Render Synthetic Fog into Interior and Exterior Photographs

Fukai Zhao[1] Ming Zeng[2] Bo Jiang[1] Xinguo Liu[1*]
[1]State Key Lab of CAD&CG, Zhejiang University
[2]Software School of Xiamen University

(a) Clear (b) Foggy

Figure 1: *Fig.(a) shows the input clear photograph, and Fig.(b) shows the foggy result produced by our method.*

Abstract

In this paper, we propose a framework for rendering synthetic fog into interior and exterior photographs. Given a single photograph as input, our approach first obtains an approximation of the illumination and geometry for the underlying scene with a few user assistance, then generates a Transmittance Map (TMap) and a Volumetric Map (VMap) for the input photograph according to a decomposition of reflective radiance and transmittance, finally synthesizes the foggy photograph by combining the input photograph, the Transmittance Map and the Volumetric Map. Our approach employs physically based rendering technique and take into account not only the effect of increasing illumination in shadow areas and desaturation of surface reflection but also glow around the area lights and light shafts for directional lights. Compared with the previous image processing method, our approach can produce more realistic results. In addition, we demonstrate in a user study that synthetic images produced by our approach are confusable with real photographs.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Display Algorithms I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Radiosity;

Keywords: fog, render, photograph

*e-mail:xgliu@cad.zju.edu.cn

VRCAI 2013, November 17 – 19, 2013, Hong Kong.

1 Introduction

Recently, virtual reality applications usually require users to add virtual scenes into photographs. Many work [Debevec 1998; Alnasser and Foroosh 2006; Cossairt et al. 2008; Lalonde et al. 2009b; Karsch et al. 2011] have been proposed to insert synthetic objects into photographs. However, insertion of participating media into photographs remains a problem due to its complexity of light transport in media. Jorge et al. proposed an image processing based method [Jorge et al. 2008] to add perceptually plausible participating media effects to a single image. This method assumes the illumination as a constant value for photographs taken under skylight, which is unsuitable for interior photographs and some other exterior lighting cases such as streetlight at night.

We propose an approach to add synthetic fog to both interior and exterior photographs with a few user assistance. Our approach takes advantage of illumination [Karsch et al. 2011; Liu et al. 2009] and geometry [Zheng et al. 2012; Hoiem et al. 2007] estimation technique and physically based rendering technique [Pharr and Humphreys 2010] to produce more realistic foggy images, including effects of surface shading desaturation, increasing the brightness in shadow area, glow around light sources and shaft of light, as shown in Fig.1. The main contribution of our work is a framework for rendering fog into both interior and exterior photographs, which consists of the following novel techniques:

- A decomposition of reflective radiance and transmittance for light transportation in participating media as a theoretical basis.
- Definition of Reflection Map, Transmittance Map and Volume Map for synthesizing foggy images.
- An objective function for sunlight direction optimization and a bounding volume constraint for light shaft rendering.

2 Related Work

Our research is closely related to realistic participating media rendering and techniques for inserting synthetic objects into real scenes.

Participating media, such as fog and haze, has a great inf uence in the light transport of a scene which introduces some complex phenomena such as glow and light shaft et al. Realistic participating media rendering have made signif cant progress in the past three decades. Many physically based stochastic methods [Lafortune and Willems 1996; Jensen and Christensen 1998; Pauly et al. 2000] have been proposed to effectively simulate various atmosphere effects. For foggy image synthesis, our approach employs Monte Carlo ray tracing and ray marching to evaluate the scattering radiance in participating media. Although other analytical and deterministic methods are also applicable, they are beyond the scope of this paper. For more detailed discussion of participating media rendering techniques, please refer to survey [Cerezo et al. 2005].

Inserting synthetic objects into real scenes has become a hot research area for its wide application in virtual reality recently. Image based rendering techniques capture the incoming radiance for relighting synthetic objects using special measurement equipment or multiple images. Debevec's work [Debevec 1998] shows that a light probe, such as a spherical mirror, can be used to capture a physically accurate radiance map for relighting. Alnasser et al. proposed a method [Alnasser and Foroosh 2006] to capture several images of the surroundings and represent the incoming radiance by Chebyshev polynomials for diffuse objects rendering. In addition to synthetic objects relighting, Cossairt et al. [Cossairt et al. 2008] devised an equipment to cast indirect light from synthetic objects to real objects with projectors. They utilized a lens array to capture and project light f eld between synthetic objects and real scene. Lalonde et al. [Lalonde et al. 2009c] combined data-driven appearance transfer techniques and physically based illumination model to transfer scene appearance from wide spread web camera to users own time-lapse sequences or single photographs.

Different from previous traditional image based rendering methods, Karsch et al. [Karsch et al. 2011] proposed a method to insert synthetic objects into existing photographs without requiring access to the scene or additional measurement. The key idea of this method is to estimate the illumination, material and geometry for an interior photograph. With the approximated 3D content of a real scene, this method renders synthetic objects using physically based renderer which captures the indirect lighting between synthetic objects and the real scene. Similarly, our approach employs the approximated illumination and geometry of the real scene for rendering fog into both interior and exterior photographs. However, our approach differs from theirs in that light transport model in media is quite different from the model between surfaces. In addition, the illumination and geometry estimation for interior and exterior photographs are also different from each other.

3 Overview

According to our theoretical derivation, the reduced ref ective radiance can be decomposed into product of ref ective radiance and transmittance. We def ne the input photograph, the radiance attenuation ratio map and the volumetric scattering radiance map as the Ref ection Map, the Transmittance Map and the Volumetric Map respectively. The theoretical derivation will be introduced in section 4.

The framework of our approach is illustrated in Fig.2. Given an clear photograph as input, we f rst estimate the approximate illumination and geometry for the underlying scene with user assistance. And then we evaluate the Transmittance and Volumetric Map according to the approximated illumination and geometry information. Finally, we synthesize the foggy photograph by combining the input photograph, the Transmittance and the Volumetric Map.

When estimating the illumination and geometry for input photographs, we handle interior and exterior photographs differently. In illumination estimation step, we adopt the method proposed in [Karsch et al. 2011] to approximate area light and sunlight through window for interior photographs, and we adopt the out door illumination model in [Liu et al. 2009] to estimate the sunlight and skylight for exterior photographs. In geometry estimation step, we approximate the geometry of interior man-made objects by cuboid proxies and employ a depth map to approximate the geometry of exterior objects.

4 Algorithm

In this section, we present the theoretical basis and derivation for the proposed method. We f rst simply review the background of light transport in participating media, then derive the formulations for decomposition of transmittance and ref ective radiance, f nally give the def nition of Ref ection Map $\mathcal{R}$, Transmittance Map $\mathcal{T}$ and Volumetric Map $\mathcal{V}$ as well as the composition equation for evaluating the f nal result.

4.1 Light Transport in Participating Media

A physically accurate light transport solution in participating media implies solving the radiative transfer equation (RTE) [Glassner 1995] which is an integro-differential equation including four types of interactions between light and media, namely emission, absorption, in-scattering and out-scattering. The equation describes the variation of radiance L at a point $\mathbf{x}$ in the direction ω. Following the notation introduced by Jensen and Christensen [Jensen and Christensen 1998], we can write it as

$$\begin{aligned}\frac{\partial L(\mathbf{x},\omega)}{\partial \mathbf{x}} = \; & \alpha(\mathbf{x})L_e(\mathbf{x},\omega) + \delta(\mathbf{x})L_i(\mathbf{x},\omega) \\ & - \alpha(\mathbf{x})L(\mathbf{x},\omega) - \delta(\mathbf{x})L(\mathbf{x},\omega), \end{aligned} \tag{1}$$

where α and δ are absorption and scattering coeff cients, and L_e and L_i are emitted and in-scattered radiance, respectively. For inhomogeneous media, both the absorption and scattering coeff cients vary throughout the media space. Def ning the extinction coeff cient as $\kappa(\mathbf{x}) = \alpha(\mathbf{x}) + \delta(\mathbf{x})$ and the in-scattering radiance L_i as the sum of incoming radiance from all possible directions ω to point $\mathbf{x}$ over the whole spherical domain Ω, we can rewrite Eq.(1) as

$$\begin{aligned}\frac{\partial L(\mathbf{x},\omega)}{\partial \mathbf{x}} = \; & \alpha(\mathbf{x})L_e(\mathbf{x},\omega) \\ & + \delta(\mathbf{x})\int_{\Omega} p(\mathbf{x},\omega,\omega')L(\mathbf{x},\omega')d\omega' \\ & - \kappa(\mathbf{x})L(\mathbf{x},\omega), \end{aligned} \tag{2}$$

where $p(\mathbf{x},\omega,\omega')$ is the phase function which describes the fraction of radiance coming from direction ω' that is in-scattering to direction ω at point $\mathbf{x}$. Solving Eq.(2) by integrating with respect to its boundary conditions, we have the integral form of RTE [Siegel and Howell 1992] given by

$$\begin{aligned}L(\mathbf{x},\omega) = \; & e^{-\tau(\mathbf{x}_0,\mathbf{x})}L(\mathbf{x}_0,\omega) \\ & + \int_{\mathbf{x}_0}^{\mathbf{x}} e^{-\tau(\mathbf{x}',\mathbf{x})}\alpha(\mathbf{x}')L_e(\mathbf{x}',\omega)d\mathbf{x}' \\ & + \int_{\mathbf{x}_0}^{\mathbf{x}} e^{-\tau(\mathbf{x}',\mathbf{x})}\delta(\mathbf{x}')L_i(\mathbf{x}',\omega)d\mathbf{x}', \end{aligned} \tag{3}$$

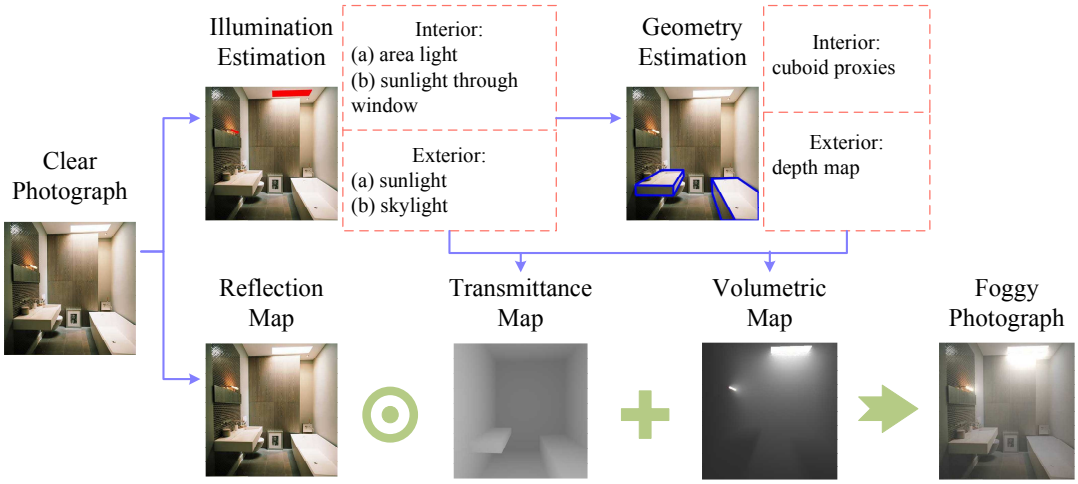


Figure 2: *Framework of the proposed method for rendering fog into photographs. The illumination and geometry approximation are evaluated from the input photograph with some user assistance. The Transmittance and Volumetric Map are computed from the approximated illumination and geometry information. The foggy photograph is synthesized by combining the input photograph, the Transmittance and the Volumetric Map, where ⊙ denotes the Hadamard product.*

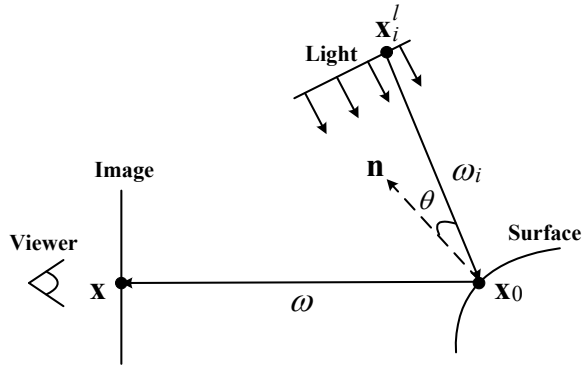


Figure 3: *Diagram showing light transport of surface ref ection in participating media with light sources and objects in the media.*

Symbol	*Meaning*
θ	Angle between incident light and surface normal
$\mathbf{n}$	Surface normal
$\mathbf{x}_0$	Surface point
$\mathbf{x}$	Point on image plane
$\mathbf{x}_i^l$	Point on light source
ω_i	Incident light direction
ω	Reflected light direction
$L(\mathbf{x}_0, \omega_i)$	Incident light
$\rho(\mathbf{x}_0, \omega, \omega_i)$	BRDF
$V(\mathbf{x}_0, \omega_i)$	Visibility function

Figure 4: *Notations used in our derivations.*

where $e^{-\tau(\mathbf{x}',\mathbf{x})}$ is the transmittance from point $\mathbf{x}'$ to point $\mathbf{x}$, which indicates the attenuation of the light along the path, and $\mathbf{x}_0$ is the boundary point. The optical length $\tau(\mathbf{x}', \mathbf{x})$ is def ned as $\tau(\mathbf{x}', \mathbf{x}) = \int_{\mathbf{x}'}^{\mathbf{x}} \kappa(\mathbf{t})d\mathbf{t}$.

4.2 Theoretical Derivation

Since our goal is to simulate the effect of fog which is optically thin (the transmittance through the entire medium is nearly one), the scattering is dominated by single scattering [Cerezo et al. 2005]. And it is known that fog is a kind of non-emissive media. Therefore, we make the following two pre-assumptions:

- When simulating fog, we only consider single scattering effect,
- and we ignore the radiance emitted by itself.

Then, Eq.(3) can be simplif ed as

$$L(\mathbf{x},\omega) = e^{-\tau(\mathbf{x}_0,\mathbf{x})}L(\mathbf{x}_0,\omega) + \int_{\mathbf{x}_0}^{\mathbf{x}} e^{-\tau(\mathbf{x}',\mathbf{x})}\delta(\mathbf{x}')L_i(\mathbf{x}',\omega)d\mathbf{x}'. \quad (4)$$

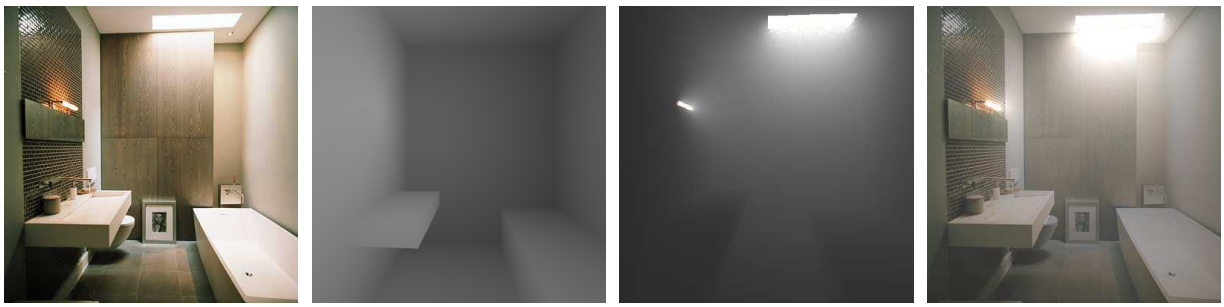
According to Eq.(4), the f nal radiance can be separated into the reduced ref ective radiance (the f rst term) and volumetric scattering radiance (the second term). Denote them by $L_r(\mathbf{x},\omega)$ and $L_v(\mathbf{x},\omega)$, then the f nal radiance can be evaluated by $L(\mathbf{x},\omega) = L_r(\mathbf{x},\omega) + L_v(\mathbf{x},\omega)$.

Reduced Refl ctive Radiance Consider the scenario illustrated in Fig.3 (the notations used are indicated in Fig.4). Denote the ref ected light radiance at a surface point $\mathbf{x}_0$ and the ref ected light radiance due to the incident light from direction ω_i by $I(\mathbf{x}_0,\omega)$ and $I_i(\mathbf{x}_0,\omega,\omega_i)$. According to rendering equation [Kajiya 1986], $I(\mathbf{x}_0,\omega)$ can be represented by

$$I(\mathbf{x}_0,\omega) = \int_\Omega L(\mathbf{x}_0,\omega_i)\rho(\mathbf{x}_0,\omega,\omega_i)V(\mathbf{x}_0,\omega_i)\cos\theta d\omega_i = \int_\Omega I_i(\mathbf{x}_0,\omega,\omega_i)d\omega_i.$$

Denote the light source position in direction ω_i by $\mathbf{x}_i^l$ and transmittance from $\mathbf{x}_a$ to $\mathbf{x}_b$ by $\mu(\mathbf{x}_a,\mathbf{x}_b) = e^{-\tau(\mathbf{x}_a,\mathbf{x}_b)}$, then the reduced ref ective radiance can be represented by

$$L_r(\mathbf{x},\omega) = \mu(\mathbf{x}_0,\mathbf{x})\int_\Omega \mu(\mathbf{x}_i^l,\mathbf{x}_0)I_i(\mathbf{x}_0,\omega,\omega_i)d\omega_i.$$



(a) Clear (b) TMap (c) VMap (d) Foggy

Figure 5: *Rendering fog into an bathroom photograph. When estimating illumination, we annotate two light sources, they are all approximated by area light. When estimating geometry, we construct two cuboid proxies for the platform and the bathtub respectively.*

Denote the area of a light source by S_{light}. According to the mean value theorem of integrals, there exist a point $\tilde{\mathbf{x}}^l \in S_{light}$ giving

$$\begin{aligned} L_r(\mathbf{x},\omega) =& \mu(\mathbf{x}_0,\mathbf{x})\mu(\tilde{\mathbf{x}}^l,\mathbf{x}_0)\int_\Omega I_i(\mathbf{x}_0,\omega,\omega_i)d\omega_i \\ =& \mu(\mathbf{x}_0,\mathbf{x})\mu(\tilde{\mathbf{x}}^l,\mathbf{x}_0)I(\mathbf{x}_0,\omega). \end{aligned} \tag{5}$$

In our implementation, we take the geometric middle point (the intersection of two diagonals of a quadrilateral) of light area as $\tilde{\mathbf{x}}^l$ for each light source. According to experiments, the middle point approximation is sufficient for participating media rendering. The validation experiment is discussed in section 6.1.

As we take a 2D photograph as input, we have the information of the reflected light radiance $I(\mathbf{x}_0,\omega)$ for each pixel. And we have decomposed the reflective radiance and transmittance from reduced reflective radiance in Eq.(5). Therefore, we define the input image as a Reflection Map $\mathcal{R}$ and the transmittance along the Light-Surface-View path , $\mu(\tilde{\mathbf{x}}^l,\mathbf{x}_0)\mu(\mathbf{x}_0,\mathbf{x})$, as a Transmittance Map $\mathcal{T}$. In consequence, the reduced reflective radiance can be evaluated by $\mathcal{R}\odot\mathcal{T}$ where $\odot$ denotes the Hadamard product.

Volumetric Scattering Radiance When considering single scattering effect, the volumetric scattering radiance $L_v(\mathbf{x},\omega)$ can be represented as

$$\begin{aligned} L_v(\mathbf{x},\omega) &= \int_{\mathbf{x}_0}^{\mathbf{x}} \mu(\mathbf{x}',\mathbf{x})\delta(\mathbf{x}')L_i(\mathbf{x}',\omega)d\mathbf{x}', \\ L_i(\mathbf{x}',\omega) &= \int_{S_{light}} \mu(\mathbf{x}_i^l,\mathbf{x}')L(\mathbf{x}',\omega_i)p(\mathbf{x}',\omega,\omega_i)d\mathbf{x}_i^l. \end{aligned}$$

As the approximate illumination and geometry are provided, given a participating media density distribution it can be evaluated directly. We define the result of volumetric scattering radiance as a Volume Map $\mathcal{V}$ and the final result as a Composite Map $\mathcal{C}$. Then, the final result can be represented by a composition equation, $\mathcal{C}=\mathcal{R}\odot\mathcal{T}+\mathcal{V}$.

5 Implementation

In this section, we describe the implementation details of the proposed method. Since the input is a 2D photograph, we employ user annotations to indicate light sources and estimate geometry of the underlying scenes. As the implementation for interior and exterior photographs are different, we discuss them separately.

5.1 Interior Photograph

Illumination Estimation We adopt the interior illumination estimation method proposed in [Karsch et al. 2011] to estimate the intensity of each light source. Generally, we approximate ordinary interior light sources by area light sources and the sunlight through windows or doors by directional light. For point and spot light sources, we manually set the light source depth and directions.

Geometry Estimation Zheng et al.[Zheng et al. 2012] proposed a method to approximate interior man-made objects, e.g., tables and chairs, etc., by 3D cuboid proxies. We employ this method to build approximate proxies for objects in input photographs. According to our experiments, cuboid proxy approximation can preserve the depth information of most interior objects, which is sufficient for participating media rendering. The validation experiments and user study are discussed in section 6.1.

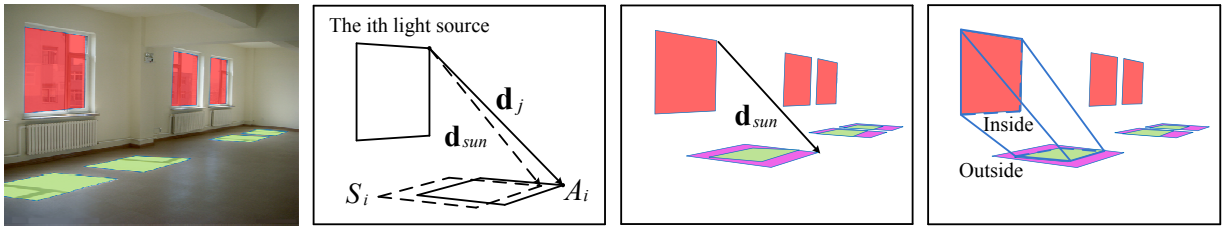
With the approximated illumination and geometry, we can produce the TMap and VMap for a clear interior photograph using the physically based ray tracing engine (PBRT) [Pharr and Humphreys 2010], as shown in Fig.5(b) and 5(c). The foggy photograph can be synthesized by combining the clear photograph, TMap and VMap, as shown in Fig.5(d). Note that we employ an homogenous media for rendering in this result, but heterogenous media is also applicable.

Sunlight Shaft In order to simulate sunlight shaft effect, we first annotate the area of sunlight sources (open windows and doors) and the corresponding shaft with polygons, as shown in Fig.6(a). We then evaluate the directions from vertices of sunlight sources to vertices of the shaft and apply a non-linear least square method to optimize the sunlight direction. Finally, we construct a sunlight shaft bounding volume for sunlight shaft rendering.

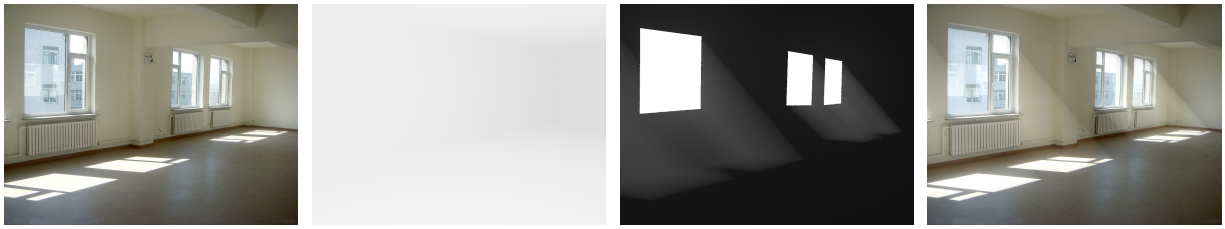
Consider the scenario illustrated in Fig.6(b). Assume there are n sunlight sources. For the ith sunlight source with m vertices, we can get m directions $\mathbf{d}_j (j\in[0,m-1])$ from vertices of sunlight source to vertices of the corresponding shaft. Denote the casted sunlight sources and corresponding shaft by $S_i \in [S_0,S_{n-1}]$ and $A_i \in [A_0,A_{n-1}]$, and denote the area of a object in image space and the sunlight direction by $\mathcal{A}(\cdot)$ and $\mathbf{d}_{sun}$. We propose the following objective function for optimizing $\mathbf{d}_{sun}$

$$\min_{arg}\sum_{i=0}^{n-1}\left(\sum_{j=0}^{m-1}\arccos(\mathbf{d}_{sun}\cdot\mathbf{d}_j)-\frac{\mathcal{A}(S_i)\cap\mathcal{A}(A_i)}{\mathcal{A}(S_i)\cup\mathcal{A}(A_i)}\right).$$

The first term represents the sum of the angle between $\mathbf{d}_{sun}$ and



(a) Annotation (b) Diagram for illustration (c) Optimized sunlight direction (d) Sunlight shaft bounding volume



(e) Clear (f) TMap (g) VMap (h) Foggy

Figure 6: *Sunlight direction optimization, sunlight shaft volume construction and rendering. The red and green polygons represent sunlight sources and sunlight shafts. In Fig.(c) and Fig.(d), purple polygons represent casted light sources on the f oor along the sunlight direction. The blue lines represents the sunlight shaft volume.*

$\mathbf{d}_j$. Since $\mathbf{d}_j$ are observed sun light directions, we consider $\mathbf{d}_{sun}$ must be approximated to all of $\mathbf{d}_j$. The second term represents the overlap area of the casted sunlight source and the corresponding shaft divided by the union area. We consider the overlap area of the casted light source and the illuminated area must be maximized.

However, due to the inaccuracy of annotation and geometry estimation, the casted sunlight area may not completely cover the shaft area, as shown in Fig.6(c). Therefore, we construct a bounding volume for each sunlight source. We insert lines from each vertex of sunlight source to the vertex of sunlight shaft, as shown in Fig.6(d). When rendering sunlight shaft, only the points inside the sunlight volume are illuminated, and others are in the shadow.

5.2 Exterior Photograph

Illumination Estimation Some methods have been proposed for automatic estimating natural illumination from a single out door image. However, most of them [Lalonde et al. 2009a; Lalonde et al. 2012] are learning based and quite complex. In our implementation, we take advantages of user assistance and adopt the approach introduced in [Xing et al. 2012] for out door illumination estimation. According to the lighting model proposed in [Liu et al. 2009], the out door illumination is comprised of the sunlight and skylight. They are assumed to be a directional and a uniformly distributed light source respectively. And all points in the scene are assumed to share the same lighting condition. Consider a pixel $\mathbf{p}$ with diffuse BRDF, the radiance at $\mathbf{p}$ can be represented by

$$\begin{aligned} I(\mathbf{p}) =& I_{sun} + I_{sky} \\ =& L_{sun}\rho(\mathbf{p})V(\mathbf{p},\omega_{sun})\cos\theta_{sun} + \\ & \int_{\Omega} L_{sky}\rho(\mathbf{p})V(\mathbf{p},\omega_i)\cos\theta d\omega_i, \end{aligned} \tag{6}$$

where L_{sun} and L_{sky} are the sunlight and skylight. ρ and V are BRDF and visibility function. When evaluating L_{sky}, we select pixels on horizontal ground in shadow leading Eq.(6) to

$$I(\mathbf{p}) = I_{sky} = \pi L_{sky}\rho(\mathbf{p}).$$

Then L_{sky} can be evaluated by $L_{sky} = s\frac{I(\mathbf{p})}{\pi\rho(\mathbf{p})}$, where s is a scaling factor. When evaluating L_{sun}, we select pixels on horizontal ground out of shadow leading Eq.(6) to

$$I(\mathbf{p}) = I_{sun} + I_{sky} = L_{sun}\rho(\mathbf{p})\cos\theta_{sun} + I_{sky}.$$

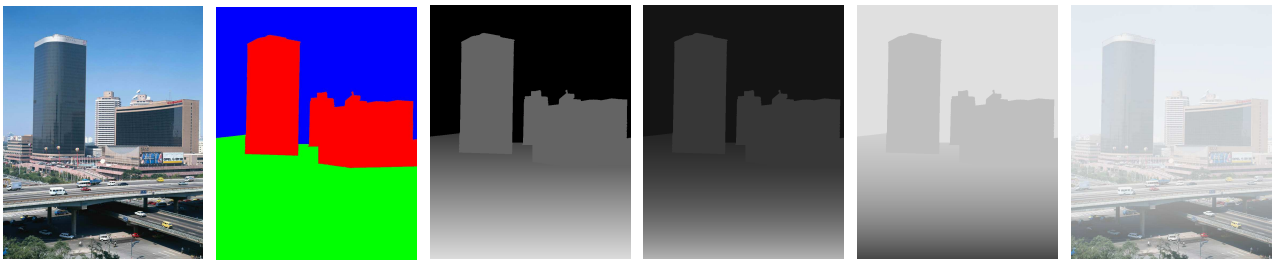
Then L_{sun} can be evaluated by $L_{sky} = s\frac{I(\mathbf{p})-I_{sky}}{\rho(\mathbf{p})\cos\theta_{sun}}$. In our implementation s, ρ and sunlight direction are set manually.

Geometry Estimation Geometry of objects in exterior photographs is more complex than that of interior man-made objects. Buildings, trees and roads are hard to approximate by simple proxies. Therefore, we employ a depth map as a geometry approximation for exterior photographs. We adopt the classif cation idea of the approach proposed in [Hoiem et al. 2007] and take advantage of user assistance in our implementation. In general, the pixels in the input photograph are classif ed into three classes: support, vertical and sky, as show in Fig.7(b). Support surfaces are roughly parallel to the ground, the examples of which includes roads, lawns, trees and lakes. The depth of support surfaces vary continuously from near to far. Vertical surfaces are steep and roughly a vertical plane, such as walls, cliffs, buildings. The depth of vertical surfaces can be approximated by a constant value. The sky is simply the region corresponding to the open air. The depth of the sky is set to the maximum depth value. In our implementation, the depth value of near and far range of support, the vertical and sky area are set manually, as show in Fig.7(c).

When rendering fog, the viewing ray direction and depth is determined by the depth map. We f rst denote the intrinsic matrix of the camera by K:

$$K = \begin{pmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{7}$$

where c_x and c_y can be set to $\frac{1}{2}$width and $\frac{1}{2}$height of the image, respectively. The focal length f can be obtained by specifying at least one vanishing points [Hartley and Zisserman 2004]. Consider a 3D surface point $\mathbf{x}$ with a 2D projection pixel $\mathbf{p}$, if we set the depth of

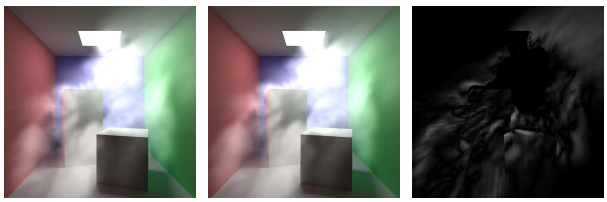


(a) Clear (b) Classif ed (c) Depth (d) TMap (e) VMap (f) Foggy

Figure 7: *Rendering fog into an exterior photograph. In Fig.(b), red, green and blue color represent vertical buildings, rough ground and sky respectively.*

(a) Ground Truth (b) Synthetic (c) Difference $\times 5$

Figure 8: *Comparison of ground truth with synthetic results rendered with homogenous media. Fig.(c) shows the difference of Fig.(a) and Fig.(b) with a scaling of f ve.*



(a) Ground Truth (b) Synthetic (c) Difference $\times 5$

Figure 9: *Comparison of ground truth with synthetic results rendered with heterogenous media. Fig.(c) shows the difference of Fig.(a) and Fig.(b) with a scaling of f ve.*

the pixel as d, then $\mathbf{x}$ can be evaluated by $\mathbf{x} = dK^{-1}[\mathbf{p}, 1]^T$. The view ray depth and direction can be evaluated by $||\mathbf{x}||$ and $\frac{\mathbf{x}}{||\mathbf{x}||}$. Fig.7(f) shows an example of rendering fog into an exterior photograph.

6 Results

We have implemented the proposed method on a PC equipped with an Intel(R) Core(TM) Duo E7400 2.8GHz CPU, 4G host memory, and an Nvidia Geforce GTX 285 graphics card.

6.1 Validation

We conduct experiments to validate the decomposition of transmittance and ref ective radiance in the proposed method. We render Cornell Box with participating media using physically based ray tracing engine as ground truth and the same scene without participating media as an input photograph of the proposed method. With the approximate illumination, geometry and the same media coeff cients, we produce the TMap, VMap and synthesize the composite result. Then, we visualize the difference of the ground truth and the composite result. We also evaluate the MSE (Mean Square Error) of the composite result compared to the ground truth for different media coeff cients.

Fig.8 and Fig.9 show examples of comparison of our results with the ground truth for homogenous and heterogenous media respectively. Fig.10(a) and Fig.10(b) show the curves of MSE with respect to α and δ respectively for both homogenous and heterogenous media. When evaluating MSE value, pixel colors are represented as a f oat point value in $[0, 1.0]$.

6.2 Comparison with Previous Method

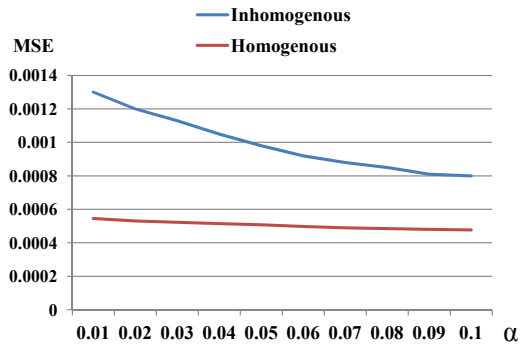
Jorge et.al [Jorge et al. 2008] proposed an image based participating media editing (IBPM) method using image processing techniques. However, they assume the illumination is constant for all pixels. According to our experiments, this assumption is unsuitable for environments in which the illumination varies largely, such as interior illumination. Fig.11 shows a comparison of the proposed method with IBPM. As can be seen, constant illumination makes the scattering effect look unnatural and even leads to loss of glow effect around light sources.

For most cases of exterior photographs taken in daytime, the illumination can be simplif ed to constant skylight. However, there are still exceptions, e.g., the foggy scenery illuminated by streetlights and headlight of cars at rainy night. Fig.12 shows a comparison of the proposed method with IBPM for exterior photographs taken at night.

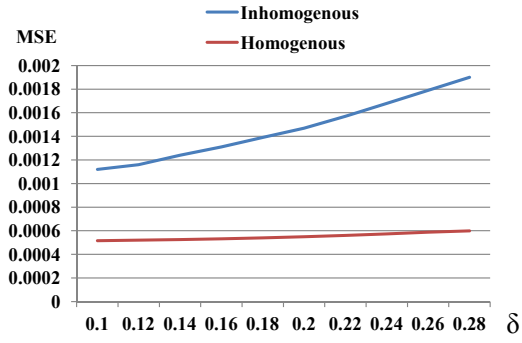
6.3 User Study

We also devised a user study to measure how well users can differentiate between real images and synthetic images of the same scene. A total number of 11 subjects participated in the experiment: 10 male and 1 female. All of them are computer graphics researchers with normal or rectif ed to normal eye sight. We showed each of them the clear photograph, ground truth photograph with fog and synthetic fog result for different scenes. Images were shown on a 20 inch Dell LCD monitor one at a time. Participants were asked to pick the ”real one” from the two given foggy Images.

Stimulus Images Fig.13 shows an example of stimulus images, in which ”Clear” and ”Real Fog” are real photographs and ”Synthetic Fog” are produced by the proposed method using ”Clear” photographs as input. We also simulate thin and heavy foggy effect respectively with the proposed method and provide real foggy photographs in user study, as shown in Fig.14.



(a) curve of MSE varying with respect to α



(b) curve of MSE varying with respect to δ

Figure 10: *Curves of MSE varying with respect to α and δ respectively for both homogenous and inhomogenous media.*

(a) Clear (b) IBPM (c) Ours

Figure 11: *Comparison of our result with IBPM for interior photographs.*

(a) Clear (b) IBPM (c) Ours

Figure 12: *Comparison of our result with IBPM for exterior photographs taken at night.*

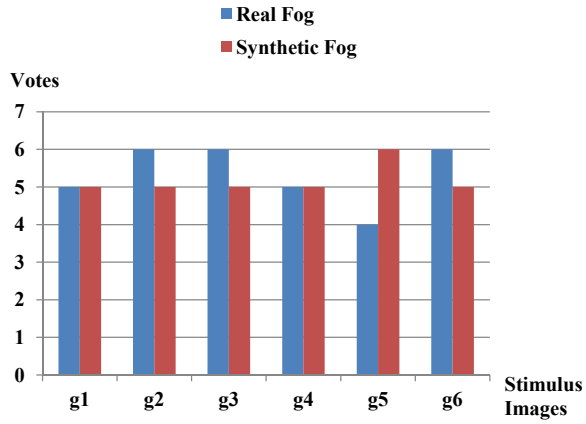


Figure 15: *Statistic result of user study. "g1" to "g6" represent 6 different groups of stimulus images, including real captured photographs and synthetic images, viewed by participants.*

Result Fig.15 shows the result of the study. We provided 6 different groups of stimulus images to participants. They were asked to observe each group of images and choose the "real" one. The time to observe and answer was limited to 15 seconds.

6.4 Timing

Table 1 shows time statistics of each stage of two typical examples. The first is a interior scene (see Fig. 5(a)). On the stage of illumination estimation, it requires manual annotation for light area and sunlight region on the ground, which costs about 30 seconds, and the sunlight shaft optimization costs about 10 seconds. It takes about 30 seconds to specify cuboid proxies on the geometry estimation. The rendering of transmittance map and volumetric map cost about 20 and 35 seconds, respectively.

The second is an exterior scene (see Fig. 7(a)). In this example, illumination estimation costs about 20 seconds (only need to specify a non-shadow region and a shadow region, refer to [Xing et al. 2012]), and geometry estimation costs about 50 seconds for scene labeling and depth generation. It takes about 25 and 40 seconds to render transmittance map and volumetric map, respectively.

Table 1: *Time statistics of our method (second)*

Img.	Res.	Illu. Est.	Geo. Est.	Tra. Ren.	Vol. Ren.
bathroom	480 × 480	∼ 40	∼ 30	∼ 20	∼ 35
building	560 × 700	∼ 20	∼ 50	∼ 25	∼ 40

7 Conclusion

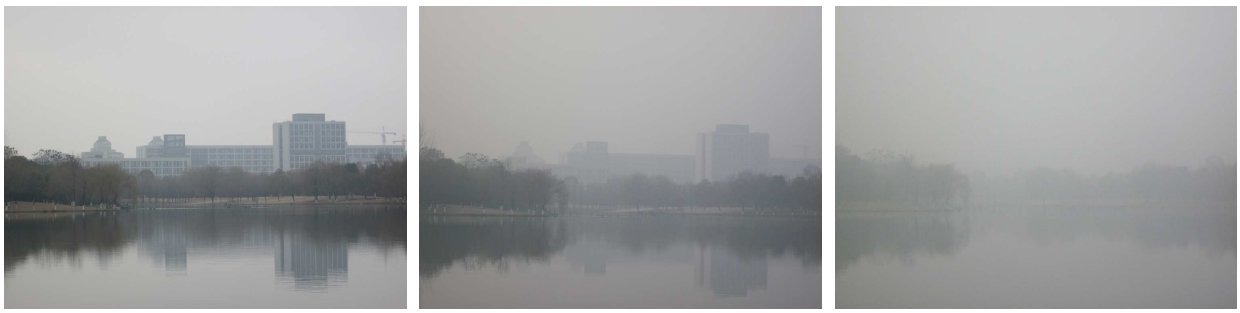
We have presented a framework for rendering synthetic fog into interior and exterior photographs with some user assistance. Compared with the previous image processing method, our approach can produce better result for taking advantage of physically based rendering techniques. Moreover, our approach takes into account the glow and light shaft effect which improves realism of the synthetic results. Through a user study, we show that our approach can produce synthetic images confusable with real photographs.

Limitation and Future Work Since our approach simulates foggy effect with single scattering, it can not capture volumetric caustic effect. Especially for input photographs containing refractive objects, it is difficult to use a simple geometry approximation

(a) Clear (b) Real Fog (c) Synthetic Fog

(d) Clear (e) Real Fog (f) Synthetic Fog

Figure 13: *Example of stimulus images provided in user study.*



(a) Clear (b) Real Thin Fog (c) Real Heavy Fog

(d) Synthetic Thin Fog (e) Synthetic Heavy Fog

Figure 14: *Example of stimulus images for simulating different extent of fog provided in user study.*

for caustics evaluation. For future work, we would like to introduce multiple scattering rendering and caustics. On the other hand, we want to model participating media density from real photographs. We consider it would be more realistic to transfer participating media from a real scene to another.

Acknowledgements

This work was partially supported by NSFC (No. 61379068).

References

ALNASSER, M., AND FOROOSH, H. 2006. Image-based rendering of synthetic diffuse objects in natural scenes. In *Proceedings of the 18th International Conference on Pattern Recognition - Volume 04*, IEEE Computer Society, Washington, DC, USA, ICPR '06, 787–790.

CEREZO, E., PEREZ-CAZORLA, F., PUEYO, X., SERON, F., AND SILLION, F. X. 2005. A Survey on Participating Media Rendering Techniques. *Visual Computer*.

COSSAIRT, O., NAYAR, S., AND RAMAMOORTHI, R. 2008. Light f eld transfer: global illumination between real and synthetic objects. *ACM Trans. Graph. 27*, 3 (Aug.), 57:1–57:6.

DEBEVEC, P. 1998. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, ACM, New York, NY, USA, SIGGRAPH '98, 189–198.

GLASSNER, A. 1995. *Principles of digital image synthesis*. The Morgan Kaufmann Series in Computer Graphics and Geometric Modeling. Morgan Kaufmann.

HARTLEY, R. I., AND ZISSERMAN, A. 2004. *Multiple View Geometry in Computer Vision*, second ed. Cambridge University Press, ISBN: 0521540518.

HOIEM, D., EFROS, A. A., AND HEBERT, M. 2007. Recovering surface layout from an image. *Int. J. Comput. Vision 75*, 1 (Oct.), 151–172.

JENSEN, H. W., AND CHRISTENSEN, P. H. 1998. Eff cient simulation of light transport in scences with participating media using photon maps. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, ACM, New York, NY, USA, SIGGRAPH '98, 311–320.

JORGE, L.-M., ANGEL, C., AND DIEGO, G. 2008. Image-based participating media. In *CEIG 08*, 179–187.

KAJIYA, J. T. 1986. The rendering equation. *SIGGRAPH Comput. Graph. 20*, 4 (Aug.), 143–150.

KARSCH, K., HEDAU, V., FORSYTH, D., AND HOIEM, D. 2011. Rendering synthetic objects into legacy photographs. *ACM Trans. Graph. 30*, 6 (Dec.), 157:1–157:12.

LAFORTUNE, E. P., AND WILLEMS, Y. D. 1996. Rendering participating media with bidirectional path tracing. In *Proceedings of the eurographics workshop on Rendering techniques '96*, Springer-Verlag, London, UK, UK, 91–100.

LALONDE, J.-F., EFROS, A. A., AND NARASIMHAN, S. G. 2009. Estimating natural illumination from a single outdoor image. In *ICCV*, IEEE, 183–190.

LALONDE, J.-F., EFROS, A. A., AND NARASIMHAN, S. G. 2009. Webcam clip art: appearance and illuminant transfer from time-lapse sequences. *ACM Trans. Graph. 28*, 5, 131:1–131:10.

LALONDE, J.-F., EFROS, A. A., AND NARASIMHAN, S. G. 2009. Webcam clip art: appearance and illuminant transfer from time-lapse sequences. *ACM Trans. Graph. 28*, 5 (Dec.), 131:1–131:10.

LALONDE, J.-F., EFROS, A. A., AND NARASIMHAN, S. G. 2012. Estimating the natural illumination conditions from a single outdoor image. *International Journal of Computer Vision 98*, 2, 123–145.

LIU, Y., QIN, X., XU, S., NAKAMAE, E., AND PENG, Q. 2009. Light source estimation of outdoor scenes for mixed reality. *Vis. Comput. 25*, 5-7 (Apr.), 637–646.

PAULY, M., KOLLIG, T., AND KELLER, A. 2000. Metropolis light transport for participating media. In *Proceedings of the Eurographics Workshop on Rendering Techniques 2000*, Springer-Verlag, London, UK, UK, 11–22.

PHARR, M., AND HUMPHREYS, G. 2010. *Physically Based Rendering, Second Edition: From Theory To Implementation*, 2nd ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

SIEGEL, R., AND HOWELL, J. 1992. *Thermal Radiation Heat Transfer*. Hemisphere.

XING, G., LIU, Y., QIN, X., AND PENG, Q. 2012. A practical approach for real-time illumination estimation of outdoor videos. *Computers & Graphics 36*, 7, 857–865.

ZHENG, Y., CHEN, X., CHENG, M.-M., ZHOU, K., HU, S.-M., AND MITRA, N. J. 2012. Interactive images: Cuboid proxies for smart image manipulation. *ACM Transactions on Graphics 31*, 4, 99:1–99:11.