

Stochastic Calculus Analysis of Optical Time-of-Flight Range Imaging and Estimation of Radial Motion

LEE STREETER^{1,*}

¹School of Engineering, University of Waikato, Hamilton 3240, New Zealand

*Corresponding author: streeter@waikato.ac.nz

Compiled November 30, 2017

Time-of-flight range imaging is analysed using stochastic calculus. Through a series of interpretations and simplifications, the stochastic model leads to two methods for estimating linear radial velocity: maximum likelihood estimation on the transition probability distribution between measurements; and a new method based on analysing the measured correlation waveform and its first derivative. The methods are tested in a simulated motion experiment from (-40) — $(+40)$ m/s, with data from a camera imaging an object on a translation stage. In tests maximum likelihood is slow and unreliable, but when it works it estimates the linear velocity with standard deviation of 1 m/s or better. In comparison the new method is fast and reliable but works in a reduced velocity range of (-20) — $(+20)$ m/s with standard deviation ranging from 3.5 m/s to 10 m/s. © 2017 Optical Society of America

OCIS codes: (280.7250) Velocimetry; (280.3640) Lidar; (110.6880) Three-dimensional image acquisition.

<http://dx.doi.org/10.1364/ao.XX.XXXXXX>

INTRODUCTION

Time-of-flight (ToF) range imaging [1] is a form of full field optical lidar for measuring, at each pixel, the distance to the imaged point in the scene. ToF range imaging is an active multiframe imaging method. The light source and sensor shutter are modulated in homodyne, and the phase offset between the backscattered light return and the sensor is linearly proportional to the ToF of the light. A second electronic phase shift is added between the light source and sensor modulations which is stepped in sequence between frame integrations, effectively sampling the correlation waveform. The phase of the correlation waveform then is directly determined by the ToF of the light. The phase of the correlation waveform is computed, from which the ToF, and subsequently the distance, is computed. If objects move during the data capture process, then the ToF changes which spoils the data. Thus, the fact that ToF requires multiple frames acquired over time places a restriction on the application of ToF range imaging to static scenes. Motion causes a critical error to ToF range imaging, one which is not trivially solved, requiring a complete reevaluation of how the data is acquired and processed. The motion information is inherently encoded by the ToF range imaging data acquisition modality, and the topic of this work is to analyse the effect of motion find new efficient methods for motion estimation.

Early investigations into motion in ToF range imaging examined the detection of motion and use of a dynamic algorithm that reduced the number of frames [2] when motion was detected.

These ideas were later extended to a detection and restoration method that reconstructs the distance based on frames that are automatically deemed not corrupted by motion [3]. Modelling the effect of radial motion on the measured correlation waveform lead to a method for reducing motion artefacts, and application of optical flow on the amplitude of the correlation waveform was used to facilitate correction of motion error [4]. Range flow, the application of optical flow to range data, was applied to ToF range image distance measurements [5]. Range flow was subsequently used as the foundation for motion segmentation from ToF range data [6]. In these works motion was either treated as something to be mitigated, or estimated using computer vision techniques that assumed that the effect of motion on the range image data is small to begin with.

More recent work broached motion in ToF range imaging as something to be measured. Coded exposure using specialist camera hardware, in conjunction with optical flow, was used to estimate and correct motion to a subframe level [7]. Continuous wave modulation without phase stepping encoded radial velocity as the frequency of the correlation waveform [8]. Concurrent multi-camera ToF ranging alleviates interframe effects, and measurement of motion follows by operating one camera in homodyne mode and the other in heterodyne mode [9, 10]. Development of the theory of heterodyne ToF ranging revealed that the integration process intrinsic to the camera operation exposed motion as a coefficient. Data from the second camera running in homodyne with the light source normalised for other factors, such as brightness and distance, leading to an estimate

of radial motion.

Motion in ToF ranging can be viewed as a manifestation of the Doppler effect. The classical method for estimating instantaneous frequency is to use the Wigner distribution, e.g. [11, 12], a practical method but one that requires a comparatively large number of measurements to what is typically used in ToF ranging. Stochastic calculus [13, 14] has been applied to the estimation of the instantaneous frequency of pure sinusoids [15, 16] using statistical bootstrapping, but under restrictive assumptions on the amplitude and the nature of the noise.

In this paper we undertake a first foray into rethinking the problem of motion in ToF range imaging. We do so by applying stochastic calculus to amplitude modulated continuous wave ToF range imaging to derive a general model of the measurement of noisy ToF data. The model is then used to solve the problem of radial velocity estimation from a single camera operating in homodyne mode. The emphasis is on practical methods for velocity estimation that are efficient in terms of computational power. Hence multiple approximations and simplifications are made in the development to lead to a simple and practical solution. In Section 2 we briefly review ToF range imaging and stochastic calculus, and then apply stochastic calculus to the ToF ranging equations to derive a general model. Then, in Section 3 we make a series of simplifying assumptions and apply parameter inference techniques to derive methods for estimating linear radial motion from ToF ranging data. In Section 4 experimental tests are described and in 5 the results are presented and discussed.

OPTICAL RANGE IMAGING AND STOCHASTIC CALCULUS

Time-of-Flight Range Imaging

ToF range imaging is an active multiframe method of imaging lidar. A ToF range imaging camera measures, at each pixel, the time taken by light to travel to the scene, backscatter from it and return to the camera. The ToF of light is measured indirectly by periodically amplitude modulating the light source at frequency f , typically at rates of tens up to the low hundreds of megahertz. The travel distance d , hence travel time, induces a phase shift, ϕ in the light return,

$$\phi = \frac{4\pi fd}{c}, \quad (1)$$

where c is the speed of light in air or a vacuum. The sensor is similarly homodyne modulated with the light source, which then demodulates the returning light signal, and the demodulated light return is integrated by the pixel.

There are three unknowns in ToF range imaging, the phase, the brightness of the return α , and the background light β . Setting the phase between the light source and the sensor to θ_n , the total integrated intensity I_n is

$$I_n = \alpha \cos(\phi + \theta_n) + \beta. \quad (2)$$

Thus, the total integrated intensity is dependent on the ToF of the light. Stepping θ_n to several discrete phase steps, we acquire a set of phase step images. Typically the phase steps, θ_n , are evenly spaced over a span of 2π rad, i.e. $\theta_n = n\Delta\theta$, where $\Delta\theta = 2\pi/N$ and $n = 0, \dots, N-1$. The profile of I_n over n at each pixel is called the correlation wavefunction, and the phase, ϕ , is found from I_n by Fourier analysis.

The effect of linear radial motion is to stretch or compress the correlation wavefunction. This effect is represented one of two

ways. The first, more physically accurate, is to represent ϕ as a linear function in time from initial phase ϕ_0 and with rate of change $\Delta\phi$, viz $\phi(t) = \phi_0 + n\Delta\phi$. The second, more pragmatic, representation is to multiply θ_n by parameter k ,

$$I_n = \alpha \cos(\phi_0 + k\theta_n) + \beta. \quad (3)$$

The two representations are equivalent, and describe a change in frequency of the correlation waveform. Conversion between the two is straightforward: equating the arguments of the cosine for each parameterisation, $\phi_0 + n\Delta\phi + \theta_n = \phi_0 + k\theta_n$, we see that $k\theta_n = n\Delta\phi + \theta_n$. Consequently $k - 1 = n\Delta\phi/\theta_n$, a fact that will be useful later.

Stochastic Calculus

Stochastic calculus is concerned with the analysis of noisy processes that vary in time [13, 14]. Stochastic calculus is an extension of the familiar deterministic calculus to continuous functions of random variables, where random variables are useful for modelling noise. The rules provided by stochastic calculus give meaningful understanding of integration and differentiation of functions that contain variables that are random in nature. It is the differentiation rules that are of particular interest in this work, through which we will understand the effect that radial motion has on the optical ToF measurements, and subsequently estimate k from analysis of the I_n in Eq. (3). Furthermore, as the name suggests, stochastic calculus leads to statistical modelling, specifically finding probability distribution functions describing the transition of time series data from one state to the next. These statistical models form powerful tools for determining interesting model parameters from data such as k .

We recall the two types of functions that stochastic calculus operates on [13, 14]. The first is the class of deterministic functions that classical calculus operates on. In this work the first and second derivatives are assumed to exist and be continuous. The second class of functions are the so called Wiener processes $W(t)$, which are useful for describing noise. A Wiener process is random in time with mean 0 and the differences $W(s) - W(t)$, $s > t$, have variance $s - t$. The derivative $dW(t)$ does not exist in the classical sense, instead it is understood to be a process such that samples from $dW(t)$ are identified with samples from a Gaussian distributed random variable with variance 1. From these facts it should be apparent that generating a sequence of Gaussian pseudo-random data of variance 1 in the computer, and then taking the cumulative sum, simulates a Wiener process.

It must be emphasised that a stochastic differential, such as $dW(t)$, is shorthand for the integration required to find $W(t)$. Bearing in mind this shorthand representation, an important result arises from products of the differentials dt and $dW(t)$, namely $(dW(t))^2 = dt$, $dt^2 = 0$, $dW(t)dt = 0$, and all higher order powers are zero (see e.g. [13] for a more complete justification). This is the so called box calculus, a set of rules which together with the differential form shorthand representation is useful for simplifying functions of Wiener processes.

In general, a stochastic process is one that is modeled by a stochastic differential equation (SDE), that contains a differential of a random process. A time dependent stochastic process $X(t)$ has the form [13]

$$dX(t) = \mu(X(t), t) dt + \sigma(X(t), t) dW(t), \quad (4)$$

where the functions μ and σ represent, respectively, the drift and diffusion coefficients of X . For simplicity, in what follows we will suppress the notational dependence on t , unless required for

clarity. Like the differentials above, a SDE should be interpreted as shorthand for the integral required to find X , given some initial condition $X(0)$. If μ and σ do not depend on X then dX is called an Itô process, and we can take the integral directly, if not then some substitution must be made to remove the explicit dependence via Itô's theorem, below.

Central to modern stochastic calculus is a class of stochastic processes called martingales [13, 14]. A martingale is a stochastic process for which the conditionally expected value of future values of $X(t)$ is the most recently observed one. So, given a series of observations of a martingale, X_0, \dots, X_n , the expected value of the next observation,

$$\langle X_{n+1} | X_0, \dots, X_n \rangle = X_n, \quad (5)$$

where $\langle \cdot \rangle$ denotes expectation. Accordingly, a supermartingale (submartingale) has nonincreasing (nondecreasing) conditional expectation,

$$\langle X_{n+1} | X_0, \dots, X_n \rangle \leq X_n, \quad (\langle X_{n+1} | X_0, \dots, X_n \rangle \geq X_n). \quad (6)$$

The relevance of martingales to this work is in the statistical method of martingale estimation for finding parameters [17, 18], the principles of which we later leverage to estimate radial motion.

Differentiation in deterministic calculus has an important extension in stochastic calculus known as Itô's theorem [13, 14], which can be intuitively understood as the chain rule for functions of stochastic variables. However, it is more properly understood as the application of Taylor's series expansion to functions of variables perturbed by random fluctuations. Consider a multivariate function $f(x)$ of vector x of length J . The Taylor series expansion is, explicitly to second order,

$$f(x + \Delta x) = f(x) + \sum_{i=1}^J \frac{\partial f}{\partial x_i} \Big|_{x_i} \Delta x_i + \frac{1}{2} \sum_{j=1}^J \sum_{i=1}^J \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_{x_i, x_j} \Delta x_i \Delta x_j + \text{HOT}, \quad (7)$$

where HOT stands for higher order terms. Now, from the box calculus

$$dX^2 = (\mu dt + \sigma dW)^2 = (\mu dt)^2 + (\sigma dW)^2 + 2\mu\sigma dt dW = \sigma^2 dt. \quad (8)$$

Likewise, for two different processes X_i and X_j with covariance $\Sigma_{i,j}^2$ we have

$$dX_i dX_j = \Sigma_{i,j}^2 dt. \quad (9)$$

Substituting X_i for the x_i into the Taylor series and taking the limit $\Delta X_i \rightarrow dX_i$ one arrives at the main result of Itô's theorem

$$dF = \sum_{i=1}^J \frac{\partial f}{\partial x_i} \Big|_{x_i} dX_i + \frac{1}{2} \sum_{j=1}^J \sum_{i=1}^J \Sigma_{i,j}^2 \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_{x_i, x_j} dt. \quad (10)$$

The result of applying Itô's theorem to a function of a stochastic variables results in a SDE, and therefore is shorthand for the function $F(X)$ which is the Itô integral of Eq. (10) (see [13] for more detail).

Itô's theorem is an elegant result that tidily summarises functions of stochastic processes, such as those that describe optical modulation. Therefore, it is useful for analysis of ToF range imaging in the presence of noise, as we do below.

A Stochastic Description of Time-of-Flight Ranging

Noise enters the ToF ranging measurement process through jitter [19, 20], the statistical nature of light and light detection, thermal noise in the electronics, and discretisation error in the conversion of analog signals to digital data. In short, every stage of the measurement process is touched by the influence of random effects. Analysis of ToF ranging for change in the parameters in time therefore requires careful consideration of the influence of the noise. Stochastic calculus is the natural tool for this analysis, opening the door to the use of a rich and mature body of analysis tools [17, 18].

The parameters of the ToF ranging process are henceforth enumerated in the vector $x = (\phi, \alpha, \theta, \beta)$, where we assume time dependence on each element of x . Let $\Sigma = (\sigma_\phi, \sigma_\alpha, \sigma_\theta, \sigma_\beta)$, and assume zero covariance between the parameters. The stochastic processes that govern the ToF ranging process are assumed to follow the drift diffusion model (so called Itô processes [13])

$$dX_i = \frac{dx_i}{dt} dt + \sigma_i dW_i, \quad i = 1 \dots 3, \quad (11)$$

where the W_i are time-dependent Wiener processes. The X_i are each Wiener process with deterministic drift parameters, thus each are Gaussian random variables with mean x_i and variance $t\sigma^2$ [13].

The noise associated with β is dominated by electronic thermal noise, with contribution from discretisation noise, so the use of a Wiener process is a close approximation. For the other parameters, the use of Wiener processes is an approximation based on an appeal to the central limit theorem. Such approximation is easily justified for α (related to the light intensity) at large values due to the Poisson nature of photon statistics [21], but breaks down for low light levels. The distribution for random jitter is unknown but is commonly modelled as Gaussian [22].

The stochastic observation model is of F where (c.f Eq. (2))

$$f(x) = \alpha \cos(\phi + \theta) + \beta \Rightarrow F(X) = X_2 \cos(X_1 + X_3) + X_4. \quad (12)$$

Itô's theorem for multivariate functions and the zero covariance between variables yields the SDE

$$dF = \cos(X_1 + X_3) \frac{d\alpha}{dt} dt - X_2 \sin(X_1 + X_3) \left[\frac{d\phi}{dt} + \frac{d\theta}{dt} \right] dt - X_2 \cos(X_1 + X_3) \left[\sigma_\theta^2 + \sigma_\phi^2 \right] dt + \frac{d\beta}{dt} dt - X_2 \sin(X_1 + X_3) \left[\sigma_\phi dW_1 + \sigma_\theta dW_3 \right] + \cos(X_1 + X_3) \sigma_\alpha dW_2 + \sigma_\beta dW_4. \quad (13)$$

Eq. (13) is the most general form of the ToF ranging stochastic process we present. This model describes the rate of change of samples of the correlation wavefunction with variation in brightness and phase over time. Arguably the most surprising part of Eq. (13) is the second cosine term arising from the second derivatives in Eq. (10) which implies that random fluctuations in ϕ or θ will bias the range imaging process.

In Fig. 1 we show examples of simulations of the stochastic ToF ranging correlation wavefunction model for motion towards the camera and varying proportions between the noise variance variances. In each, 1000 simulations are performed of 1001 phase steps. The starting phase is $\pi/4$ and the velocity is approximately 5 m/s towards the camera. The shaded region represents

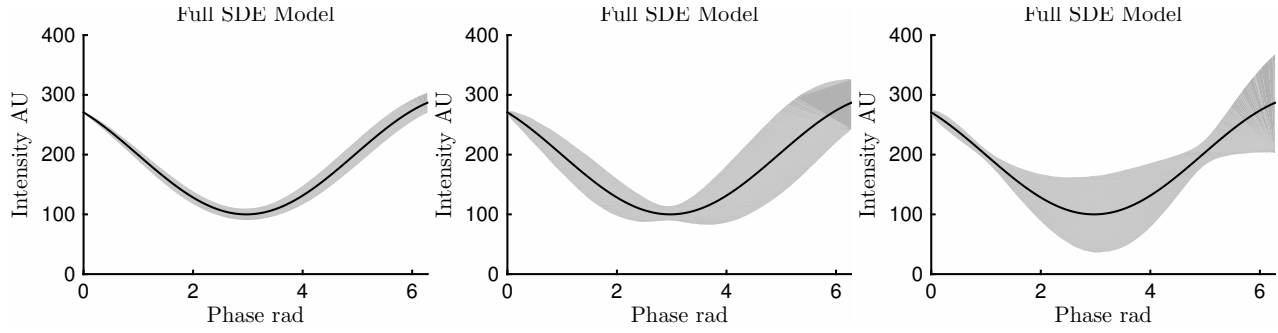


Fig. 1. Example stochastic processes based on the ToF model. The solid line is the deterministic model $f(x)$. The left graph is of small noise, specifically $\sigma_\phi = \sigma_\theta = 0.0016$ rad, $\sigma_\alpha = 0.01$ AU (where $\alpha = 100$ AU), and $\sigma_\beta = 0.1$ AU. The middle graph has comparatively large jitter to the left graph of $\sigma_\phi = 0.008$ rad, whereas the right graph has comparatively large amplitude noise of $\sigma_\alpha = 1$ AU.

three standard deviation margins, and the solid lines are $f(x)$ in Eq. (12) above. The increasing width of the shaded regions with phase represents the increasing uncertainty in predicting future intensity values, assuming exact knowledge at zero phase of the intensity and correlation waveform parameters. When all noise variances are small there is no distinct shape to the noise variance. When either the jitter (σ_ϕ or σ_θ), or the amplitude noise, σ_α , is large then there is distinctive shape to the noise variance.

Eq. (13) is a one-dimensional process with multiple time-varying parameters. The objective of this work is to find just one of those parameters, namely $d\phi/dt$, and subsequently in what follows we make successive approximations leading to tractable solutions for the estimation of radial motion.

RADIAL LINEAR MOTION ESTIMATION

The parameters in Eq. (13) describe the rate of change of ϕ and α , therefore inference of these terms from a sampling of the correlation waveform using the SDE leads to estimation of radial motion.

Consider the following parameterisation of Eq. (13)

$$dF = k_1 X_2 \cos(X_1 + X_3) dt - k_2 X_2 \sin(X_1 + X_3) dt + \frac{d\beta}{dt} dt - X_2 \sin(X_1 + X_3) [\sigma_\phi dW_1 + \sigma_\theta dW_3] + \cos(X_1 + X_3) \sigma_\alpha dW_2 + \sigma_\beta dW_4. \quad (14)$$

where

$$k_1 = \frac{1}{X_2} \frac{d\alpha}{dt} - k_3, \quad (15)$$

$$k_2 = \frac{d\phi}{dt} + \frac{d\theta}{dt}, \quad (16)$$

$$k_3 = \sigma_\phi^2 + \sigma_\theta^2. \quad (17)$$

The parameter k_3 is interpreted as the total variance due to jitter, k_1 depends on the amplitude, the rate of change of the amplitude, and jitter, and k_2 is the sum of radial motion and the controlled change of phase used to construct the correlation waveform. The parameter k_2 is equivalent to k in Eq. (2), hence estimation of k_2 leads to an estimate of the radial velocity. The assumption of linear radial motion implies to a close approximation

$$\frac{d\alpha}{dt} = 0 \Rightarrow k_1 = -k_3. \quad (18)$$

Finally, from Fig. 1 we see that F is essentially a function that fluctuates randomly about the deterministic model, f . The simplest,

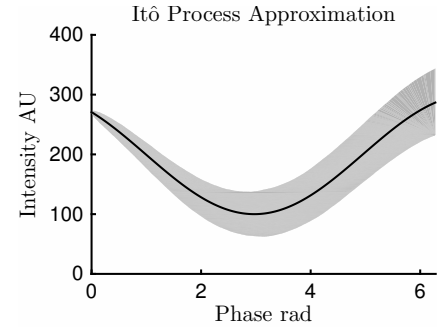


Fig. 2. The Itô process approximation of the ToF ranging stochastic model. The solid line is the deterministic model $f(x)$. No distinct shape is present in the noise.

most pragmatic example of such a process is an Itô process [13] with constant variance, σ^2 , and deterministic drift, viz.

$$dF \approx \left[k_1 \alpha \cos(\phi + \theta) - k_2 \alpha \sin(\phi + \theta) + \frac{d\beta}{dt} \right] dt + \sigma dW. \quad (19)$$

In Fig. 2 an example of the simplified model is shown according to Eq. (19) with constant β , $\alpha = 100$, and $\sigma = 0.6$. The deterministic drift and constant variance in dF leads to no distinct shape in the noise variance with phase. Therefore we regard Eq. (19) a Gaussian additive noise dominated model, c.f Fig. 1.

Modelling of Probability and Statistical Estimation of Parameters

The solution to the SDEs above is the function $F(X)$, or approximations thereof, obtained by inserting stochastic variables for the deterministic one into $f(x)$. Such a solution is called a strong solution [13] and describes the individual paths that satisfy the SDE. The strong solution leads us back to the traditional signal processing based approaches for estimating parameters such as ϕ and the motion k_2 . Here we are interested in statistical approaches to parameter inference that follow naturally from the modelling of stochastic variables.

The first step to statistical inference is the modelling of the statistical distributions. In stochastic calculus this amounts to finding the transition probability function (TPF). The TPF describes the probability that a stochastic variable F will transition from one value $F(t)$ to another value $F(s)$, $t < s$. Given a TPF $p(f, t, y, s)$ denoting the probability of transition between value

y at time s and value f at time t , and $g(y; T) = Pr(F = y; T)$ denoting the probability that F is measured equal to y at some time T , then $p(f; t) = Pr(F = f, t)$, the probability distribution for earlier measurements, is

$$p(f; t) = \int_{y \in \mathcal{R}} g(y; T) p(f, t, y, T) dy. \quad (20)$$

Solutions in terms of the TPF capture important information about the mean and variance of the stochastic variable in time. The use of statistical distributions lends itself naturally to inferential techniques which we use to validate our stochastic ToF model.

Derivation of the TPF is performed by solving the backward equation, a deterministic partial differential equation which, for a SDE, e.g. Eq. (4), is defined by the solution of [13]

$$\frac{1}{2} \sigma^2 \frac{\partial^2 p}{\partial f^2} + \mu \frac{\partial p}{\partial f} = -\frac{\partial p}{\partial t}. \quad (21)$$

Fortunately, we do not need to solve Eq. (21) directly, instead the Feynman-Kac formula provides an indirect route via the conditional expectation [13]

$$\begin{aligned} p(f; T) &= \langle g(F(T); T) | F = f(t) \rangle \\ &= \int_{F=-\infty}^{\infty} g(F(T); T) g(F = f(t); t) dF, \end{aligned} \quad (22)$$

where $\langle \cdot \rangle$ denotes expectation over X .

To facilitate the derivation of p , we use Eq. (19), which has integral

$$F \approx \frac{k_1}{k_2} \alpha \sin(\phi_0 + k_2 \theta(t)) + \alpha \cos(\phi_0 + k_2 \theta(t)) + \beta + \sigma W, \quad (23)$$

and expected value

$$\langle F \rangle \approx \frac{k_1}{k_2} \alpha \sin(\phi_0 + k_2 \theta(t)) + \alpha \cos(\phi_0 + k_2 \theta(t)) + \beta. \quad (24)$$

The TPF for a drift-diffusion equation with deterministic mean is well known to be a Gaussian distribution [13], therefore we choose $g(\cdot) = G(\cdot; m, \sigma)$, the Gaussian probability density function with mean m and standard deviation σ , and make the simplifying assumption of homoscedasticity, leading to the following result for the Feynman-Kac formula

$$\begin{aligned} p(f; T) &= \int_{F=-\infty}^{\infty} G(F; F(T), \sigma) G(F; f, \sigma) dF \\ &= \frac{1}{2\pi\sigma^2} \int_{F=-\infty}^{\infty} e^{-\frac{(F-F(T))^2}{2\sigma^2}} e^{-\frac{(F-f)^2}{2\sigma^2}} dF, \end{aligned} \quad (25)$$

where $f = \langle F \rangle$. Eq. (25) is a convolution. The convolution of two Gaussians is also a Gaussian with variances summed, therefore,

$$p(f; T) = \frac{1}{2\pi\sigma^2} e^{-\frac{(f-F(T))^2}{4\sigma^2}}. \quad (26)$$

Eq. (26) represents the probability that $f = F$ at time T . Therefore, given such a noisy measurement, the optimal model $f = \langle F \rangle$ is one which maximises p at $t = T$. In other words, it constrains the function f only at T . There are six unknown variables in Eq. (19), therefore infinitely many such functions that maximise p . To fully constrain f more measurements are required leading to a problem of maximising a joint probability.

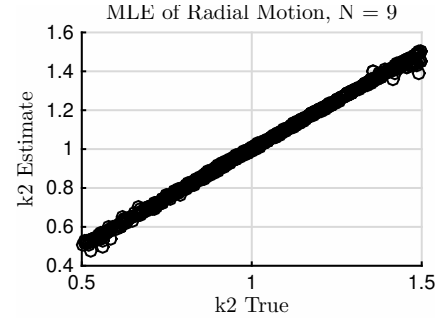


Fig. 3. Simulation of the estimation of the radial velocity parameter by MLE.

Taking a series of measurements at phase steps, $T = \theta_i$, the joint probability $p_J(f)$ is then

$$p_J(f) = \frac{1}{(2\pi\sigma^2)^N} \prod_{i=1}^N e^{-\frac{(f-F(\theta_i))^2}{4\sigma^2}}. \quad (27)$$

Maximum likelihood estimation (MLE) [23] is used to find k_2 , see the simulation results in Fig. 3. The data were generated with $\alpha = 1$ and standard deviation of noise 0.01. Failure in MLE occurred due to local solutions. In numerical simulations MLE failed in about one third of all phase step sets. The nonlinear nature of Eq. (27) means that there is no known closed form solution, necessitating an iterative algorithm. Over the six variables, $\mathcal{O}(36N)$ operations are required at each iteration, but the total number of iterations is indeterminate. Therefore, MLE is also slow, requiring 582 seconds to process 1000 phase step sets in the simulation. Faster noniterative processing methods are clearly required, and is the topic of the following section.

Motion Estimation

To derive a fast motion estimation procedure we require a simplified SDE with one motion dependent parameter. *Prima facie* martingale estimation [17, 18] seems appropriate, but dF has non stationary mean, therefore is not a martingale. It is both a local super and submartingale, each in non-overlapping regions, but those exact regions depend on ϕ which is unknown *a priori*. Therefore we need to derive an alternative, albeit biased, method to estimate the motion.

To such end, we make a small jitter approximation by neglecting k_1 . Doing so also neglects the biasing effect seen in the second cosine term in Eq. (13). But, we will see that accepting some jitter bias leads to a tractable solution. Now, consider the Itô process

$$dF(\theta) = \frac{\partial f}{\partial \theta} dt + \sigma dW, \quad (28)$$

where

$$f(\theta) = \alpha \cos(\phi_0 + k\theta) + \beta, \quad (29)$$

and here k is a radial velocity factor, like Eq. (3), where $k > 1$ indicates motion away from the camera and $k < 1$ motion towards the camera. Combining Eqs. (28) and (29) results in

$$dF(\theta) = -k\alpha \sin(\phi_0 + k\theta) + \sigma dW. \quad (30)$$

Clearly, given measurements $F(\theta_i)$, the differences $F(\theta_i) - F(\theta_{i-1})$ have expected value $\langle F(\theta_i) - F(\theta_{i-1}) \rangle = -k\alpha \sin(\phi_0 +$

$k\theta_i$) [17]. Therefore, an appropriate method to find the estimate, \hat{k} , of k from N measurements of F_i is to compute the following

$$\hat{k} = \frac{\sum_{i=1}^{N-1} F(\theta_i) - F(\theta_{i-1})}{-\sum_{i=0}^{N-1} \alpha \sin(\phi_0 + k\theta_i)}. \quad (31)$$

We are left with the problem of finding $\alpha \sin(\phi_0 + k\theta_i)$.

A modified measurement procedure to compute \hat{k} is as follows. Define the phase steps θ_i and offset $\Delta\theta$. For each i take three measurements in the following order: $F(\theta_i)$, $F(\theta_i + \Delta\theta)$, $F(\theta_i + \pi/2)$ ¹. We therefore have the given approximate expected values,

$$\langle F(\theta_i + \Delta\theta) - F(\theta_i) \rangle \approx -2\Delta\theta k \alpha \sin(\phi_0 + k\theta_i), \quad (32)$$

$$\langle F(\theta_i + \pi/2) \rangle \approx \alpha \sin(\phi_0 + k\theta_i) + \beta. \quad (33)$$

The process of taking three measurements means that two measurement integration time intervals are required between $F(\theta_i)$, and $F(\theta_i + \pi/2)$, which gives rise to the factor of 2 in Eq. (33). Finally, to eliminate β we apply the discrete Fourier transform (DFT) and utilise the first order bin, leading to the following estimate

$$\hat{k} = \frac{1}{2\Delta\theta} \left| \frac{\sum_{i=1}^{N-1} (F(\theta_i) - F(\theta_{i-1})) e^{j2\pi(i-1)/N}}{\sum_{i=0}^{N-1} F(\theta_i + \pi/2) e^{j2\pi i/N}} \right|. \quad (34)$$

An appealing feature of Eq. (34) is it requires $\mathcal{O}(N)$ operations to compute \hat{k} .

In practice, the ToF camera modulation signals for the light source and sensor are both generated using digital logic (for simplicity and reduced cost). Assuming the camera bandwidth is adequate, then the square wave modulation is preserved in the operation of the sensor and light source, and the correlation waveform is triangular in shape. Therefore Eq. (29) becomes

$$f(\theta) = \alpha \sum_{m=0}^{\infty} \frac{1}{(2m+1)^2} \cos((2m+1)(\phi_0 + k\theta)) + \beta. \quad (35)$$

Consequently

$$\frac{\partial f}{\partial \theta} = -\alpha k \sum_{m=0}^{\infty} \frac{1}{(2m+1)} \sin((2m+1)(\phi_0 + k\theta)), \quad (36)$$

is a square wave and Eq. (34) is prone to severe harmonic error. An alternative procedure becomes apparent when we note that the radial velocity parameter affects the amplitude of the expected value of the differences, $\langle F(\theta_i + \Delta\theta) - F(\theta_i) \rangle \approx \partial f / \partial \theta$. Therefore, let A_F and $A_{\Delta F}$ be respectively the amplitudes of the measured noisy signals $F(\theta_i)$ and $F(\theta_i + \Delta\theta) - F(\theta_i)$ in i , and find the estimate of the velocity as

$$\hat{k} = c_1 \frac{A_{\Delta F}}{2\Delta\theta A_F}, \quad (37)$$

where c_1 is some constant determined by the bandwidth of the ToF range image camera system, hence number of harmonics present in the measurements. Eq. (37) also only requires $\mathcal{O}(N)$ operations. From Eq. (29), $k = 1$ corresponds to no motion, therefore in general the calculation $\hat{k} - 1$ is proportional to the change in position per raw frame. If the camera runs at N raw frames per raw frame set times M raw frame sets per second,

¹Here the measurements are dependent on θ_i , a programmable parameter on the camera, which are not chosen to strictly increase in time.

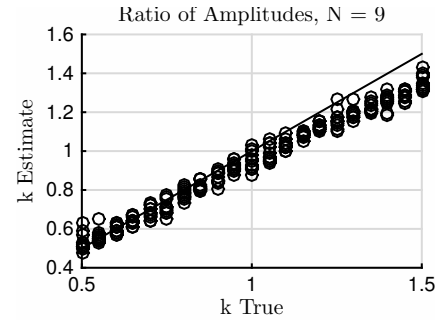


Fig. 4. Simulation of the estimation of radial motion by the ratio of amplitudes.

then there are NM raw frames per second, therefore the final estimate of velocity is

$$MNc_2 (\hat{k} - 1). \quad (38)$$

The coefficients c_1 and c_2 are determined by calibration. We call Eq. (37) the ratio of amplitudes method.

In simulation we test the ratio of amplitudes method with $\alpha = 100$, noise with variance $\sigma^2 = 1$, and the third and fifth harmonics. Results of the simulation are shown in Fig. 4. Comparing Fig. 4 with Fig. 3 the variance in the estimate of the motion parameter is increased and there is some bias for larger values of k . However, unlike MLE, ratio of amplitudes returned an estimate in every test, and in $\mathcal{O}(N)$ operations.

EXPERIMENTATION

We test the radial velocity estimation using a prototype proprietary ToF range imaging camera [8] on which we can set the phase step sequence. A white foam board target with diffuse reflectance is affixed to a 3 meter translation stage (macro Dynamics Inc, Croydon, PA, USA) and moved in coordination with the camera data acquisition. The translation stage has maximum speed of 2.3 m/s. The hardware is controlled and coordinated using MatLab (The Mathworks, Natick, Massachusetts, United States). We require much faster speeds, so the motion is simulated by moving the target to discrete positions 0.005 m apart and acquiring the raw frames. In processing we then select the required frames from the data set to reconstruct a phase step set with motion. Pixels in an 11×11 pixel (121 pixels) square window on the white foam board were processed. Like in [8] we assume that the motion is sufficiently slow such that α is constant during each individual raw frame.

Because of harmonic content we only test the MLE and ratio of amplitude methods. MLE was performed using the `mle` routine in MatLab with the `fmincon` optimisation function. In the ratio of amplitude method the coefficients c_1 and c_2 are found by inspection. While Eq. (34) was useful for the exposition, it is not tested with real data for the reasons given above, and left for future work when we have a device that can produce a correlation waveform free of harmonic content (see for e.g [24]).

There are two parameters to test for in the motion estimation: the number of phase steps and the phase range that the phase steps span. The phase steps are then set evenly distributed over the phase range. For motion towards the camera, the phase steps are effectively compressed. Therefore, for motion towards the camera, if we choose the usual phase range of 2π then the measurements will not span one full period of the correlation

waveform. The ratio of amplitudes depends on the observation of at least one peak and one trough of the correlation waveform, so the phase range must be greater than 2π . In our experiments we test 3π and 4π .

The modulation frequency of the camera is set to 70 MHz (ambiguity distance of 2.143 m) and the phase range is divided into eighteen even steps. These are pragmatic choices: the modulation frequency must be high enough that one ambiguity distance fits within the 3 m translation stage; but low enough that we have as many phase steps available as possible, as determined by the ratio of the internal system reference clock frequency of the camera to the desired modulation frequency. We test various numbers of phase steps by decimating the eighteen phase steps into multiples, viz. six, nine and eighteen. Three provides inadequate coverage of the correlation waveform, therefore three steps is not tested.

RESULTS AND DISCUSSION

In Fig. 5 we show the results for estimation of linear radial motion by MLE. Velocity indicates the displacement of the translation stage per phase step frame multiplied by the frames per second. Assuming thirty phase step sets per second, that is, thirty range images per second, and N phase steps we have $30N$ phase step images per second. Therefore, a displacement of 0.005 m per frame (the minimum tested) gives a velocity of 0.15N m/s in simulation. Sign indicates direction of the motion, negative is towards the camera and positive is away.

Linearity in the estimates is herein assessed as deviation from the reference line with respect to the standard deviation as indicated by the error bars. For six phase steps spread over 3π , MLE estimates the radial velocity between (-10) – 0 m/s, $(+20)$ – $(+30)$ m/s, and nowhere else. The MLE velocity estimation grossly failed for six phase steps over 4π due to insufficient sampling of the correlation waveform. Overall MLE radial motion estimation from six phase steps is not recommended. For nine phase steps the linearity is within uncertainty in the range (-20) – $(+20)$ m/s, but underestimates the speed above $+20$ m/s towards the camera. For eighteen phase steps the linearity degrades above velocities of 20 m/s in both directions, but is good overall.

The estimation by MLE is required on average 0.95 s to process each phase step set. MLE is sensitive to initialisation values and hence difficult to perform consistently. Therefore MLE estimation of velocity is only recommended for non-realtime applications where initialisation parameters are well known. Overall the MLE algorithm failed in about 21% of tests. Missing points in Fig. 5 represent velocities where no estimate is obtained over the 121 pixels tested (3π , 18 steps), despite attempts on a best effort basis. However, where estimates were obtained, the standard deviation in the velocity estimates were typically less than 1 m/s, and frequently were better than 0.3 m/s, as seen in the small error bars in Fig. 5 for nine and eighteen steps.

As the speed increases the change in brightness as seen by the camera increases, therefore the validity of the approximation made in Eq. (18) decreases. The results in Fig. 5 show that the linearity decreases with speed which coincides with the increased brightness change. However, the results herein indicate that the approximation is adequate for speeds up to about 20 m/s (72 km/h or 44.7 mph). Surpassing this limit requires explicit modelling of the brightness change, and is left for future work.

The use of the Gaussian distribution in MLE arises from the use of Wiener processes in the stochastic model. The Gaus-

sian distribution is an adequate approximation of the statistics assuming large photon numbers for the Poisson photon and detector shot noise statistics [21]. The Gaussian also models electronic thermal noise. However, in physical reality the noise is the sum of two random variables drawn from, respectively, a Gaussian and a Skellam distribution (the Skellam distribution being the model for the difference of two Poisson random variables). There is no known solution to the distribution of a sum of Gaussian and Poisson/Skellam random variables. In imaging poorly reflecting objects the MLE approach taken here may break down and more careful modelling of the noise distributions may be required [25], in future work we will investigate this issue further.

In Fig. 6 we show the results for estimation by the ratio of amplitudes, viz. Eq. (37). The intensity change with distance was a confounding factor for high velocities, and no good estimates were obtained beyond 20 m/s, hence the change in axis limits between Figs. 5 and 6. Also, a small jitter approximation was made in the derivation of the ratio of amplitudes method, the understanding of the full impact of which is left for future work. The error bars are up to ± 6 m/s for 18 steps over 3π and up to ± 10 m/s over 4π , whereas for 9 steps the error bars are less than ± 3.5 m/s over 3π and less than ± 4 m/s over 4π . Comparing the error bars in Figs. 5 and 6 it is immediately apparent that MLE is more robust to noise than the ratio of amplitudes. Unlike MLE, however, the ratio of amplitudes always returns an estimate and in fixed efficiency of $\mathcal{O}(N)$ operations per pixel. On the real data the average processing time was 4 ms, a 243 times speed increase over MLE. Increasing N decreased $\Delta\theta$, which lead to less signal in the estimate of $A_{\Delta F}$ and greater noise in \hat{k} . From the F-test under regression analysis the estimated velocity significantly predicted the radial velocity ($p = 0$ in each case for $N = 9$ and 18), but note that the sample sizes are in the thousands, and the standard deviations are a better indicator of prediction power.

For the six phase steps distributed over 3π the ratio of amplitudes method provides good linearity for motion away from the camera but failed for motion towards the camera. The ratio of amplitudes method requires adequate sampling to obtain measurements of the extrema of the correlation waveform. Motion towards the camera decreases the range phase which causes an apparent reduction in the phase step size. The risk here is that the measurements may not cover both one peak and trough of the correlation waveform, in which case the estimate of A_F is inaccurate. In contrast, over 4π the six phase steps worked well for motion towards the camera but, confoundingly, not away. A reasonable explanation for why six phase steps fails over 4π for motion away from the camera is that the sampling density of the correlation waveform is inadequate to reliably provide samples of both extrema of the correlation waveform, and thereby provide good estimates of A_F and $A_{\Delta F}$. Thus we observe a tradeoff in using the proposed ratio of amplitudes method for estimating radial velocity with a small number of samples that leads to one of two failure cases. Namely, that reducing the span increases the density of the samples but risks the inability to observe both a peak and a trough, and vice versa.

For nine and eighteen phase steps the ratio of amplitudes performed better over a span of 3π than 4π in terms of both linearity and standard deviation. Over 4π , nine steps did not have adequate sampling density to accurately estimate the velocity of motion away from the camera. Overall, in our tests nine steps over a span of 3π provided the best tradeoff of accuracy and precision.

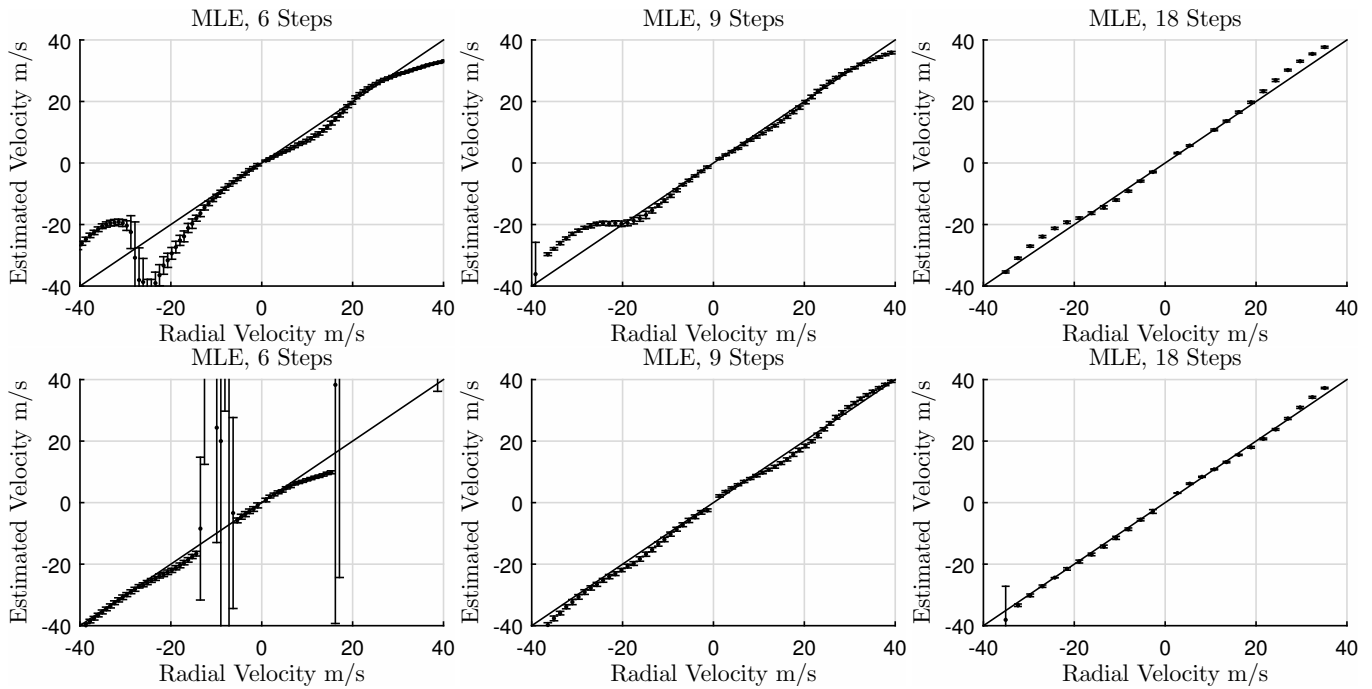


Fig. 5. Estimation of radial velocity by maximum likelihood estimation for various numbers of phase steps. Top row are evenly spaced over 3π , bottom row are over 4π . The solid line is for reference. Motion towards the camera corresponds to negative velocity. Error bars indicate one standard deviation in estimation. The linearity improves with the number of phase steps.

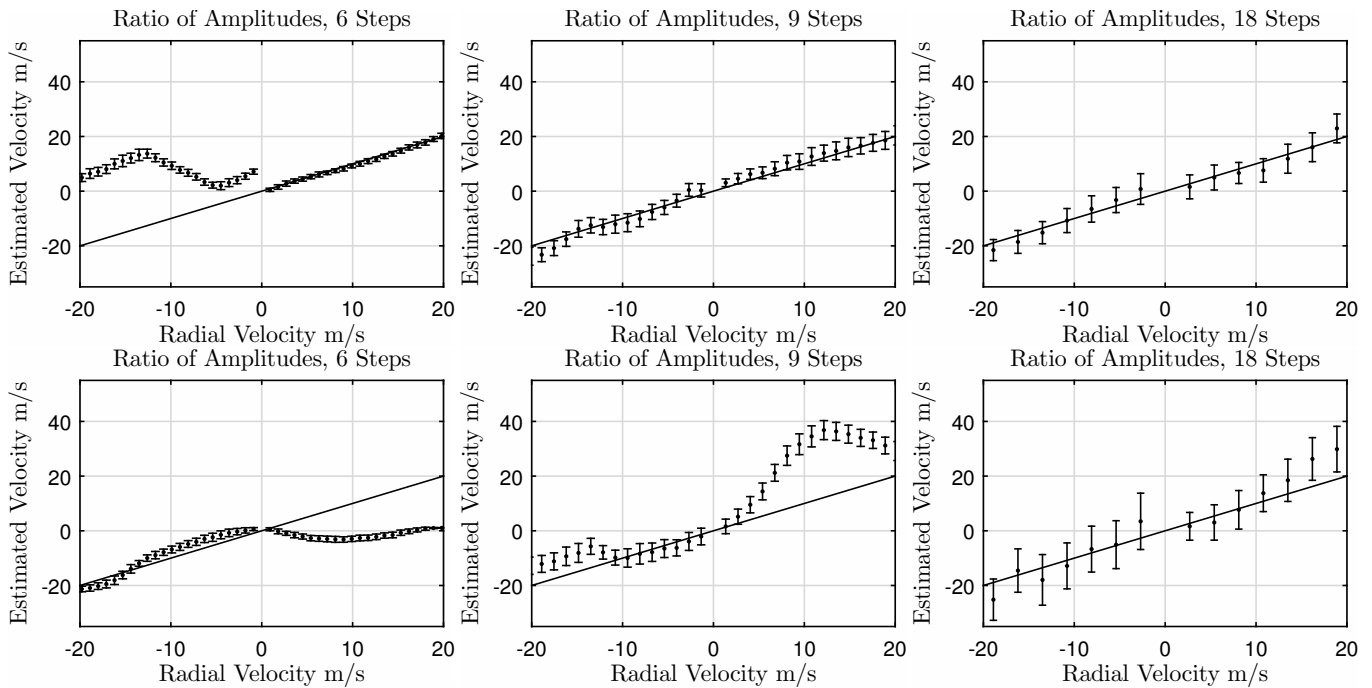


Fig. 6. Radial velocity estimation by ratio of amplitudes, Eq. (37), for various numbers of phase steps. Top row are evenly spaced over 3π , bottom row are over 4π . The solid line is for reference. Error bars indicate one standard deviation. The reduced phase step size with increased number of steps leads to more noise in the estimate of $A_{\Delta F}$, hence the increase in noise.

CONCLUSION

We have investigated the application of stochastic calculus to ToF range imaging and successfully used the resultant model to estimate linear radial motion. Two methods for motion estimation were developed and tested: MLE using statistical distributions derived from the stochastic model, and a new closed form method we call the ratio of amplitudes. Estimation was tested for six, nine, and eighteen phase steps, six was found to be unreliable, whereas nine and eighteen steps were more reliable than six. MLE performed better than ratio of amplitudes in terms of random noise, but MLE was more than 200 times slower and was not able to provide an estimate in every example. In contrast ratio of amplitudes is efficient and always produces an estimate of radial motion.

Two major aspects are raised for future work. The first aspect is that the general stochastic model derived implies that the variance in the transition probability between measurements should have a specific shape, depending on the relative variances of the noise sources. Also, more interestingly, the stochastic model implies that random jitter biases the estimation of phase, even for a static scene. These are important metrological questions, and future investigations we will delve into the consequences of this result. The second aspect is the extension of the results presented herein to nonlinear motion and, most challengingly, nonlinear motion due to the passage of an edge over a pixel during data acquisition. This will be extended to the measurement of transverse motion.

ACKNOWLEDGEMENTS

This work was supported by Marsden Council Fast Start Grant 15-UOW-008. The review of mathematics by Moira Steyn-Ross is acknowledged.

REFERENCES

1. D. Lefloch, R. Nair, F. Lenzen, H. Schäfer, L. Streeter, M. J. Cree, R. Koch, and A. Kolb, *Technical Foundation and Calibration Methods for Time-of-Flight Cameras* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013), pp. 3–24.
2. O. Lottner, A. Sluiter, K. Hartmann, and W. Weihs, "Movement artefacts in range images of time-of-flight cameras," in "International Symposium on Signals, Circuits and Systems," (Iasi, Romania, 2007), pp. 117–120.
3. S. Lee, "Time-of-flight depth camera motion blur detection and deblurring," *IEEE Signal Processing Letters* **21**, 663–666 (2014).
4. M. Lindner and A. Kolb, "Compensation of motion artifacts for time-of-flight cameras," in "Dynamic 3D Vision Workshop," , vol. 5742 (Jena, Germany, 2009), vol. 5742, pp. 16–27.
5. M. Schmidt, M. Jehle, and B. Jähne, "Range flow estimation based on photonic mixing device data," *International Journal of Intelligent Systems Technologies and Applications* **5**, 380–392 (2008).
6. S. Ghuffar, N. Brosch, N. Pfeifer, and M. Gelautz, "Motion segmentation in videos from time of flight cameras," in "19th International Conference on Systems, Signals and Image Processing (IWSSIP)," (Vienna, Austria, 2012), pp. 328–332.
7. L. Streeter and A. A. Dorrington, "Coded exposure correction of transverse motion in full-field range imaging," *Optical Engineering* **53**, 102109–1–102109–10 (2014).
8. R. Whyte, L. Streeter, M. J. Cree, and A. A. Dorrington, "Application of lidar techniques to time-of-flight range imaging," *Applied Optics* **54**, 9654–9664 (2015).
9. F. Heide, W. Heidrich, M. Hullin, and G. Wetzstein, "Doppler time-of-flight imaging," *Transactions on Graphics* **34**, 36:2–36:11 (2015).
10. S. Shrestha, F. Heide, W. Heidrich, and G. Wetzstein, "Computational imaging with multi-camera time-of-flight systems," *Transactions on Graphics* **35**, 33:1–33:11 (2016).
11. M. U. Bikdash and K. B. Yu, "Analysis and filtering using the optimally smoothed wigner distribution," *IEEE Transactions on Signal Processing* **41** (1993).
12. V. Katkovnik and L. Stanković, "Instantaneous frequency estimation using the wigner distribution with varying and data-driven window length," *IEEE Transactions on Signal Processing* **46**, 2315–2325 (1998).
13. F. C. Klebaner, *Introduction to Stochastic Calculus with Applications* (Imperial College Press, London, UK, 2012), 3rd ed.
14. S. N. Cohen and R. J. Elliott, *Stochastic Calculus and Applications* (Springer, New York, USA, 2015), 2nd ed.
15. A. Abutaleb, "Instantaneous frequency estimation when the amplitude is a stochastic process using stochastic calculus and bootstrapping," *Circuits, Systems and Signal Processing* **24**, 35–52 (2005).
16. A. Abutaleb, "Instantaneous frequency estimation using stochastic calculus and bootstrapping," *EURASIP Journal on Applied Signal Processing* **12** (2005).
17. B. M. Bibby and M. Sørensen, "Martingale estimation functions for discretely observed diffusion processes," *Bernoulli* **1**, 17–39 (1995).
18. M. Sørensen, *Statistical Methods for Stochastic Differential Equations* (CRC Press, Boca Raton, FL, 2012), chap. Estimating functions for diffusion-type processes, pp. 1–108.
19. L. Streeter, M. J. Cree, and A. A. Dorrington, "A strategy for the correction of effects of jitter in AMCW lidar images," in "The 28th International Conference of Image and Vision Computing New Zealand," (Wellington, New Zealand, 2013), pp. 500–505.
20. J. Seiter, M. Hofbauer, M. Davidovic, and H. Zimmermann, "Investigation of the distance error induced by cycle-to-cycle jitter in a correlating time-of-flight distance measurement system," *Optical Engineering* **53**, 073104 (2014).
21. G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet, "Stein's method and normal approximation of Poisson functionals," *The Annals of Probability* **38**, 443–478 (2010).
22. M. P. Li, *Jitter, Noise, and Signal Integrity at High-Speed* (Prentice Hall, Upper Saddle River, NJ, USA, 2008).
23. A. Stuart, J. K. Ord, and S. Arnold, *Kendall's Advanced Theory of Statistics*, vol. 2A (Arnold, London, UK, 1999).
24. C. Peters, J. Klein, M. B. Hullin, and R. Klein, "Solving trigonometric moment problems for fast transient imaging," *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)* **34**, in press (2015).
25. D. Shin, F. Xu, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "Computational multi-depth single-photon imaging," *Opt. Express* **24**, 1873–1888 (2016).