

Extending depth of field and dynamic range from differently focused and exposed images

Qinchun Qian · Bahadir K. Gunturk

Received: 19 February 2014 / Revised: 9 January 2015 / Accepted: 19 January 2015 /
Published online: 31 January 2015
© Springer Science+Business Media New York 2015

Abstract Focus stacking and high dynamic range (HDR) imaging are two paradigms of computational photography. Focus stacking aims to produce an image with greater depth of field (DOF) from a set of images taken with different focus distances; HDR imaging aims to produce an image with higher dynamic range from a set of images taken with different exposure values. In this paper, we present an algorithm which combines focus stacking and HDR imaging in order to produce an image with both extended DOF and dynamic range from a set of differently focused and exposed images. The key step in our algorithm is focus stacking regardless of the differences in exposure values of input images. This step includes photometric and spatial registration of images, and image fusion to produce all-in-focus images. This is followed by HDR radiance estimation and tonemapping. We provide experimental results with real data to illustrate the algorithm.

Keywords High dynamic range imaging · Focus stacking · Photometric and spatial registration

1 Introduction

One of the goals of computational photography is to exceed the limitations (such as spatial resolution, depth of field, field of view, and dynamic range) of cameras through some mod-

This work is supported in part by Texas Instruments.

Q. Qian · B. K. Gunturk (✉)
School of Electrical Engineering and Computer Science, Louisiana State University, Baton Rouge,
LA 70803, USA
e-mail: bahadir@ece.lsu.edu

Q. Qian
e-mail: qqian1@lsu.edu

B. K. Gunturk
Istanbul Medipol University, Istanbul, Turkey

ification of the camera and/or capturing multiple images. Focus stacking and high dynamic range (HDR) imaging are two well-known paradigms of computational photography where multiple images with different camera settings are captured and merged. In focus stacking, the goal is to extend the depth of field by using multiple images that are focused at different depths. In HDR imaging, the goal is to improve the dynamic range by using multiple images that have different exposure values.

Focus stacking is a relatively old research area; a large volume of papers have been published on this topic since early 1980s (Pieper and Korpel 1983; Sugimoto and Ichioka 1985; Burt and Kolczynski 1993; Li et al. 1995; Subbarao and Choi 1998; Valdecasas et al. 2001; Li et al. 2001; Forster et al. 2004; Antunes et al. 2005; Huang and Jing 2007; Aguet et al. 2008; Tian et al. 2011; Pertuz et al. 2013). A typical focus stacking algorithm consists of two steps: (1) application of a focus measure to each input image to determine the amount of focus at each pixel, and (2) fusion of input images based on the focus measure. Focus measures can be based on point-wise intensity values in the image stack (Pieper and Korpel 1983), spatial energy (measured using variance, gradient energy, Laplacian energy, edge pixel count, etc.) (Sugimoto and Ichioka 1985; Subbarao and Choi 1998; Li et al. 2001; Antunes et al. 2005; Huang and Jing 2007; Tian et al. 2011), and wavelet coefficients (Burt and Kolczynski 1993; Li et al. 1995; Valdecasas et al. 2001; Forster et al. 2004). Comparisons of some of these measures can be found in (Valdecasas et al. 2001; Huang and Jing 2007). The fusion process picks up the pixel value from the image with the highest focus value or does a weighted sum of pixel intensities, where the weights depend on the focus values. The fusion process can be done in spatial domain, scale-space domain, or wavelet-domain (Aguet et al. 2008; Pertuz et al. 2013).

HDR imaging from multiple differently exposed images has been receiving increasing attention in the last decade. HDR imaging algorithms produce a radiance map, which typically has much higher dynamic range than a typical display can show; therefore, a tonemapping step is required to display HDR images. Camera response function (CRF) is needed to obtain the radiance map. CRF estimation can be based on a non-parametric (Debevec and Malik 1997) or a parametric (Mann and Mann 2001; Mitsunaga and Nayar 1999) model. When merging multiple images to obtain the radiance map, a reliability function, e.g. a “hat” function (Debevec and Malik 1997) or CRF derivative (Mann and Mann 2001), is used. For tone-mapping, global (Reinhard et al. 2002) and local (Fattal et al. 2002; Pattanaik et al. 2000; Durand and Dorsey 2002) operators have been developed.

Both focus stacking and HDR imaging in general require image registration before the fusion process. The registration problem is easier for focus stacking because the input images have the same exposure values; thus standard motion estimation algorithms can be used. Whereas, in HDR imaging standard optical flow algorithms cannot be used directly because brightness constancy assumption no longer holds. The problem is further complicated due to saturation when the exposure difference is large. There are several approaches addressing the image registration problem in HDR imaging. Global translational motion estimation can be achieved through phase correlation, which is robust to illumination changes. Global parametric transformations can be estimated through interest point extraction and matching (Gevrekli and Gunturk 2007; Gunturk and Gevrekli 2006; Gevrekli and Gunturk 2007). While well-known Harris or SIFT interest point detectors can be used for this purpose, there are also techniques specifically designed to extract features robust to illumination variations (Gevrekli and Gunturk 2009). Another method to estimate global parametric transformation is the Median Threshold Bitmap (MTB) method (Ward 2003); it binarizes the image in a way to eliminate the effect of illumination. While these methods are limited to parametric motion, a two-step procedure we presented in (Hossain and Gunturk 2011) allows estimating

dense flow: a histogram-based intensity mapping function estimation (Grossberg and Nayar 2003) is first used for photometric registration, followed by an optical flow method for spatial registration. In our paper, we also adopted this approach. The preliminary results of this paper were presented at a conference (Qian et al. 2013); compared to the conference version, the current manuscript includes detailed algorithm, analysis, discussion, and experiments; also, alternative approaches are compared against.

Although many sophisticated algorithms have been proposed for either focus stacking or HDR image creation individually, extending both DOF and dynamic range simultaneously has not been investigated. In this paper, we show that by introducing diversity in both focus and exposure settings in the images that are captured, we can at the end produce an image that has larger DOF and dynamic range than any of the input images. Our approach is first to achieve focus stacking regardless of the exposure settings of the input images. We should note that standard focus stacking algorithms cannot be used even after photometric/spatial registration because there might be image regions that do not have any correspondence in the registered image due to saturation. We will present a method to handle saturation as well as registration errors. The focus stacking process is repeated for different exposure levels, and followed by radiance map estimation and tonemapping to produce an all-in-focus and HDR image.

The paper is organized as follows. In Sect. 2, we will first present the proposed algorithm, including photometric and spatial image registration, then describe extension to color images. HDR image creation is explained in Sect. 3. In Sect. 4, we will provide experimental results with real data to demonstrate the proposed method. We will conclude our paper in Sect. 5.

2 Focus stacking under exposure diversity

In this section we explain our approach to achieve focus stacking when the input images do not necessarily have the same exposure settings. We first present the method assuming the input images are grayscale, with the purpose of notational simplicity; and at the end we explain how to treat color images. The method is illustrated in Fig. 1, and briefly works as follows. The input images I_1, \dots, I_N are first spatially registered using the reference image

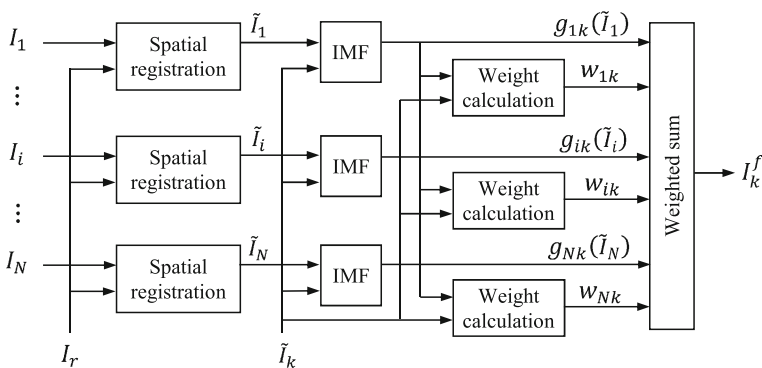


Fig. 1 Focus stacking from differently exposed images. I_1 to I_N are input images. I_r is the reference image for spatial registration. $\tilde{I}_i, i = 1, \dots, N$, are registered images. \tilde{I}_k is the reference image for photometric registration and focus stacking. $g_{ik}(\tilde{I}_i), i = 1, \dots, N$, are photometrically registered images. I_k^f is the focus stacked image

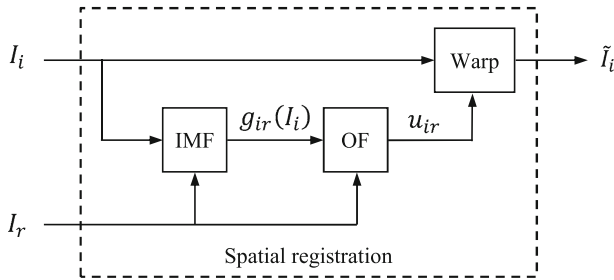


Fig. 2 Spatial registration of differently exposed images. An input image I_i is first photometrically registered to the reference image I_r by applying the intensity mapping function (IMF). The motion field u_{ir} between the photometrically mapped image $g_{ir}(I_i)$ and the reference image I_r is then estimated by applying an optical flow (OF) estimation method. The input image I_i is finally warped to achieve spatial registration

I_r , which is chosen among the input images. The registered images $\tilde{I}_1, \dots, \tilde{I}_N$ are then photometrically mapped to the reference image \tilde{I}_k , which is chosen among the registered images. Next, the weight maps are calculated. The weight maps take two things into account: local sharpness and registration errors. Local sharpness is an indicator of focus; if a pixel is in an in-focus area, its local sharpness is large. Registration error is an indicator about the accuracies of spatial and photometric registration steps. The weight of a pixel is linearly proportional with the local sharpness and inversely proportional with the registration error. The photometrically mapped images $g_{ik}(\tilde{I}_1), \dots, g_{ik}(\tilde{I}_N)$ are finally merged as a pixel-by-pixel weighted sum to produce the focus stacked image I_k^f . We now explain these steps in detail.

2.1 Spatial registration

Suppose we have N grayscale images $I_i, i = 1, \dots, N$ with different in-focus regions and exposure settings. Our first task is to estimate the motion field and register the images. We start with choosing a reference image I_r among the input images. The optical flow equation between an input image I_i and the reference image I_r should reflect the intensity mapping to take care of different exposure settings and can be formulated as

$$g_{ir}(I_i(\mathbf{x} + u_{ir}(\mathbf{x}))) = I_r(\mathbf{x}), \tag{1}$$

where \mathbf{x} is a pixel position, $u_{ir}(\mathbf{x})$ is the motion vector at pixel \mathbf{x} from the reference image I_r to input image I_i , and $g_{ir}(\cdot)$ is the intensity mapping function (IMF) from I_i to I_r . (Note that when two images have the same exposure settings, $g_{ir}(\cdot)$ is identity function and equation (1) reduces to the standard optical flow equation.) Once the IMF is applied to an input image, the photometrically mapped image $g_{ir}(I_i)$ and the reference image I_r satisfy the constant brightness assumption; and therefore, standard optical flow estimation methods can be utilized to estimate the motion field u_{ir} . In our implementation, we estimated the IMF using the histogram-based method (Mitsunaga and Nayar 2000), which is robust to small misalignments, and the motion field using (Liu 2009), which has a robust data fidelity term and a discontinuity-preserving total-variation regularization term. Using the estimated motion field, we warp the input image I_i onto the reference image I_r . The warped images $\tilde{I}_i, i = 1, \dots, N$ are now expected to be spatially registered. The flowchart of the spatial registration process is given in Fig. 2.

2.2 Photometric registration and focus stacking

The images $\tilde{I}_i, i = 1, \dots, N$ are spatially registered but with possibly different exposure settings and focus areas. The next step is to produce an image with extended depth of field. Standard focus stacking methods cannot be applied to $\tilde{I}_i, i = 1, \dots, N$ because of different exposure settings. One may suggest to apply a standard focus stacking to photometrically mapped images $g_{ir}(\tilde{I}_i)$; however, this approach would not generally work either because saturated regions in a long-exposure image cannot be mapped photometrically to corresponding unsaturated regions in shorter-exposure images. To handle complications due to different exposure settings, we propose the following approach.

We choose a reference image \tilde{I}_k among the spatially registered images $\tilde{I}_i, i = 1, \dots, N$. We estimate the IMF $g_{ik}(\cdot)$ and obtain photometrically registered images $g_{ik}(\tilde{I}_i)$. As mentioned earlier, the focus stacked image I_k^f is obtained as a weighted sum of $g_{ik}(\tilde{I}_i), i = 1, \dots, N$. The weights reflects two things at each pixel: (1) local sharpness, which is expected to be correlated with focus, and (2) spatial and photometric registration errors. We would like to have a large weight for a pixel that is in an in-focus area and has low spatial/photometric registration error.

The use of local sharpness as an indicator of focus is common in focus stacking. The local sharpness $\tilde{s}_{ik}(\mathbf{x})$ at pixel \mathbf{x} of image $g_{ik}(\tilde{I}_i)$ is defined as

$$\tilde{s}_{ik}(\mathbf{x}) = \sum_{\mathbf{y} \in N_h(\mathbf{x})} \|\nabla g_{ik}(\tilde{I}_i(\mathbf{y}))\|, \tag{2}$$

where $\nabla g_{ik}(\tilde{I}_i)(\mathbf{y})$ is the gradient vector at pixel \mathbf{y} obtained by applying Sobel filter, $\|\cdot\|$ denotes the gradient magnitude, and $N_h(\mathbf{x})$ is an $h \times h$ window around pixel \mathbf{x} . (In our experiments, $h = 3$.)

Now we have spatially and photometrically registered images $g_{ik}(\tilde{I}_i)$ and the corresponding weight maps \tilde{s}_{ik} indicating in-focus regions for each image. Before fusing the images, we need to take care of two possible issues. The first one is the saturation issue. If there are saturated pixels in an input image, then we cannot photometrically map them to the reference image. If used, these pixels cause artifacts in the final focus stacked image. The second issue is registration errors due to occlusion or inaccurate motion vectors. It is necessary to eliminate pixels that are saturated or misregistered from the fusion process to avoid artifacts. We decided to use a reliability mask for each input image and use pixels that are reliable during fusion. The reliability mask $M_{ik}(\mathbf{x})$ has two components, one to eliminate saturated pixels and the other to eliminate misregistered images:

$$M_{ik}(\mathbf{x}) = H(T_s - \tilde{I}_i(\mathbf{x}))H(T_m - |g_{ik}(\tilde{I}_i(\mathbf{x})) - \tilde{I}_k(\mathbf{x})|), \tag{3}$$

where $H(\cdot)$ is a step function, outputting 1 when the input is greater than or equal to zero, 0 otherwise. The first term $H(T_s - \tilde{I}_i(\mathbf{x}))$ outputs 1 when $\tilde{I}_i(\mathbf{x})$ is less than or equal to the saturation threshold T_s . The second term $H(T_m - |g_{ik}(\tilde{I}_i(\mathbf{x})) - \tilde{I}_k(\mathbf{x})|)$ outputs 1 when the absolute difference $|g_{ik}(\tilde{I}_i(\mathbf{x})) - \tilde{I}_k(\mathbf{x})|$ between the registered input image and the reference image is less than or equal to the misregistration threshold T_m . $M_{ik}(\mathbf{x})$ is 1 when both terms are 1; that is, when there is no saturation and no misregistration. (Note that $M_{kk}(\mathbf{x}) = 1$ with this definition as it should be.) In our experiments, we set the saturation threshold $T_s = 253$, eliminating pixels with values 254 and 255. This threshold selection is a relatively easy decision; whereas, for the misregistration threshold T_m , we had to test different values as there is no obvious choice. We cannot set the threshold too low because the input images have different focus distances and the absolute difference may be large even when the images

are correctly registered. We noticed that when T_m is too low, we start to eliminate pixels from sharp regions. We cannot set the threshold too high, which would result in artifacts due to misregistration. After extensive experiments, we set $T_m = 60$, which provides a good overall performance.

The reliability mask is then applied on the sharpness map $\tilde{s}_{ik}(\mathbf{x})$ to get the weights of the pixels:

$$\tilde{w}_{ik}(\mathbf{x}) = M_{ik}(\mathbf{x})\tilde{s}_{ik}(\mathbf{x}). \tag{4}$$

Before we do the image fusion, we need to normalize the calculated weight maps:

$$w_{ik}(\mathbf{x}) = \frac{\tilde{w}_{ik}(\mathbf{x})}{\sum_{i=1}^N \tilde{w}_{ik}(\mathbf{x})}. \tag{5}$$

Note that when all weights are zero at a pixel, then division by zero occurs in the above normalization. We considered two options to address this issue. The first option is to assign $1/N$ to each weight. The second option is to assign 1 to the weight associated with the reference image, and 0 to all other target images. In our paper, we choose the second option because we trust the reference image more than other input images. This second option is incorporated by adding a small scalar to the weight associated with the reference image before the normalization: $\tilde{w}_{kk}(\mathbf{x}) = \tilde{s}_{kk}(\mathbf{x}) + \epsilon$, where ϵ is a small scalar number, which is set to 0.01 in our paper. As the final step of focus stacking, we fuse all images to get the all-in-focus image I_k^f :

$$I_k^f(\mathbf{x}) = \sum_{i=1}^N w_{ik}(\mathbf{x})g_{ik}(\tilde{I}_i(\mathbf{x})). \tag{6}$$

2.3 Extension to color images

So far, we discussed the method for grayscale images. We extend it to the color images as follows. For spatial registration, we use the luminance channel of input images to estimate the motion vectors, which are then used to warp all three color channels. For photometric registration, we estimate the IMF for each color channel separately because the IMF may differ from one channel to another. After photometric registration comes the fusion step. We do not want to have different weights for different channels to avoid color artifacts. Therefore, the green channel is used to get the weight maps. Red, green, and blue channels are finally fused using the same weight maps obtained from the green channel.

3 HDR radiance estimation and tone mapping

I_k^f is the all-in-focus image where the k th image and therefore its exposure level is used as the reference. If we have K with $K \leq N$ different exposure levels, then we may repeat the focus stacking process for each of those K different exposure levels. The resulting all-in-focus images with different exposure levels can then be processed to obtain an HDR image using a standard HDR creation algorithm. In our paper, we used (Debevec and Malik 1997) to estimate the HDR radiance map. After the HDR radiance map is obtained, we should do tonemapping for low dynamic range displays. In this paper, we use the local tonemapping method given in (Fattal et al. 2002) to display the images.

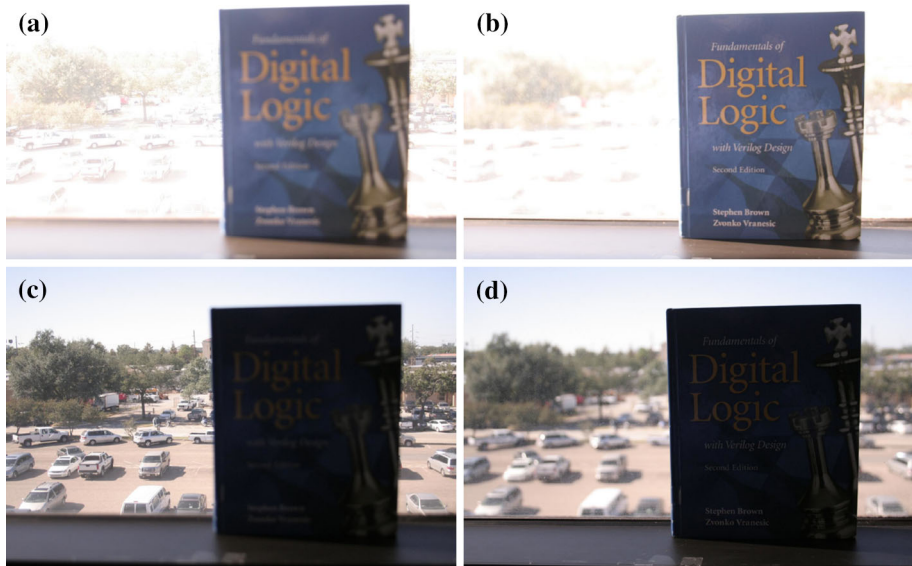


Fig. 3 Original four images with different focus regions and exposure levels. **a** Far focused with long (1/200 s) exposure. **b** Near focused with long exposure. **c** Far focused with short (1/1600 s) exposure. **d** Near focused with short exposure

4 Experimental results

We provide experimental results with real data to illustrate our proposed method for extending DOF and dynamic range from differently exposed and focused images. We captured four images with a hand-held DSLR camera. Each image has a different focus area and exposure time combination. These images are given Fig. 3. Figure 3a is near focused with long (1/200 sec) exposure. Figure 3b is far focused with long (1/200 sec) exposure. Figure 3c is far focused with with short (1/1600 s) exposure. Finally, Fig. 3d is near focused with with short (1/1600 s) exposure. All other settings are identical.

Figure 4 shows the spatial registration process. The first row shows the luminance channels of the input images given in Fig. 3. The first image I_1 is set as the reference image; and the other images are the target images. The second row shows the estimated intensity mapping functions (IMFs). The third row shows the photometrically registered images. And finally, the fourth row shows the estimated motion fields between the photometrically registered target images and the reference image. The motion fields are displayed with the colormap given in [<http://vision.middlebury.edu/flow/>].

Spatial registration is done by warping all three color channels using the estimated motion fields. The spatially registered images are shown in Fig. 5. (Note that the first image is the reference image, therefore, it is not warped but shown for convenience.)

The accuracy of the photometric and spatial registration may not be obvious from figs. 4 and 5. Therefore, we include Fig. 6, which shows the absolute differences between the reference image and the input images for the luminance channel. Figure 6a1–a3 show the differences between the input reference image and each of the three target images without any photometric or spatial registration. Figure 6b1–b3 show the differences after photometric registration. The residuals are reduced compared to Fig. 6a1–a3; however, it is obvious that

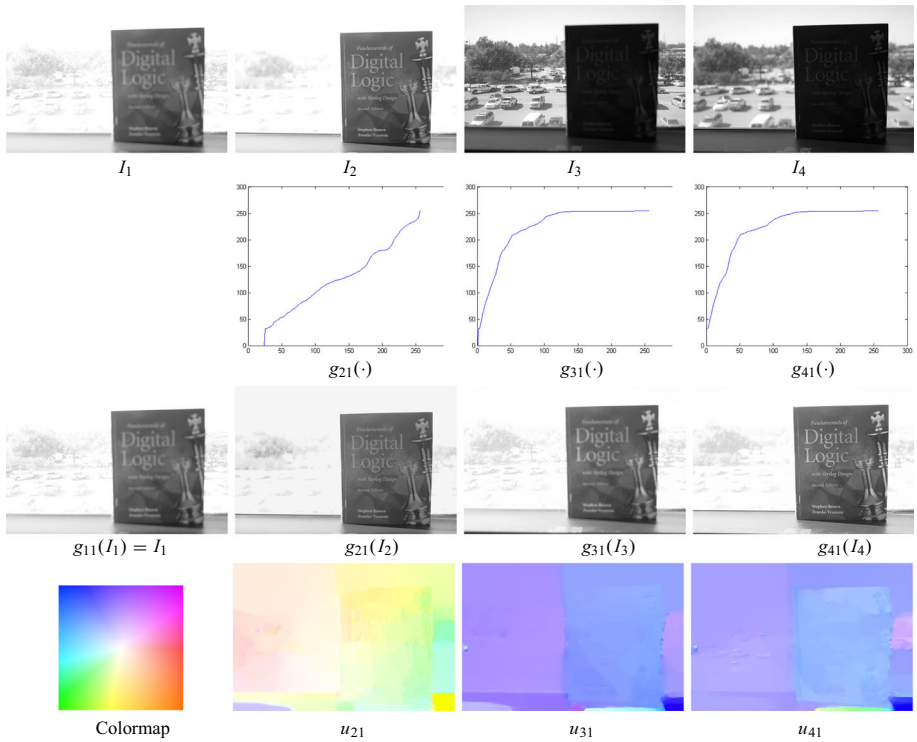


Fig. 4 Photometric and spatial registration. *First row* Input images, where I_1 is set as the reference image. *Second row* Estimated intensity mapping functions (IMFs) from target images (I_2 , I_3 , and I_4) to reference image I_1 . *Third row* Photometrically registered input images. *Fourth row* Estimated motion fields, displayed with the shown color coding



Fig. 5 Spatially registered images. **a** Original reference image from Fig. 3a; **b–d** Spatially registered target images from using the estimated motion vectors given in Fig. 4

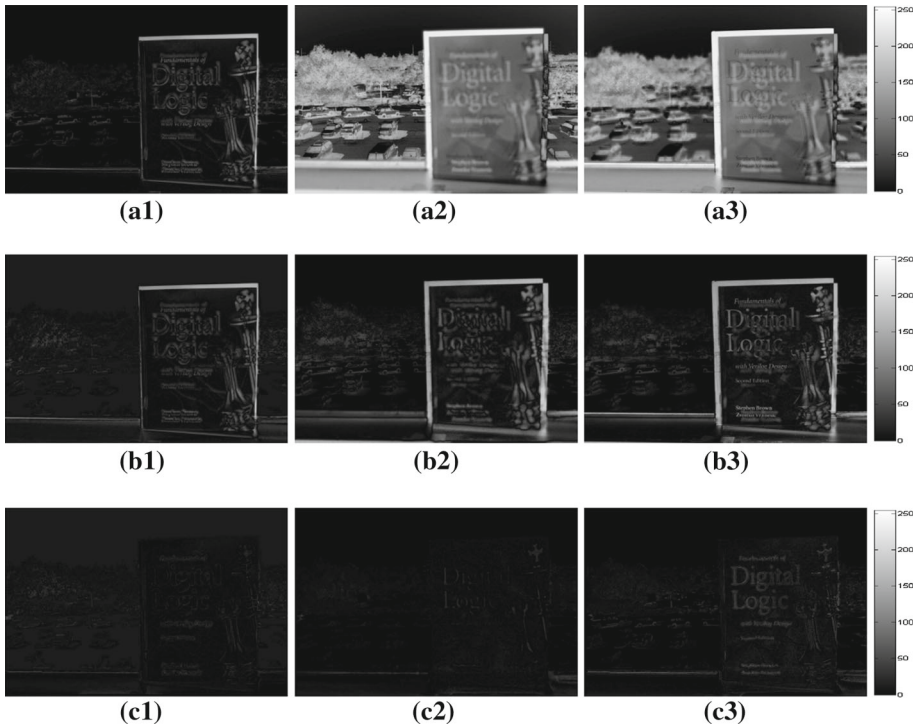


Fig. 6 Absolute difference values between the reference image and the target images before and after photometric/spatial registration for the luminance channel. **a1–a3** Difference between the reference image and three target images. The mean absolute differences (MADs) are 9.4, 82.6 and 80.6, respectively. **b1–b3** Difference between the reference image and photometrically registered target images. (MADs are 12.8, 18.0 and 17.3, respectively). **c1–c3** Difference after photometric and spatial registration. (MADs are 6.9, 3.0 and 5.1, respectively)

there is some movement between the images. Figure 6c1–c3 show the differences after both photometric and spatial registration. The residuals are reduced significantly, demonstrating the effectiveness of the registration process. In Fig. 7, we include registration results of two commonly used registration software. One is the MTB alignment method (Ward 2003), and the other is the Hugin alignment method [http://hugin.sourceforge.net/]. Both methods are available as a part of the Luminance HDR software [http://qtpfsgui.sourceforge.net/]. As seen in the results, these methods fail compared to the proposed registration method.

In Figs. 8, 9, 10, we demonstrate focus stacking from differently exposed images. Figure 8 shows two input images selected from the registered images in Fig. 5. Figure 8a1 is far focused with short exposure, and Fig. 8b1 is near focused with long exposure. Zoomed-in regions from these images are also included. To show the details, the zoomed-in region in Fig. 8a3 includes a version where the brightness is increased. Note that Fig. 8a3 is out of focus and noisy, whereas Fig. 8a2 is in focus. When we look at the corresponding regions in the second image, we note that the near field (Fig. 8b3) is well focused and exposed, but the far field (Fig. 8b2) is out of focus and over exposed.

Now, we would like to achieve focus stacking from these two images. First, Fig. 8a1 is set as the reference image. The resulting focus stacked image is given in Fig. 9a1. The exposure



Fig. 7 Absolute difference values between the reference image and the target images after spatial registration with alternative registration methods. **a1–a3** MTB alignment. **b1–b3** Hugin alignment. These alignment methods are included in the Luminance HDR software [<http://qtpfsgui.sourceforge.net>], and the results are obtained directly using this software

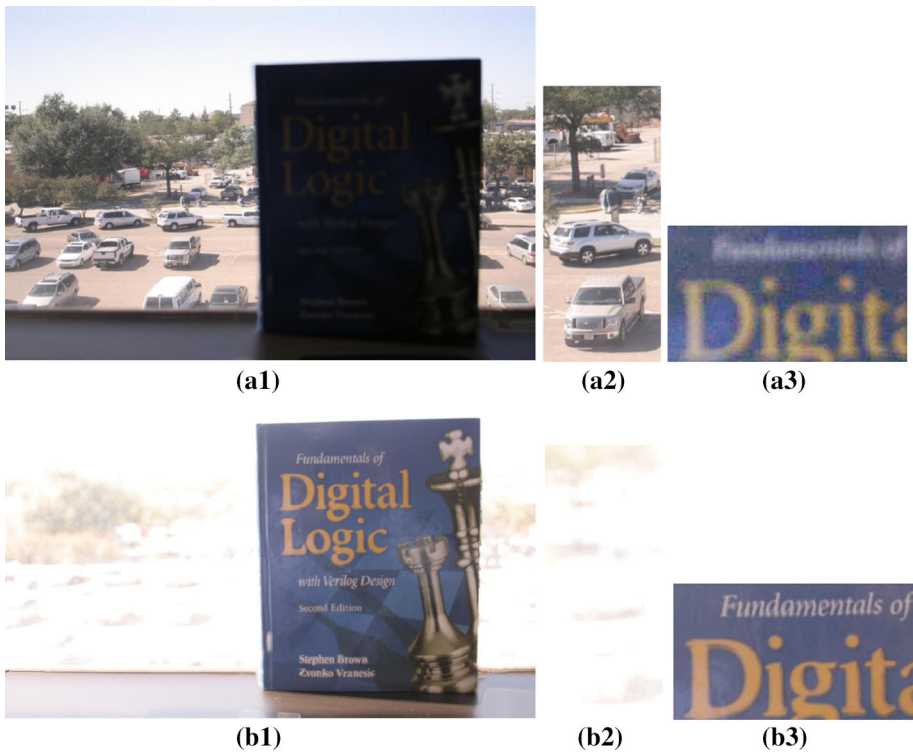


Fig. 8 **a1** First input image. **a2** A zoomed-in region. **a3** Another zoomed-in region (with increased brightness for visibility purposes). **b1–b3** Second input image and the zoomed-in regions

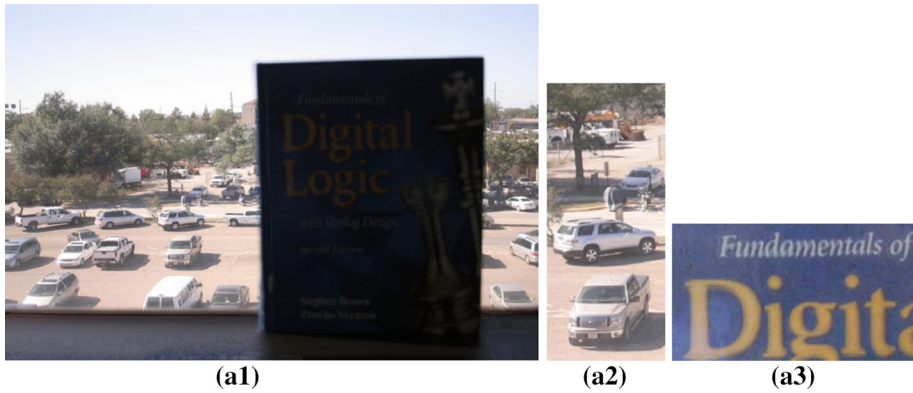


Fig. 9 **a1** Focus-stacked image when Fig. 8a1 is the reference. **a2** Zoomed-in region. **a3** Zoomed-in region (with increased brightness for visibility purposes)

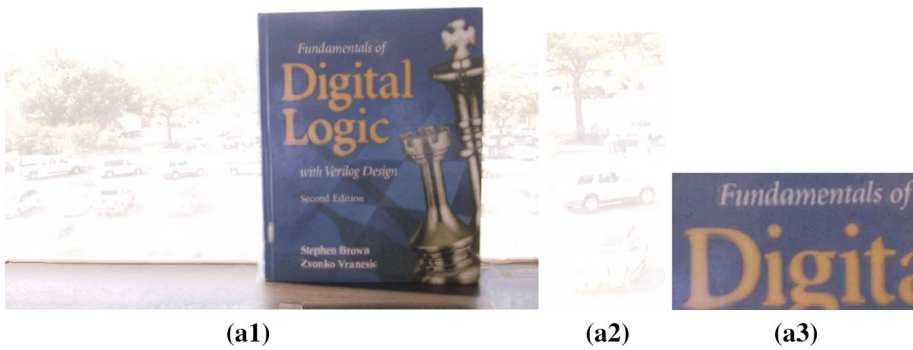


Fig. 10 **a1** Focus-stacked image when Fig. 8b1 is the reference. **a2–a3** Corresponding zoomed-in regions

level of Fig. 8a1 is preserved, but when we look at the zoomed-in regions, we note that both near and far fields are in focus. Specifically, when we compare Fig. 9a3 with the original Fig. 8a3, we note the improvement, which is coming from the second image.

Second, we set Fig. 8b1 as the reference image. The resulting focus stacked image is given in Fig. 10a1. This time, the exposure level of Fig. 8b1 is preserved but the entire image is in focus. This is clearly observed when Fig. 10a2 and the original Fig. 8b2 are compared.

These results clearly show that the proposed focus stacking from differently exposed imaging algorithm is working effectively regardless of which (short or long exposure) image is set as the reference. It is also possible to create an all-in-focus and HDR image using the procedure explained in the previous section: For each exposure level, choose one image as the reference, and use all others to form the all-in-focus image; and then apply a standard HDR imaging algorithm to these all-in-focus images to form the all-in-focus and HDR image. Figure 11 shows the HDR image for the dataset that we have, tone-mapped with (Fattal et al. 2002).

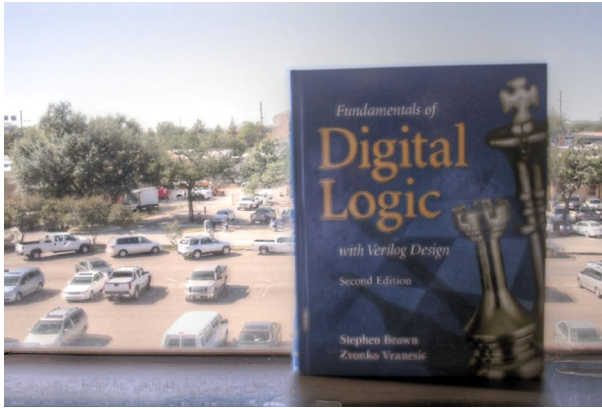


Fig. 11 Tonemapped HDR image

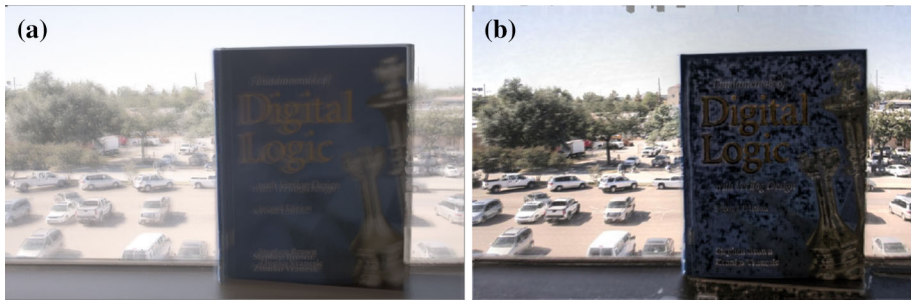


Fig. 12 Alternative approaches. **a** Output of the Luminance HDR software with Hugin alignment. **b** Output with the cascade application of Helicon Focus software followed by Luminance HDR software

In Fig. 12a, we provide the result obtained by Luminance HDR software using the Hugin alignment method. We see that the result is far from satisfactory because this particular alignment fails as we showed previously. In Fig. 12b, we tested an alternative approach: we first aligned the images using our proposed method; then we applied a commercially available focus stacking software (Helicon Focus, www.heliconsoft.com) to images of same exposure value. These focus stacked images are then merged using Luminance HDR software. Because our registration method works quite well, we do not see any registration errors, however, there are still objectionable artifacts in the final result.

Figures 13 to 16 demonstrate the algorithm for another dataset. Figure 13 shows the two input images. Figure 14 shows the estimated IMF and motion field. Figure 15 shows the differences between the reference and target images before and after registration. And finally, Fig. 16 shows the focus stacking result. Figure 16a1 is the reference image, with zoomed-in regions in Fig. 16a2. Figure 16b1 is the reference image, with zoomed-in regions in Fig. 16b2. Figure 16c1 is the focus stacked image, where both near and far fields are in focus.

Figure 17 shows another dataset consisting of four images with saturation and reflection regions. Figure 18 shows that the registration works robustly. And, finally, Fig. 19 shows the tonemapped HDR image, with good sharpness in all regions.

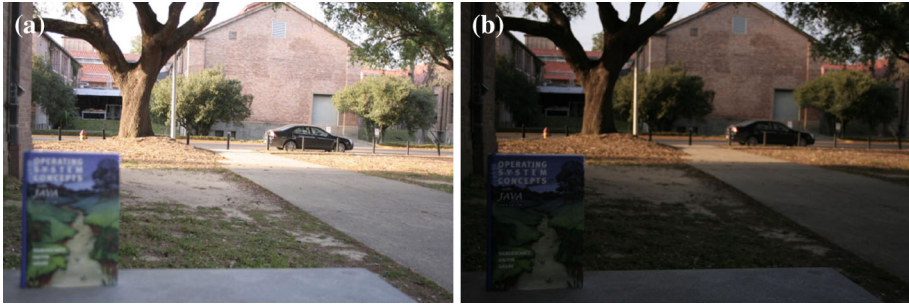


Fig. 13 Original images with different focus regions and exposure levels. **a** Far focused with long (1/20 s) exposure. **b** Near focused with short (1/33 s) exposure

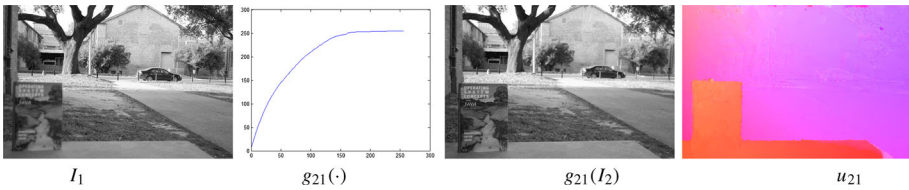


Fig. 14 Photometric and spatial registration. I_1 is the reference image. $g_{21}(\cdot)$ is estimated IMF. $g_{21}(I_2)$ is the photometrically registered target image. u_{21} is the estimated motion field



Fig. 15 Absolute difference between the reference and target image before/after registration. **a** Difference before registration. **b** Difference after photometric registration. **c** Difference after photometric and spatial registration

5 Conclusions

In this paper, we presented an algorithm for extending depth of field and dynamic range from differently focused and exposed images. The core process is focus stacking regardless of the exposure value; it requires photometric and spatial registration, and includes pixel-by-pixel weight calculation (involving sharpness, saturation, and registration errors) for fusion. For HDR imaging, we proposed to do focus stacking for each exposure level and fuse the focus-stacked imaging using a standard HDR imaging algorithm. We have done experiments with real data and obtained satisfactory results. There is some future work that may improve results and provide further understanding. The most important one is threshold selection. We have done threshold selection empirically; one may further investigate this, and come up with an optimal threshold selection that may depend on, for example, spatial location, exposure value, and IMF. It is known that photometric registration error is higher when mapping from a long exposure image to a short exposure image compared to mapping in the opposite direction. This could be used to modify the algorithm and adjust the threshold value. Other than the threshold selection the choice of the optical flow algorithm is also

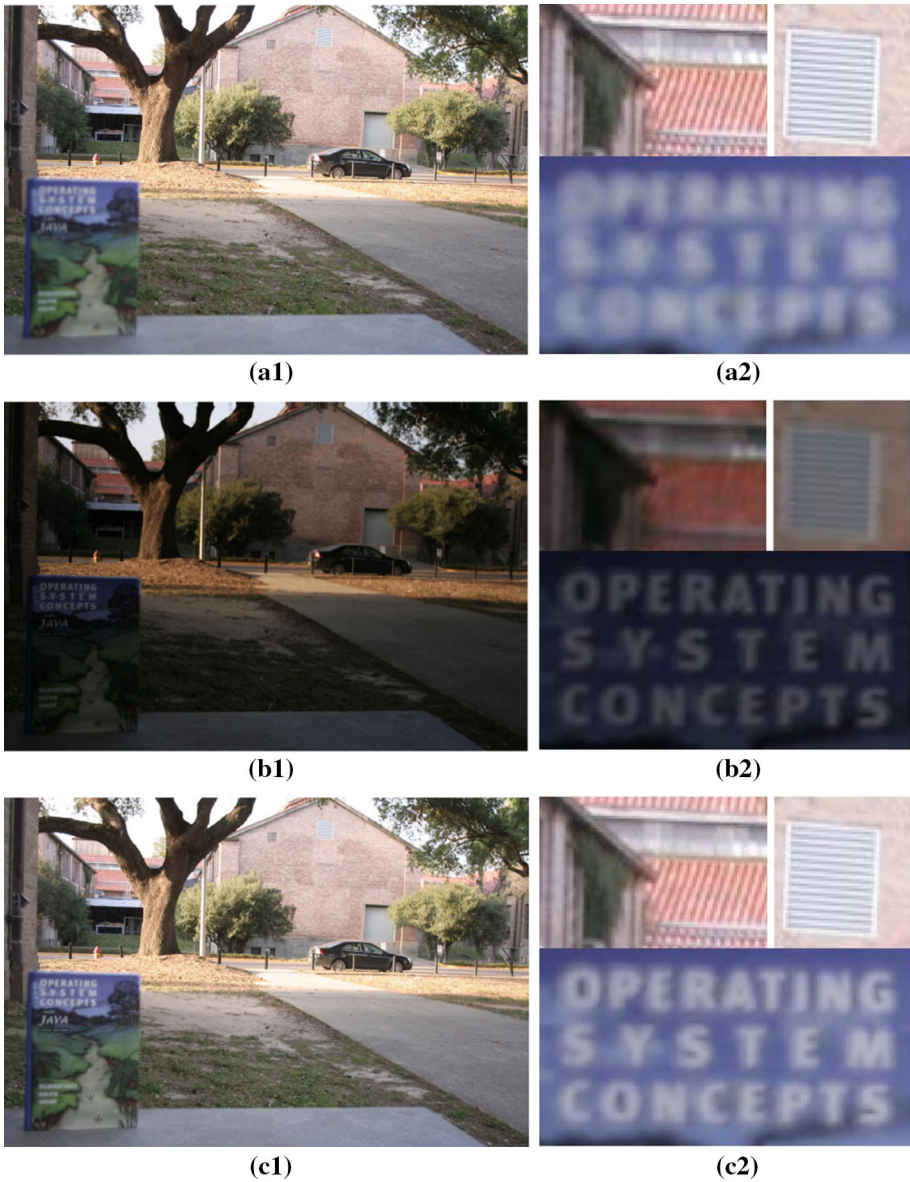


Fig. 16 a1–a2 First input image and zoomed-in regions. b1–b2 Second input image and zoomed-in regions. c1–c2 Focus-stacked image with Fig. 13a1 as the reference, and zoomed-in regions

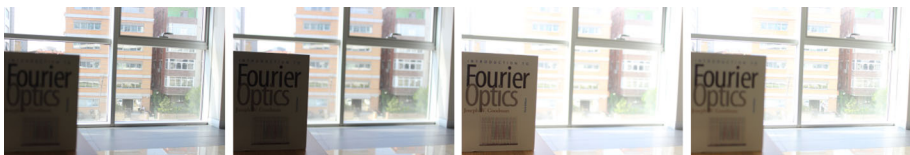


Fig. 17 Input images. The exposure times are 1/200 and 1/80s for *dark* and *bright* images, respectively

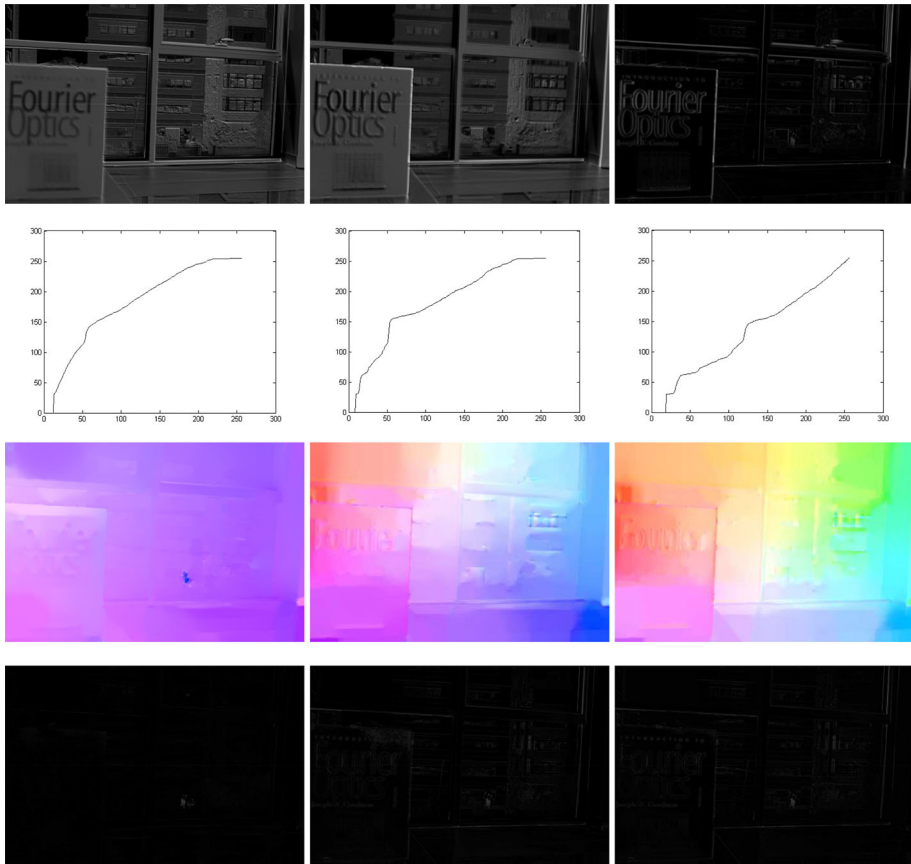


Fig. 18 Photometric and spatial registration of images, with fourth image is as the reference image. *First row* Residuals before registration. *Second row* Estimated intensity mapping functions. *Third row* Estimated motion fields. *Fourth row* Residuals after photometric and spatial registration

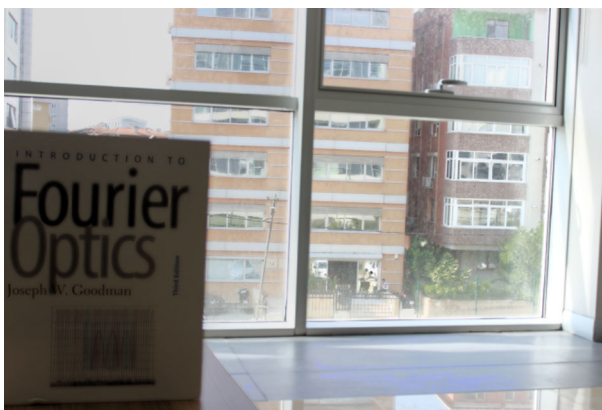


Fig. 19 Tonemapped HDR image

critical and should be investigated further as spatial registration errors may lead to artifacts. Another possible future work is the use of other focus measures and fusion methods within the proposed framework.

References

- Aguet, F., Van De Ville, D., & Unser, M. (2008). Model-based 2.5-d deconvolution for extended depth of field in brightfield microscopy. *IEEE Transactions on Image Processing*, 17(7), 1144–1153.
- Antunes, M., Trachtenberg, M., Thomas, G., & Shoa, T. (2005). All-in-focus imaging using a series of images on different focal planes. *Proceedings of International Conference on Image Analysis and Recognition*, 3656, 174–181.
- Burt, P. J., & Kolczynski, R. J. (1993). Enhanced image capture through fusion. In: *Proceedings of International Conference on Computer Vision*, May 1993, pp. 173–182.
- Debevec, P., & Malik, J. (1997). Recovering high dynamic range radiance maps from photographs. In: *Proceedings of ACM Conference on Computer Graphics and Interactive Techniques*, pp. 369–378.
- Durand, F., & Dorsey, J. (2002). Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics*, 21, 257–266.
- Fattal, R., Lischinski, D., & Werman, M. (2002). Gradient domain high dynamic range compression. *ACM Transactions on Graphics*, 21, 249–256.
- Forster, B., Van De Ville, D., Berent, J., Sage, D., & Unser, M. (2004). Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images. *Microscopy Research and Technique*, 65, 33–42.
- Gevrekci, M., & Gunturk, B. K. (2007). On geometric and photometric registration of images. *Proceedings of IEEE International Conference on Acoustics, Speech, Signal Processing*, 1, 1261–1264.
- Gevrekci, M., & Gunturk, B. K. (2007). Superresolution under photometric diversity of images. *EURASIP Journal on Advances in Signal Processing*, 360761, 1–12.
- Gevrekci, M., & Gunturk, B. K. (2009). Illumination robust interest point detection. *Computer Vision and Image Understanding*, 113(4), 565–571.
- Grossberg, M. D., & Nayar, S. K. (2003). Determining the camera response from images: What is knowable? *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 25, 1455–1467.
- Gunturk, B. K., & Gevrekci, M. (2006). High-resolution image reconstruction from multiple differently exposed images. *IEEE Signal Processing Letters*, 13(4), 197–200.
- Hossain, I., & Gunturk, B. K. (2011). High dynamic range imaging of non-static scenes. *Proceedings of SPIE Electronic Imaging*, 7876, pp. 78 760P1–78 760P9.
- Huang, W., & Jing, Z. (2007). Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters*, 28(4), 493–500.
- Li, H., Manjunath, B. S., & Mitra, S. K. (1995). Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 57(3), 235–245.
- Li, S., Kwok, J. T., & Wang, Y. (2001). Combination of images with diverse focuses using the spatial frequency. *Information Fusion*, 2(3), 169–176.
- Liu, C. (2009). Beyond pixels: Exploring new representations and applications for motion analysis. In: *Doctoral Thesis*. Massachusetts Institute of Technology, 2009.
- Mann, S., & Mann, R. (2001). Quantigraphic imaging: Estimating the camera response and exposures from differently exposed images. *Proceedings of International Conference on Computer Vision*, 1, 842–849.
- Mitsunaga, T., & Nayar, S. K. (1999). Radiometric self calibration. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2, 374–380.
- Mitsunaga, T., & Nayar, S. (2000). High dynamic range imaging: Spatially varying pixel exposures. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1, pp. 472–479.
- Pattanaik, S. N., Tumblin, J., Yee, H., & Greenberg, D. P. (2000). Time-dependent visual adaptation for fast realistic image display. In: *Proceedings of ACM Conference on Computer Graphics and Interactive Techniques*, pp. 47–54.
- Pertuz, S., Puig, D., Garcia, M. A., & Fusiello, A. (2013). Generation of all-in-focus images by noise-robust selective fusion of limited depth-of-field images. *IEEE Transactions on Image Processing*, 22(3), 1242–1251.
- Pieper, R. J., & Korpel, A. (1983). Image processing for extended depth of field. *Applied Optics*, 22(10), 1449–1453.
- Qian, Q., Gunturk, B. K., & Batur, A. U. (2013). Joint focus stacking and high dynamic range imaging. In: *Proceedings of SPIE Electronic Imaging*, pp. 8 660 041–8 660 047.

- Reinhard, E., Stark, M., Shirley, P., & Ferwerda, J. (2002). Photographic tone reproduction for digital images. *ACM Transactions on Graphics*, 21, 267–276.
- Subbarao, M., & Choi, T. (1998). Selecting the optimal focus measure for autofocusing and depth-from-focus. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 20(8), 864–870.
- Sugimoto, S. A., & Ichioka, Y. (1985). Digital composition of images with increased depth of focus considering depth information. *Applied Optics*, 24(14), 2076–2080.
- Tian, J., Chen, L., Ma, L., & Yu, W. (2011). Multi-focus image fusion using a bilateral gradient-based sharpness criterion. *Optics Communications*, 284(1), 80–87.
- Valdecasas, A. G., Marshall, D., Becerra, J. M., & Terrero, J. J. (2001). On the extended depth of focus algorithms for bright field microscopy. *Micron*, 32(6), 559–569.
- Ward, G. (2003). Fast, robust image registration for compositing high dynamic range photographs from handheld exposures. *Journal of Graphics Tools*, 8, 17–30.



Qinchun Qian received his B.S. degree in Mathematics from Shandong University, Jinan, China in 2006, and the M.S. degree in Control Engineering from Nankai University, Tianjin, China in 2009. He obtained his M.S. and Ph.D. degree both in Electrical and Computer Engineering at Louisiana State University, Baton Rouge, LA, US in 2011 and 2014, respectively. His research interests include image/video processing, computer vision, machine learning, pattern recognition and their parallel and web applications. Specific research areas include inverse problems in image processing (super-resolution image restoration and image deblurring), computational photography (high dynamic range imaging, extending depth of field), mobile imaging on smart-phone.



Bahadır K. Gunturk received his B.S. degree from Bilkent University, Turkey in 1999, and his Ph.D. degree from Georgia Institute of Technology in 2003, both in electrical engineering. Between 2003 and 2014, he was with the Department of Electrical and Computer Engineering at Louisiana State University, first as an assistant professor, then as an associate professor. Since 2014 he has been the Department of Electrical and Electronics Engineering at Istanbul Medipol University, where he is currently an associate professor. He has published more than 50 peer-reviewed journal/conference papers in the areas of image processing and computer vision.