

Shape Recovery Using Stochastic Heat Flow

Vinay P. Namboodiri and Subhasis Chaudhuri
Department of Electrical Engineering
Indian Institute of Technology, Bombay
Powai, Mumbai 400076, India.
{vinaypn,sc} @ ee.iitb.ac.in

Abstract

We consider the problem of depth estimation from multiple images based on the defocus cue. For a Gaussian defocus blur, the observations can be shown to be the solution of a deterministic but inhomogeneous diffusion process. However, the diffusion process does not sufficiently address the case in which the Gaussian kernel is deformed. This deformation happens due to several factors like self-occlusion, possible aberrations and imperfections in the aperture. These issues can be solved by incorporating a stochastic perturbation into the heat diffusion process. The resultant flow is that of an inhomogeneous heat diffusion perturbed by a stochastic curvature driven motion. The depth in the scene is estimated from the coefficient of the stochastic heat equation without actually knowing the departure from the Gaussian assumption. Further, the proposed method also takes into account the non-convex nature of the diffusion process. The method provides a strong theoretical framework for handling the depth from defocus problem.

1 Introduction

The limited depth of field introduces a defocus blur in images captured with conventional lenses based on the range of depth variation in a scene. This artifact has been used in computer vision for estimating depth in the scene. As discussed in [7], this method of shape recovery is particularly relevant for complex scenes which have a large amount of geometric detail and complex self occlusion relationships which make it difficult to estimate the shape using stereo based methods. In this paper we introduce a new technique for recovering the structure based on the defocus blur. The principal idea that enables our work is that the defocus effect can be modeled in terms of inhomogeneous diffusion (e.g., spatially varying coefficients) of heat using the heat equation. This is because the defocus blur can be modeled as a Gaussian blur, which forms a temporally evolving kernel for the isotropic heat equation. This approach was explored by Favaro *et al.*[4]. Their method had two main shortcomings. First, it could not handle departure from Gaussian assumption in case of self-occlusions. Second, it made an assumption that the diffusion coefficient is a convex function and the solution was based on conjugate gradient based method. In this paper we address both these shortcomings. Here, we propose a model wherein the heat equation is perturbed stochastically. In this approach the departure from the Gaussian blur model is implicitly accounted for in the stochastic perturbation of diffusion. The mathematical existence for the stochastically perturbed heat equation, which is used

here, is analysed by Yip[15] and he has used it to model the dendritic growth of crystal structures. Here we adapt the model for solving the depth from defocus problem by correlating the stochastic heat equation to the defocus blurring process.

The research in depth from defocus was initially introduced by Pentland [12] in which the problem of DFD was posed as an estimation of linear space variant blur. Subsequently, there has been substantial work done using deterministic [6] and statistical techniques [3] in the spatial domain and also by solving in the frequency domain [14]. However, most of the works done assume that the observations do not suffer from self-occlusion. The handling of occlusion effects in depth computation has been addressed in [1],[2],[5]. The extent of departure from the Gaussian shape depends on the nature of depth discontinuity in the scene, which is unknown. Any imperfection in the lens aperture would also change the shape of the blur kernel. Unlike earlier methods, the proposed method can handle such an effect under a unified framework without having to estimate the departure from the assumed model. An interesting recent work has been by Hasinoff and Kutulakos [7], where the authors consider depth from focus as a pixel matching operation. However, this method requires many high resolution images.

2 Defocus as a Stochastically Perturbed Diffusion

In this section we discuss the mathematical basis of stochastic perturbation of the heat equation as a tool to analyse defocused images.

2.1 Defocus as a Diffusion Process

Consider the image formation process in a real aperture camera employing a thin lens [3]. When a point light source is in focus, all light rays that are radiated from the object point and intercepted by the lens converge at a point on the image plane. When the point is not in focus, it is imaged as a circular patch instead of a point. The point spread function (PSF) of the camera describes the image intensity caused by a single point light source. Geometric optics approximates the PSF to a circular disk. However, as discussed in the literature [3],[12] due to diffraction it will be roughly a circular blob with the brightness falling off gradually at the border rather than sharply. The resultant PSF has the general shape of a 2-D Gaussian function [3],[12]. For an equifocal plane the resultant image formed is then given by

$$I(x) = \int f(\tau)h(x, \tau)d\tau, \quad (1)$$

where we adopt x to denote the 2D space co-ordinates in an image, $f(x)$ is the focused image of the scene and h is the space-varying PSF. Here $h(x)$ is given by a circularly symmetric 2-D Gaussian function

$$h(x) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right), \quad (2)$$

where σ is a function of depth at a given point. It is quite well-known that, for a scene with constant depth the imaging model in eqn(1) can be formulated in terms of the isotropic heat equation [9] given by

$$\frac{\partial u(x;t)}{\partial t} = c(\Delta u(x;t)), \quad u(x,0) = f(x)$$

where Δu is the Laplacian operator. Here the solution $u(x, t)$ taken at a specific time $t = \tau$ plays the role of an image $I(x) = u(x, \tau)$ and $f(x)$ corresponds to the initial condition, i.e. the pin-hole equivalent observation of the scene. Note that we have used $u(x, t)$ to represent the evolution of heat everywhere in the paper. The blurring parameter σ is related to the diffusion coefficient by the following relation [4]

$$\sigma^2 = \frac{2tc}{\gamma} \quad (3)$$

where t is the time variable in the diffusion equation, c is the diffusion coefficient, and γ is a proportionality constant relating the blur radius to the spread (σ) of the blur kernel that can be determined via initial calibration. In the depth from defocus problem, the depth in the scene varies over the image and hence the constant c will actually be $c(x)$, i.e., it will vary over the image. This corresponds to a heat equation in an inhomogeneous medium.

2.2 Stochastic Perturbation

The stochastic form of the isotropic heat equation can be given by

$$du = \alpha(x, t)dW(x, t). \quad (4)$$

where $W(x, t)$ is the brownian motion of a particle ω located at position x and time t . $\alpha(x, t)$ is the diffusion coefficient. The stochastic form of the heat equation corresponds to an Ito diffusion without drift [11].

The addition of stochastic perturbation to the deterministic diffusion equation can be physically thought of as adding thermal fluctuations to the heat diffusion equation. The issues like existence and regularity of the evolution arise by such an addition. These were rigorously studied and proved by Yip[15]. They were studied in the context of crystal growth. However, the same formulation is valid for the defocus problem. The form of the stochastically perturbed diffusion or the stochastic heat equation is given by

$$du = (c(x)\Delta u)dt + \alpha(x, t)dW(t) \quad (5)$$

where $W(t)$ is a spatially correlated infinite dimensional Brownian motion, $dW(t)$ is the Ito's differential and Δu corresponds to the Laplacian of u in space. The spatial correlation of W is essential for proving the Gibbs-Thomson condition [15]. This implies that the movement of each particle is not stochastic in space but in time. The Gibbs-Thomson condition is related to the regularity and existence of the solution of eqn(5). Gibbs-Thomson relation is a function which relates the temperature and curvature values in equilibrium for the interface of evolution. Loosely speaking the Gibbs-Thomson condition essentially prescribes an equality between the variation of the energy of the interface and the total divergence of the Gibbs-Thomson relation. These are discussed in detail by Yip in his work[15] where he gives a proof of the Gibbs-Thomson condition for eqn(5).

2.3 Defocusing as a Stochastically Perturbed Diffusion

The defocusing phenomenon has a specific space varying characteristic at surface edges and occluding edges. Consider the particular case as shown in fig. 1. Here we consider the specific case of a surface edge discontinuity which results in self-occlusion. In depth from

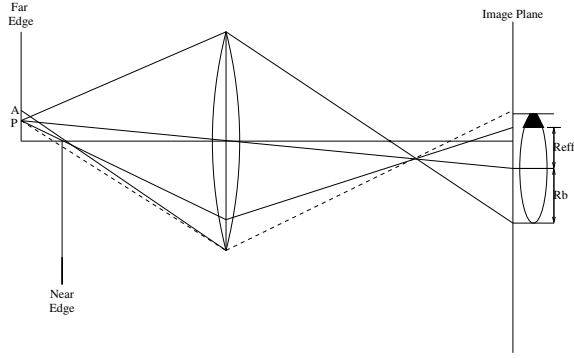


Figure 1: Illustration of the self-occlusion on account of surface discontinuity. For the point P, the point spread function (PSF) is the circular region devoid of the darkened region. For the point A in the scene, the PSF is circular as there is no self-occlusion.

defocus, self occlusion results when a continuum of rays is partially occluded and results in the blur kernel being modified [13]. This is illustrated in fig.1. Here, the rays emanating from the point P are partially blocked due to the surface discontinuity. The image plane is at a distance from the focus point and so the observation of point P results in a blur with radius R_b . However, due to the partial occlusion due to the near edge, the resultant blur instead of being circular is deformed (being R_{eff}). This artifact is present for all points in the observation from the surface edge to the point A. From point A onwards, the blur kernel is unaffected. A similar effect can be observed in the case of an occluding edge as well [1].

There have been a few approaches [1], [2], [5] where the authors have tried to address this problem by explicit modeling of this phenomenon or by adding a post-processing step. However, in our model, due to the stochastically perturbed curvature driven motion along the level sets, it is possible to incorporate this variation implicitly. This is particularly important in correctly estimating the blur kernels along discontinuities like surface edges and occluding edges. This is depicted in fig.1. As shown in the figure, along the surface edge, the contributions from the near and far surface are inhomogeneously mixed and this results in an anisotropic nature to the resultant blur kernel. So, when one does a stochastic curvature driven motion along the level sets, the blur contribution along the surface edge can be appropriately estimated. The non-uniformity of the kernel is implicitly handled in this model. There exists a similar effect when one has an occluding edge as well[1].

2.4 Evolution Equation

We now proceed to obtain an explicit evolution equation. In order to do this we first obtain an expression for the stochastic perturbation part of eqn(5). Here we consider the recent work done in stochastic level sets [8] and stochastic curvature driven motion [10]. As discussed by Yip[15], the stochastic perturbation of the eqn(5) can be seen to be given by

$$du(x,t) = \mathbf{n}(x,t)dW(t). \quad (6)$$

where $\mathbf{n}(x, t)$ is the normal to the surface interface $u(t)$ (i.e. the interface $u(x, t) \forall x$). The equivalent deterministic evolution using the level set framework for the geometric heat equation is given by the following equation

$$du(x, t) = \mathbf{n}(x, t) \kappa(x, t) \quad (7)$$

where $\kappa(x, t)$ is the mean curvature of the level set and $\mathbf{n}(x, t)$ is the normal to the level set. Here κ is given by

$$\kappa(x) = \frac{u_{x_1}^2 u_{x_2 x_2} - 2u_{x_1} u_{x_2} u_{x_1 x_2} + u_{x_1 x_1} u_{x_2}^2}{(u_{x_1}^2 + u_{x_2}^2)^{3/2}}, \quad (8)$$

where $x = x_1, x_2$ i.e. two-dimensional space and u_{x_i} refers to $\frac{\partial u}{\partial x_i}$, $i = 1, 2$. The normal $\mathbf{n}(x)$ is given by

$$\mathbf{n}(x) = \frac{\nabla u(x)}{\|\nabla u(x)\|},$$

where $\nabla u(x) = u_{x_1} + u_{x_2}$. The geometric heat equation is similar to the linear heat equation except that it diffuses orthogonal to its gradient and does not diffuse along the direction of the gradient. As a result the stochastic perturbation mainly affects the level set curves and does not affect the homogeneous regions. This is appropriate since any kernel variation for instance due to self occlusion would mainly occur along edges and would be reflected in the stochastic perturbation. The effect of the perturbation is further spread on the homogeneous regions through the deterministic diffusion component.

Now, the stochastic formulation of the above deterministic formulation according to eqn(6) could be written as

$$du(x, t) = \mathbf{n} dW(t), \quad (9)$$

The differential in eqn(9) is the *Ito differential*. This suffers from problems like it is not invariant to the parameterization of the curve, i.e., the evolution depends on the implicit representation of the initial curve and ill posedness, i.e., under certain conditions it approaches the inverse heat equation which is unstable[8],[10]. These difficulties are overcome by introducing the *Stratonovich differential*[11] given by

$$du(x, t) = \mathbf{n} \circ dW(t). \quad (10)$$

The Stratonovich form is in an implicit form and converting it to the explicit Ito form results in an added second order term. This is because of the difference in estimating Ito and Stratonovich differentials. In Ito diffusion the integration happens at the left end point whereas in the Stratonovich case the integration happens at the mid-point while evaluating the integration of the differential[11]. With a single Gaussian perturbation in space, the eqn(10) is written as

$$du(x, t) = \mathbf{n} dW(t) + \frac{1}{2} \Delta u(x, t) \left[\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right]. \quad (11)$$

The numerical implementation of the scheme for evolution is done by considering a step δt in time and δx in space and is given by[8]

$$u(x, t + \delta t) = u(x, t) + \mathbf{n} \sqrt{\delta t} \mathcal{N}_{(0,1)}(t) + \frac{1}{2} \Delta u(x, t) \left[\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right]. \quad (12)$$

where \mathcal{N} is the noise term and it denotes a standard Gaussian random variable, and the second order term is introduced because of the Stratonovich differential component. This term is a kind of smoothing term and is nothing but the degenerate diffusion component along the edges with the stochastic term corresponding to the diffusion component across the edges. Hence the complete stochastic heat equation would then be

$$u(x, t + \Delta t) = u(x, t) + \mathbf{n}\sqrt{\Delta t} \mathcal{N}_{(0,1)}(t) + \frac{1}{2}\Delta u(x, t) \left[\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right] + c(x)\Delta u(x, t).$$

Since the stochastic perturbation appropriately handles the deformation of the kernel, the diffusion coefficient c is taken to be only a single inhomogeneous coefficient value and not a diffusion tensor.

3 Depth Estimation

We consider the case when we are given two images $I_1(x)$, $I_2(x)$ with different defocus blurs. Then the resultant formulation is

$$\begin{aligned} u(x, t) &= I_1(x) \\ u(x, t + m\delta t) &= I_2(x), \end{aligned} \quad (13)$$

and where the term $u(x, t + m\delta t)$ is obtained from $u(x, t)$ by the evolution in eqn(13) and m is the number of iterations in going from image I_1 to I_2 . The evolution equation in eqn(13) blurs the image I_1 with a space-variant blur till it approximates the image I_2 closely enough which is tracked by a discrepancy measure ϕ . The blur parameter σ is related to the diffusion coefficient by the eqn(3) and the blur parameter is directly proportional to the depth in the scene[3]. In order to estimate the depth in the scene one therefore has to estimate the diffusion coefficient for the evolution equation. In a deterministic case one would obtain the following minimization problem:

$$\hat{c}(x) = \arg \min_{c(x) \geq 0} \int \int \phi(u(x, t + dt), I_2(x)) dx dt. \quad (14)$$

where $\phi(\cdot)$ is a discrepancy measure and $\hat{c}(x)$ is the diffusion coefficient for the deterministic diffusion equation. However in the stochastically perturbed case, the resultant diffusion coefficient is a combination of deterministic and stochastic diffusion. The deterministic diffusion coefficient is obtained from the contribution from the following part of the evolution equation:

$$du_{\text{det}} = (c_{\text{det}}(x)\Delta u)dt \quad (15)$$

which is the deterministic part of eqn(5). The stochastic diffusion coefficient contribution is obtained by normalizing the stochastic perturbation component in the evolution equation. We recall that the stochastic perturbation component is given by

$$\begin{aligned} du_{\text{st}} &= \mathbf{n}\sqrt{\delta t} \mathcal{N}_{(0,1)}(t) + \frac{1}{2}\Delta u(x, t) \left[\frac{\nabla u(x, t)}{|\nabla u(x, t)|} \right] \\ &= \mathbf{n} \circ dW(t) \end{aligned} \quad (16)$$

The stochastic diffusion coefficient is then given by normalizing the stochastic contribution by the corresponding deterministic evolution:

$$c_{\text{st}}(x) = \frac{\mathbf{n} \circ dW(t)}{\kappa \mathbf{n}} \quad (17)$$

where κ is the curvature and \mathbf{n} is the normal.

Thus the combined diffusion coefficient is given by

$$d(x) = c_{\text{det}}(x) + \eta c_{\text{st}}(x) \quad (18)$$

where η is the weight factor which determines the relative weight of the stochastic perturbation. The depth in the scene is obtained by solving for $d(x)$ in a minimization problem of the form

$$\hat{d}(x) = \arg \min_{d(x) \geq 0} \int \int \phi(u(x, t + dt), I_2(x)) dx dt. \quad (19)$$

We adopt a Euclidean distance measure for ϕ . Here the image $I_2(x)$ is assumed to be the more defocused image. However, that may not always be the case, and one can have sections in an image which are more in focus and other sections which are more defocused compared to the corresponding sections in the second image. In that case as an initial step the images are preprocessed and the regions which are more in focus are identified. The diffusion always happens in a forward direction to avoid instabilities that may arise due to backward diffusion. The method used to ensure this is similar the one suggested in [4]. The minimization in eqn.(19) cannot be done using conjugate gradient descent algorithm due to the stochastic perturbation. We adopt a simple simulated annealing scheme to perform the stochastic optimization. The various steps for the algorithm for depth estimation are as follows:

STEP 1: Given the initial images $I_1(x)$, and $I_2(x)$ divide them into sections such that the diffusion is always in the forward direction using the preprocessing step discussed earlier.

STEP 2: Compute u_{n+1} from u_n using the formula for du given in eqn(13).

STEP 3: Compute the discrepancy measure ϕ_n

STEP 4: Accept u_{n+1}

- if $\phi_{n+1} < \phi_n$
- otherwise, accept u_{n+1} with probability $\exp\left(\frac{-(\phi_{n+1} - \phi_n)}{T(n)}\right)$.

STEP 5: Loop back to STEP 2 till the stopping criterion is satisfied.

Here $T(n)$ is a time-dependent function that plays the same role as a decreasing temperature. Its choice is crucial. If the temperature decreases too fast the process may get stuck in a local minimum, else if it decreases slowly the convergence is delayed. Here we adopt $T(n) = T_0/\sqrt{n}$ as suggested by Juan *et al.*[8]. The stopping criterion is based on the Euclidean distance measure approaching zero.

The depth estimate is then obtained by considering the deterministic and the stochastic parts separately. For the deterministic part, we assume a constant diffusion coefficient and relate the blur to the time of evolution. The blur cannot be related directly in the stochastic part due to the non-uniform nature of evolution. Hence, in each iteration we normalize

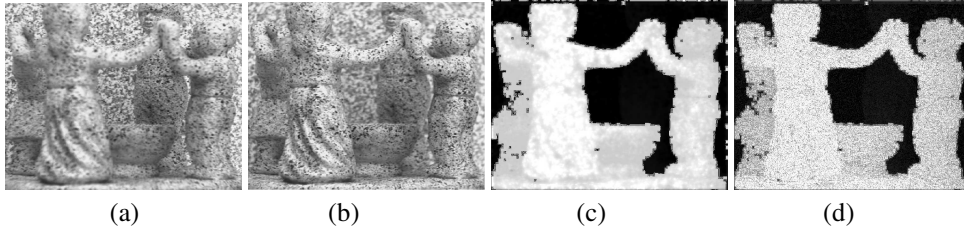


Figure 2: Here (a,b) are two real data sets showing the dolls placed at different depths (Images courtesy Favaro [4]). (c) shows the resultant depth map for the deterministic method[4]. (d) shows the corresponding result from the proposed method.

the stochastic perturbation with the corresponding orthogonal diffusion component. We then integrate the corresponding contributions over time to obtain the contribution of the stochastic perturbation to the blurring process. The final depth estimate is obtained as the joint contribution of the deterministic and stochastic components.

The depth obtained in this method has a space-varying characteristic, i.e., the problem solved is equivalent to space varying point spread function (PSF) estimation. Further due to the stochastic nature, the self occlusion effects and other imperfections are implicitly handled by the method when it does a stochastic perturbation of the blur model.

4 Experiments

The algorithm has been tested with real images and the results are compared with state of the art techniques. The method works quite well on all these test data sets.

The experimental setup shown in fig.2 is the “dolls” data set[4]. The images were taken with varying lens to image plane distances to obtain different amount of defocus in different observations. The Fig.2(c) shows the depth map estimated by the deterministic method[4] and Fig.2(d) shows the depth map obtained by the proposed method. Once again we can clearly identify the depth boundaries from the recovered depth map, justifying the usefulness of the proposed algorithm. The different dolls are clearly visible to be at different depths.

A challenging data set is the “hair” data set used in [7]. The data set is of a wig with a messy hairstyle surrounded by several artificial plants. This data poses challenging self occlusion and complex structure issues. Fig. 3(a,b) shows the 2 input images used. Fig. 3(c) shows the depth map obtained for the deterministic method [4]. As can be seen, the method does not handle the self occlusion and non-convex diffusion coefficient issues efficiently. Fig. 3(d) shows the depth map obtained from the confocal stereo method [7]. However, they have images from 13 aperture settings each with 61 focal settings. Fig. 3(e) shows the depth map obtained from the proposed method using two input images which is comparable to the depth map in [7] from many images. The results are illustrated clearly and additional results are provided at <http://vinaypn.googlepages.com/stochdfd>.

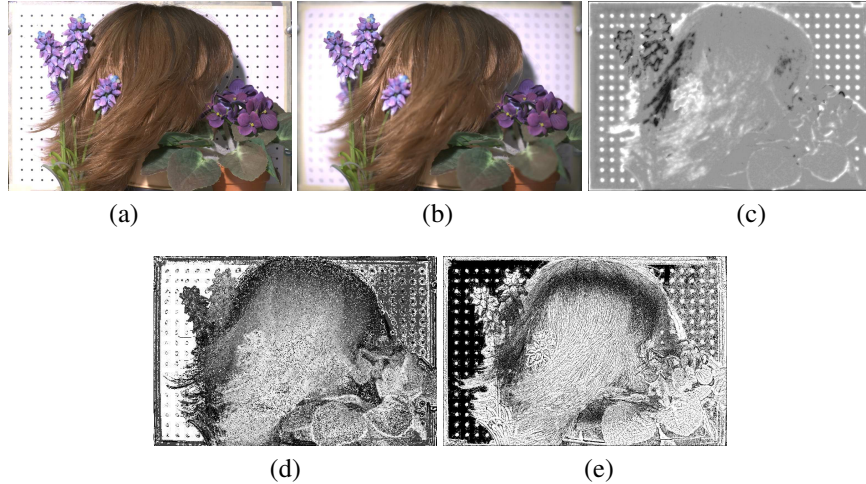


Figure 3: Here (a,b) are two real data sets showing a wig and flowers (Images courtesy [7]). (c) shows the resultant depth map for the deterministic method[4]. (d) shows the corresponding result from [7] (which uses images from 13 aperture, each with 61 focus settings) and (e) depicts the result from the proposed method (using only 2 images).

5 Conclusion

In this paper we have proposed a method based on stochastic perturbation of diffusion for solving the depth from defocus problem. The main contribution here has been in incorporating a stochastic formulation of the blur model which can effectively handle variations in the blur from the standard Gaussian blur model. The variations arise in the real world due to aberrations in the lenses and aperture and are experimentally too elaborate to measure. Further the problem of deformation of the Gaussian kernel due to self occlusion is also implicitly handled. We demonstrate that improved results can be obtained using the proposed technique. The proposed method also takes into account the non-convex nature of diffusion propagation.

It may be noted that most researchers in the area of structure recovery have pointed out the need for regularization of the recovered surface. The proposed method does not impose any such constraint while recovering the depth. Currently we are investigating the possibility of incorporating a spatial smoothness constraint during diffusion propagation to improve accuracy.

6 Acknowledgments

Financial assistantship under the *Swarnajayanti* fellowship scheme of Department of Science and Technology, India is gratefully acknowledged.

References

- [1] N. Asada, H. Fujiwara, and T. Matsuyama. Seeing behind the scene: Analysis of photometric properties of occluding edges by the reversed projection blurring model.

IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(2):155–167, February 1998.

- [2] S.S. Bhasin and S. Chaudhuri. Depth from defocus in presence of partial self occlusion. In *Proc. Eighth IEEE Int'l Conf. on Computer Vision*, volume 1, pages 488–493, July 2001. Vancouver, Canada.
- [3] S. Chaudhuri and A. N. Rajagopalan. *Depth From Defocus: A Real Aperture Imaging Approach*. Springer Verlag, New York, 1999.
- [4] P. Favaro, S. Osher, S. Soatto, and L. Vese. 3d shape from anisotropic diffusion. In *Proceedings of IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 179–186, 2003. Madison, Wisconsin, USA.
- [5] P. Favaro and S. Soatto. Seeing beyond occlusions (and other marvels of a finite lens aperture) learning shape from defocus. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR 03)*, volume 2, pages 579–586, 2003. Madison, Wisconsin, USA.
- [6] Paolo Favaro and Stefano Soatto. *3-D Shape Estimation and Image Restoration: Exploiting Defocus and Motion-Blur*. Springer-Verlag, London, 2007.
- [7] S.W. Hasinoff and K.N. Kutulakos. Confocal stereo. In *Proc. Ninth European Conference on Computer Vision*, pages 620–634, May 2006. Graz, Austria.
- [8] O. Juan, R. Keriven, and G. Postelnicu. Stochastic motion and the level set method in computer vision: Stochastic active contours. *International Journal of Computer Vision*, 69(1):7–25, August 2006.
- [9] J. J. Koenderink and A. J. van Doorn. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.
- [10] P.-L. Lions and Panagiotis E. Souganidis. Fully nonlinear stochastic partial differential equations. *C. R. Acad. Sci. Paris, t. 331, Serie I*, pages 1085–1092, 1998.
- [11] B. Oksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer Verlag, New York, sixth edition, 2004.
- [12] A. P. Pentland. A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):523–531, July 1987.
- [13] Y. Y. Schechner and N. Kiryati. Depth from defocus vs. stereo: How different really are they? *International Journal of Computer Vision*, 39(2):141–162, September 2000.
- [14] Y. Xiong and S.A. Shafer. Moment filters for high precision computation of focus and stereo. In *Proc. Intl Conf. Intelligent Robots and Systems*, pages 108–113, August 1995. Pittsburgh, Pennsylvania, USA.
- [15] N.K. Yip. Existence of dendritic crystal growth with stochastic perturbations. *Journal of Nonlinear Science*, 8:491–579, 1998.