www.lytro.com/imaging/power-tools, (Accessed on Aug., 2017).

[18] Raytrix R29 Camera, https://www.raytrix.de/produkte/#r29series, (Accessed on Aug., 2017).

[19] S. Wanner, S. Meister, and B. Goldluecke, "Datasets and Benchmarks for Densely Sampled 4D Light Fields," In Vision, Modelling and Visualization (VMV), 2013.

Author Biography

Basel Salahieh received his BS in electrical engineering from the University of Aleppo – Syria (2007), his 1st MS in electrical engineering from the University of Oklahoma (2010), his 2nd MS in optical sciences and PhD in electrical and computer engineering from the University of Arizona (2015). Since then he has worked as Research Scientist at Intel Labs in Santa Clara, CA. His work has focused on computational imaging, light fields, autostereoscopic displays, and virtual reality systems. He has published 15+ technical papers and has 5+ pending patents.

Yi Wu received her Ph.D. in electrical and computer engineering from the University of California, Santa Barbara (UCSB) in 2005. Yi Wu is now a Computer vision Researcher at Houzz inc. Before joining Houzz, she was a Senior Research Scientist at Intel Labs, Santa Clara, CA. Her research interests include computer vision, machine learning, augmented reality, image/video processing. She owns multiple US patents and published 40+ technical publications. She has served a program committee member and organizing committee member of 10+ international conferences and workshops on multimedia, video and image processing.

Oscar Nestares is a Principal Engineer at Intel Labs working on computational imaging algorithms, light field processing, and video enhancement. Before he was a tenured Research Scientist (Institute of Optics, CSIC) working on visual system models and biologically inspired image and video processing and analysis, a Fulbright Scholar at Stanford University, and consultant at Xerox PARC. He received his M.S. (1994) and Ph.D. (1997) in Telecommunication Engineering from Universidad Politécnica de Madrid. He has published 30+ papers in international journals and conferences and 30+ patents.