Depth Estimation from Defocus Images Based on Oriented Heat-flows

Liu Hong, Jia Yu, Cheng Hong, Wei Sui
Education Ministry Key Laboratory of Intelligent Computing and Signal Processing
Anhui University
Hefei, China
liuhong@ahu.edu.cn

Abstract—The amount of blur on the defocus image depends on the depth information of the scene. So depth of the scene can be estimated by calculating the blur with the knowledge of the lens parameters. A novel depth estimation mode based on oriented heat-flows is proposed in this paper. In this model, the process of image defocusing is described using oriented heat-flows diffusion. The diffusion can be seen as the coupling of two weighted heat flows along orthonormal directions. The diffusion directions are defined by the local coherence geometry of image and the diffusion strength is the function of blur amount. Experimental results show that the model is quite effective and the emergence of artificial depth information in edge can be avoided by using this model.

Keywords-depth estimation; oriented heat-flows; orthonormal directions; diffusion strength

I. INTRODUCTION

Depth estimation is one of the most important tasks in computer vision, applied to three dimension object recognition, scene interpretation and robotics. Pentland has proposed a depth-from-defocus (DFD) method, where the basic principle is to use the characteristics of the imaging system [1]. The image will be focused when object is located at the focus plane, whereas the image will be blurred when object is front or behind the focus plane. As the distance between the object and the surface of exact focus increases, the imaged object becomes more defocused. Due to defocus images provide a depth cue, the depth of a point object can be recovered, by measuring the amount of blur of the point object in the defocus images. DFD only requires two or more defocus images obtained by a single camera, providing a method of measuring depth, avoiding the correspondence problem of stereo vision.

The depth estimation algorithms calculates the amount of blur from either the image's power spectrum in the frequency domain, or from the image's point spread function in the spatial domain. For examples, Pentland recovered the depth by using the deconvolution in the frequency domain, imposing the constraint of one image being formed with the pinhole aperture [1]. Subbarao proposed a more general method in which he removed the constraint of one image being formed with the pinhole aperture using two defocus images recorded with different camera settings [2]. Watanabe and Nayar suggested the use of broadband rational filters [3]. Subbarao proposed a S-Transform which does deconvolution in the spatial domain [4]. Ziou and Deschenes proposed a technique based on image decomposition using the Hermite polynomial basis [5]. Favaro minimized a measure of information divergence between blurred images [6]. Favaro and Soatto used a functional singular value decomposition to compute the point spread function [7]. Namboodiri and Chaudhuri explored the linear diffusion principle for depth estimation [8]. Subsequently they proposed the defocus space, which is the set of all possible observations by varying the lens parameters for a given scene [9].

However, these techniques either introduce artifacts in the solution, such as edge bleeding and windowing, or simplify of the imaging model, such as, some assume that the scene contains "sharp edges", some assume that the scene can be locally approximated by polynomials, others that the surface of scene can be locally approximated by a plane paralleled the image plane.

In this paper, a novel depth estimation model is proposed based on oriented heat-flows. The retrieval of focus image and the excess restrictions are avoided. Experimental results show that better depth results can be obtained and depth leakage is prevented in edge.

II. DEFOCUS IMAGING

Single convex lens defocus imaging is shown in Fig. 1.

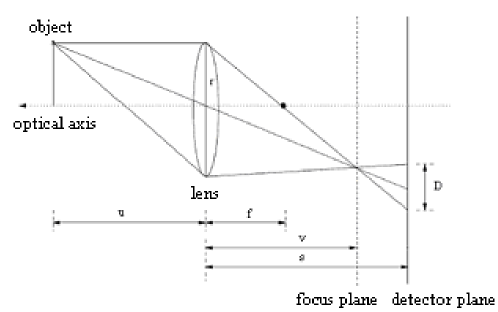

Figure 1. Single convex lens defocus imaging.

In Fig. 1, r is the diameter of lens, D is the diameter of blur circle, f is the focal length of the lens, u is the distance between the object and the lens, v is the distance between the focus image plane and the lens, s is the distance between the image detector plane and the lens. When the image

978-0-7695-3944-7/10 $26.00 © 2010 IEEE
DOI 10.1109/ICMV.2009.32

detector plane is located at v, the image of object is focused, then u , v and f certainly satisfy the thin lens law:

$$\frac{1}{u}+\frac{1}{v}=\frac{1}{f}. \tag{1}$$

When the image detector plane is located at s, $s \neq v$, the image is defocused. The diameter of blur circle is D. By introducing geometric optics in our analysis, the diameter D is given by [2]:

$$D = rs\left|\frac{1}{f}-\frac{1}{s}-\frac{1}{u}\right|. \tag{2}$$

The defocus image of point light source is the point spread function (PSF), using the notation $h(x,y)$ which is a constant intensity of the blur circle. $h(x,y)$ can be well approximated with the circularly symmetric 2D Gaussian [2] , as follow:

$$h(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{3}$$

where σ is the standard deviation, $\sigma = kD/2$, k is common approximated with $1/\sqrt{2}$ [2]. Thus, substituted (2) and k into $\sigma = kR$ can obtain:

$$\sigma=\frac{rs}{2\sqrt{2}}\left|\frac{1}{f}-\frac{1}{s}-\frac{1}{u}\right|. \tag{4}$$

Equation (4) points out that the relation between multiplicative inverse of depth u and standard deviation σ. In (4), f and r are camera parameters known. s can be obtained through the camera calibration. The depth u of a 3D object can be obtained by estimating the standard deviation of the defocus image.

III. Oriented Heat-flows Diffusion

A. Oriented Heat-flows

$I(x,y)$ is the gray of 2D image. The oriented heat-flows have the following partial differential equation (PDE) form [10]:

$$\frac{\partial I}{\partial t}=c_1 I_{nn}+c_2 I_{mm}, \tag{5}$$

where $n,m \in R^2$ and $n \perp m$, $c_1, c_2 \in R$ and $c_1, c_2 \geq 0$. I_{nn} and I_{mm} is the second-order derivative of $I(x,y)$ respectively along the direction n and m. The above PDE describes 1D oriented heat-flows diffusion which is along two orthogonal direction n and m respectively with the weight c_1 and c_2. At this time, the diffusion is entirely decided via the direction n and m and the weight c_1 and c_2. If $c_1 = c_2$, the diffusion degenerates to the isotropic heat diffusion.

The equation can also be rewritten based on the trace of diffusion tensor as follow form:

$$\frac{\partial I}{\partial t}=c_1 I_{nn}+c_2 I_{mm}=trace(\mathrm{TH}), \tag{6}$$

where H is the Hessian matrix of I and $\mathrm{T}=c_1 nn^{\mathrm{T}}+c_2 mm^{\mathrm{T}}$ is diffusion tensor which is a 2×2 semi-definite symmetric matrix with the eigenvalue c_1 , c_2 corresponding the eigenvector n , m . It is shown that the diffusion tensor T describes the local diffusion characteristics of the oriented heat-flows PDE.

Introducing the above oriented heat-flows PDE to information diffusion, how to choose the diffusion orientation and the corresponding diffusion coefficient is a key, that is how to design diffusion tensor T .

B. Design Diffusion Orientation

The image structure tensor is defined as:

$$\mathrm{S}=\nabla I \nabla I^{\mathrm{T}}. \tag{7}$$

Its significance lies in that the two eigenvectors of the structure tensor are respectively the gradient direction and its vertical direction. The corresponding eigenvalue is the square of amplitude along that direction.

Let μ_1 and μ_2 is respectively the eigenvalue of structure tensor S , corresponding eigenvector η and ξ. If $\mu_1 \geq \mu_2$, η is the correlative gradient direction and the vertical gradient direction ξ is the correlative isophote direction.

C. Design Diffusion Coefficient

After determining the diffusion direction, we choose diffusion coefficient.

At the small $|\nabla u|$, it should make the diffusion approximation for the isotropic. Namely, the intensity of diffusion along the two directions is approximately equal:

$$\lim_{\kappa\to\infty} c_1 = \lim_{\kappa\to\infty} c_2 = \gamma(u), \tag{8}$$

where $\gamma(u)$ denotes diffusion function [11].

At the large $|\nabla u|$, the diffusion is anisotropic. The diffusion along the vertical edge direction should tend to 0 which intensity should be much smaller than the intensity along the isophote direction:

$$\lim_{\kappa\to\infty} c_1 = 0,\quad \lim_{\kappa\to\infty} c_1 \ll \lim_{\kappa\to\infty} c_2 \leq \gamma(u). \tag{9}$$

The diffusion direction and corresponding diffusion coefficient which is satisfied above conditions can choose to:

$$\begin{cases} n = \eta \\ m = \xi \end{cases}, \quad \begin{cases} c_1 = \dfrac{\gamma(u)}{1+|\nabla u|^2} \\ c_2 = \gamma(u) \end{cases}. \tag{10}$$

IV. OPTIMIZATION

Adjusting the distance of the lens and the image detector plane which is v_1 and v_2 separately, we can obtain two defocus images $I_1(x,y)$ and $I_2(x,y)$ corresponding to blurring parameters σ_1 and σ_2. Set $\sigma_1 < \sigma_2$, then $I_2(x,y)$ can be diffused with the anisotropic manner from $I_1(x,y)$.

The depth estimate problem is converted into the cost functional minimum problem as follow:

$$\min_u \left(\int_\Omega \left(I_2(x,y) - I(x,y,t) \right)^2 d\Omega + \lambda \int_\Omega |\nabla u(x,y)| d\Omega \right), \tag{11}$$

where $\forall x, y \in \Omega$, $u(x,y)$ is depth of the scene at point (x,y), λ is Lagrange multiplier. The first term is the fidelity term, to ensure the solution and the fidelity between the measured data. The second term is the total variation regularization term, to choose from a number of candidate solutions of an optimal solution. The interaction between the regularization term and the fidelity ensure that the algorithm automatically converges to the optimal solution.

Denote $\Psi = \int_\Omega \left(\mathrm{I}_2(x,y) - \mathrm{I}(x,y,t) \right)^2 d\Omega$, then (11) with minimum is satisfied the following Euler_Lagrange equation:

$$\Psi'(\gamma(u)) \cdot \gamma'(u) - \lambda \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) = 0. \tag{12}$$

By introducing the additional time parameter, the solution can be gotten via gradient descend method:

$$\frac{\partial u}{\partial t} = \Psi'(\gamma(u)) \cdot \gamma'(u) - \lambda \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right). \tag{13}$$

V. EXPERIMENTAL RESULTS

To verify the validity and accuracy of this model, this paper performs the simulated images and the real images on the MATLAB R2008a. And the result is compared to the least square method. Experimental results show that the proposed model has higher accuracy.

A. Simulation Image Experiment

In this experiment, simulate three depth defocus images. The size of image is 125×125, shown in Fig. 2 (a). The parameters of camera are set as follows: the focal length f is 0.012m, F-number is 2. The distance between the lens and the image plane are respectively 52mm and 85mm. The two different blur defocus images are shown in Fig. 2 (b) and Fig. 2 (c). This camera parameter setting and the two defocus images are used as input to estimate the depth. The true depth map is shown by gray-level in Fig. 2 (d). The estimated depth image using least square algorithm is shown in Fig. 2 (e). The estimated depth image using the proposed algorithm is shown in Fig. 2 (f). Gray-level images are used to indicate the depth, white means near to camera while black means far from camera. In the estimated depth (e) and (f), the proposed model avoid the information in the depth jump edge leaking, making the depth in edge is better maintained than that of the least square [7].

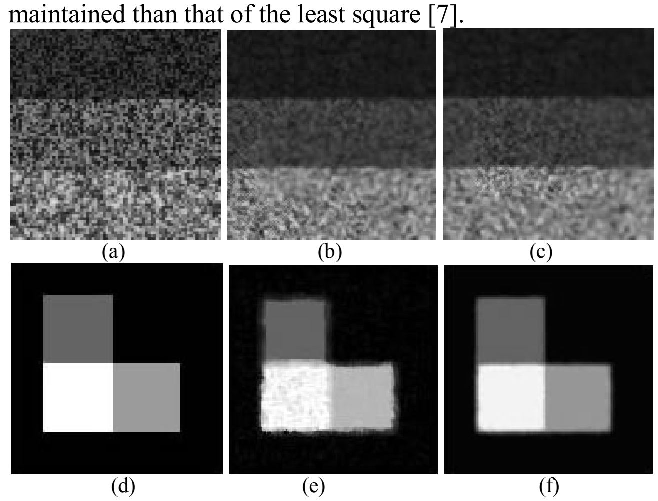

Figure 2. Simulation image experiment: (a) simulation focused image, (b) and (c) two different defocus images, (d) true depth, where white means near to camera and black means far from camera, (e) estimated depth using least square algorithm, (f) estimated depth using proposed algorithm.

Define the root mean square relative error ε using (14) to quantify the accuracy between the true depth and estimated depth, given in Table 1.

$$\varepsilon = \frac{1}{M*N} \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} \left(\frac{u_\gamma(i,j)}{u_t(i,j)} - 1 \right)^2}, \tag{14}$$

where $u_t(i,j)$ is the true depth, $u_\gamma(i,j)$ is the estimated depth, M, N is the row and the column of image.

TABLE I. ERROR ANALYSIS

Error	Algorithm	
	Least square	***Total variation***
Root mean square relative error	6.35%	3.85%

From Table 1, the proposed model in this paper is more accurate than that of the least square. The depth recovery error almost reduces about 39% compare with the least square.

B. *Real Images Experiment*

Experiments use the EOS 450D Camera which lens is EF-S18-55mm f/3.5-5.6.

In the real experiment, to demonstrate the feasibility of the proposed model, we choose two arbitrary shape objects. One is a plum, another is an apricot. Adjust the distance between the lens and the image plane, obtain two different defocus images and cut them, shown in Fig. 3 (a) and (b). The size of images is 625×707, the focal length f is 42mm and F-number is 5. The distance between the lens and objects are respectively 0.65m and 0.79m. Fig. 3 (c) is the estimated depth map of experimental images which is gray-level using the white corresponding near to camera and black corresponding far from camera.

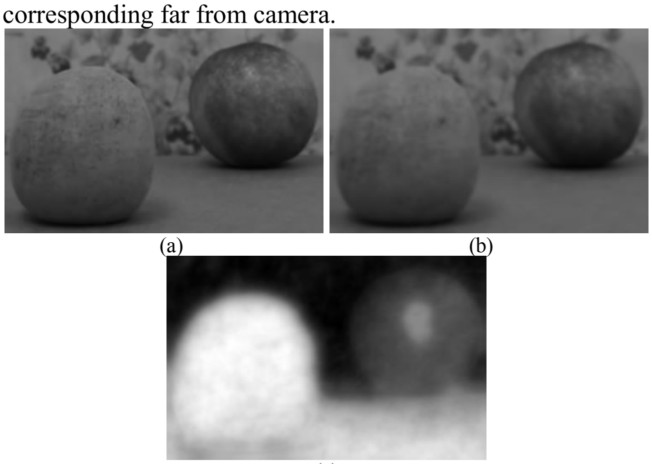

Figure 3. Real experiment: (a) and (b) two different defocus images, (c) the estimated depth using the proposed model

VI. CONCLUSION

This paper presents a new depth estimation model using oriented heat-flows on the basis of optical defocus. This model voids the matching problem which is inevitable in stereo vision and motion vision. The experimental results show that the model is quite effective and the emergence of artificial depth information in edge can be avoided by using this model. Compared with others, the model is stability with small error. So it has good application prospects.

ACKNOWLEDGMENT

This paper is supported by the National Natural Science Foundation of China under Grant No. 60603083 and No. 60872106 and the Research Fund for the Doctoral Program of Higher Education 200703570003.

REFERENCES

[1] A.P. Pentland, “A New Sense for Depth of Field”, IEEE Trans. Pattern Anal. Machine Intell., Vol.9, No.4, 1987, pp.523-531, doi: 10.1109/TPAMI.1987.4767940.

[2] M. Subbarao, “Parallel Depth Recovery by Changing Camera Aperture”, In: Proc. Internat. Conf. on Computer Vision, Florida, USA, 1988, pp. 149–155.

[3] M.Watanabe and S.K.Nayar, “Rational Filters for Passive Depth from Defocus”, *Internat. J. Comput. Vision* 27 (3), 1998, 203–225, doi: 10.1023/A:1007905828438.

[4] M. Subbarao and G. Surya, “Depth from Defocus: a Spatial Domain Approach”, *Internat. J. Computer Vision*, Vol. 13, 1994, pp. 271–294, doi: 10.1007/BF02028349.

[5] D. Ziou and F. Deschenes, “Depth from Defocus Estimation in Spatial Domain”, *Computer Vision and Image Understanding*, Vol. 81, No. 2, February 2001, pp. 143–165, doi: 10.1006/cviu.2000.0899.

[6] P. Favaro, A. Mennucci and S. Soatto, “Observing Shape from Defocused Images”, *Internat. J. Comput. Vision* 52 (1), 2003, 25–43, doi: 10.1023/A:1022366408068.

[7] P. Favaro and S. Soatto, “A Geometric Approach to Shape from Defocus”, *IEEE Trans. Pattern Anal. Machine Intell.* 27 (3), 2005, 406–417, doi: 10.1109/TPAMI.2005.43.

[8] V.P.Namboodiri, and S.Chaudhuri, “Use of Linear Diffusion in Depth Estimation Based on Defocus Cue”, *In: Proc. 4th Indian Conf. on Computer Vision, Graphics and Image Processing*, Kolkata, India, December 2004, pp. 133–138.

[9] V.P. Namboodiri and S. Chaudhuri, “On Defocus, Diffusion and Depth Estimation”, *Pattern Recognition Letters* 28, 2007, pp.311–319, doi: 10.1016/j.patrec.2006.04.001.

[10] Lei Qu, Sui Wei, Dong Liang and Nian Wang, “Inpainting Model which Based on Curvature Driven Oriented Heat-flows”, *Computer Engineering ans Applications*, 2008, Vol. 44, No.4, pp.31–33.

[11] P. Favaro, S. Soatto, M. Burger and S. J. Osher, “Shape from Defocus via Diffusion”, *IEEE Trans. PRMI*, March 2008, doi: 10.1109/TPAMI.2007.1175.