Global Depth from Defocus with Fixed Camera Parameters

Yangjie Wei,
Graduate School of Chinese Academy of Sciences, & the State Key Laboratory of Robotics
Shenyang Institute of Automation, Chinese Academy of Sciences
Shenyang, Liaoning Province 110016, China
weiyangjie@sia.cn

Zaili Dong
the State Key Laboratory of Robotics
Shenyang Institute of Automation, Chinese Academy of Sciences
Shenyang, Liaoning Province 110016, China
dzl@sia.cn

Chengdong Wu
School of Information Science & Engineering
Northeast University
Shenyang, Liaoning Province 110016, China
wuchengdong@ise.neu.edu.cn

Abstract - Reconstruction depth from 2D images is an important research issue in computer vision, and depth from defocus (DFD) is an effective way which takes the blurring degree of the region images whose depth of field is limit as the tool of computing depth. Now though there are many DFD methods, they all need to change camera parameters in order to attain blurred images, such as the focal length of the lens, the radius of the lens. If cameras with high level of amplification are used, it is inhibitory to change camera parameters. Therefore, in this paper a novel DFD method is proposed. First, two different blurred images are captured through changing depth. Second, the blurred imaging model is constructed with the relative blurring and the diffusion equation, and the relation between depth and blurring is discussed from two aspects. Finally, the problem of computing depth is transformed into an optimization issue. The method proposed in this paper does not need to change camera parameters, so the process is very simple and can be used in some special applications. The simulation results show that this method can attain depth with high precision.

Index Terms –Depth from Defocus, Diffusion Equation, 3D Reconstruction

I. INTRODUCTION

Depth reconstruction, as an important research field in computer vision, is one of the ways of attaining useful 3D information from 2D images, and now it has been used in all kinds of fields, such as Medicine, Robotics, Remote Sensing. In recent years, there are various kinds of methods proposed in depth reconstruction, such as depth from stereo (DFS), depth from focus (DFF) and depth from defocus (DFD)[1].

DFD, firstly introduced by Pentland[5], is an effective way which takes the blurring degree of region images whose depth of field is limit as the tool of computing depth[6-8]. As is well known, blurring has an inherent relation with camera parameters and depth information, so if camera parameters and the blurring degree are given, it is easy to compute depth. The general way using DFD is to capture two or more images with different blurring by varying camera parameters, compute blurring degree of every point, and estimate depth from the point spread function. Now, DFD has become so attractive, because it needs little images; it avoids matching and masking problems; it is effective both in the frequency domain and in the spatial domain[9-10].

Measuring the blurring degree of every point is not only the base of DFD estimation, but also an important factor of determining the precision of it. However, blurring is not a point property of an image, but a region property. Therefore, there are two methods, including focal and global, used in DFD. As for focal DFD, a predefined window around every pixel point is needed, and the blurring of it is equal to the blurring of the window[5][11]. Although this method is simple, the size of the window is difficultly defined, because there is a trade-off between having a window that is as large as possible to average out noise, but as small as possible to guarantee that within it[12-13]. However, without assuming a window in which points have the same depth, depth estimation is on the base of losing the radiance and the depth just using two defocus images. The instinctive method is to construct the depth model and the radiance model simultaneously[12][15-16], but the computational cost is huge. One of the simplification methods is to simplify the imaging model, for example, assuming the scene contains “sharp edges”, that is there are discontinuities in the scene[17]. Another way is that assuming the radiance can be approximated by a cubic function[18], or structured light[19]. However, the application environment of these methods is limit.

In this paper a novel global DFD method is proposed, in which the blurred imaging model is constructed with the relative blurring and the diffusion equation, and the relation between depth and blurring is discussed from two aspects.
The method proposed in this paper does not need changing camera parameters, so the process is simple, and the simulation results show that this method can attain depth with high precision.

The contents of this paper are organized as follows. Firstly, in section II, the imaging model and the diffusion equation are proposed. Then, the relative blurring and the novel DFD method are introduced respectively in section III and section IV. In section V, the simulation results based on the new method are given. Finally, section VI is the conclusion of this paper.

II. THE IMAGING MODEL FOR DEFOCUS

A. The Imaging Theory for Defocus

\[
\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \tag{1}
\]

Based on the optical theory, when the focal length \( f \), the distance of the object from the principal \( u \), and the distance of the focused image from the lens plane \( v \) satisfy the standard function (Eq. (1)), the image is focused, that means the image of a source point is also a focused point, as is shown in Fig. 1.

![Fig.1 the standard theory](image)

If the relation above is changed, the image of a source point becomes a blurred round whose distribution can be denoted with a 2D Gaussian function which is written as,

\[
h(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{2}
\]

Where, \( x \) and \( y \) are the horizontal and the vertical axis respectively; \( \sigma \) denotes the spread of the Gaussian kernel. Therefore, a blurred image can be considered as the summation of many blurred points, and this process can be denoted by the following convolution function,

\[
E(x, y) = I(x, y) * h(x, y) \tag{3}
\]

Where, \( E(x, y) \) is the blurred image, \( I(x, y) \) denotes the focused image, and the radius satisfies,

\[
b = \frac{Dv}{2} \frac{1}{f} \frac{1}{v} - \frac{1}{s} \tag{4}
\]

Where, \( s \) denotes depth of the blurred point, \( D \) denotes the radius of the lens.

B. The imaging model denoted by the diffusion function

Suppose the point spread function Eq. (2) is not shift-invariant, that is

\[
h(R) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{R^2}{2\sigma^2}\right)
\]

\[
R = \sqrt{x^2 + y^2}
\]

Eq. (3) can be denoted as an isotropic heat equation,

\[
\begin{align*}
\tilde{u}(x, y, t) &= a \Delta u(x, y, t) \quad a \in [0, \infty) \quad t \in (0, \infty) \\
u(x, y, 0) &= r(x, y)
\end{align*} \tag{6}
\]

Consider \( u(x, y, 0) \) is a focused image, the solution of the heat equation can be obtained in terms of convolution of the image with a temporally evolving Gaussian kernel, so the heat equation can be introduced into the process of the defocus imaging. Where, \( a \) is called the diffusion coefficient,

\[
\Delta u = \frac{\partial^2 u(x, y, t)}{\partial x^2} + \frac{\partial^2 u(x, y, t)}{\partial y^2}
\]

If the depth map is an equifocal plane, \( a \) is a constant, or \( a \) is shift-variant. It is also easy to verify that the variance \( \sigma \) is related to the diffusion coefficient \( a \) via

\[
s^2 = 2ta \tag{8}
\]

If the depth map is not an equifocal plane, the diffusion equation is defined as,

\[
\begin{align*}
\tilde{u}(x, y, t) &= \nabla \cdot (a(x, y) \nabla u(x, y, t)) \quad t \in (0, \infty) \\
u(x, y, 0) &= r(x, y)
\end{align*} \tag{9}
\]

Where \( \nabla \) denotes the gradient operator, \( \nabla \cdot \) is the divergence operator,

\[
\nabla \cdot = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \tag{10}
\]

III. THE RELATIVE BLURRING

To construct the diffusion equation, we need know \( r(x, y) \), that is the radiance. However, on some situations, construction of the radiance is a very complicated process. Moreover, the resolution is always very low without the restoration model. Favaro \cite{14} has introduced a concept of the relative blurring. Suppose there are two images \( E_1(x, y) \) and \( E_2(x, y) \) for two different focus setting. Also, \( \sigma_i < \sigma_z \) (that is, \( E_1 \) is more defocused than \( E_2 \)), \( E_2 \) can be written as,

\[
E_2(x, y) = \int \int \frac{1}{2\pi\sigma_y^2} e^{\frac{(x-u)^2+(y-v)^2}{2\sigma_y^2}} r(u, v) \, du \, dv
\]

\[
= \int \int \frac{1}{2\pi\sigma_y^2 - \sigma_z^2} e^{\frac{(x-u)^2+(y-v)^2}{2\sigma_y^2 - \sigma_z^2}} \, du \, dv \tag{11}
\]

\[
= \int \int \frac{1}{2\pi\sigma_y^2} e^{\frac{(x-u)^2+(y-v)^2}{2\sigma_y^2}} r(x, y) \, dx \, dy
\]

\[
= \int \int \frac{1}{2\pi\sigma_y^2} e^{\frac{(x-u)^2+(y-v)^2}{2\sigma_y^2}} \Delta \sigma^2 \, du \, dv
\]

Where, \( \Delta \sigma^2 \Delta = \sigma_z^2 - \sigma_i^2 \), is called the relative blurring. Thus Eq. (11) can be explained as the solution of Eq. (6), but the initial value is \( E_i \), rather than \( r \). So, Eq. (6) can be written as,
The coefficient can be defined as,

\[ s = \arg \min f(u(x,y,t) - E(x,y))^2 \]  

where, \( f \) is the objective function, \( u(x,y,t) \) is the original image, and \( E(x,y) \) is the desired image. The optimization process can be formulated as,

\[ s = \arg \min f(u(x,y,t) - E(x,y))^2 \]  

where, \( s \) is the solution of the optimization problem.

Based on Eq.(17), the diffusion equation can be written as,

\[ \Delta \sigma^2 = 2\sigma_0 \]  

where, \( \Delta \sigma^2 \) is the relative blurring, \( \sigma_0 \) is a constant between the blurring radius and the degree of blurring.

Thus, the relation between the relative blurring and the depth map can be denoted as,

\[ \Delta \sigma^2 = \|b_i \|^2 - \|b_j \|^2 \]  

where, \( b_i \) is the blurring radius and \( b_j \) is the degree of blurring.

Thus, the depth map is,

\[ s(x,y) = \frac{1}{f} \left( \frac{1}{v} s(x,y) - \frac{1}{f} \right) \]  

where, \( f \) is the focal length of the camera and \( v \) is the depth of the object.

Thus, the new DFD method can be formulated as,

\[ \bar{s} = \arg \min \int (u(x,y,t) - E(x,y))^2 \]  

where, \( \bar{s} \) is the solution of the optimization problem.

However, the optimization process above is ill posed, that is, the minimum may not exist, even if it exists, it may not be stable with respect to data noise. A common way to regularize the problem is by adding a Tikhonov Penalty,

\[ F(s) = \int \int (u(x,y,t) - E(x,y))^2 dxdy \]  

where, \( F(s) \) is the cost energy.

Thus, the solution process is equal to the following,

\[ \bar{s} = \arg \min \int (u(x,y,t) - E(x,y))^2 \]  

s.t. Eq.(17), Eq.(20)
Eq. (24) is a dynamic optimization which can be solved by the gradient flow, the algorithm can be divided into the following steps, (the detailed process can be seen in literature [14]).

Step-I. Given camera parameters \( f, D, \gamma, v, s_0 \); two defocused images \( E_1, E_2 \); a threshold \( \varepsilon \); \( \alpha \) and optimization step \( \beta \).

Step-II. Initialize the depth map with a plan \( s_0 \).

Step-III. Compute Eq. (18), (19);

Step-IV. Compute Eq. (17);

Step-V. Compute Eq. (23) with the solution \( u(x, y, \Delta t) \) of last step. If the cost energy is below \( \varepsilon \), stop; or, compute the following equation with step \( \beta \),

\[
\frac{\partial s}{\partial t} = -F'(s)
\]

(25)

Step-VI. Compute Eq. (20), update the depth map, and return to Step-III.

B. The DFD with known variation value of depth

Suppose \( E_1(x, y) \) attained before variation, whose depth \( s_1(x, y) \) is unknown, and \( E_2(x, y) \) attained after variation, we estimate the depth map \( s_1(x, y) \) through increasing or decreasing depth which is known, the theory is shown in Fig. 2.

Suppose \( s_0 \) is the focused depth, then

\[
\frac{1}{f} \left( \frac{1}{v} - \frac{1}{s_0} \right) = 1
\]

(26)

\[
s_1(x, y) = s_0 + \Delta s(x, y)
\]

(27)

Where \( s_2(x, y) \) is the depth map after variation. Based on Eq. (4), the blurring radius \( b \) is,

\[
b = \frac{Dv}{2} \left( \frac{1}{f} - \frac{1}{v} - \frac{1}{s_0} \right) = \frac{Dv}{2} \left| \frac{1}{s_0} - \frac{1}{s} \right|
\]

(28)

Where, \( ds = |s - s_0| \).

According to Eq. (16), we have,

\[
-2\Delta s \cdot s_1^2 + (2\Delta s^2 - 2\Delta s \cdot s_0) \cdot s_1 + 2\Delta s^2 \cdot s_0
\]

\[
= \frac{4\Delta s^2}{\gamma^2 D^2 v^2} s_0 \cdot s_0 \cdot s_0
\]

(29)

Eq. (29) can be changed into the following form,

\[
as^4 + bs^3 + cs^2 + ds + e = 0
\]

(30)

Where,

\[
a = \frac{4\Delta s^2}{\gamma^2 D^2 v^2} s_0, b = -\frac{8\Delta s^2}{\gamma^2 D^2 v^2} s_0 \cdot \Delta s
\]

\[
c = \frac{4\Delta s^2}{\gamma^2 D^2 v^2} s_0 \cdot \Delta s^2 + 2\Delta s
\]

\[
d = 2\Delta s(s_0 - \Delta s), e = -\Delta s^2 \cdot s_0
\]

The solution process is as follows.

First, normalize the coefficients,

\[
a = 1, \quad b = \frac{b}{a}, \quad c = \frac{c}{a}, \quad d = \frac{d}{a}, \quad e = \frac{e}{a}
\]

Then, transform the quartic equation into a cubic equation,

\[
y^3 + ky^2 + my + n = 0
\]

(31)

Where, \( k = -c, \ m = bd - 4e, \ n = -d^2 - b^2e + 4ce \).

The solution of Eq. (31) is,

\[
y = \sqrt[3]{\frac{2k^3}{27} + \frac{km}{3} + n} - \frac{2k}{3}
\]

(32)

Where,

\[
S = \sqrt[3]{\frac{(2k^3 - km + n)^2}{27} - \frac{k^2}{3} + \frac{m}{3}} - \frac{2k}{3}
\]

Finally, the depth map is equal to,

\[
s = \frac{1}{2} \sqrt{\frac{(1 - b + \bar{s})^2}{2} - 4 \left( \frac{1}{2} y - \bar{s} \right)^2}
\]

(33)

Where,

\[
\bar{s} = \sqrt{\frac{1}{4} b^2 - c + y \bar{s}} = \sqrt{\frac{1}{4} y^2 - c}
\]

The steps are the same as that in section 4.1 except that Eq. (20) is replaced by Eq. (33).

V. SIMULATION RESULTS

In order to validate the new algorithms, we use a number of synthetic images to test, where \( f=12\text{mm}, \ v=12.2\text{mm}, \ s_0=850\text{mm} \) which is the distance of the object from the principal when the image is focused, F-number 2, \( D=ft2, \ \gamma=0.002 \). In experiment, we test the new algorithm on a smooth plane, a slope, a cosine, and a wave. Furthermore in
order to test the influence of the texture sharpness, we add three levels of texture sharpness along the horizontal axis.

A. Simulation with unknown variation value of depth

In this section, we test the proposed algorithm in section 4.1, and the results are shown in Fig.3 to Fig.8. Fig.3 to Fig.5 are the simulation results of a cosine, where Fig.3 is two defocused images (the left is the image before variation; the right is after variation); the depth maps of gray are shown in Fig.4 (the left is the computed depth map with our algorithm; the right is the true depth map); Fig.5 is the depth maps of curve (the left is the computed map with our algorithm; the right is the true map). From these figures, we can conclude that the new algorithm can attain good results in constructing depth map, and at the same time the influence of the texture sharpness is very little.

B. Simulation with known variation value of depth

In this section, we test the proposed algorithm in section 4.2, and the results are shown in Fig.9 to Fig.14. Fig.9 to Fig.11 is the simulation results of a wave, and Fig.12 to Fig.14 are the simulation results of a smooth plane. From these figures, we can conclude that our new algorithm is effective in all kinds of planes, and the precision is very high.
VI. CONCLUSION

On some special situations in computer vision, changing the camera parameters, such as $f$, $v$ or $D$, to attain two defocused images is very difficult. In this paper a new DFD method is proposed based on changing the depth of the scene. The new algorithm is not only simple, but also can be used in some high amplification applications. Finally, the simulations with respect to all kinds synthetic images show that the feasibility and validity of the new DFD.

REFERENCES