Photo-consistent synthesis of motion blur and depth-of-field effects with a real camera model

Huei-Yung Lin *, Kai-Da Gu, Chia-Hong Chang

Department of Electrical Engineering and Advanced Institute of Manufacturing with High-tech Innovations, National Chung Cheng University, Chiayi 621, Taiwan

1. Introduction

In the past few years, we have witnessed the convergence of computer vision and computer graphics [1]. Although traditionally regarded as inverse problems of each other, image-based rendering and modeling share the common ground of these two research fields. One of its major topics is how to synthesize computer images for graphic representations based on the knowledge of a given vision model. This problem commonly arises in the application domains of computer animation, virtual reality and augmented reality [2,3]. For computer animation and virtual reality, synthetic images are generated from existing graphical models for rendering purposes. Augmented reality, on the other hand, requires the image composition of virtual objects and real scenes in a natural way. The ultimate goal of these applications is usually to make the synthetic images look as realistic as possible, compared to the photographs actually filmed by the cameras [4,5].

For the computer graphics and visualization applications, the choice of camera models plays an important role for the development of image synthesis algorithms. Most commonly used approaches adopt a perspective projection model to simulate a real camera system. Although the ray tracing techniques can be readily applied for real-time rendering, several important characteristics of practical imaging systems are not explicitly taken into account. For example, geometric and chromatic aberrations of the optics, the aperture shape and size of the camera system, and nonlinear intensity response of the CCD image sensors cannot be derived from a simple pinhole camera model [6]. Furthermore, it is not possible to generate the defocus blur phenomenon present in the human vision and real optical systems, which is de facto an important visual cue used for depth perception.

To achieve a more realistic scene generation, it is clear that the imaging model with a real lens should be used for image synthesis [7,8]. One of the major successes in the past few decades is the simulation of depth-of-field (DOF) effect due to the finite aperture size of a real camera system [9]. To create the DOF effect for a given 3D scene, different amounts of defocus blur are generated for each image pixel based on the distance to the focused position in the scene. The existing techniques in the literature include distributed ray tracing, image blending using the accumulation buffer, blurring with layered depth, post processing by image filtering [10–12], etc. Although some of the above approaches focused on accurate DOF simulation and others emphasized the capability of real-time rendering, all of them used ideal camera models for synthetic image generation. The defocus blur computations are not based on the actual DOF of practical optical systems. Thus, the realism of the rendered scenes is usually limited to the pure computer generated virtual environments.

To simulate the DOF effects for the applications with mixture of real scenes and synthetic objects, the visual inconsistency might become prominent in the rendered image due to different imaging models and camera parameters used for the real and virtual cameras. In addition to the global illumination issues commonly addressed in the augmented reality (AR) research field [13–15], a more realistic camera model should also be applied for photo-consistent image synthesis. More specifically, the synthetic image blur due to the optical defocus should be derived based on the parameter settings of a real camera system. Currently, most popular approaches use the so-called “circle of confusion” (CoC) introduced by out-of-focus of a thin lens camera model to calculate the defocus blur [16,17]. However, the simulation results are
still approximations from an ideal image formation model with oversimplified parameter settings.

On the other hand, when modeling a scene containing a fast moving object during the finite camera exposure time, it is not possible to insert the object directly into the scene by simple image composition [18,13]. In addition to the geometric and photometric consistency constraints imposed on the object for the given viewpoint, motion blur or temporal aliasing due to the relative motion between the camera and the scene usually have to be taken into account. It is a very important visual cue to human perception for the illusion of object motion, and thus commonly used in photography to illustrate the dynamic features in the scene. For computer generated or stop motion animations with limited temporal sampling rate, unpleasant effects such as jerky or strobing appearance might be present in the image sequence if motion blur is not modeled appropriately.

Early research on the synthesis of motion blur suggested a method by convolving the original image with the linear optical system-transfer function derived from the motion path [19,20]. The uniform point spread function (PSF) was demonstrated in their work, and high-degree resampling filters were later adopted to further improve the results of temporal anti-aliasing [21,22]. More recently, Sung et al. introduced the visibility and shading functions in the spatial–temporal domain for motion blur image generation [23]. Brostow and Essa proposed a frame-to-frame motion tracking approach to simulate motion blur for stop motion animation [24]. Egan et al. performed frequency domain analysis and presented a motion blur rendering algorithm using a sheared reconstruction filter [25]. Besides the generation of realistic motion blur effect, there are also some researchers focusing on real-time rendering using hardware acceleration for interactive graphic applications [26,27]. Although the results are smooth and visually consistent, they are only approximations due to the oversimplified image formation model and not suitable for rendering virtual objects into real scenes.

It is commonly believed that the image acquisition process can be approximated by a linear system, and motion blur is a result of image convolution with a given PSF of the camera. However, the nonlinear behavior of image sensors becomes prominent when the light source changes rapidly during the exposure time [28,13]. In this case, the conventional method using a simple box filter cannot create a photo-realistic or photo-consistent motion blur phenomenon. This fact might not be a problem in purely computer-generated animation, but inconsistency will certainly be noticeable in the image synthesized by combining virtual objects with a real scene.

In this paper, we present a photo-realistic image synthesis technique in terms of depth-of-field and motion blur modeling. For a given focus setting of a real camera system, the relationship between the blur extent and the object distance is first derived by a defocus estimation process. A simple yet robust blur extent identification method using a solid circle pattern is proposed. The amount of defocus blur used for synthetic DOF generation is based on the calibrated blur–distance curve of the actual camera system. The DOF effect simulation for a given focused image is then achieved by spatial convolution with a distance dependent circular Gaussian mask. By creating the individual camera profiles with defocus blur characteristics, synthetic virtual objects can be inserted into the real scenes with photo-consistent DOF generation. As for the motion blur modeling, we have proposed a nonlinear imaging model for synthetic image generation. Image formation is modified and incorporated with nonlinear intensity response of the CCD sensor. More photo-consistent results for both space invariant and space variant motion blur syntheses are then obtained by using the calibrated parameters of given camera settings. Experiments have demonstrated that the proposed method is capable of photo quality rendering for the real scene images.

The rest of this paper is organized as follows. Section 2 introduces the image formation model and nonlinear camera response function. Section 3 describes the characteristics of a real camera system and defocus blur synthesis. Synthetic motion blur generation for both space invariant and variant degradation is given in Section 4. Section 5 shows the implementation and experimental results, followed by the conclusion of this work.

2. Image formation model

The process of image formation is generally determined by the optical parameters of the lens, geometric parameters of the camera projection model, and photometric parameters associated with the environment and the CCD image sensor. For the synthesis from an image from the same viewpoint of a real scene image, a synthetic image created by a simple linear pinhole camera model can be used as a starting point for the post-processing. From basic radiometry, the relationship between scene radiance \( L \) and image irradiance \( E \) is given by

\[
E = L \frac{\pi}{4} \left( \frac{D}{f} \right)^2 \cos^4 \alpha
\]

where \( D, f \) and \( \alpha \) are the aperture diameter, focal length and the angle between the optical axis and the line of sight, respectively. Since the image intensity is commonly used to represent the image irradiance, it is in turn assumed proportional to the scene radiance for a given set of camera parameters. Thus, most existing algorithms adopt a simple pinhole camera model for synthetic image generation of real scenes.

For a general optical system, the image formation process is usually divided into two separate steps: the light rays collected by the lens and the image intensities generated by the photosensitive elements. Thus, the defocus blur introduced by a digital camera system can be modeled by the geometric optics and the nonlinear intensity response of the image sensor.

### 2.1. Geometrical optics and defocus blur

As shown in Fig. 1, an optical system consisting of a single convex lens with focal length \( f \) is used to derive some fundamental characteristics of focusing based on geometric optics. Let \( p \) represent the focused distance for a given set of camera parameters, then the diameter of the CoC for the scene point located at the distance \( z \) from the camera is given by

\[
d = \frac{Dpf}{p - f} \left( \frac{1}{p} - \frac{1}{z} \right)
\]

where \( D \) is the diameter of the lens.

It is clear that the size of the CoC depends only on the depth \( z \) if fixed camera settings of \( D, f \) and \( p \) are given. Thus, Eq. (2) can be written as

\[
d = c \left[ \frac{1}{p} - \frac{1}{z} \right]
\]

where \( c \) is a constant.

![Fig. 1. Camera model for the circle of confusion.](image-url)
where \( c \) is a camera constant represented by

\[
c = \frac{D_f}{p - f} \tag{4}
\]

From the above equations, the diameter of the CoC is linearly related to the inverse distance of the object. Furthermore, the constant \( c \) given by Eq. (4) represents the maximum size of the CoC for \( z > p \) since, from Eq. (3), \( d \rightarrow c \) as the object distance \( z \) approaches infinity. Note that the size of the CoC approaches infinity as \( z \rightarrow 0 \), which is usually out of practical interests for image synthesis.

In most cases, it is not possible to obtain the accurate camera parameters such as aperture diameter and focal length without an elaborate calibration process. Thus, the constant \( c \) will be given by the blur extent estimated from the scene at the infinity, instead of the direct computation using Eq. (4). It is shown in the experiments that the relationship between the object distance and the corresponding CoC basically follows the curve described by Eq. (3), except for a scale factor given by different values of the constant \( c \). The proposed model is therefore suitable for deriving the blur–distance characteristics of a practical optical system.

2.2. Motion blur model

Linear motion blur is generated by convolving the original image with a box filter or a uniform PSF. Although image synthesis or composition is relatively easy to implement based on the above image formation, the results are usually not satisfactory when compared to the real images captured by a camera. Consequently, photo-realistic scene modeling cannot be accomplished by this simplified imaging model.

Motion blur arises when the relative motion between the scene and the camera is fast during the exposure time of the imaging process. The most commonly used model for motion blur is given by

\[
g(x, y) = \frac{1}{T} \int_{t_0}^{t} f(x - x_0(t), y - y_0(t)) dt \tag{5}
\]

where \( g(x, y) \) and \( f(x, y) \) are the blurred and ideal images, respectively, \( T \) is the duration of the camera exposure, \( x_0(t) \) and \( y_0(t) \) are the time varying components of the relative motion in the \( x \) and \( y \) directions, respectively [19]. If only the uniform linear motion in the \( x \)-direction is considered, the motion blurred image can be generated by taking the line integral average along the motion direction. That is,

\[
g(x, y) = \frac{1}{R} \int_{-R}^{R} f(x - \rho, y) d\rho \tag{6}
\]

where \( R \) is the extent of the motion blur. Eq. (6) essentially describes that the blurred image is the convolution of the original (ideal) image and a uniform PSF [29]

\[
h(x, y) = \begin{cases} 
1/R, & 0 \leq x \leq R \\
0, & \text{otherwise}
\end{cases}
\tag{7}
\]

Fig. 2. Nonlinear behavior of the intensity response curves with different F-numbers. (a) Intensity versus exposure time for the black (left) and gray (right) patterns. (b) Small exposure range clearly shows the nonlinear behavior of the intensity.
This model is used for the most widely adopted approach for generating motion blur images. Its discrete counterpart used for computation is given by

\[ g(m, n) = \frac{1}{K} \sum_{i=0}^{K-1} f[m-i, n] \]

where \( K \) is the number of blurred pixels.

2.3. Nonlinear camera response function

During the image acquisition process, it is usually assumed that the image intensity increases linearly with the camera exposure time for any given scene point. For a practical camera system, however, nonlinear sensors are generally adopted to have the output voltage proportional to the logarithm of the light energy for high dynamic range imaging [30,31]. Furthermore, the intensity response function of the image sensors is also affected by the F-number (or aperture size) of the camera from our observation. Thus, the defocus and motion blur of an out-of-focus image cannot be simply characterized with only the lens system. The nonlinear behavior of the image sensor response should also be taken into account.

To illustrate this phenomenon, an image printout with white, gray and black stripes is used as a test pattern. Image intensity values under different camera exposures are calibrated for various F-number settings. The plots of intensity value versus exposure time for both the black and gray image stripes are shown in Fig. 2. The figures demonstrate that, prior to saturation, the intensity values increase nonlinearly with the exposure times. Although the nonlinear behaviors are not severe for large F-numbers (i.e., small aperture diameters), they are conspicuous for smaller F-numbers. Another important observation is that, even with different scene radiances, the intensity response curves for the black and gray patterns are very close if they are scaled by a constant along the axis of exposure time. Fig. 3 shows the intensity response curves for several F-numbers normalized with respect to the gray and black image patterns. The results suggest that the intensity values of a scene point under different exposure times are governed by the F-number.

To establish a more realistic image formation model from the above observations, a monotonically increasing function with nonlinear behavior determined by additional parameters should be adopted. Since the intensity response curves shown in Fig. 2 cannot be completely fitted

---

1 The F-number is a commonly used term in optics design and photography community. It is defined as \( F\text{-number} = \frac{f}{D} \) where \( f \) and \( D \) are the focal length and aperture diameter, respectively.

2 The intensity response of the white image pattern is not shown because it is saturated in a small exposure range.
by gamma or logarithm functions with various F-numbers, we model the intensity accumulation versus exposure time using an inverse exponential function of the integration time $t$ given by

$$I(t) = I_{\text{max}} \left(1 - e^{-\frac{kD^2}{\rho t}}\right) \quad 0 \leq t \leq T$$

(9)

where $T$, $I_{\text{max}}$, $D$, $k$ and $\rho$ are the exposure time, the maximum possible output response for the pixel, aperture diameter, a camera constant, and a parameter related to the reflectance property of the object surface, respectively. If all the parameters in Eq. (9) are available, then it is possible to determine the intensity value of the image pixel for any exposure time less than $T$.

For a general 8-bit grayscale image, the maximum intensity $I_{\text{max}}$ is 255 and $I(t)$ is always less than $I_{\text{max}}$. The aperture diameter $D$ is defined as the F-number divided by the focal length, and can be obtained from the camera settings. The parameters $k$ and $\rho$ are constants for any fixed scene point in the image. Thus, Eq. (9) can be rewritten as

$$I(t) = I_{\text{max}} \left(1 - e^{-k^\prime t}\right)$$

(10)

for a given set of camera parameters. The only parameter $k^\prime$ can then be determined by an appropriate calibration procedure with different camera settings. To verify Eq. (10), we first observe that $I(0) = 0$ as expected for any camera settings. The intensity value saturates as $t \to \infty$, and the larger the parameter $k^\prime$ is, the faster the intensity saturation occurs. This result is consistent with the physical model: $k^\prime$ contains the reflectance of the scene point and thus represents the irradiance of the image point. Therefore, the most important aspect of the equation is to characterize the image intensity accumulation versus integration time based on the fixed camera parameters.

For a given intensity value, it is not possible to determine the exposure time since the image irradiance also depends on the object’s reflectance property. However, it is possible to calculate the intensity of a scene point under any exposure if an intensity–exposure pair is given and the normalized response curve is known for specific camera parameter settings. This is one of the requirements for generating space invariant and space variant motion blur as described in the following sections.

To obtain the normalized intensity response function up to an unknown scale factor along the time axis, the images of the calibration patterns are captured with various exposures followed by least-squared fitting to find the parameter $k^\prime$ for different F-numbers. As shown in Fig. 3, the resulting fitting curves (black dashed lines) for any given F-number provide good approximation to the actual intensity measurements for both the black and gray patterns. This curve fitting and parameter estimation process can be referred to as photometric calibration for the intensity response function. It should be noted that only the shape of the intensity response curve is significant, the resulting function is normalized in the time axis by an arbitrary scale factor. Given the intensity value of an image pixel with known camera exposure, the corresponding scene point under different amounts of exposure can be calculated by Eq. (10).

3. Camera profiling and defocus blur synthesis

Based on the image formation model proposed in the previous section, synthetic image generation with depth-of-field effect includes the defocus blur identification and modeling for different

---

**Fig. 4.** The real scene images (left figures) and graylevel histograms (middle figures) used for blur identification. The red and green dots in the right figures represent the pixels with intensity thresholds $T_1$ and $T_2$, respectively. In Fig. 4(a), $T_1 = 204$ and $T_2 = 37$; in Fig. 4(b), $T_1 = 226$ and $T_2 = 41$. The camera focus range is set as 30 cm. (a) The pattern is placed at 200 cm from the camera. (b) The pattern is placed at 400 cm from the camera.
focused positions and scene distances, the derivation of intensity response curve for the nonlinear image sensors, and the defocus blur synthesis for different object depths.

3.1. Blur identification and modeling

For any focus range setting given by a camera system, the amount of out-of-focus blur or CoC in the image is a function of scene distance according to Eq. (2). To establish the relationship between the distance and the defocus blur, a white solid circle pattern is used to identify the blur extent for various object positions. Different from the previous approaches which estimate the blur parameter based on a Gaussian point-spread function [32], the blur extent is measured directly from the defocused black and white image in this work. An image graylevel transition identification technique based on histogram analysis is used to identify the blur extent robustly.

Fig. 4 shows the real images captured with different degrees of defocus blur (left figures) and the corresponding graylevel histograms (middle figures). It is clear that two clusters on both sides of the histograms represent the black and white regions in the images. Furthermore, the blur pixels in the images correspond to the histogram areas with intensity values between the two clusters, since the optical defocus process introduces a gradual intensity transition for the black and white images. Let $T_1$ and $T_2$ be the upper and lower bounds of the left and right clusters in the histogram, respectively.

Then the number of pixels with intensity values below $T_1$ and $T_2$ represents the non-white and black regions in the image, respectively. Ideally, these two regions should be bounded by an inner and an outer circle with radii, say $R_1$ and $R_2$, respectively. Thus, the blur extent defined by the intensity values between $T_1$ and $T_2$ can be derived from

$$ b = \Delta R = R_2 - R_1 = \frac{1}{\sqrt{\pi}} \left( \sqrt{A_2} - \sqrt{A_1} \right) $$

where $A_1$ and $A_2$ are the number of pixels below thresholds $T_1$ and $T_2$, respectively.

In the above algorithm for blur extent estimation, $T_1$ and $T_2$ in the histogram are the only parameters to be identified. These two parameters can be obtained by finding the abrupt pixel count changes for two consecutive graylevels with a given threshold. Since $T_1$ and $T_2$ correspond to the upper and lower bounds of the blur regions, they are identified by searching from the middle of the histogram to the left and right, respectively. In practice, due to the acquisition noise and digitization, the image pixels with the intensities $T_1$ and $T_2$ might not be perfect circles, as illustrated in Fig. 4 (right figures). Thus, a circle fitting algorithm based on the Hough transform is used to derive the radii $R_1$ and $R_2$ for blur extent computation. It should be noted that the proposed method does not need to identify

![Fig. 5. The blur extent versus object distance. (a) The camera is focused at 20 cm. (b) The camera is focused at 30 cm. (c) The camera is focused at 50 cm. (d) The camera is focused at 100 cm.](image-url)


the centers of the circles explicitly, which is usually not possible for the defocused images.

Fig. 5 shows the blur extent (in pixels) versus object distance obtained from real images with different camera focus settings. For all cases, the blur extent approaches an upper limit as the object moves beyond a certain range. The blur extent at the infinity and the distance with minimum defocus blur represent the camera constant $c$ and the focus range $p$ as suggested by Eq. (3), respectively. These two parameters can be derived from the least-squares fitting using the equation and the calibrated blur–distance pairs. The red and green curves illustrated in Fig. 5 represent the identified blur extents using the above algorithm and the curve fitting results, respectively.

3.2. Derivation of camera response functions

To model the nonlinear behavior of the image sensor and obtain the intensity response function for a given camera parameter setting, an image printout with black, gray and white patterns is used as a test object. The average intensity values for a small image region under different camera exposures are shown in Fig. 6. The red, green and blue curves correspond to the intensities of the white, gray and black image patterns, respectively. It is clear that, prior to saturation, the intensity value increases nonlinearly with the exposure time. Since the intensity response curves are equivalent up to an exposure scaling, all of the curves can be modeled by Eq. (10). Fig. 7(a) shows the intensity response function of the gray image pattern (marked in green) and the curve fitting result with $k = 0.009$ (marked in red).

Most image processing algorithms use the intensity value to represent the scene radiance under the assumption of a linear camera response function. To incorporate the nonlinearity derived from the photometric calibration, an intensity input–output mapping is created based on the fitting curve as shown in Fig. 7(a). Since the linear model assumes that the intensity increases linearly with the exposure time, the intensity correction can be derived from the full range mapping as illustrated in Fig. 7(b). The horizontal and vertical axes represent the input and output intensities, respectively. The curve means that the intensity value should be modified with the nonlinear effect of the image sensor prior to the subsequent image processing tasks. In the implementation, a lookup table is created based on the intensity input–output relation and applied to the image manipulation pipeline.

If the nonlinear behavior of the image sensor is modeled for the image formation process, the intensities of the defocus blur introduced by the optics are further adjusted according to Eq. (10). Thus, an intensity or irradiance correction step using the corresponding inverse function should be applied prior to image blurring or deblurring. This nonlinear modification of image intensities, however, does not affect the amount of blur extent derived from the geometric optics. The diameter of the circle of confusion for any out-of-focus distance is not altered by this nonlinear intensity adjustment. After the blurring or deblurring process, the synthesized image should recover the nonlinear property of the image sensor using Eq. (10).

3.3. Defocus blur image synthesis

Our algorithm for the synthesis of DOF effect is a post-processing technique. Similar to most previous approaches, a convolution mask is applied to the intensity of each pixel with appropriate point-spread function. However, we have also taken several important characteristics of the real camera model into account, rather than applying simple image filtering with arbitrary parameter settings. The calibrated blur–distance relation is used to create a depth dependent convolution mask. For the nonlinear behavior of the intensity accumulation, the image is pre-processed and post-processed based on the calibrated camera response function.

For the synthetic defocus blur generation process, the input is an image captured by a real camera system with a small aperture setting, and the output is a synthetic image with depth-of-field effect for a given focus range. At the first stage of the pipeline, the intensity conversion used to model the non-linear sensor response is applied to the input image to create an irradiance corrected image. At the second stage, image segmentation is carried out to separate the image
Fig. 8. Motion blur synthesis of an ideal edge image. (a) Ideal step edge and the corresponding intensity profile. (b) Motion blur edge generated using Eq. (8). (c) Real motion blur edge captured by a camera. (d) Motion blur edge synthesized by the proposed method.
regions with different scene distances for depth dependent defocus blur generation. The scene distance for each image pixel can be assigned manually for image synthesis if the depth information is not available. At the third stage, the blur extent for a given distance is derived from Eq. (3) or the curves shown in Fig. 5. The circular Gaussian convolution masks with the sizes of identified blur extents are then created and applied to each image region for the blurred counterpart generation. At the fourth stage, image composition and blending are performed with the derived image regions to synthesize an image with DOF effect. The last stage is the reverse process of the first stage, which is used to restore the nonlinear irradiance phenomenon present in the image.

For the DOF generation of a virtual scene, the Z-buffer value of each pixel is used to derive the level of blurring. The synthetic image can then be rendered by ray tracing techniques. If a well-focused image and the layered depth information are given, one major challenge for creating the DOF effect is to synthesize the defocus blur on the depth discontinuities. To avoid the artifact caused by the convolution operations with different mask sizes on both sides of region boundaries, a gradual transition of blur extents is adopted for the image regions near the depth discontinuities. More precisely, suppose the blur extents for two adjacent regions are assigned as \( d_1 \) and \( d_2 \), respectively. Then the diameter of the circular Gaussian kernel for any image pixel is given by

\[
\frac{d_1 + d_2}{2} = n
\]

where \( n = 0, 1, 2, \ldots \) represents the distance between the pixel and the region boundaries.

The above method uses a single pinhole image and the depth information of each object to synthesize the DOF effect. In case that the well-focused images for each object are available, then the DOF effect can be created by blurring each object separately with suitable amount and compositing them with a transparency weight on each pixel. When blurring an object, each image pixel is spread to its neighbors according to a Gaussian kernel, instead of conventional smoothing operation. This approach is based on the point spread characteristics of optics and capable of modeling the “see beyond occlusions” effect.

---

\(^3\) According to the geometric optics described in Section 2.1, a cylindrical filter should be applied to create the defocus blur. However, due to diffraction and non-idealities of lenses, a two-dimensional Gaussian model is usually suggested for the intensity distribution of a real camera system [33,34].
4. Synthetic motion blur image generation

As an example of using the conventional image degradation model, motion blur of an ideal step edge can be obtained by performing the spatial domain convolution with the PSF given by Eq. (7). The synthetic result is therefore a ramp edge with the width of the motion blur extent \( R \). If this motion blur model is applied on a real edge image, however, the synthetic result is generally different from the recorded motion blur image. Fig. 8(a), (b) and (c) illustrate the images and intensity profiles of an ideal step edge, motion blur edge created using Eq. (8) and real motion blur edge captured by a camera, respectively. As shown in Fig. 8(c), the intensity profile indicates that there exists non-uniform weighting on the pixel intensities for the real motion blur. Since the curve is not symmetric with respect to its midpoint, this nonlinear response phenomenon is clearly not due to the optical defocus of the camera and cannot be described by a simple Gaussian process.

4.1. Space invariant motion blur composition

In this work, motion blur is modeled using nonlinear intensity response of the image sensor as discussed in the previous section. For an image position under uniform motion blur, its intensity value is given by the integration of image irradiance associated with different scene points during the exposure time. Every scene point in the motion path thus contributes the intensity for smaller yet equal exposure time. Although these partial intensity values can be derived from the static image with full exposure by linear interpolation in the time domain, nonlinear behavior of the intensity response should also be taken into account. Suppose the monotonic response function is \( I(t) \), then the motion blur image with horizontal uniform motion is given by

\[
g(x, y) = I \left( \int_{R} I^{-1} (f(x - \rho, y)) d\rho \right)
\]

where \( R \) is the motion blur extent and \( I^{-1}(\cdot) \) is the inverse function of \( I(t) \). The discrete counterpart of Eq. (13) is represented by

\[
g[m, n] = I \left( \frac{1}{K} \sum_{i=0}^{K-1} I^{-1}(f[m-i, n]) \right)
\]

where \( K \) is the number of blurred pixels. If we consider the special case that the intensity response function \( I(t) \) is linear, then the integration (or summation) and \( I^{-1}(\cdot) \) are interchangeable on the right hand sides of the above equations. Consequently, Eqs. (13) and (14) are simplified to Eqs. (6) and (8), respectively.

Fig. 8(d) shows the synthetic motion blur edge of Fig. 8(a) and the corresponding intensity profile of the image scanlines generated using Eq. (14). The intensity response function \( I(t) \) is given by Eq. (10) with \( F-5 \) and \( k = 1.5 \times 10^{-5} \). Comparing the generated images and intensity profiles with those given by the real motion blur, the proposed model clearly provides more photo-consistent results than the one synthesized using a uniform PSF. The fact that the iso-contours of intensity are shifted to the left and expand the brighter region, as shown in Fig. 8(c), is successfully modeled by the nonlinear response curve.

4.2. Space variant motion blur composition

The motion blur model described in the previous section assumes that the image degradation is space invariant. It can be used to synthesize the motion blur effect of a static scene captured by a fast moving camera. If an object is moving in front of a still background and captured by a static camera, however, the space variant degradation model has to be adopted. Motion blur in this case consists of the blur caused by the mixture of the moving object and the still background around the boundary of the object, and the blur induced by the motion inside the object region. These two types of motion blur are defined as partial blur and total blur, respectively. The purpose of space variant image synthesis model is to simultaneously consider both the partial and total blur for motion blur effect generation.

Fig. 10. Experimental results of multiple objects. The real scene images are captured by an Olympus C-4040Z with F-1.8. (a) The real scene images captured with focus setting of 30 cm, 50 cm and 100 cm (from the left to the right), respectively. (b) The synthetic images with focus setting of 30 cm, 50 cm and 100 cm (from the left to the right), respectively.
If we consider a general case that an object moves a distance $R$ along the direction $\theta$ from the horizontal scanlines of the image, then the blurred image $g(x, y)$ is given by

$$g(x, y) = I \left\{ \frac{1}{R} \int_0^R \int_0^{R_1}(f(x-\rho \cos \theta, y-\rho \sin \theta) \, d\rho) \right\}$$

if the point $(x, y)$ is in the total blur region, and

$$g(x, y) = I \left\{ \frac{1}{R} \left\{ \int_0^{R-R'} \int_0^{R_1}(f_b(x, y)) + \int_0^{R'} \int_0^{R_1}(f(x-\rho \cos \theta, y-\rho \sin \theta)) \, d\rho \right\} \right\}$$

if the point $(x, y)$ is in the partial blur region, where $R'$ specifies the distance of the foreground region for integration and $f_b(x, y)$ is the background scene for image composition. It is clear that if the point $(x, y)$ is not in the motion blur regions, then $g(x, y)$ is identical to $f(x, y)$ and nothing needs to be done for the motion blur generation.

Since the background image can be rotated according to the angle $\theta$ to achieve horizontal object motion, the above equations can be simplified to one-dimensional cases and used for motion blur generation.

Fig. 11. Motion blur synthesis of a computer generated image. (a) The original image. (b) Space invariant motion blur generated using Eq. (8). (c) Space invariant motion blur synthesized by the proposed method.

Fig. 12. Experimental results of a real scene. (a) The static image, (b) the real motion blur image, (c) the motion blur generated using Eq. (8), and (d) the motion blur synthesized by the proposed method.

Fig. 13. Real and synthetic images of space variant motion blur due to the object moving in front of a static background. The real motion blur image is captured by Olympus E-20N with the F-number at F-5. (a) Real motion blur image. (b) Motion blur generated using the linear model. (c) Motion blur synthesized by the proposed method.
along the image scanlines. After rotation of the coordinate system, the discrete counterparts of the above equations are given by

\[ g[m, n] = I \left( \frac{1}{K} \sum_{i=0}^{K-1} f[m-i, n] \right) \]

for the total blurred region, and

\[ g[m, n] = I \left( \frac{1}{K} \left( \sum_{i=0}^{L-1} f[m-i, n] + (K-L)I^{-1}f_b(m, n) \right) \right) \]

for the partial blur region, where \( L < K \) represents the extent of foreground mixed into the final intensity value.

5. Results

The proposed method for depth-of-field and motion blur generation has been implemented on both synthetic and the real scene images. For the first experiment, a planar object is placed at 120 cm in front of the camera. An Olympus C-4040Z camera with F-1.8 is used to capture all of the real scene images. Fig. 9 represents the images captured or synthesized with focus settings of 30 cm (left), 50 cm (middle) and 100 cm (right). The real images captured by the camera, synthetic images with and without irradiance correction are shown in Fig. 9(a), (b) and (c), respectively. The lack of irradiance correction, as illustrated in Fig. 9(c), makes the synthetic images appear darker and much more different from the real photographs.

For the second experiment, two objects and a planar background are placed at 50 cm (left object), 100 cm (right object) and 110 cm in front of the camera, as shown in Fig. 10. A focused image captured by an Olympus C-4040Z with a large F-number (F-10) and used to simulate the depth-of-field effect. Fig. 10(a) and (b) illustrate the real scene images and the synthetic images with depth-of-field effect generation using our approach for the focus setting of 30 cm, 50 cm and 100 cm (from the left to the right), respectively.

The experiment of motion blur synthesis is carried out first using an image frame of a computer animation sequence. The original images, and the motion blur images generated using the conventional linear model and the proposed method are shown in Fig. 11. Color image synthesis is achieved by processing the red, green and blue channels separately. Motion blur images are first created for each channel using the same intensity response curve, and then combined to form the final result. There are clearly noticeable differences between the synthetic motion blur images generated by the linear and nonlinear intensity response models. Since the bright radiance values contribute more intensity in the actual motion blurred photographs [30], the proposed method provides more photo-realistic results compared to the conventional linear motion blur model.

To demonstrate the improvement of visual consistency using the proposed method, a real motion blur image is captured in the indoor environment for comparison. Fig. 12 shows the experimental results.
with the camera placed at about 1 m in front of the object (a tennis ball). The static image shown in Fig. 12(a) is taken by an Olympus E-20N at F-2.4 with camera exposure time of 1/8 s. Fig. 12(b) shows the motion blur image taken under 300 mm/s lateral motion of the camera using the same set of camera parameters. The blur extent in the image is 180 pixels, which is used for synthetic motion blur image generation. Fig. 12(c) illustrates the motion blur synthesized using the widely adopted uniform PSF for image convolution. The result generated using the proposed nonlinear intensity response function is shown in Fig. 12(d). By examining the images shown in Fig. 12, it is not difficult to find that Fig. 12(d) gives more photo-realistic motion blur than Fig. 12(c) does. The image scanline intensity profiles of Fig. 12(d) are very close to those exhibited in the real motion blur image. Thus, it is clear that the nonlinear behavior becomes prominent and has to be considered for more realistic motion blur synthesis.

An object moving in front of a static scene and captured by a still camera (Olympus E-20N) is used to illustrate the space variant motion blur effect. Fig. 13 shows the results of a real motion blur image (Fig. 13(a)), and the synthesized images using the conventional linear model (Fig. 13(b)) and the proposed method (Fig. 13(c)), respectively. Nonlinear intensity response function of the proposed model successfully eliminates the artifacts imposed by the foreground object boundaries.

In the last experiments, a virtual object is rendered into a real scene image to generate space variant motion blur effect. Fig. 14 shows the results consisting of a moving object with real motion blur (around the center) and another object with synthesized motion blur (at the top). Image composition using linear motion blur model and the proposed nonlinear intensity response function with F-5 are illustrated in Fig. 14(a) and (b), respectively, for different amounts of blur extent. The motion blur effect illustrated in Fig. 14(b) looks more consistent in terms of brightness and smoothness. Fig. 15 shows another example of rendering a virtual object into a real scene photograph with synthesized motion blur effect. The real scene images are captured by Olympus SP-500 UZ with the F-number at F-3.2. The results demonstrate that our method is capable of modeling synthetic objects with a complex background scene.

6. Conclusion

Image synthesis or composition with depth-of-field and motion blur phenomena has many applications in computer graphics and...
visualization. Most of the existing techniques for the simulation of DOF and motion blur effects adopt a simple camera model and do not consider the characteristics of practical imaging systems. Thus, the realism of the rendered scenes is usually limited to the pure computer generated virtual environments. In the paper, we have presented a method for the generation of defocus and motion blur images based on the parameters of a real camera. More photo-realistic DOF and motion blur effects can be obtained and rendered into the real scenes with least visual inconsistency. Experimental results demonstrate that our approach has taken one step towards photo-realistic synthesis of real scene images.

Acknowledgment

This work is supported in part by the National Science Council of Taiwan, R.O.C, under Grant NSC-96-2221-E-194-016-MY2.

References