3D Shape and Motion Analysis from Image Blur and Smear: A Unified Approach

Y. F. Wang
Department of Computer Science
University of California
Santa Barbara, CA 93106

Ping Liang
College of Engineering
University of California
Riverside, CA 92521-0425

Abstract

This paper addresses 3D shape recovery and motion estimation using a realistic camera model with an aperture and a shutter. The **spatial blur** *and* **temporal smear** *effects induced by the camera's finite aperture and shutter speed are used for inferring both the* **shape** *and* **motion** *of the imaged objects.*

1 Introduction

In this paper, we address the problem of 3D shape recovery and motion estimation using a realistic camera model. Even though the pin-hole camera model has been prevalent in computer vision and graphics for its simplicity and mathematical tractability, this model hides certain shape cues that can be very powerful in 3D analysis. For example, if a more realistic camera model is used—which takes into consideration the aperture and shutter speed settings of the camera—the depth of field and motion blur effects become significant in the image formation. Depth of field induces *spatial blur*, where the amount of image blur is effected by the imaged object's apparent depth. Motion blur induces *temporal smear*, the effect of which is affected by the speed of the imaged object.

Traditional shape and motion analysis utilizing a pin-hole camera model does not consider the image blur and smear effects induced by a finite aperture and/or shutter speed. However, these effects often times cannot be ignored. Unless the imaged object is at a known distance to be sharply focused or the sensor is truly a pin-hole camera, *spatial blur* due to a finite aperture and depth of field will occur. Furthermore, unless the shutter speed is such that during the time the shutter opens the scene object's motion is negligible, *temporal smear* due to object motion need be considered. These degradation effects do exist in practical applications. For examples, often times in navigation an obstacle's location (depth) may not be known in advance (hence the object may not be focused correctly). A camera may not be equipped with a motorized lens with computer-controlled zoom or focus rings, hence, adaptation to objects at different depths might not be possible. The illumination level might be low, which requires extended exposure. These effects should be considered in designing robust depth inference and motion estimation algorithms.

Algorithms that utilize image blur for image analysis have recently been proposed. Most notable are the "shape-from-focus" and "shape-from-defocus" approaches [2, 4, 5, 6] for 3D shape recovery. In "shape-from-defocus," a single camera, or a pair of cameras with a specially-designed beam-splitter placed in front of them [5] used to produce a pair of identical images except for the aperture size and therefore the depth of field. The amount of image blur in the two images can be shown to be a simple function of one variable: the distance between the viewer and the imaged point. Hence, to estimate the depth of an object, we need only compare the corresponding points in the two images and measure the change in the image blur.

Image motion estimation can be roughly classified into gradient-based, feature-matching, and spatial-temporal filtering approaches. Recently, a "motion-from-smear" framework [1] was proposed as another alternative. [1] studies motion estimation from the *smear effect in an image sequence*, but does not consider possible image blur due to defocusing. The proposed algorithm unifies these two approaches and takes into account both *finite-aperture spatial blurring* and *finite-shutter-speed temporal smear* effects for shape and motion computation.

It is nontrivial to extend shape and motion analysis to allow both spatial blur and temporal smear. As shown later in Sec. 2, *indistinguishable visual effects can be produced non-incidentally by an object which is out of focus but otherwise stationary, or by a moving object which is in focus, or by objects with different combinations of the degree of out-of-focusness*

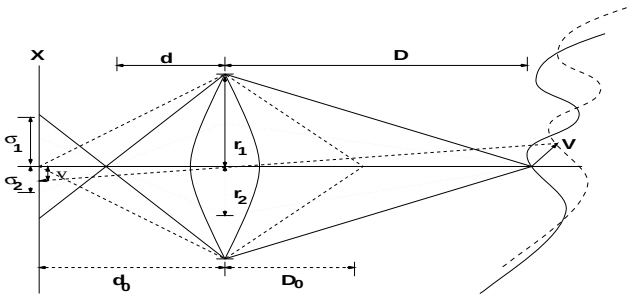

Figure 1: Image configuration with spatial blur and temporal smear

and speed of motion in between. Hence, image interpretation in the presence of spatial blur and temporal smear can be highly ambiguous. This ambiguity in a sense is similar to the "aperture problem" well known in the optical flow analysis [3]. However, the problem is much more severe in a general shape and motion analysis allowing blur and smear, as *it may not be possible to uniquely identify the source of the image blur even at image locations with a multitude of gradient directions (e.g., a corner).* Hence, it is important to isolate the image blur induced by a finite aperture and shutter speed to enable shape and motion analysis, respectively.

2 Technical Description

Fig. 1 depicts the image configuration where blur due to finite aperture size and smear due to non-zero exposure time are shown. D_0 and d_0 denote the object and image distances that achieve a perfect focus, assuming that all lens aberration effects are ignored. For objects located at a distance D other than D_0, image blur is induced by a spatial "spillage" of pixel values into adjacent pixels, and the amount of image blur is affected by the lens aperture and the object's distance. Furthermore, we assume that an imaged object can undergo a general motion, denoted by V in Fig. 1. Object motion performs a temporal "integration" of pixel values along the trajectory when the shutter opens and induces image smear as a result.

The problem we address in this paper is this: Given that the camera is modeled as a finite aperture sensor instead of a pin-hole model, and that the shutter may be open long enough for the imaged object's motion to be registered (e.g., in a low light scenario), the recorded images are subject to both spatial blur and temporal smear degradation. How can we reliably estimate the shape and motion of the imaged object from both image blur and smear? In the next section, we will first review the "shape-from-focus" approach for shape recovery using static, finite-aperture images. Then we will discuss the difficulty in analyzing both blurred and smeared images, and present our formulation for deducing shape and motion from such images.

2.1 Shape-from-Defocus using Static, Finite-Aperture Images

To simplify the analysis, we assume that the point spread function of the optical system is approximated by a Gaussian function of standard deviation σ.[1] As shown in Fig. 1, two images, taken of the same object but with different aperture opening r_1 and r_2 will show different degrees of blurring denoted by σ_1 and σ_2. It can be shown that the amount of image blur is related to the object's perceived distance D by [5]

$$D = \frac{Fd_0}{d_0 - F - k\sigma f} , \qquad (1)$$

where F is the focal length and f the f-number of the lens. k is a proportionality constant which can be determined through camera calibration. The blurred image, $g(x,y)$, can be represented as the convolution of an un-blurred image, $f(x,y)$ (taken using a pin-hole camera), and a Gaussian kernel of variance σ:

$$g(x,y) = \int_{x'}\int_{y'} f(x-x', y-y')\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x'^2+y'^2}{2\sigma^2}} dx'dy' .$$

Fourier transform of the above equation is

$$G(\omega_x,\omega_y) = \frac{1}{\sqrt{2\pi}} F(\omega_x,\omega_y) e^{-\frac{\sigma^2(\omega_x^2+\omega_y^2)}{2}} .$$

If two images are taken of the same scene using different aperture settings of σ_1 and σ_2, then the ratio of their Fourier coefficients reveals the relative blurring:

$$\frac{G_1(\omega_x,\omega_y)}{G_2(\omega_x,\omega_y)} = \frac{e^{-\frac{\sigma_1^2(\omega_x^2+\omega_y^2)}{2}}}{e^{-\frac{\sigma_2^2(\omega_x^2+\omega_y^2)}{2}}} , \text{or}$$

$$ln\frac{G_1(\omega_x,\omega_y)}{G_2(\omega_x,\omega_y)} = \frac{1}{2}(\sigma_2^2 - \sigma_1^2)(\omega_x^2+\omega_y^2) , \qquad (2)$$

which provides one constraint on σ_1 and σ_2. Furthermore, observe that the two aperture settings are constrained by the perceived depth as shown in Eq. 1, hence, we have:

$$ln\frac{G_1(\omega_x,\omega_y)}{G_2(\omega_x,\omega_y)} = \frac{1}{2k^2}((\frac{Fd_0}{f_2 D} - \frac{d_0 - F}{f_2})^2 - (\frac{Fd_0}{f_1 D} - \frac{d_0 - F}{f_1})^2)(\omega_x^2+\omega_y^2) .$$

If the camera parameters are known, the only unknown in the above equation is D. After D is solved for, we can recover σ_1 and σ_2 using Eq. 1. In some implementations [5], one of the cameras is assumed to be a pin-hole camera and produces images with a negligible blur (or $\sigma = 0$). Eq. 2 can then be used to derive the other σ.

[1] Researchers have used different approximations to the PSF, such as Gaussian [5] and pillbox functions [6].

2.2 Shape and Motion using Blurred and Smeared Images

A more complicated scenario is when the shutter is open long enough for object motion to be registered. The images thus produced suffer from both *spatial blur* and *temporal smear* degradation. We establish the following fact:

Proposition: *For a single blurred and smeared image, the same visual image blur effects can be non-incidentally produced by a multitude of combinations of finite-aperture blur and motion smear, with different degrees of out-of-focusness and object motions. This is true for image areas with a single gradient direction (e.g., an edge) and multiple gradient directions (e.g., a corner).*

The proposition leads to the following observations:

• It is not beneficial to analyze a single blurred and smeared image by itself. Multiple images, or multiple sequences of images will be needed.

• The aperture problem in optical flow [3] is exacerbated by the presence of blur and smear. As the presence of multiple gradient directions (e.g., a corner) in a small neighborhood theoretically allows the recovery of local motion using the local flow analysis, which is not the case for blurred and smeared images.

• Mechanisms probably will be needed to isolate the image blur effects induced by a finite aperture and shutter speed. This will lead to a more robust estimation of shape (based on the finite-aperture blur effect) and motion (based on the nonzero-exposure-time smear effect).

To illustrate, we will assume that the effect of object motion, if any, over a small image neighborhood and a short time period can be characterized by a constant, nonzero flow velocity denoted by (u, v). Now consider two extreme scenarios: degradation in one comes entirely from finite aperture blur (i.e., no object motion) and in the other entirely from nonzero exposure time smear (i.e., a pin-hole camera is used with infinite depth-of-field). Denote the images in these two cases as g_b (blur only) and g_s (smear only). Then these two images can be related to the un-blurred and un-smeared image (f) by:

$$\begin{aligned} g_b(x,y) &= \int_x' \int_y' f(x-x', y-y')\mathcal{B}(x',y')dx'dy' \quad \text{and} \\ g_s(x,y) &= \int_t f(x-ut, y-vt)\mathcal{S}(t)dt \end{aligned} \tag{3}$$

where $\mathcal{B}$ and $\mathcal{S}$ denote the blurring and smear mechanisms. The proposition above suggests that g_b and g_s can be indistinguishable even at places with a high information content (e.g., edges and corners). We will illustrate first for the case where the the image neighborhood contains a single step edge.

Without loss of generality, we will assume the step edge aligns with the y-axis, or

$$f(x,y) = \begin{cases} a & x < 0 \\ b & x \geq 0 \end{cases} .$$

Then we can simplify the expressions of g_s and g_b into

$$g_b(x) = \int_x f(x-x')\bar{\mathcal{B}}(x')dx', \quad \text{and}$$

$$g_s(x) = \int_t f(x-ut)\mathcal{S}(t)dt \quad \text{where}$$

$$\bar{\mathcal{B}}(x') = \int_y \mathcal{B}(x',y')dy' .$$

We consider the following $\mathcal{B}$ and $\mathcal{S}$ functions: $\mathcal{B}$ can be either a 2D Gaussian (e.g., in [5]) or a pillbox function (e.g., in [6]), and $\mathcal{S}$ can either a box (a fast shutter with negligible open and close times) or a trapezoid function (a slow shutter with nonzero rising and trailing edges). Their expressions are

	$\mathcal{B}$ or $\mathcal{S}$	$\mathcal{B}$
Gaussian	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$	$e^{-\frac{x^2}{2\sigma^2}}$
pillbox	$\frac{1}{\pi r^2}\Pi(\frac{\sqrt{x^2+y^2}}{r})$	$\frac{2}{\pi r^2}\sqrt{r^2 - x^2}$
box	$\Pi(\frac{t}{T})$	
trapezoid	$\Lambda(\frac{t}{R}) + \Pi(\frac{t-R}{T-R-F}) + \Lambda(-\frac{t-T}{F})$	

where σ denotes the variance of the Gaussian function, r denotes the radius of the pillbox function, T is the length of time the shutter opens, and R and F are the rise and fall times in a trapezoid function, respectively. Π is the unit-length rectangular function (1 between 0 and 1 and 0 otherwise) and Λ is the unit-length ramp function ($\Lambda = \int \Pi dt$).

Sample 1D Gaussian, pillbox, box, and trapezoid functions are plotted in Fig. 2(a) and edge profiles result from Gaussian or pillbox blurring, and from box or trapezoid smear are shown in Fig. 2(b). As can been seen that the four edge profiles are extremely alike. The reason for the similarity is that both out-of-focus blur and motion smear induce an "accumulation" of neighboring pixel values. The accumulation results from either a spatial "spillage" by blurring or a temporal "integration" by smear. The degradation filters have very similar shapes as shown in Fig 2(a) and produce similar accumulation effects.

Even in a neighborhood with multiple gradients, it can still be impossible to make such an assertion. The reason is that consider, for example, a 90^o corner with edges aligned with the coordinate axes. A 2D motion

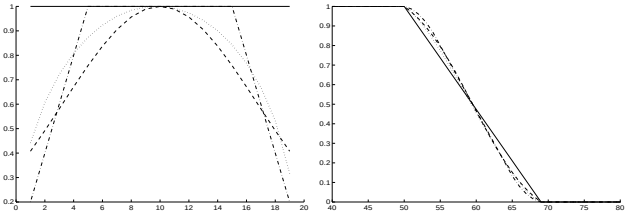

Figure 2: (a) Different image blur and smear filters, and (b) blurred and smeared 1D edge profiles. Solid: box, dashdot: trapezoid, dot: pillbox, and dashed: Gaussian.

in a 45^o direction in between the x and y axes will induce the same smear effect along both the x and y edges, which in turn can be attributed to out-of-focus blur. Furthermore, many combinations of out-of-focus blur and motion smear along the bisecting direction of the corner can explain the phenomenon equally well. Hence, unlike in an optical flow analysis, a reliable estimate of motion is not possible even in the presence of multiple gradient directions in a neighborhood.

To enable a robust shape and motion estimation in the presence of both blur and smear, it is important that the degradation effects induced by out-of-focus blur and motion smear be identified and isolated. We present such a framework below.

2.3 Mathematical Formulation

Combining the two expressions in Eq. 3 taken into consideration both blur and smear, a blurred and smeared image (g) is related to the un-blurred and un-smeared one (f) by:

$$g(x,y) = \int_T^{T+\Delta t} \int_{x'} \int_{y'} f(x-ut-x', y-vt-y') \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x'^2+y'^2}{2\sigma^2}} dx'dy'dt ,$$

where (u,v) denote an object pixel's image velocity and Δt denote the length of time the shutter opens. Fourier transform of the above equation is

$$\begin{aligned}
&G(\omega_x,\omega_y) \\
&= \int_x \int_y \{\int_T^{T+\Delta t} \int_{x'} \int_{y'} f(x-ut-x', y-vt-y') \\
&\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x'^2+y'^2}{2\sigma^2}} dx'dy'dt\} e^{-i(\omega_x x+\omega_y y)} dxdy \\
&= \int_T^{T+\Delta t} \int_{x'} \int_{y'} \int_x \int_y f(x-ut-x', y-vt-y') e^{-i(\omega_x x+\omega_y y)} dxdy \\
&\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x'^2+y'^2}{2\sigma^2}} dx'dy'dt \\
&= F(\omega_x,\omega_y) \int_T^{T+\Delta t} \int_{x'} \int_{y'} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x'^2+y'^2}{2\sigma^2}} e^{-i(\omega_x x'+\omega_y y')} dx'dy' \\
&e^{-i(\omega_x u+\omega_y v)t} dt \\
&= F(\omega_x,\omega_y) \frac{1}{\sqrt{2\pi}} e^{-\frac{\sigma^2(\omega_x^2+\omega_y^2)}{2}} \int_T^{T+\Delta t} e^{-i(\omega_x u+\omega_y v)t} dt \\
&= F(\omega_x,\omega_y) \frac{1}{\sqrt{2\pi}} e^{-\frac{\sigma^2(\omega_x^2+\omega_y^2)}{2}} e^{-i(\omega_x u+\omega_y v)(T+\frac{\Delta t}{2})} \\
&sinc((\omega_x u+\omega_y v)\frac{\Delta t}{2})\Delta t .
\end{aligned} \tag{4}$$

As shown in the above equation, two degradation mechanisms due to both spatial blur and temporal smear are present. Suppose that two image sequences are recorded of the same scene using two different aperture settings. Then we can identify four different combinations: (I) images taken over the same time interval with the same aperture setting, (II) images taken over the same time interval with different aperture settings, (III) images taken over different time intervals with the same aperture setting, and (IV) images taken over different time intervals with different aperture settings. (I) produces only a single image and is not very interesting. (II) has been shown to facilitate 3D shape inference. It will be shown below that (III) and (IV) can be used to recover object motion from blurred and smeared images.

Taking the ratio of the Fourier transform of two images taken with (possibly) different apertures over (possibly) different times using Eq. 4, we have

$$\begin{aligned}
&\frac{G_1(\omega_x,\omega_y)}{G_2(\omega_x,\omega_y)} = \\
&\frac{e^{-\frac{\sigma_1^2(\omega_x^2+\omega_y^2)}{2}} e^{-i(\omega_x u+\omega_y v)(T_1+\frac{\Delta t_1}{2})} sinc((\omega_x u+\omega_y v)\frac{\Delta t_1}{2})\Delta t_1}{e^{-\frac{\sigma_2^2(\omega_x^2+\omega_y^2)}{2}} e^{-i(\omega_x u+\omega_y v)(T_2+\frac{\Delta t_2}{2})} sinc((\omega_x u+\omega_y v)\frac{\Delta t_2}{2})\Delta t_2} .
\end{aligned} \tag{5}$$

For scenario (iii) above, if two images are taken with the same aperture setting but over different time periods then the ratio is simplified to

$$\frac{G_1(\omega_x,\omega_y)}{G_2(\omega_x,\omega_y)} = \frac{e^{-i(\omega_x u+\omega_y v)(T_1+\frac{\Delta t_1}{2})} sinc((\omega_x u+\omega_y v)\frac{\Delta t_1}{2})}{e^{-i(\omega_x u+\omega_y v)(T_2+\frac{\Delta t_2}{2})} sinc((\omega_x u+\omega_y v)\frac{\Delta t_2}{2})} .$$

Assume that the shutter speed is the same and the image velocity stays constant over adjacent image frames, we have:

$$ln\frac{G_1(\omega_x,\omega_y)}{G_2(\omega_x,\omega_y)} = i(\omega_x u+\omega_y v)(T_2-T_1) .$$

The blurring effect is canceled out and (u,v) can be recovered from the above equation.

For the scenario (iv) above, observe from Eq. 4 that blurring affects only the magnitude of the system impulse response, while smear affects both the magnitude and the phase. Hence, if only the phase component is considered, again with the same shutter speed Δt, the ratio in Eq. 5 is simplified and different image blur σ is eliminated, or

$$phase\{\frac{G_1(\omega_x,\omega_y)}{G_2(\omega_x,\omega_y)}\} = (\omega_x u+\omega_y v)(T_2-T_1) .$$

In either case, (u,v) can be isolated and estimated from the image sequence.

2.4 Theoretical Analysis

If an image is subject to both spatial blur and temporal smear, how accurate can we recover the shape

and motion of the imaged objects? As will be shown in this section that these degradation mechanisms have a detrimental effect in the shape and motion estimation. Even though their effects can be isolated as shown in the previous subsection, in reality, the degradation effects can not be completely eliminated even with a suitable windowing operation. Our analysis is for a 1D image using the algorithm outlined above.

Consider an ideal (not yet blurred or smeared) 1D signal $f(x)$ of length n. To generate smeared signals, assume that the motion vector is u pixel/sec and the shutter opens for Δt second. Two 1D images $g_1(x)$ and $g_2(x)$ are taken at time 0 and $T \geq \Delta t$, then

$$\begin{aligned} g_1(x) &= \int_0^{\Delta t} f(x-ut)dt \\ g_2(x) &= \int_T^{T+\Delta t} f(x-ut)dt \ . \end{aligned}$$

It is easily shown that

$$\begin{aligned} g_2(x+uT) &= \int_T^{T+\Delta t} f(x+uT-ut)dt \\ &= \int_0^{\Delta t} f(x+uT-u(t'+T))dt' \ \ t'=t-T \\ &= \int_0^{\Delta t} f(x-ut')dt' = g_1(x) \ . \end{aligned}$$

Denote uT as s. Without loss of generality, assume that $u \geq 0$ and $s \geq 0$, then we have:

$$g_2(x) = \begin{cases} \phi(x) & 0 \leq x < s \\ g_1(x-s) & s \leq x < n \end{cases} ,$$

where $\phi(x)$ denotes the (extraneous) signal included in $g_2(x)$ through motion. We can rewrite the above equation as

$$g_2(x) = g_1((x-s)\%n) + \varphi(x) \ ,$$

where

$$\varphi(x) = \begin{cases} \phi(x) - g_1((x-s)\%n) & 0 \leq x < s \\ 0 & s \leq x < n \end{cases} ,$$

and % denotes the modular operator. Then taking the Fourier transform of the above equation, we get:

$$G_2(\omega) = G_1(\omega)e^{-i\omega s} + \Phi(\omega) \ ,$$

where $\Phi(\omega) = \mathcal{F}(\varphi(x))$ is the Fourier transform, and

$$\frac{G_2(\omega)}{G_1(\omega)} = \frac{G_1(\omega)e^{-i\omega s} + \Phi(\omega)}{G_1(\omega)} = e^{-i\omega s} + \frac{\Phi(\omega)}{G_1(\omega)} \text{ or,}$$

$$\begin{aligned} ln(\frac{G_2(\omega)}{G_1(\omega)}) &= -i\omega s + ln(1 + \frac{\Phi(\omega)}{G_1(\omega)}e^{i\omega s}) \\ &\approx -i\omega s + \frac{\Phi(\omega)}{G_1(\omega)}e^{i\omega s} \ . \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbf{imag}(-\frac{1}{\omega}ln(\frac{G_2(\omega)}{G_1(\omega)})) &\approx s - \mathbf{imag}(\frac{1}{\omega}\frac{\Phi(\omega)}{G_1(\omega)}e^{i\omega s}) \\ |\mathbf{imag}(-\frac{1}{\omega}ln(\frac{G_2(\omega)}{G_1(\omega)})) - s| &\leq |\frac{1}{\omega}||\frac{\Phi(\omega)}{G_1(\omega)}| \ . \end{aligned}$$

Hence, the accuracy of the analysis depends on the relative strength of the signal and the extraneous signal included through shift. Intuitively speaking, if the extraneous signal is short relative to the signal itself (or the amount of movement is small comparing to the length of the processing window), the energy packed in the extraneous signal is much smaller than that of the signal and a reliable motion estimation is possible. By Parseval's theorem we have

$$\sum g_1^2(x) = \frac{1}{2\pi}\sum G_1^2(\omega) \ .$$

Neglecting the constant term $\frac{1}{2\pi}$, we have

$$|G_1(\omega)| = \sqrt{G_1^2(\omega)} \approx \sqrt{\frac{1}{n}\sum G_1^2(\omega)} = \sqrt{\frac{1}{n}\sum g_1^2(x)} \approx \sqrt{E(g_1^2)} \ .$$

Similarly,

$$|\Phi(\omega)| \approx \sqrt{\frac{s}{n}E(\varphi^2)} \ .$$

We will assume that the signal is ergodic, under the ergodic assumption, we have

$$R_n(\tau, x) = \sum_0^{n-1} g_1(x)g_1(x+\tau) \approx R_n(\tau, x')$$

for all τ, x, and x'. If the ergodic condition holds for sample length of s or longer, it is easily shown that

$$R_s(\tau, x) = \frac{s}{n}R_n(\tau, x') \ .$$

Setting $\tau = 0$, we have:

$$|G_1(\omega)| \approx \sqrt{\frac{1}{n}R_n(0,x)} = \sqrt{\frac{1}{s}R_s(0,x)} = \sqrt{\frac{n}{s}}|\Phi(\omega)| \ , \text{ or}$$

$$\frac{|\Phi(\omega)|}{|G_1(\omega)|} = \sqrt{\frac{s}{n}} \ .$$

Accuracy analysis with both blur and smear is similar and will not be repeated. Hence, to ensure that the

motion estimation is accurate, we need to ensure that $s << n$. Furthermore, the actual energy distribution of $\Phi(\omega)$ and $G_1(\omega)$ on different ω can not be predicted in advance. It is possible that the extraneous signal φ packs all its energy in a few frequency components where the strength of the real signal f is low and hence $\frac{|\Phi(\omega)|}{|G_1(\omega)|} >> 1$. This means that s may not be reliably estimated at all frequency components. The "outlier" frequency components (where $\frac{|\Phi(\omega)|}{|G_1(\omega)|} >> 1$) may totally invalid the estimation process if a simple average scheme is used to obtain the motion parameters by averaging over all frequency components. Our solution is to (1) use Gaussian pre-filtering to suppress the strength of the extraneous signal at the edges of a signal, and (2) employ a median type weighting of all frequency components, which is much more reliable in getting rid of the outliers in a few frequency components and arrives at a faithful estimation.

3 Experimental procedures

Since "shape-from-blurring" formulation has been extensively studied in the literature, our experiments instead concentrated on the "motion-from-smear" formulation. To verify the correctness of the proposed motion estimation algorithms, experiments using both synthetic and real images were conducted. Due to the page limitation, we will present here only the results on real blurred and smeared images.

Our experimental setup comprised a stationary camera viewing objects on a mobile platform. The objects were shifted to the right and downward slightly at each position, and two images with two different aperture stops ($f/1.8$ and $f/2.8$) were taken at each position. Three adjacent images from a single aperture stop were added to simulate the smearing effect.

Figs. 3(a) and (b) show the two-image sequence with $f/1.8$ aperture setting while Figs. 3(c) and (d) the two-image sequence with $f/2.8$ setting. Four types of combinations in motion computation are possible: Using the pair of images from the $f/1.8$ sequence, using the pair of images from the $f/2.8$ sequence, using the first image from the $f/1.8$ sequence and the second image from the $f/2.8$ sequence, and using the first image from the $f/2.8$ sequence and the second image from the $f/1.8$ sequence. Because no ground truth was available, we instead computed the histogram and verified the consistency. The combined results from the four analysis are displayed in Fig. 4.

As seen from these figures, motion estimation in the regions occupied by the 3M box and the cylinder was fairly consistent and correctly predicted the right and downward motion. The motion estimation in the upper right corner, which imaged a far-away

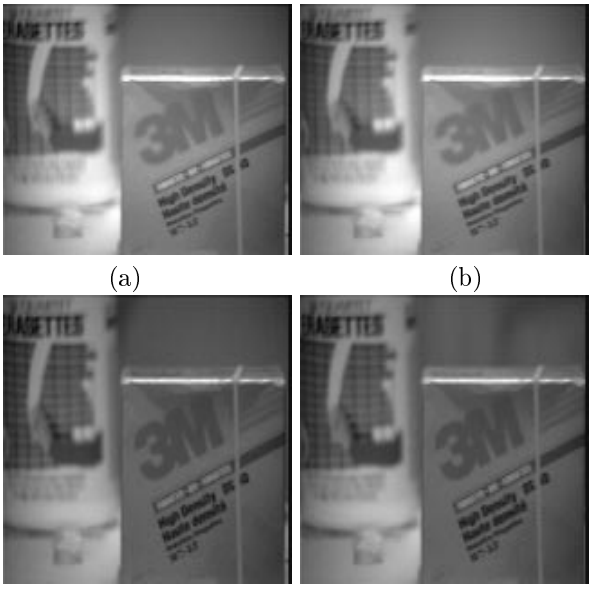

Figure 3: Real spatially-blurred and temporally-smeared image sequences used in the analysis

wall, was less consistent. The inconsistency was probably caused by that the wall lacked any discernible features and was at a large distance that its motion relative to the camera was small. Motion estimation could easily be fooled by background noise and varying lighting condition.

4 The Concluding Remarks

In this paper, we broaden the scope of "shape-from-blurring" to allow the camera to have not only a finite aperture but also a non-zero exposure time. Hence both spatial blur and temporal smear were used for inferring the shape and motion of objects.

References

[1] W. G. Chen, N. Nandhakumar, and W. Martin. Image Motion Estimation from Motion Smear — A New Computational Approach. *IEEE Trans. Pattern Analy. Machine Intell.*, 18(4):412–425, 1996.

[2] B. K. P. Horn. Focusing. Technical Report 160, MIT, 1968.

[3] B. K. P. Horn. *Robot Vision*. The MIT Press, Cambridge, MA, 1986.

[4] S. K. Nayar and Y. Nakagawa. Shape from Focus. *IEEE Trans. Pattern Analy. Machine Intell.*, 16(8):824–831, 1994.

[5] A. Pentland. A New Sense for Depth of Field. *IEEE Trans. Pattern Analy. Machine Intell.*, 9(4):523–531, 1987.

[6] M. Watanabe and S. K. Nayar. Telecentric Optics for Computational Vision. In *Proc. European Conf. Comput. Vision*, Apr. 1996.

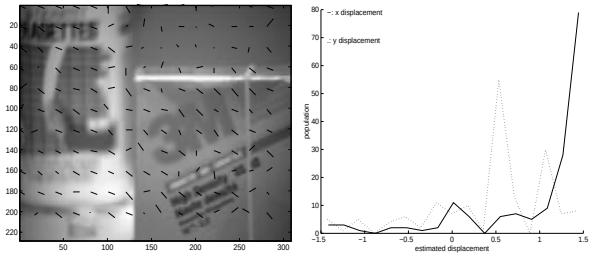

Figure 4: Final results by combining those from four different estimation procedures