Scene and Motion Reconstruction from Defocused and Motion-Blurred Images via Anisotropic Diffusion

Paolo Favaro[1], Martin Burger[2], and Stefano Soatto[1]

[1] Computer Science Department, UCLA, Los Angeles, CA 90095, USA,
{favaro,soatto}@cs.ucla.edu
[2] Mathematics Department, UCLA, Los Angeles, CA 90095, USA,
martinb@math.ucla.edu

Abstract. We propose a solution to the problem of inferring the depth map, radiance and motion of a scene from a collection of motion-blurred and defocused images. We model motion-blur and defocus as an anisotropic diffusion process, whose initial conditions depend on the radiance and whose diffusion tensor encodes the shape of the scene, the motion field and the optics parameters. We show that this model is well-posed and propose an efficient algorithm to infer the unknowns of the model. Inference is performed by minimizing the discrepancy between the measured blurred images and the ones synthesized via forward diffusion. Since the problem is ill-posed, we also introduce additional Tikhonov regularization terms. The resulting method is fast and robust to noise as shown by experiments with both synthetic and real data.

1 Introduction

We consider the problem of recovering the motion, depth map and radiance of a scene from a collection of defocused and motion-blurred images. Defocus is commonly encountered when using cameras with a finite aperture lens, while motion-blur is common when the imaging system is moving. To the best of our knowledge, we are the first to address the above problem. Typically, this problem is approached by considering images that are affected either by defocus or by motion-blur alone. The first case is divided into two fields of research depending on which object one wants to recover. When we are interested in recovering the radiance from defocused (and possibly downsampled) images, we are solving a *super-resolution* problem [2]. If we are interested in recovering the depth map of the scene (and possibly the radiance), then we are solving the so-called problem of *shape from defocus* [8,12,15,17,19,6].The second case corresponds to the problem of *motion deblurring*, where one is mainly interested in reconstructing the radiance, which can be thought of as the *unblurred* or *ideal* image, of a scene under the assumptions of Lambertian reflection and uniform illumination [3,4,14]. Motion deblurring is a problem of blind deconvolution [5] or blind image restoration [21], and, therefore, is related to a large body of literature [20].

T. Pajdla and J. Matas (Eds.): ECCV 2004, LNCS 3021, pp. 257–269, 2004.

1.1 Contributions of This Paper

The contribution of this paper is twofold: to link the estimation of the depth map of a scene to the recovery of the radiance and to introduce a simple and computationally efficient imaging model for images that are both defocused and motion-blurred. We model motion-blur via the depth map of the scene and the rigid motion of the camera, which requires at most 6 scalar numbers for 2 images (see section 2.2). This model avoids the artifacts of employing oversimplified motion models (e.g. each point on the image plane moves with the same constant velocity) and yields better estimates than motion models where the motion field is completely unconstrained, due to its lower dimensionality. The second contribution of this paper is the introduction of a novel model for defocused and motion-blurred images in the framework of anisotropic diffusion, in the spirit of [10]. The literature on anisotropic diffusion is quite substantial and, therefore, this work relates also to [13,18,16].

We pose the inference problem as the minimization of the discrepancy between the data and the model (i.e. the final value of the anisotropic diffusion for several different focal settings). The problem is ill-posed, it consists in finding a diffusion tensor and an unknown initial value from final values of parabolic equations. For this sake we introduce Tikhonov-type regularization, which also remedies an unwanted effect with respect to motion-blur, where a local minimum would be attained for zero motion in the absence of suitable regularization (see section 3).

2 A General Model for Defocus and Motion-Blur

2.1 An Imaging Model for Space-Varying Defocus

Images captured with a camera are measurements of energy emitted from the scene. We represent an image with a function $J : \Omega \subset \mathbb{R}^2 \mapsto [0, \infty)$, that maps pixels on the image plane to energy values. We assume that Ω is a bounded domain with piecewise smooth boundary $\partial\Omega$. The intensity of the measured energy depends on the distance of the objects in the scene from the camera and the reflectance properties of their surfaces. We describe the surfaces of the objects with a function $s : \mathbb{R}^2 \mapsto [0, \infty)$, and the reflectance with another function $r : \mathbb{R}^2 \mapsto [0, \infty)$; s assigns a depth value to each pixel coordinate and it is called *depth map*. Similarly, r assigns an energy value to each point on the depth map s and it is called, with an abuse of terminology[1], *radiance*. Furthermore, we

[1] In the context of *radiometry*, the term *radiance* refers to a more complex object that describes energy emitted along a certain direction, per solid angle, per foreshortened area and per time instant. However, in our case, since we do not change vantage point and the size of the optics and the CCD are considerably smaller than the size of the scene, each pixel will collect energy mostly from a single direction, and the change in the solid angle between different pixels is approximately negligible. Hence, a function of the position on the surface of the scene, which is the one we use, suffices to describe the variability of the radiance.

usually know lower and upper bounds $0 < s^{min} < s^{max}$ for the depth map s, which we may incorporate as an additional inequality constraint of the form

$$s^{min} \leq s(x) \leq s^{max}, \qquad \forall x \in \Omega. \tag{1}$$

The energy measured by an image J also depends on the optics of the camera. We assume the optics can be characterized by a function $h : \Omega \times \mathbb{R}^2 \mapsto [0, \infty)$, the so-called *point spread function* (PSF), so that an image J can be modeled by

$$J(y) = \int h(y, x) r(x) dx. \tag{2}$$

Although we did not write it explicitly, the PSF h depends on the surface s and the parameters of the optics (see section 3 for more details).
Under the assumption that the PSF is Gaussian and that the surface s is smooth, we can substitute the above model with a PDE whose solution $u : \mathbb{R}^2 \times [0, \infty) \mapsto \mathbb{R}$, $(x, t) \mapsto u(x, t)$, at each time t represents an image with a certain amount of blurring. In formulas, we have that $J(y) = u(y, T)$, where T is related to the amount of blurring of J. We use the following *anisotropic diffusion* equation:

$$\begin{cases} \dot{u}(y,t) = \nabla \cdot (D(y) \nabla u(y,t)) \; t > 0 \\ u(y,0) = r(y) \qquad\qquad\qquad\quad \forall y \in \Omega \\ D(y) \nabla u(y,t) \cdot n = 0 \end{cases} \tag{3}$$

where $D \doteq \begin{bmatrix} d_{11} \; d_{12} \\ d_{21} \; d_{22} \end{bmatrix}$ with $d_{ij} : \mathbb{R}^2 \mapsto \mathbb{R}$ for $i, j = 1, 2$ and $d_{12} \equiv d_{21}$, is called *diffusion tensor*. We assume that $d_{ij} \in C^1(\mathbb{R}^2)$ (i.e. the space of functions with continuous partial derivatives in $\mathbb{R}^2$) for $i, j = 1, 2$, and[2] $D(y) \geq 0 \; \forall y \in \mathbb{R}^2$. The symbol ∇ is the gradient operator $\left[\frac{\partial}{\partial y_1} \; \frac{\partial}{\partial y_2} \right]^T$ with $y = [y_1 \;\; y_2]^T$, and the symbol $\nabla \cdot$ is the divergence operator $\sum_{i=1}^{2} \frac{\partial}{\partial y_i}$. n denotes the unit vector orthogonal to $\partial\Omega$. Notice that there is a scale ambiguity between the time T and the determinant of the diffusion tensor D. We will set $T = \frac{1}{2}$ to resolve this ambiguity.

When the depth map s is a plane parallel to the image plane, the PSF h is a Gaussian with constant covariance σ^2, and it is easy to show that $2tD = \sigma^2 I_d$, where I_d is the 2×2 identity matrix. In particular, at time $t = T = \frac{1}{2}$ we have $D = \sigma^2 I_d$. This model is fairly standard and was used for instance in [10].

2.2 An Imaging Model for Motion-Blur

On the image plane we measure projections of three dimensional points in the scene. In other words, given a point $X(t) = [X_1(t) \; X_2(t) \; X_3(t)] \in \mathbb{R}^3$ at a time instant t, we measure

$$x(t) \doteq [x_1(t) \; x_2(t)]^T \doteq \left[\frac{X_1(t)}{X_3(t)} \; \frac{X_2(t)}{X_3(t)} \right]^T. \tag{4}$$

[2] Since D is a tensor, the notation $D(y) \geq 0$ means that $D(y)$ is positive semi-definite.

Using the projections of the points on the image plane $x(t)$, we can write the coordinates of a point $X(t)$ as

$$X(t) = [x(t)^T \ 1]^T \ s(x(t)). \tag{5}$$

We denote with $V = [V_1(t) \ V_2(t) \ V_3(t)]^T \in \mathbb{R}^3$ the translational velocity and with $\omega \in \mathbb{R}^3$ the rotational velocity of the scene. Then, it is well known that the time derivative of the projection x satisfies (see [11] for more details):

$$\dot{x}(t) = \frac{1}{s(x(t))} \begin{bmatrix} 1 & 0 & -x_1(t) \\ 0 & 1 & -x_2(t) \end{bmatrix} V + \begin{bmatrix} -1 - x_2^2(t) & x_1(t)x_2(t) & -x_2(t) \\ -x_1(t)x_2(t) & 1 + x_1^2(t) & x_1(t) \end{bmatrix} \omega. \tag{6}$$

We define $v \doteq \dot{x}(t)$ and call it the *velocity field*.

As we have anticipated, we restrict ourselves to a crude motion model that only represents translations parallel to the image plane, i.e.

$$v(t) = \frac{V_{1,2}(t)}{s(x(t))} \tag{7}$$

where $V_{1,2}$ is the velocity in focal length units. Now, recalling eq. (2), we have that $J(x+vt)$ denotes an image captured at time t. If the camera shutter remains open while moving the camera with velocity V for a time interval ΔT, then the image I we measure on the image plane can be written as:

$$I(x) = \frac{1}{\Delta T} \int_{-\frac{\Delta T}{2}}^{\frac{\Delta T}{2}} J(x + vt)dt \simeq \int \frac{1}{\sqrt{2\pi\gamma^2}} e^{-\frac{t^2}{2\gamma^2}} J(x + vt)dt \tag{8}$$

where γ depends on the time interval ΔT. The parameter γ can be included in the velocity vector v since there is an ambiguity between the duration of the integration time and the magnitude of the velocity. Therefore, we have

$$I(x) = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} J(x + vt)dt. \tag{9}$$

For simplicity, the above model has been derived for the case of a sideway translational motion, but it is straightforward to extend it to the general case of eq. (6).

2.3 Modeling Motion-Blur and Defocus Simultaneously

In this section, we consider images where defocus and motion-blur occur simultaneously. In the presence of motion, a defocused image J measured at time t can be expressed as

$$J(y + vt) = \int h(y + vt, x)r(x)dx. \tag{10}$$

Following eq. (9), we obtain

$$I(y) = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int \frac{1}{2\pi\sigma^2} e^{-\frac{(y-x+vt)^T(y-x+vt)}{2\sigma^2}} r(x)dxdt. \tag{11}$$

If we now interchange the integration order, we can write the previous equation in a more compact way as

$$I(y) = \int \frac{1}{2\pi|C|^{\frac{1}{2}}} e^{-\frac{(y-x)^T C^{-1}(y-x)}{2}} r(x)dx \tag{12}$$

where $C = \sigma^2 I_d + vv^T$.
Eq. (12) is also the solution of the anisotropic diffusion PDE (3) with initial condition the radiance r and diffusion tensor $D = \frac{C}{2t}$. Hence, a model for defocused and motion-blurred images is the following:

$$\begin{cases} \dot{u}(y,t) = \nabla \cdot (D\nabla u(y,t)) & t > 0 \\ u(y,0) = r(y) & \forall y \in \Omega \\ D\nabla u(y,t) \cdot n = 0 \end{cases} \tag{13}$$

where at time $t = T = \frac{1}{2}$, $D = C = \sigma^2 I_d + vv^T$. Now, it is straightforward to extend the model to the space-varying case, and have that

$$D(y) = \sigma^2(y) I_d + v(y)v(y)^T. \tag{14}$$

In particular, when eq. (7) is satisfied, we have

$$D(y) = \sigma^2(y) I_d + \frac{V_{1,2}V_{1,2}^T}{s^2(y)}. \tag{15}$$

Notice that the diffusion tensor just defined is made of two terms: $\sigma^2(y)I_d$ and $\frac{V_{1,2}V_{1,2}^T}{s^2(y)}$. The first term corresponds to the isotropic component of the tensor, and captures defocus. The second term corresponds to the anisotropic component of the tensor, and it captures motion-blur. Furthermore, since both of the terms are guaranteed to be always positive semi-definite, the tensor eq. (15) is positive semi-definite too. We will use eq. (13) together with eq. (15) as our imaging model in all subsequent sections.

2.4 Well-Posedness of the Diffusion Model

A first step in the mathematical analysis is to verify the well-definedness of the parameter-to-output map $(r, s, V_{1,2}) \mapsto u(., T)$, which corresponds to a well-posedness result for the degenerate parabolic initial-boundary value problems

$$\begin{cases} \dot{u}(y,t) = \nabla \cdot (D(y)\nabla u(y,t)) & t > 0 \\ u(y,0) = r(y) \\ D(y)\nabla u(y,t) \cdot n = 0 \end{cases} \tag{16}$$

for diffusion tensors of the form $D(y) = \sigma(y)^2 I_d + \frac{V_{1,2}V_{1,2}^T}{s(x)^2}$. n denotes the unit vector orthogonal to the boundary of Ω. The following theorem guarantees the existence of weak solutions for the direct problem:

Theorem 1. *Let $r \in L^2(\Omega)$ and $s \in H^1(\Omega)$ satisfies* (1)*. Then, there exists a unique weak solution $u \in C(0,T;L^2(\Omega))$ of* (16)*, satisfying*

$$\int_0^T \int_\Omega \lambda(y) |\nabla u(y,s)|^2 \, dy \, ds \leq \int_\Omega r(y)^2 \, dy, \tag{17}$$

where $\lambda(y) \geq 0$ denotes the minimal eigenvalue of $D(y)$.

Proof. See technical report [9].

3 Estimating Radiance, Depth, and Motion

In section 2.1 we introduced the variance σ^2 of the PSF h to model defocus. The variance σ^2 depends on the depth map via $\sigma^2(x) = \left(\frac{d}{2}\right)^2 \left(1 - p\left(\frac{1}{F} - \frac{1}{s(x)}\right)\right)^2$, where d is the aperture of the camera (in pixel units), p is the distance between the image plane and the lens plane, F is the focal length of the lens and s is the depth map of the scene. We simultaneously collect a number N of defocused and motion-blurred images $\{I_1, \ldots, I_N\}$ by changing the parameter $p = \{p_1, \ldots, p_N\}$. Notice that the parameters p_i lead to different variances $\sigma_i^2(x)$, which affect the isotropic component of the diffusion tensor D, but not its anisotropic component $\frac{V_{1,2}V_{1,2}^T}{s^2(x)}$. As shown in section 2.3 we can represent an image I_i by taking the solution u_i of eq. (13) at time $t = T = 1/2$ with a diffusion tensor $D_i(x) = \sigma_i^2(x) I_d + \frac{V_{1,2}V_{1,2}^T}{s^2(x)}$, and with initial condition $u_i(y,0) = r(y) \quad \forall i = 1 \ldots N$.

We pose the problem of inferring the radiance r, the depth map s and the motion field v of the scene by minimizing the following least-squares functional with Tikhonov regularization (cf. [7])

$$\hat{r}, \hat{s}, \hat{V}_{1,2} = \arg \min_{r,s,V_{1,2}} \sum_{i=1}^{N} \int_\Omega \left(u_i(x,T) - I_i(x)\right)^2 dx + \alpha \left\|r - r^*\right\|^2 + \beta \left\|\nabla s\right\|^2 + \\ + \gamma \left(\left\|V_{1,2}\right\| - M\right)^2, \tag{18}$$

where α, β, and γ are positive regularization parameters, r^* is a prior[3] for r and M is a suitable positive number[4]. One can choose the norm $\|\cdot\|$ depending on the desired space of solutions. We choose the L_2 norm for the radiance and the components of the gradient of the depth map and the ℓ_2 norm for the velocity vector $V_{1,2}$. In this functional, the first term takes into account the discrepancy between the model and the measurements; the second and third term are classical regularization functionals, imposing some regularity on the estimated depth map

[3] We do not have a preferred prior for the radiance r. However, it is necessary to introduce this term to guarantee that the estimated radiance does not diverge. In practice, one can use as a prior r^* one of the input images, or a combination of them, and choose a very small α.

[4] Intuitively, the constant M is related to the maximum degree of motion-blur that we are willing to tolerate in the input data.

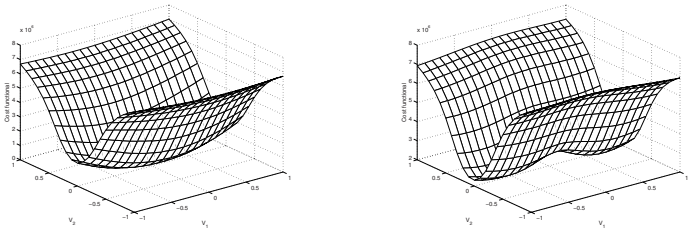

Fig. 1. Left: cost functional for various values of V_1 and V_2 when $\gamma = 0$ or $M = 0$. Right: cost functional for various values of V_1 and V_2 when $\gamma \neq 0$ and $M \neq 0$. In both cases the cost functional eq. (18) is computed for a radiance $\hat{r}$ and a depth map $\hat{s}$ away from the true radiance r and the true depth map s. Notice that on the right plot there are two symmetric minima for $V_{1,2}$. This is always the case unless the true velocity satisfies $V_{1,2} = 0$, since the true $V_{1,2}$ can be determined only up to the sign.

and penalizing large deviations of the radiance from the prior. The last term is of rather unusual form, its main objective being to exclude $V_{1,2} = 0$ as a stationary point. One easily checks that for $\gamma = 0$ or $M = 0$, $V_{1,2} = 0$ is always a stationary point of the functional in (18), which is of course an undesirable effect. This stationary point is removed for positive values of M and γ (see Figure 1).

3.1 Cost Functional Minimization

To minimize the cost functional (18) we employ a gradient descent flow. For each unknown we compute a sequence converging to a local minimum of the cost functional, i.e. we have sequences $\hat{r}(x,\tau), \hat{s}(x,\tau), \hat{V}_{1,2}(\tau)$, such that $\hat{r}(x) = \lim_{\tau \mapsto \infty} \hat{r}(x,\tau)$, $\hat{s}(x) = \lim_{\tau \mapsto \infty} \hat{s}(x,\tau)$, $\hat{V}_{1,2} = \lim_{\tau \mapsto \infty} \hat{V}_{1,2}(\tau)$. At each iteration we update the unknowns by moving in the opposite direction of the gradient of the cost functional with respect to the unknowns. In other words, we let $\partial\hat{r}(x,\tau)/\partial\tau \doteq -\nabla_{\hat{r}} E(x)$, $\partial\hat{s}(x,\tau)/\partial\tau \doteq -\nabla_{\hat{s}} E(x)$, $\partial\hat{V}_{1,2}(\tau)/\partial\tau \doteq -\nabla_{\hat{V}_{1,2}} E(x)$. It can be shown that the above iterations decrease the cost functional as τ increases. The computation of the above gradients is rather involved, but yields the following formulas, that can be easily implemented numerically:

$$\begin{aligned}
\nabla_r E &= \sum_{i=1}^{N} w_i(x,0) \\
\nabla_s E &= 2\sum_{i=1}^{N} \int_0^T \left(\sigma_i(x) \frac{p_i}{s^2(x)} I_d + \frac{V_{1,2} V_{1,2}^T}{s^3(x)} \right) \nabla u_i(x,t) \cdot \nabla w_i(x,t)\, dt \\
\nabla_{V_{1,2}} E &= -\sum_{i=1}^{N} \int_0^T \int_\Omega \left(\frac{V'_{1,2} V_{1,2}^T + V_{1,2} V'^T_{1,2}}{s^2(x)} \nabla u_i(x,t) \cdot \nabla w_i(x,t) \right) dx\, dt
\end{aligned} \tag{19}$$

where w_i satisfies the following adjoint parabolic equation (see [9] for more details):

$$\begin{cases} \dot{w}_i(y,t) = -\nabla \cdot (D_i(y)\nabla w_i(y,t)) \\ w_i(y,T) = u_i(y,T) - I_i(y) \\ (D_i(y)\nabla w_i(y,t)) \cdot n = 0. \end{cases} \tag{20}$$

4 Experiments

The algorithm presented in section 3.1 is tested on both synthetic (section 4.1) and real (section 4.2) data. In the first case, we compare the estimated unknowns with the ground truth and establish the performance of the algorithm for different amounts of noise. In the second case, since we do not have the ground truth, we only present a qualitative analysis of the results. We implement the gradient flow equations in section 3.1 with standard finite difference schemes (see [18,1]).

4.1 Synthetic Data

In this first set of experiments, we consider a scene made of a slanted plane (see the leftmost image in Figure 4), that has one side at $0.52m$ from the camera and the opposite side at $0.85m$ from the camera. The slanted plane is painted with a random texture. We define the radiance r to be the image measured on the image plane when a pinhole lens is used (see first image from the left in Figure 2). The second image from the left in Figure 2 has been captured when the scene or the camera are subject to a translational motion while the camera shutter remains open. Notice that the top portion of the image is subject to a more severe motion-blur than the bottom part. This is due to the fact that in this case points that are far from the camera (bottom portion of the image) move at a slower speed than points that are close to the camera (top portion of the image).

We simulate a camera that has focal length $0.012m$ and F-number 2. With these settings we capture two images: one by focusing at $0.52m$, and the other by focusing at $0.85m$. If neither the camera nor the scene are moving, we capture the two rightmost images shown in Figure 2. Instead, if either the camera or the scene are moving sideway, we capture the two leftmost images shown in Figure 3. The latter two are the images we give in input to our algorithm. In Figure 3 we show the recovered radiance when no motion-blur is taken into account (third image from the left) and when motion-blur is taken into account (rightmost image). As one can notice by visual inspection, the latter estimate of the radiance is sharper than the estimate of the radiance when motion-blur is not modeled. The improvement in the estimation of the radiance can also be evaluated quantitatively since we have ground truth. To measure the accuracy of the estimated radiance, we compute the following normalized RMS error:

$$NRMSE(\phi_{estimated}, \phi_{true}) = \frac{\|\phi_{estimated} - \phi_{true}\|}{\|\phi_{true}\|} \tag{21}$$

where $\phi_{estimated}$ is the estimated unknown, ϕ_{true} is the ground truth and $\|\cdot\|$ denotes the L_2 norm. We obtain that the NRMSE between the true radiance

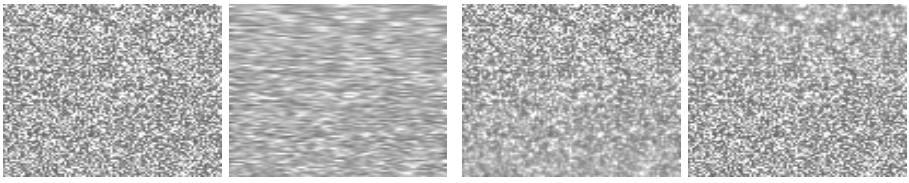

Fig. 2. First from the left: synthetically generated radiance. Second from the left: motion-blurred radiance. This image has been obtained by motion-blurring the synthetic radiance on the left. Third and fourth from the left: defocused images from a scene made of the synthetic radiance in Figure 2 (leftmost) and depth map in Figure 4 (leftmost) without motion-blur.

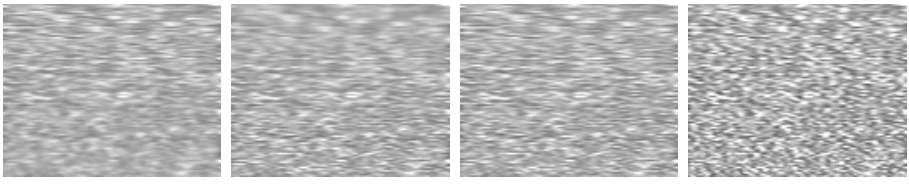

Fig. 3. First and second from the left: defocused and motion-blurred images from a scene made of the synthetic radiance in Figure 2 (leftmost) and depth map in Figure 4 (leftmost). Third from the left: recovered radiance from the two defocused and motion-blurred images on the left when no motion-blur is taken into account ($V_{1,2} = 0$). Fourth from the left: recovered radiance from the two defocused and motion-blurred images on the left when motion blur is taken into account ($V_{1,2} \neq 0$).

Fig. 4. Left: true depth map of the scene. Middle: recovered depth map. Right: profile of the recovered depth map. As can be noticed, the recovered depth map is very close to the true depth map with the exception of the top and bottom sides. This is due to the higher blurring that the images are subject to at these locations.

and the motion-blurred radiance (second image from the left in Figure 2) is 0.2636. When we compensate only for defocus during the reconstruction, the NRMSE between the true radiance and the recovered radiance is 0.2642. As expected, since the motion-blurred radiance is the best estimate possible when we do not compensate for motion-blur, this estimated radiance cannot be more accurate than the motion-blurred radiance. Instead, when we compensate for both defocus and motion-blur, the NRMSE between the true radiance and the recovered radiance is 0.2321. This shows that the outlined algorithm can restore

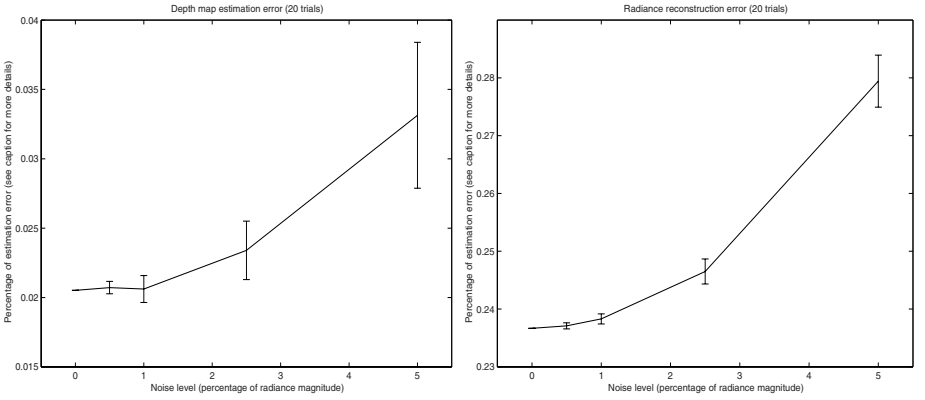

Fig. 5. Left: depth map estimation for 5 levels of additive Gaussian noise. The plot shows the error bar for 20 trials with a staircase depth map and a random radiance. We compute the RMS error between the estimated depth and the true depth, and normalize it with respect to the norm of the true depth (see eq. (21)). Right: radiance estimation for 5 levels of additive Gaussian noise. As in the previous error bar, we compute the RMS error between the true radiance and the reconstructed radiance and then normalize it with respect to the norm of the true radiance.

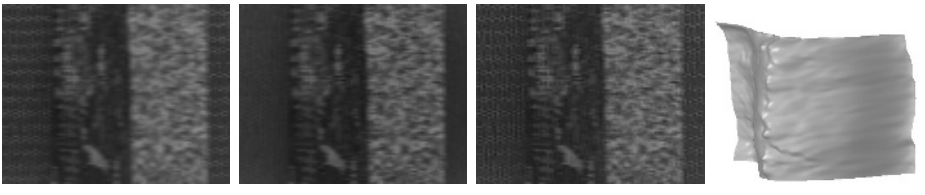

Fig. 6. First and second from the left: input images of the first data set. The two images are both defocused and motion-blurred. Motion-blur is caused by a sideway motion of the camera. Third from the left: recovered radiance. Fourth from the left: recovered depth map.

images that are not only defocused, but also motion-blurred. The recovered depth map is shown in Figure 4 on the two rightmost images together with the ground truth for direct comparison (left). The true motion is $V_{1,2} = [0.8\ 0]^T$ and the recovered motion is $[0.8079\ \ -0.0713]^T$ in focal length units.

To test the performance and the robustness of the algorithm, we synthetically generate defocused and motion-blurred images with additional Gaussian noise. We use a scene made of a staircase depth map with 20 steps, with the first step at $0.52m$ from the camera and the last step at $0.85m$ from the camera. As in the previous experiment, we capture two images: one by focusing at $0.52m$ and the other by focusing at $0.85m$. To each of the images we add the following 5 different amounts of Gaussian noise: 0%, 0.5%, 1%, 2.5% and 5% of the radiance magnitude. For each noise level we run 20 experiments from which we compute the mean and the standard deviation of the NRMSE. The results are shown in Figure 5.

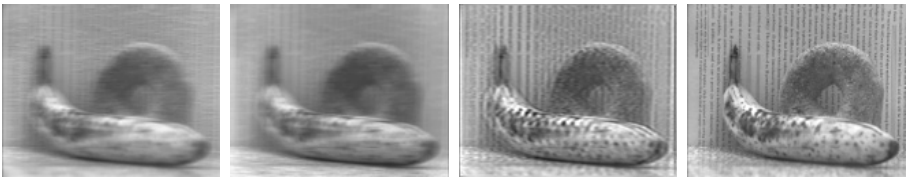

Fig. 7. First and second from the left: input images of the second data set. The two images are both defocused and motion-blurred. Motion-blur is caused by a sideway motion of the camera. Third from the left: recovered radiance. Fourth from the left: an image taken without motion-blur.

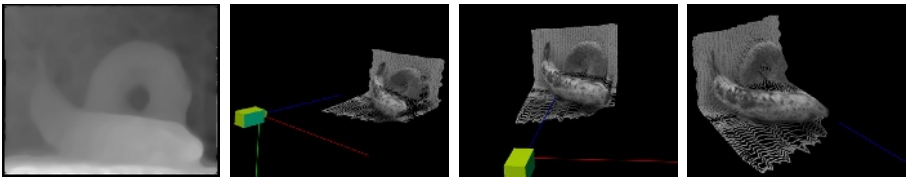

Fig. 8. First from the left: estimated depth map visualized as a gray level intensity image. Second, third and fourth from the left: visualization of the estimated depth map from different viewing angles. The depth map is also texture mapped with the estimated radiance.

4.2 Real Images

We test the algorithm on two data sets. The first data set is made of the two real images shown in Figure 6. The scene is made of a box that is moving sideway. We simultaneously capture two images with a multifocal camera kindly lent to us by S. K. Nayar. The camera has an AF NIKKOR 35mm Nikon lens, with F-number 2.8. We capture the first image by focusing at 70mm from the camera and the second image by focusing at 90mm from the camera. The scene lies entirely between 70mm and 90mm. The estimated radiance is shown in Figure 6, together with the recovered depth map. The estimated motion is $V_{1,2} = [0.5603\ 0.0101]^T$ in units of focal length. In the second data set we use the two defocused and motion-blurred images in Figure 7 (first and second image from the left) captured with the same camera settings as in the first data set. The scene is composed of a banana and a bagel and the scene is moving sideways. The estimated radiance is shown in the third image from the left of the same figure. To visually compare the quality of the estimated radiance, we also add the fourth image from the left in Figure 7. This image has been obtained from about the same viewing point when neither the camera nor the scene was moving. Hence, this image is only subject to defocus. The reconstructed depth map is shown in Figure 8. The first image from the left is the depth map visualized as a gray level image. Light intensities correspond to points that are close to the camera and dark intensities correspond to points that are far from the camera. The next three images are visualizations of the depth map from different viewing angles with the estimated radiance texture mapped onto it. The estimated velocity for this data set is $V_{1,2} = [0.9639\ \ -0.0572]^T$, that corresponds to a sideway motion.

5 Summary and Conclusions

In this manuscript we proposed a solution to the problem of inferring the depth, radiance and motion of a scene from a collection of motion-blurred and defocused images. First, we presented a novel model that can take into account both defocus and motion-blur (assuming motion is pure sideway translation), and showed that it is well-posed. Motion-blurred and defocused images are represented as the solution of an anisotropic diffusion equation, whose initial conditions are defined by the radiance and whose diffusion tensor encodes the shape of the scene, the motion field and the optics parameters. Then, we proposed an efficient algorithm to infer the unknowns of the model. The algorithm is based on minimizing the discrepancy between the measured blurred images and the ones synthesized via diffusion. Since the inverse problem is ill-posed, we also introduce additional Tikhonov regularization terms. The resulting method is fast and robust to noise as shown in the experimental section.

Acknowledgements. This research has been supported by funds AFOSR F49620-03-1-0095, ONR N00014-03-1-0850, NIH PRE-NDEBC, NSF RHA 115-0208197.

References

1. G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing.* Springer-Verlag, 2002.
2. Simon Baker and Takeo Kanade. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1167–1183, September 2002.
3. B. Bascle, A. Blake, and A. Zisserman. Motion deblurring and super-resolution from an image sequence. In *European Conference on Computer Vision*, volume 2, pages 573–582, 1996.
4. M. Ben-Ezra and S.K. Nayar. Motion deblurring using hybrid imaging. In *Computer Vision and Pattern Recognition*, volume 1, pages 657–664, 2003.
5. M. Bertero and P. Boccacci. Introduction to inverse problems in imaging. *Institute of Physics Publishing, Bristol and Philadelphia*, 1998.
6. S. Chaudhuri and A. Rajagopalan. *Depth from defocus: a real aperture imaging approach.* Springer Verlag, 1999.
7. H. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publishers, Dordrecht, 1996.
8. J. Ens and P. Lawrence. An investigation of methods for determining depth from focus. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15:97–108, 1993.
9. P. Favaro, M. Burger, and S. Soatto. Scene and motion reconstruction from defocused and motion-blurred images via anisotropic diffusion. In *UCLA - Math. Dept. Technical Report (cam03-63)*, November 2003.
10. P. Favaro, S. Osher, S. Soatto, and L. Vese. 3d shape from anisotropic diffusion. In *Intl. Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 179–186, 2003.

11. D.J. Heeger and A.D. Jepson. Subspace methods for recovering rigid motion i. *Int. J. of Computer Vision*, 7(2):95–117, 1992.
12. A. Pentland. A new sense for depth of field. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9:523–531, 1987.
13. P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–39, 1990.
14. A. Rav-Acha and S. Peleg. Restoration of multiple images with motion blur in different directions. In *IEEE Workshop on Applications of Computer Vision (WACV)*, Palm-Springs, 2000.
15. M. Subbarao and G. Surya. Depth from defocus: a spatial domain approach. *Intl. J. of Computer Vision*, 13:271–294, 1994.
16. D. Tschumperle and R. Deriche. Vector-valued image regularization with pde's: A common framework for different applications. In *CVPR03*, pages I: 651–656, 2003.
17. M. Watanabe and S. Nayar. Rational filters for passive depth from defocus. *Intl. J. of Comp. Vision*, 27(3):203–225, 1998.
18. J. Weickert. *Anisotropic Diffusion in Image Processing*. B.G.Teubner Stuttgart, 1998.
19. Y. Xiong and S. Shafer. Depth from focusing and defocusing. In *Proc. of the Intl. Conf. of Comp. Vision and Pat. Recogn.*, pages 68–73, 1993.
20. Y. Yitzhaky, R. Milberg, S. Yohaev, and N. S. Kopeika. Comparison of direct blind deconvolution methods for motion-blurred images. In *Applied Optics-IP*, volume 38, pages 4325–32, July 1999.
21. Y. You and M. Kaveh. Blind image restoration by anisotropic diffusion. *IEEE Trans. on Image Processing*, 8(3):396–407, 1999.