https://doi.org/10.2352/ISSN.2470-1173.2021.7.ISS-067

Under Display Camera Quad Bayer Raw Image Restoration using Deep Learning

Irina Kim, Yunseok Choi, Hayoung Ko, Dongpan Lim, Youngil Seo, Jeongguk Lee, Geunyoung Lee, Eundoo Heo, Seongwook Song, Sukhwan Lim; Samsung Electronics Ltd; Hwaseong, South Korea

Abstract

Can a mobile camera see better through display? Under Display Camera (UDC) is the most awaited feature in mobile market in 2020 enabling more preferable user experience, however, there are technological obstacles to obtain acceptable UDC image quality. Mobile OLED panels are struggling to reach beyond 20% of light transmittance, leading to challenging capture conditions. To improve light sensitivity, some solutions use binned output losing spatial resolution. Optical diffraction of light in a panel induces contrast degradation and various visual artifacts including image ghosts, yellowish tint etc. Standard approach to address image quality issues is to improve blocks in the imaging pipeline including Image Signal Processor (ISP) and deblur block. In this work, we propose a novel approach to improve UDC image quality - we replace all blocks in UDC pipeline with all-in-one network – UDC d^Net. Proposed solution can deblur and reconstruct full resolution image directly from non-Bayer raw image, e.g. Quad Bayer, without requiring remosaic algorithm that rearranges non-Bayer to Bayer. Proposed network has a very large receptive field and can easily deal with large-scale visual artifacts including color moiré and ghosts. Experiments show significant improvement in image quality vs conventional pipeline – over 4dB in PSNR on popular benchmark - Kodak dataset.

Introduction

Under Display Camera (UDC) enabling infinity display, with larger screen to body ratio, is the most awaited and challenging feature in mobile market in 2020.

What are the challenges we face with UDC imaging? First of all, we are struggling to achieve desired image quality. Current Organic Light-Emitting Diode (OLED) panels can only provide transmittance ratio of up to 20%, so UDC image is captured in lowlight conditions. To improve light sensitivity, conventional solution uses binned output combining adjacent pixels, i.e., it can only support less than a quarter of original image sensor resolution.

Another issue is optical diffraction from non-transparent panel that can induce not only contrast loss, but also various visual artifacts including ghosts, flare, color tints etc. (See Fig. 2). The question is how can we overcome those challenges? Can we keep full resolution and get artifacts-free UDC image?

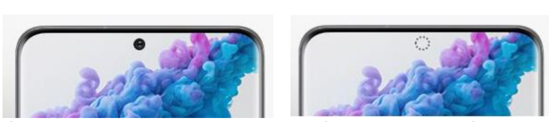

Fig. 1. Illustration of notch-type and Under Display Cameras.

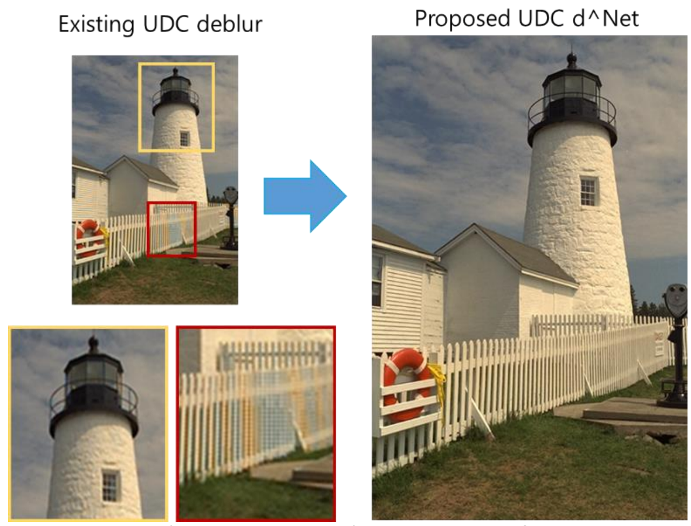

Fig. 2. Illustration of proposed solution: left – conventional UDC solution output; right – proposed method output. One can clearly see that proposed method can achieve superior artifacts-free details restoration at full resolution.

Recently, to enhance sensitivity with high resolution image sensor, various types of non-Bayer pattern, such as Quad Bayer, shown in Fig.3, are adopted. Such sensors can easily obtain high sensitivity images under low-light condition using binning method. Under normal light conditions, it requires remosaic function that rearranges non-Bayer to Bayer, to achieve full resolution imaging. The remosaic block is usually followed by Bayer Image Signal Processor (ISP) and deblur block, all operating in different domains. However, when applying this sequential hand-crafted pipeline, we end up with blur image and more visual artifacts induced by remosaic and ISP due to lack of spatial information. Image quality can be improved by retuning ISP or deblur algorithm, however it gives only marginal improvement in image quality while keeping large-scale artifacts including zippering, ghosts, false colors etc.

In this work, we propose an end-to-end **one-shot UDC imaging solution,** that can deblur and reconstruct raw UDC image using one neural network – UDC d^Net. Due to hierarchical structure, designed network has a large receptive field allowing to extract more features from source image and handle large-scale artifacts. Experiments show that proposed method outperforms conventional algorithms by a large margin – by 4dB in PSNR on Kodak dataset, for Quad Bayer CFA. Visual evaluation confirms that it can reconstruct details better and does not produce visual artifacts. Using proposed method, we can model any non-UDC ISP pipelines with better image quality and less artifacts.

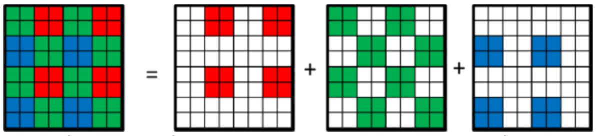

Fig. 3. Quad Bayer Color Filter Array decomposition into R, G, B components.

Contributions. To the best of our knowledge, this is first work addressing a novel problem of Quad Bayer Raw Image reconstruction for Under Display Cameras on mobile devices using deep learning. Recently, there is a trend in high-resolution CMOS sensors to adopt novel CFAs, for example Quad Bayer CFA. We solve this challenging problem using deep learning in raw domain by joint deblurring and demosaicing. Proposed solution achieves better image quality, outperforming conventional methods by 4dB in CPSNR and producing less visual artifacts.

Related works

In this work we focus mostly on demosaic and deblurring algorithms, so we will briefly review related works.

Demosaicing of Bayer Color Filter Array (CFA) has been extensively studied for several decades [1], [2], [6]. Various demosaicing approaches are exploited, such as color difference based interpolation [24], [25], edge directional interpolation [26], frequency domain filtering [3], [4], [5], and reconstruction methods [27], [28]. However, when it comes to new patterns, such as Quad Bayer depicted in Fig. 3, there are only a few works that can be applied to it. Unfortunately, hand-crafted algorithms are mainly designed for most commonly used until recently Bayer CFA, so they have to be redesigned to support any other CFA pattern. Universal demosaicing algorithm, that can be easily extended to new type of CFA pattern, was proposed in Zhang [7].

Deep learning approach to image demosaicing has been applied in 2016 - see [8], [9], [10], [11]. Early works were also designed for Bayer CFA, however recently there are two articles addressing Quad Bayer CFA - Kim [12] and Stojkovic [13] . Compared to hand-crafted algorithms, deep learning methods do not need to be completely redesigned and can be adopted to support various CFA patterns. However, for Quad Bayer CFA, learning method and network architecture shall be also changed, to avoid artifacts that cannot be seen in standard Bayer CFA.

Deep learning methods to image deblurring have been actively studied in the last five years [12], [15]. Shuler used several networks to model iterative optimization process in a "coarse-to-fine" manner [16]. Nah [17] and Tao [18] introduced a multi-scale cascade of networks that sequentially restores downscaled images. Zhang adopted a method of dividing images into patches, performing sequential restoration of small to large patches, using several networks [20]. Most state-of-the-art methods uses a cascade of networks to deblur a single image.

Image restoration for Under Display Camera, was studied in [21] for Bayer CFA, for two types of OLED displays: T-OLED and P-OLED and proposed to restore image from raw domain, showing better performance compared to Wiener Filter and plain residual networks.

Problem formulation

Demosaicing of each color channel can be treated as an interpolation problem of each channel, however due to phase shift during color subsampling and inter-channel dependency, demosaicing normally comes with various visually disturbing artifacts: color moiré, false colors or zippering along edges. For irregularly subsampled Quad Bayer pattern, aliasing increases and cause severe artifacts compared to Bayer – see [12] for more detailed analysis.

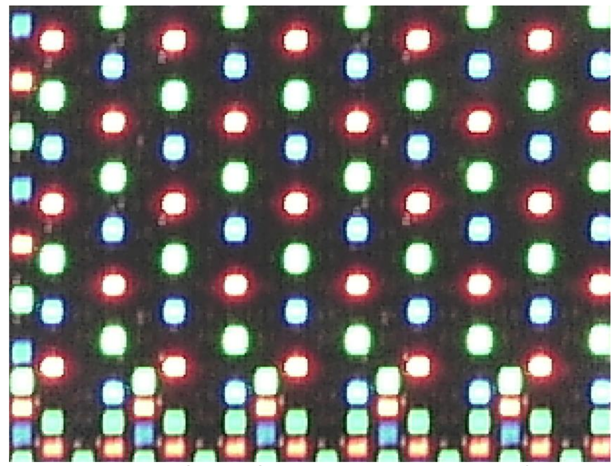

Fig. 4. Under Display Camera OLED Panel pixel arrangement (from ZTE Axon mobile phone).

When camera is placed under display, the amount of incident light is considerably decreased. Due to specific UDC display structure, the light cannot propagate freely —instead it has to pass through tiny holes and smaller open area compared to conventional cameras, see Fig. 4 for example. In addition, UDC display area is not transparent—depending on materials, transparency can vary. Insufficient amount of transmitted light creates challenging low-light capturing conditions, non-transparent display often induces specific color artifacts, like yellowish tint [21].

On top of that, since typical distance between pixels in UDC panel is very small (100~200um), when light travels through slits between pixels, diffraction of light occurs. Interference between display pixels and light wave results in the higher order peaks of light intensity, as depicted in Fig. 5. It affects captured image quality by having excessive blur, loss of details and contrast and ghosts.

Depending on particular UDC display structure, Point Spread Function (PSF) may have various sizes and shapes. In real applications, we are limited in the size of support of corresponding PSF, so in this work we used small size PSF.

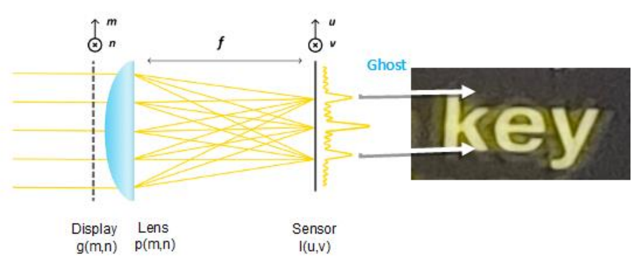

Fig. 5. Under Display Camera Image formation: light propagates through tiny holes in the UDC panel, undergoing diffraction [21].

Proposed method

In conventional solution, we first apply remosaicing on raw non-Bayer image, following by Bayer ISP processing and then apply deblurring in RGB domain. The process is shown in the Fig. 6.

Our hypothesis is that deblurring in raw domain will be more efficient and will allow to keep more details and avoid artifacts propagation. In addition to this, we propose to perform deblurring simultaneously with demosaicing.

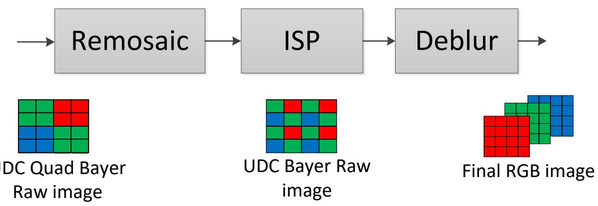

Fig. 6. Block diagram of conventional solution.

We propose to replace one or more blocks in ISP and deblur block with all-in-one neural network – UDC d^Net. Here d^ stands for de-blur, de-mosaic, de-noise, etc. Block diagram of the proposed method is depicted in Fig. 7.

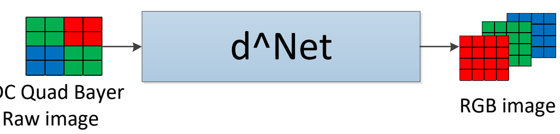

Fig. 7. Block diagram of the proposed method.

In this work, we use standard image formation model (1).

$$Y = K * M * X + n, \tag{1}$$

where N – noise, K – blur with kernel K, M – mosaic, X – original image, Y – distorted observation.

Using deep learning, we can solve this ill-posed inverse problem efficiently, without domain conversion. By using power of big data, we can address this ill-posed inverse problem more efficiently.

We learn end-to-end mapping function F from training samples pairs by taking RGB images as ground truth and mosaicked blurred images as observed images. We estimate model parameters Ω by minimizing following loss function:

$$Loss(\Omega) = L_2 \tag{2}$$

where L_2 is an Euclidean norm:

$$L_2 = \frac{1}{n}\sum_{i=1}^{n}\left\|F\left(X_{i,}\Omega\right) - Y_i\right\|_2 . \tag{3}$$

Network architecture

We design a neural network for image reconstruction – UDC d^Net, inspired by the Duplex Pyramid Network architecture by Kim [12], showing superior image quality for Quad Bayer image demosaicing compared to other deep learning methods and outperforming conventional algorithms.

Inspired by hierarchical structure of network in [12], we can extract image features at various levels, that will allow us to reconstruct image details at various levels, as well recover large scale artifacts. We assume, that such a hierarchical network structure will allow to capture blur at various levels, avoiding using a cascade of neural networks as in the state-of-the-art image deblurring algorithms [17], [18].

We adopt backbone skeleton architecture and FEB (Feature Extraction Block) that includes local residual block. D^2 network architecture is shown in Fig. 6. Here FRB is Feature Reconstruction Block, it includes one convolution and one ReLU. We use 3x3 convolutions at all levels.

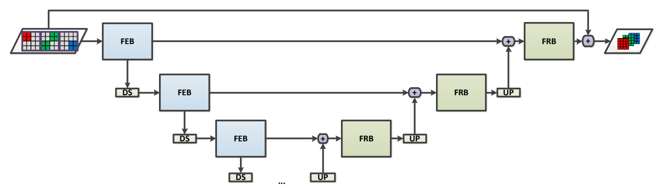

Fig. 8. Block diagram of the proposed network - d^Net.

Experimental results

We conduct experiments by preparing pairs of distorted input and ground truth images. The network was trained on MIT dataset [8] on patches of size 128x128. We augmented input data with random flip and rotation.

We trained our model with ADAM optimizer [22] with the following settings: $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$, weight decay = 10^{-8}. We set initial learning rate as 10^{-4} and schedule learning rate decrease at milestones [5, 10, 20], with decay = 0.1. The model was implemented using Pytorch and trained on NVIDIA Volta GPUs.

We used distortion model with estimated PSF of 11x11 from the sample mobile OLED panel and Quad Bayer mosaic operator. We tested our algorithm on Kodak dataset [23], in Color PSNR (CPSNR), defined as below.

$$CPSNR = \frac{1}{3}\sum_{i=1}^{3} PSNR(C_i), where \tag{4}$$

$$PSNR(C_i) = 10log_{10}\frac{2^{2b}}{\sum_{i=1}^{n}\|X_i - Y_i\|_2} \tag{5}$$

where C – color components, b – image bitwidth, n - number of pixels, X – ground truth image, Y – reconstructed image.

Objective image quality evaluation results are provided in the Table 1. For reference, we used universal demosaicing algorithm[7] applied to Quad Bayer CFA pattern and iterative deblur algorithm [19]. We can see that current algorithm outperforms conventional approach by a large margin.

Table 1. Image quality evaluation results, CPSNR [dB].

Algorithm	Conventional	Proposed
PSF1 [3x3]	36.3	40.0
PSF2 [11x11]	35.6	39.6

Subjective evaluation of experimental results show that we could reconstruct image with more details and less artifacts as compared to conventional approach. As illustrated in the Fig. 9, we observed significant loss of high frequency details and various artifacts: zippering and false colors – when using conventional method, while we could achieve almost perfect image reconstruction when using proposed method.

a) Conventional b) Proposed c) Ground truth

Fig. 9. Examples of visual quality evaluation from Kodak dataset. One can clearly see that proposed approach outperforms reference and can restore details and textures close to ground truth.

Next, we applied our method on real UDC image, captured by sample mobile OLED panel. We trained our network using same training dataset and performed visual evaluation of the results as shown in Fig. 11. Proposed method efficiently removes blur and ghosts, so it can be applied to real life scenario. Despite of having only one neural network in the proposed solution as opposed to prior state-of-the-art, we could achieve desired deblurring performance. Proposed network architecture is computationally efficient so can be deployed in mobile devices after optimization.

a) Captured UDC image *b) Result of proposed method*

Fig. 10. Examples of visual quality evaluation on real UDC images

Conclusion

In this work, we proposed a novel imaging solution for UDC that can be adopted for mobile phones with infinity displays at full resolution. Experiments show that it achieves better image quality, reduced blur and other visually disturbing artifacts, outperforming conventional solution for Quad Bayer image sensor. Proposed solution can be adopted to any non-UDC pipeline and can improve quality of any imaging pipeline for any image sensor, Bayer or non-Bayer.

References

[1] X. Li, B. Gunturk and L. Zhang, "Image demosaicing: A systematic survey," in Proc. of SPIE, 6822, 68221J, 2008.

[2] D. Menon and G. Calvagno, "Color image demosaicking: An overview," Sig. Proc.: Image Comm., vol. 26, pp. 518-533, 2011.

[3] D. Alleysson, S. Susstrunk and J. Hérault, "Linear demosaicing inspired by the human visual system," IEEE Transactions on Image Processing, vol. 14(4), pp. 439-449, 2005

[4] E. Dubois, "Frequency-domain methods for demosaicking of Bayer-sampled color images," IEEE Signal Processing Letters, Vol. 12(12), pp.847-850, 2005

[5] P. Hao, Y. Li, Z. Lin and E. Dubois, "A geometric method for optimal design of color filter arrays," IEEE Transactions on Image Processing (TIP), Vol. 20(3), 709-722, 2010.

[6] B. Bruce, "Color imaging array," US patent 3,971,065, 1976.

[7] C. Zhang, Y. Li, J. Wang and P. Hao, "Universal Demosaicking of Color Filter Arrays," IEEE Transactions on Image Processing, 25, No. 11, pp. 5173-5186, 2016.

[8] M. Gharbi, et al. "Deep Joint Demosaicking and Denoising." ACM Transactions on Graphics. Vol. 35, pp. 191:1--191:12, 2016.

[9] R. Tan, K. Zhang, W. Zuo and L. Zhang, "Color image demosaicking via deep residual learning," in Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), 2017.

[10] D. S. Tan, W. Chen and K. Hua, "Deep Demosaicking: Adaptive Image Demosaicking via Multiple Deep Fully Convolutional Networks," IEEE Transactions on Image Processing, vol. 27, pp. 2408-2419, 5 2018.

[11] N.-S. Syu, Y.-S. Chen and Y.-Y. Chuang, "Learning Deep Convolutional Networks for Demosaicing," CoRR, vol. abs/1802.03769, 2018.

[12] I. Kim et al, "Deep image demosaicing for submicron image sensors," Journal of Imaging Science and Technology (JIST), vol. 63, No.6, pp. 060410-1--060410-12, 2019.

[13] A. Stojkovic et al., " The Effect of the Color Filter Array Layout Choice on State-of-the-Art Demosaicing," Journal MDPI Sensors, vol. 19(14): 3215, 2019. (DOI: 10.3390/s19143215)

[14] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1964–1971, 2009.

[15] R. Wang and D. Tao, "Recent progress in image deblurring." ArXiv, abs/1409.6838, 2014.

[16] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Sch¨olkopf, "Learning to deblur," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38(7), pp.1439–1451, 2016.

[17] S. Nah, T.H. Kim and K.M. Lee," Deep multi-scale convolutional neural network for dynamic scene deblurring," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[18] X. Tao et al., "Scale-recurrent network for deep image deblurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[19] W. Zuo et al., "A Generalized Iterated Shrinkage Algorithm for Non-convex Sparse Coding," In Proc. of IEEE International Conference on Computer Vision, Sydney, NSW, pp. 217-224, 2013. doi: 10.1109/ICCV.2013.34.

[20] H. Zhang, et al, "Deep stacked hierarchical multi-patch network for image deblurring," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[21] Y. Zhou, et al, "Image restoration for under-display camera," arXiv preprint arXiv:2003.04857, 2020.

[22] D.-P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in Proc. of 3rd International Conference for Learning Representations, San Diego, 2015 (arXiv:1412.6980).

[23] A. Loui et al., "Kodak's Consumer Video Benchmark Data Set: Concept Definition and Annotation," in Proc. of International Workshop on Workshop on Multimedia Information Retrieval, New York, NY, USA, 2007.

[24] D.R. Cok, "Signal processing method and apparatus for producing interpolated chrominance values in a sampled color image signal," U.S. Patent 4 642 678, 1986.

[25] J.E. Adams, "Interactions between color plane interpolation and other image processing functions in electronic photography," Proc. SPIE, vol. 2416, pp. 144–151, 1995.

[26] R. Kimmel, "Demosaicing: Image reconstruction from CCD samples," IEEE Trans. Image Process., vol. 8, no. 6, pp. 1221–1228, 1999.

[27] J. Mukherjee, R. Parthasarathi, S. Goyal, " Markov random field processing for color demosaicing," IEEE Pattern Recognition Lett., vol. 22(3–4), pp. 339–351, 2001.

[28] D. Keren, M. Osadchy, "Restoring subsampled color images," in Proc. of Mach.Vision Appl., vol. 11(4), pp. 1197–202, 1999.

Author Biography

Irina Kim received her B.S. and M.S. degrees with honors in applied mathematics, in 2002 and 2004, respectively, and the Ph.D. course in mathematics from the National Research University of Electronic Technology, Moscow, in 2005. Since 2001, she developed Nano microscopic and satellite image analysis and video surveillance algorithms, before joining Samsung Electronics in 2005, where she worked on Image Signal Processor (ISP) algorithms, face detection and vision engines. Her recent research is focused on deep learning for latest CMOS sensor for mobile devices.

Yunseok Choi received his B.S., M.S., and Ph.D. degrees in Electrical Engineering and Computer Science from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1996, 1998, and 2005, respectively. Since 2004, he has been with Samsung Electronics where he developed 3D graphics engine, ISP, CIS sensor, and so on. His main research interests include imaging processing systems and imaging sensors.

Hayoung Ko received the B.S. and M.S. in Electronic Engineering from Korea University, Korea, in 2003 and 2012, respectively. She worked on Image and Video signal processing. She is currently working on 2D and 3D camera calibration and Image Enhancement after joining Samsung Electronics in 2013.

Youngil Seo received his B.S in Electrical Engineering from Hanyang University and M.S. in Electrical Engineering and Computer Science from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2001 and 2003, respectively. Since 2009, he has been with Samsung Electronics where he developed Video codec, GPU, Sensor IP and so on. His main research interests include imaging processing systems and deep learning now.

Seongwook Song received his B.S. and M.S. degrees in electrical engineering from Seoul National University, in 1997 and 1999, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana–Champaign, in 2004. He has been with Samsung Electronics since 2003, to develop 2G, 3G and 4G chipsets. His main research interests include advanced signal processing for digital communications, multimedia and deep learning systems for digital cameras.

SukHwan Lim received his B.S. degree (with honors) in electrical engineering from Seoul National University, Korea, in 1996, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, US, in 1998 and 2003, respectively. He has worked in HP Laboratories, Apple and Google before joining Samsung Electronics in 2019. His research interests include image sensors, digital cameras and image/video processing architectures and algorithms including denoising, deblurring, and superresolution.