Video-Based Tracking of Single Molecules Exhibiting Directed In-Frame Motion

M. Yavuz Yüce,[1,*] Alexandr Jonáš,[1] Alper Kiraz,[1] and Alper T. Erdoğan[2]

[1]Department of Physics, Koç University, Istanbul, Turkey
[2]Department of Electrical Engineering, Koç University, Istanbul, Turkey

Abstract: Trajectories of individual molecules moving within complex environments such as cell cytoplasm and membranes or semiflexible polymer networks provide invaluable information on the organization and dynamics of these systems. However, when such trajectories are obtained from a sequence of microscopy images, they can be distorted due to the fact that the tracked molecule exhibits appreciable directed motion during the single-frame acquisition. We propose a new model of image formation for mobile molecules that takes the linear in-frame motion into account and develop an algorithm based on the maximum likelihood approach for retrieving the position and velocity of the molecules from single-frame data. The position and velocity information obtained from individual frames are further fed into a Kalman filter for interframe tracking of molecules that allows prediction of the trajectory of the molecule and further improves the precision of the position and velocity estimates. We evaluate the performance of our algorithm by calculations of the Cramer-Rao Lower Bound, simulations, and model experiments with a piezo-stage. We demonstrate tracking of molecules moving as fast as 7 pixels/frame (12.6 μm/s) within a mean error of 0.42 pixel (37.43 nm).

Key words: single molecule microscopy, single molecule tracking, maximum likelihood estimation, Kalman filtering

INTRODUCTION

The ability to follow in space and time selected individual molecules in complex environments, surrounded by millions of other molecules, has opened up new paths toward understanding the basic principles of fundamental chemical and biological processes. Current single-molecule experimental work covers a broad spectrum of research areas ranging from the study of intracellular and membrane transport in living cells and model systems (Pralle et al., 2000; Schutz et al., 2000; Kues et al., 2001) through the single-molecule chemistry (Schmidt et al., 1996; Ho, 2002) to the analysis of mechanical properties of molecules (Evans, 2001; Pampaloni et al., 2006). The key advantage of single-molecule observations is the possibility to obtain the distribution of various behaviors within a certain molecular population. This enables clear identification of functionally distinct molecular subpopulations that would be otherwise masked by an ensemble-averaged measurement (Yorulmaz et al., 2009; Hoffmann et al., 2011).

High-sensitivity and high-resolution fluorescence microscopy is one of the very few techniques that can provide details about the organization and dynamics of complex material systems (e.g., cell cytoplasm and membranes or semiflexible polymer networks) on the single molecule level under relevant environmental conditions (Saxton & Jacobson, 1997; Moerner, 2007; Joo et al., 2008). It involves recording of a movie from a sample, where single fluorescent molecules serve as probes of their environment or

tracers of other chemically attached molecules. Upon determining locations of the single molecules in each individual frame, and relating these to the position information from other frames, one can obtain the trajectories of single molecules in the sample. This procedure is commonly referred to as single molecule tracking (SMT).

The main advantage of SMT over other single molecule techniques lies in the combined spatial and temporal information it provides. The temporal part of this information is fairly reliable with contemporary electronics. The spatial part, on the other hand, requires careful analyses, due to limited optical resolution and signal levels. This has been addressed either in the more general framework of particle tracking (Cheezum et al., 2001; Thompson et al., 2002; Anthony & Granick, 2009), or by specifically considering single molecules with an emphasis on accuracy (Ober et al., 2004), or both accuracy and speed (Smith et al., 2010). Among algorithms suitable for SMT, direct Gaussian fitting (Cheezum et al., 2001) with a maximum likelihood estimator (MLE) (Abraham et al., 2009, 2010) was found to be the most accurate one.

The estimated performance can be enhanced by a better representation of the data generation process. A good example of this has been demonstrated in Ober et al. (2004) and Smith et al. (2010), where the discrete and independent nature of the photon emission event was put into use to develop a novel imaging model based on Poisson statistics. Their analyses, however, considered stationary molecules only.

In this article we apply the same approach to molecules in motion. To the best of our knowledge, all of the present

tracking algorithms ignore a molecules' motion within one-frame time and fit the data with the point spread function of a stationary emitter. We follow an alternative method and use a new imaging model that also takes into account the motion of the molecule during the exposure time of a frame. Since this extended model better represents the real imaging process, it is expected to yield more accurate position estimates for a moving molecule. As inherently including motion, the model also allows the molecules' velocity to be estimated from single frame data.

Incorporating the in-frame motion of molecules into the image formation model has been independently addressed in a very recent study through Cramer Rao lower bound (CRLB) calculations (Wong et al., 2011). Our work is different than this study mainly in recognizing the significance of the mid-frame-time position of a molecule. Using CRLB calculations and simulations, we demonstrate that the mid-frame-time position is a very convenient parameter for accurately describing the trajectory of a molecule. We subsequently exploit this observation to propose an algorithmic framework for tracking.

The tracking algorithm feeds the position and velocity estimates of individual frames into a Kalman filter, as has been suggested by others in studies with particles (Genovesio & Olivo-Marin, 2003; Smal et al., 2008; Wu et al., 2010). The addition of Kalman filtering into the SMT algorithm allows an optimal combination of the information extracted from multiple frames for more accurate prediction of the tracked molecule position in the next frame of the recorded image stack. This is important for a correct identification of the selected molecule among other molecules that might be present in the field of view and that might possibly cross the trajectory of the selected molecule. Moreover, Kalman filtering can improve the precision of the single-frame estimates of the molecule position and velocity if appropriate values of the process-associated and measurement-associated noise levels are provided (Wu et al., 2010).

Using simulated and experimental data, we demonstrate that our extended imaging model indeed improves the precision of the SMT in comparison with the standard algorithms that neglect possible directed motion of the molecule during the image acquisition.

MATERIALS AND METHODS

Imaging Model

We extend the imaging model introduced in (Ober et al., 2004) to molecules in motion. Our model is illustrated in Figure 1. The motion of a molecule within 1 frame exposure time, T, is assumed to be a uniform translational motion from an initial position (x_0, y_0) at the beginning of the exposure to a final position at the end of the exposure. The position at an arbitrary time t is denoted by x, y, and the (constant) velocity components of the uniform motion are denoted by (v_x, v_y). The whole motion of a molecule throughout the movie, the molecule trajectory, is thus represented as a piecewise linear function of x, y versus t. For
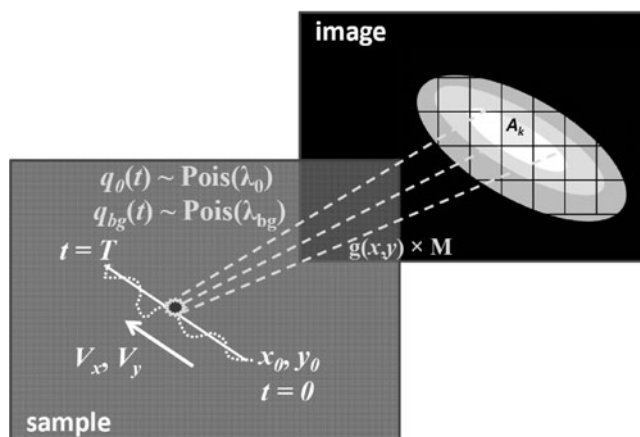

Figure 1. Description of the assumed image formation model.

reasons to be explained in the Algorithm section, we also define the midframe position components as $x_c = x_0 + v_x T/2$, and $y_c = y_0 + v_y T/2$. All of these position and velocity components are defined in the sample plane. Their corresponding values in the image plane are obtained by multiplying by the magnification, M.

The image of the sample develops progressively as the emitted photons land on the camera chip. The photons are assumed to be originating either from the imaged molecule, or the background. Both sources are assumed to emit in the form of homogeneous Poisson processes denoted by $q_0(t)$ and $q_{bg}(t)$ with respective (constant) rates of λ_0 and λ_{bg}. Landing of emitted photons on the camera chip is described by the point spread function $g(x, y)$, which we approximate by a two-dimensional Gaussian function as is well established in tracking of subwavelength particles.

Parameters of the described model and their units are tabulated in Table 1. We use two alternative unit systems corresponding to the sample and image spaces. When expressed in sample units, the parameter values remain unchanged under different exposure times. The model parameters are estimated from single frame data as described in the Algorithm section.

Simulations

We characterized the performance of our imaging model first by simulations, as they provide fully controlled imaging conditions. Simulated image of a moving molecule was generated by accumulating individual photons emitted by that molecule in the camera pixels. Each photon had three accompanying data: (1) random emission time, (2) emission

Table 1. Parameters of the Described Imaging Model and Their Units in the Sample and Image Spaces.

Parameter	Sample	Image
x_0, y_0	m	pixel
v_x, v_y	m/s	pixel/frame
λ_0	photons/s/molecule	counts/frame/molecule
λ_{bg}	photons/s/pixel	counts/frame/pixel

position in the sample space, and (3) landing position in the image space. Random numbers corresponding to times elapsed between subsequent emitted photons were drawn from an exponential distribution with parameter $1/\lambda_0$. The unique emission times of the photons were then given by the cumulative sum of the times between subsequent photons. This was continued until the emission times reached T. The emission positions were given by the position of the molecule at the time of emission. This depends on the trajectory of the molecule. For each photon, its landing position on the camera chip was determined by randomly drawing a single value from a normal distribution with mean equal to the emission position and width to the theoretical point spread function. The landing positions were then assigned to pixels to yield the final image.

Experiments

To test our model experimentally, we carried out tracking of single molecules immobilized in a polymer film. To this end we spin-coated a thin film of poly(methyl methacrylate) (Aldrich 201033; Aldrich, St. Louis, MO, USA) doped with well-separated terrylene dye molecules onto glass coverslips. Prior to spin coating, the coverslips were cleaned with a UV-O cleaner [Jelight Company (Irvine, CA, USA), Model: 42] for 10 min. This sample was then mounted onto a piezo scanner [Physical Instruments (Karlsruhe, Germany), stage: P-733.3DD, controller: E-509.C3A] and moved with a known trajectory generated by a multipurpose data acquisition (DAQ) card [National Instruments (Austin, TX, USA) PCIe-6363]. The same card was also used to read the position feedback of the stage (at 10 kHz). Wide-field imaging was done in the standard way, with a NA = 1.49 oil objective [Nikon (Tokyo, Japan), TIRF Apochromat], 178× total magnification, and an electron multiplying charge-coupled device (EM-CCD) camera with pixel size of 16 μm [Hamamatsu (Hamamatsu City, Japan) C9100-13]. The point spread function of the setup was determined from stationary single molecule images and corresponded to $\sigma = 1.2$ pixel ≈ 108 nm, when approximated with a Gaussian function. This σ value has been used in analyzing recorded movies. The camera was operated at its 11 MHz/pixel clock rate mode, with an EM gain of 255, and was externally triggered by the DAQ card, so that driving the piezo, reading its sensors, and image acquisition were all done in synchronization. Excitation was done with a green diode laser of ≈ 10 mW power and a suitable dichroic mirror (Chroma z532/1064rpc; Chroma Technology Corp., Bellows Falls, VT, USA). Fluorescence was further separated from excitation using notch (Chroma ZET532NF) and band-pass (Chroma HQ605/90M) filters.

Since our imaging model is based on photon numbers, obtaining its parameters from an image requires a conversion from pixel counts to photon numbers. This is done according to the relation:

$$N_{Photons} = \frac{(N_{counts}^{image} - N_{counts}^{dark}) \times Conversion\,factor}{Analog\,gain \times EM\,gain \times Q.E.}, \quad (1)$$

where $N_{Photons}$ is the number of photons that arrives to the pixels, N_{counts}^{image} are the pixel values of the image, N_{counts}^{dark} are the pixel values in the absence of illumination, and other quantities are the fixed or adjustable specification parameters of the camera (Hamamatsu, 2009). The distribution of the pixel values in a dark image was so narrow that we took N_{counts}^{dark} as a constant equal to the mean dark pixel value.

ALGORITHM

Our SMT algorithm can be considered as a low- and high-level analysis of a recorded movie. The low-level analysis works on an individual frame and determines the position of the tracked molecule in that frame. The high-level analysis fuses the position information from individual frames to improve the localization precision and accuracy, and to obtain the resulting trajectory.

We treat the low-level analysis as an estimation problem, where the unknown parameters of the imaging model (described in Table 1 in the Methods section) are estimated from the image of the tracked molecule. Since there are typically many molecules in a frame, a region of interest, that we call window, has to be defined around the tracked molecule. The estimation algorithm accepts the window as an input and estimates the model parameters using the maximum likelihood approach.

Defining the window of the estimation algorithm is handled by the high-level analysis. For the first frame, the window is readily defined when the user specifies the molecule to be tracked. For subsequent frames, the window position (size is kept fixed) must be renewed so as to enclose the moving molecule. This requires that the high-level analysis must have a predicted estimate of the molecule position, before the low-level analysis can actually estimate it. This results in a recursive estimation-prediction cycle, for a prediction must also depend on previous estimates. The Kalman filtering technique of statistical estimation theory is a perfect match for such recursive data processing problems and forms the basis of our high-level algorithm. Besides being a prediction tool for placing the window in the next frame, Kalman filtering also corrects the noisy position estimate of the current frame.

In the following subsections, we provide the details of our in-frame estimation (low-level) and Kalman filtering (high-level) algorithms. We also discuss the performance of the in-frame estimation.

Maximum Likelihood Approach for In-Frame Parameter Estimation

The MLE is one of the most popular approaches in estimation, which has some desired optimality properties (Kay, 1993; Abraham et al., 2009). The application of this approach requires that the joint probability mass function (pmf) of the measured pixel values $m_1, \ldots, m_N$ should be derived as a function of the unknown parameters $x_0, y_0, v_x, v_y, \lambda_0, \lambda_{bg}$ of the model described in the Methods section.

The stochastic nature of a photon detection event brings an unavoidable uncertainty to the mapping between the image and sample planes, and requires a probabilistic interpretation. The probability p_k, that a photon emitted at x', y' in the sample plane lands on a pixel k of the chip, is defined as

$$p_k(x', y') = \int_{A_k} g(x - x', y - y') \, dx \, dy, \qquad (2)$$

where $g(x, y)$ is the point spread function, A_k is the projection of the pixel area on the sample plane defined as: $A_k = \{(x, y) : x_{k1} \leq x \leq x_{k2}, y_{k1} \leq y \leq y_{k2}\}$, and the detection efficiency is assumed to be 1.

For a moving molecule, the photon emission positions x', y' and hence the p_k values change with time. The assumed linear motion of our imaging model can be included into p_k, by substituting $x' = x_0 + v_x t$, and $y' = y_0 + v_y t$ in equation (2). This yields a time-varying p_k written in terms of the model parameters:

$$p_k(t) = \int_{A_k} g(x - x_0 - v_x t, y - y_0 - v_y t) \, dx \, dy. \qquad (3)$$

To arrive from $p_k(t)$ at the pmf of the pixel values, we define the photon arrival process to pixel k, denoted by $\{q_k(t), t \geq 0\}$. Having contributions from two Poisson processes (emissions of the molecule and the background), $\{q_k(t)\}$ is also a Poisson process. Its rate can be written as

$$\lambda_k(t) = \lambda_0 p_k(t) + \lambda_{bg}, \qquad (4)$$

where the first term on the right side is the contribution from the molecule, given by the product of the photon generation rate and the landing probability. The second term is the contribution from the constant and homogeneous background, for which, by definition, the arrival probability is the same for all times and pixels. We note from equation (4) that $\{q_k(t)\}$ has a time-varying rate and is therefore a nonhomogeneous Poisson process. The nonhomogeneity comes from the motion of the molecule.

The pmf of $q_k(t)$, denoted by $p_{q_k(t)}(m)$, gives the probability of having m photons arrived at pixel k until time t and can be written as

$$p_{q_k(t)}(m) = \frac{\Lambda_k(t)^m}{m!} e^{-\Lambda_k(t)}, \qquad (5)$$

where $\Lambda_k(t) = \int_0^t \lambda_k(\tau) \, d\tau = \lambda_0 \int_0^t p_k(\tau) \, d\tau + \lambda_{bg} t$. It is worth emphasizing that the probability functions $p_k(t)$ and therefore the pmf functions $p_{q_k(t)}(m)$ are functions of the unknown model parameters, $x_0, y_0, v_x, v_y, \lambda_0, \lambda_{bg}$.

Assuming that each pixel detects photons independently from other pixels, the joint pmf of the pixel photon counts at the end of the exposure time T can be written as the product of the marginals:

$$p_{q_1(T), q_2(T), \ldots, q_N(T)}(m_1, \ldots, m_N; \boldsymbol{\Theta})$$

$$= \frac{\Lambda_1(T)^{m_1} \Lambda_2(T)^{m_2} \ldots \Lambda_N(T)^{m_N}}{m_1! m_2! \ldots m_N!} e^{-\sum_{k=1}^{N} \Lambda_k(T)}, \qquad (6)$$

where $\boldsymbol{\Theta}$ is the parameter vector defined as $\boldsymbol{\Theta} = [x_0 \; y_0 \; v_x \; v_y \; \lambda_0 \; \lambda_{bg}]^T$, with subscript T denoting the transpose.

Equation (6) is the pmf needed for the maximum likelihood algorithm. It gives the probability of obtaining a particular image $\mathbf{m} = [m_1 \; m_2 \; \ldots \; m_N]^T$, for fixed T and $\boldsymbol{\Theta}$. However, it can also be read as the likelihood function by considering $\mathbf{m}$ as the independent variable and $\boldsymbol{\Theta}$ as the dependent variable.

The log-likelihood function, L, is then just the natural logarithm of equation (6), given by

$$L(\boldsymbol{\Theta}) = \sum_{k=1}^{N} m_k \log(\Lambda_k(\boldsymbol{\Theta}))$$

$$- \sum_{k=1}^{N} \log(m_k!) - \sum_{k=1}^{N} \Lambda_k(\boldsymbol{\Theta}). \qquad (7)$$

Equation (7) is the final result of the derivations in this section. Its first- and second-order derviatives with respect to $\boldsymbol{\Theta}$ (Supplementary Article) are used both in the resolution limit computation and the iterative maximum likelihood algorithm.

Supplementary Article

A Supplementary Article, entitled "Video-Based Tracking of Single Molecules Exhibiting Directed In-Frame Motion," can be found online. Please visit journals.cambridge.org/jid_MAM.

The log-likelihood function is maximized by Newton-Raphson method, for which the updates can be written as

$$\hat{\boldsymbol{\Theta}}^{(t+1)} = \hat{\boldsymbol{\Theta}}^{(t)} - \mathbf{g}^{(t)}, \quad t = 1, \ldots, \Omega, \qquad (8)$$

where $\hat{\boldsymbol{\Theta}}^{(t)}$ is the estimate of the model parameter vector at iteration t, $\mathbf{g}^{(t)}$ is the update vector at iteration t, whose k'th entry can be written as

$$g_k = \left(\frac{\partial^2 L(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}_k^2} \right)^{-1} \left(\frac{\partial L(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}_k} \right), \qquad (9)$$

and Ω is the number of iterations.

As a final remark to this subsection, we note that it is not possible to determine the actual sign of the velocity vector from single-frame data only, i.e., there is an inherent sign ambiguity. The algorithm yields the estimates of the initial position and the velocity vector, from which one can calculate the estimate of the final position. An alternative result with these initial and final position estimates interchanged and the sign of the velocity components reversed fits the data equally well. Deciding on the correct alternative requires additional information from neighboring frames. This is done during the high-level analysis.

Fisher Information Matrix and Resolution Limits

The Fisher information matrix (FIM) is an analytical tool, widely used in parameter estimation problems. It has the utility that its inverse defines a lower bound for the error covariance of all possible (unbiased) estimators. This bound,

known as the Cramer Rao lower bound (CRLB), is therefore an important indicator for the resolution limits of our imaging model. Its square root gives us the best achievable precision, with which we can estimate the model parameters from the image of a molecule in an unbiased way.

The FIM, J, for the parameter vector Θ is defined as

$$[J(\Theta)]_{ij} = -E\left[\frac{\partial^2 L(\Theta)}{\partial \Theta_i \partial \Theta_j}\right], \qquad (10)$$

where $E[\ \ ]$ denotes expectation, and L is the log-likelihood function. Substituting L from equation (7) gives

$$[J(\Theta)]_{ij} = -E\left[\sum_{k=1}^{N}\left(\left(\frac{m_k}{\Lambda_k(\Theta)} - 1\right)\frac{\partial^2 \Lambda_k(\Theta)}{\partial \Theta_i \partial \Theta_j}\right.\right.$$
$$\left.\left. - \frac{m_k}{\Lambda_k^2(\Theta)}\frac{\partial \Lambda_k(\Theta)}{\partial \Theta_i}\frac{\partial \Lambda_k(\Theta)}{\partial \Theta_j}\right)\right]. \qquad (11)$$

Using the independence of the pixel values, and recognizing that $E(m_k) = \Lambda_k(\Theta)$, we obtain the ij entry of $J(\Theta)$ to be

$$[J(\Theta)]_{ij} = \sum_{k=1}^{N}\frac{1}{\Lambda_k(\Theta)}\frac{\partial \Lambda_k(\Theta)}{\partial \Theta_i}\frac{\partial \Lambda_k(\Theta)}{\partial \Theta_j}, \qquad (12)$$

and write $J(\Theta)$ in a more suitable form as

$$J(\Theta) = \sum_{k=1}^{N}\frac{1}{\Lambda_k(\Theta)}\frac{\partial \Lambda_k(\Theta)}{\partial \Theta}\frac{\partial \Lambda_k(\Theta)^T}{\partial \Theta}. \qquad (13)$$

The derivatives in the FIM are evaluated at the true values of the parameter vector, meaning that the CRLB resolution limits we obtain are going to be different for different parameter values. The general expressions for the derivatives are provided in the Supplementary Article. For the special case of a stationary molecule (i.e., $v_x = v_y = 0$), we get

$$\frac{\partial \Lambda_k(\Theta)}{\partial x_0} = \lambda_0 \frac{\partial p_k}{\partial x_0} T,$$

$$\frac{\partial \Lambda_k(\Theta)}{\partial y_0} = \lambda_0 \frac{\partial p_k}{\partial y_0} T,$$

$$\frac{\partial \Lambda_k(\Theta)}{\partial v_x} = \lambda_0 \frac{\partial p_k}{\partial x_0}\frac{T^2}{2},$$

$$\frac{\partial \Lambda_k(\Theta)}{\partial v_y} = \lambda_0 \frac{\partial p_k}{\partial y_0}\frac{T^2}{2},$$

$$\frac{\partial \Lambda_k(\Theta)}{\partial \lambda_0} = p_k T,$$

and

$$\frac{\partial \Lambda_k(\Theta)}{\partial \lambda_{bg}} = T.$$

Based on these derivatives, we see that for all k,

$$\frac{\partial \Lambda_k(\Theta)^T}{\partial \Theta}\begin{bmatrix} T/2 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0 \quad \text{and} \quad \frac{\partial \Lambda_k(\Theta)^T}{\partial \Theta}\begin{bmatrix} 0 \\ T/2 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} = 0 .$$
$$(14)$$

Consequently, the Fisher information matrix given by equation (13) is rank deficient, with rank 4, for the generic stationary molecule scenario. The rank deficiency implies inherent ambiguity in resolving parameters corresponding to directions within the null space of the FIM.

On the other hand, the FIM is nonsingular in the directions orthogonal to the span of the null space vectors given in equation (14). For example, the FIM is nonsingular in the directions

$$\begin{bmatrix} 1 \\ 0 \\ T/2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ T/2 \\ 0 \\ 0 \end{bmatrix}, \qquad (15)$$

corresponding to midframe x- and y-positions [i.e., $x_c = x_0 + v_x(T/2)$, and $y_c = y_0 + v_y(T/2)$], which means that there is no ambiguity in estimating these parameters.

The problem of singularity of the FIM for the stationary molecule case is transformed to the problem of ill-conditionedness in the relatively slow (near stationary) motions. This manifests itself as high estimation errors in the linear combination of model parameters nearly in the directions defined by the span of the vectors in equation (14). However, the analyses in the Results and Discussion section show that the eigenvalues of the FIM corresponding to the midframe positions are located at the high end of the eigenspread with the large disparity. This observation implies that, at almost all speeds, the midframe position vector $[x_0 + v_x(T/2)\ \ y_0 + v_y(T/2)]^T$ can almost always be estimated accurately, despite the potential inaccuracies in the individual position and velocity parameters. This argument is, of course, only valid when the in-frame motion length is less than the window size. The accuracy of the midframe location is exploited by the interframe Kalman filter based estimation/tracking algorithm described in the next subsection.

Kalman Filtering

The Kalman filter is a recursive data processing algorithm used in estimating the state of a linear dynamic system from noisy measurements. In our context, the system corresponds to the moving molecule, measurements correspond to the in-frame estimations, and state is a set of parameters

that describe molecule's motion. We describe our Kalman filtering equations, following the same notation as Kailath et al. (2000): Lower-case characters are used to show vectors and written in a bold font for random variables. Matrices are denoted by upper-case italic characters. The subscript k indicates the discrete time steps, corresponding to frame numbers.

The state-space model is given by

$$\mathbf{x}_k = F_{k-1}\mathbf{x}_{k-1} + \mathbf{n}_{k-1}$$
$$\mathbf{y}_k = H\mathbf{x}_k + \mathbf{v}_k, \qquad (16)$$

where $\mathbf{x}$ is the state vector, F is the transition matrix, $\mathbf{n}$ is the process noise, $\mathbf{y}$ is the measurement vector, H is the observation matrix, and $\mathbf{v}$ is the measurement error.

We define the state vector to be

$$\mathbf{x} = \begin{bmatrix} x_c \\ y_c \\ v_x \\ v_y \end{bmatrix}. \qquad (17)$$

The transition matrix, under the constant uniform motion assumption of the imaging model, is given by

$$F = \begin{bmatrix} 1 & 0 & (T+T_r) & 0 \\ 0 & 1 & 0 & (T+T_r) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad (18)$$

where T_r is the read-out time. We have a fully observable system, so H is the identity matrix.

The process noise $\mathbf{n}$ and the measurement error $\mathbf{v}$ are assumed to be normally distributed around 0 mean with covariance matrices Q and R, respectively. Q and R are diagonal matrices, with their elements given by the typical variations in the actual and measured values respectively of the parameters in $\mathbf{x}$. We assume these variations to remain constant throughout the motion.

In the following, we provide the operations carried at a frame k. Due to the recursive nature of the Kalman filtering, these operations involve predictions from (to) the previous (next) frame, denoted by subscript $k|k-1$ $(k+1|k)$, and related to their current estimates, denoted by subscript $k|k$, through measurement (time) update equations. The operations start with the measurement to obtain $\mathbf{y}_k$. Then the measurement innovation vector, $\mathbf{e}_k$, and its covariance matrix, $R_{\mathbf{e},k}$, and subsequently the Kalman gain matrix, $K_{f,k}$, are calculated according to

$$\mathbf{e}_k = \mathbf{y}_k - H\hat{\mathbf{x}}_{k|k-1}, \qquad (19)$$
$$R_{\mathbf{e},k} = HP_{k|k-1}H^T + R, \qquad (20)$$
$$K_{f,k} = P_{k|k-1}H^T R_{\mathbf{e},k}^{-1}. \qquad (21)$$

Next, the state vector estimate $\hat{\mathbf{x}}$ and the error covariance matrix P are updated by the measurement-update equations:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_{f,k}\mathbf{e}_k, \qquad (22)$$
$$P_{k|k} = P_{k|k-1} - K_{f,k}R_{\mathbf{e},k}K_{f,k}^T. \qquad (23)$$

And finally the time-update equations predict their values for the next frame:

$$\hat{\mathbf{x}}_{k+1|k} = F_k\hat{\mathbf{x}}_{k|k}, \qquad (24)$$
$$P_{k+1|k} = F_k P_{k|k} F_k^T + Q. \qquad (25)$$

Since $\hat{\mathbf{x}}_{k|k-1}$ and $P_{k|k-1}$ will not be present for $k = 1$, they must be initialized with typical values before the algorithm starts.

As mentioned earlier in the Maximum Likelihood Approach for In-Frame Parameter Estimation section, single frame data alone are not sufficient to determine the correct sign of the velocity vector. Therefore while updating our Kalman filter, we also have to check for the sign of the velocity vector. We do this by calculating Kalman updates of both alternatives, and comparing the Frobenius norms of the resulting windows in the next frame. If the norms are different by more than 1.5 times (a user-defined arbitrary factor), we conclude that the one with lower norm misses the molecule and therefore is the wrong choice. If the difference does not reach this threshold (i.e., the molecule still remains within the window even with wrong choice of sign, as can happen for slow motions or large windows), then we run the in-frame estimation in the next frame for both window positions and choose the alternative that yields better agreement between the time updated position and the estimation result.

RESULTS AND DISCUSSION

In this section we evaluate the performance of our tracking algorithm at the low and high levels described in the Algorithm section. In testing the low-level analysis (in-frame estimation), we first use CRLB calculations to find the smallest achievable error in estimating parameters of the proposed imaging model under typical experimental conditions. We then apply our estimation algorithm to simulated images of a moving molecule and compare its success with the theoretical outcomes of the CRLB calculations. For testing the high-level analysis (tracking), we use actual experimental data recorded as described in the Materials and Methods section. We compare the molecule trajectories obtained from analysis of the movies with the ones obtained from the piezo-stage sensor.

Low-Level Analysis

Starting with the in-frame estimation, we first consider the case where the molecule is moving with a uniform translational motion along the positive x-direction ($v_x \neq 0, v_y = 0$). The molecule speed is taken to be 3.11 μm/s and corresponds to 7 pixel/frame for $T = 0.2$ s. A window size of 15×15 pixels is used, and λ_0 and λ_{bg} are taken to be 15,000 and 300 photons/s, representing the typical values we obtained in our experiments.
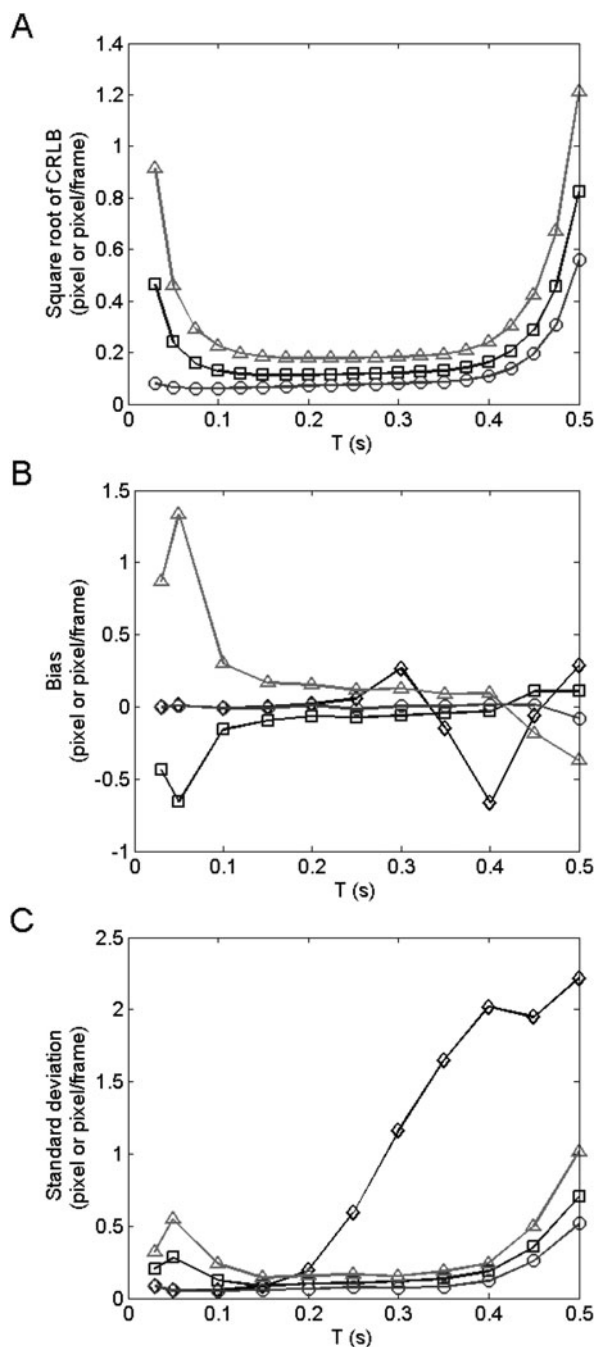
Figure 2. Performance of the proposed estimation algorithm as a function of T for parameters v_x (green triangles), x_0 (blue squares), and x_c (red circles). Tracking performance is evaluated on the basis of (**A**) CRLB limits, (**B,C**) bias and standard deviation of estimates from simulated images. For comparison, corresponding characteristics of x_m obtained from same set of simulated images are also shown (black diamonds). For all presented data, v_x is kept fixed at 3.11 μm/s (7 pixel/frame at $T = 0.2$ s).

Figure 2A shows the CRLB limits for parameters x_0, v_x, and x_c as a function of T, calculated from the FIM [equation (13)]. Except for the extreme T values, the CRLB limits increase only slightly with the increasing exposure time. The divergence on the long T end is merely a finite window size effect: As the exposure time is increased, a molecule with a

certain speed travels more pixels per frame, and its smeared image eventually extends out of the estimation window. The rise in the CRLB of x_0 and v_x, at short exposure times is a result of the ill-conditioned problem of the FIM, discussed in the Algorithm section. In the short T limit, the image of the molecule becomes less smeared, or equivalently less dependent on v_x. This causes a redundancy in parametrization that appears as a correlation between x_0 and v_x and renders their individual estimations subject to high uncertainties. On the other hand, x_c intrinsically achieves better precision by merging the two correlated parameters ($x_c = x_0 + v_x T/2$).

Figures 2B and 2C show the same type of analysis using simulated data. All the parameters of moving molecules in the simulations were set to the same values as for Figure 2A, and a set of 50 images was generated for each T to assess the accuracy and precision of the position and velocity estimations. For the parameter estimation, the window center was placed at (x_c, y_c) position, and the maximum likelihood algorithm was run for 400 iterations. Due to the uncertainty in determining direction of motion mentioned in the Materials and Methods section, the estimated v_x sometimes had the reverse sign with respect to the true molecule velocity used in the simulation. For such cases we changed the sign of the estimated velocity vector and replaced the initial position with the final one. To compare the performance of our SMT algorithm to standard tracking schemes that do not take in-frame motion of the molecule into account, simulated images of moving molecules were also processed with a modified version of the algorithm that assumed a stationary molecule and, thus, only provided estimates of the mean in-frame position of the molecule x_m.

In Figure 2B, the mean of the estimates from the 50 generated images is compared with the actual values used in the simulation, and their difference, the bias, is interpreted as the accuracy of the estimations. We see that the accuracy of x_c is notable, with <0.05 pixel (<4.4 nm) bias for all $T <$ 0.5 s, whereas estimations of both x_0 and v_x are biased. If the stationary molecule algorithm is applied, estimates of the molecule mean position x_m become biased at moderate values of $T > 0.25$ s, where smearing of the molecule image due to its motion becomes significant. The observed bias of the estimates oscillates between positive and negative values due to the combination of increasing distortion and decreasing signal-to-noise ratio (SNR) of smeared images (Cheezum et al., 2001).

Similarly, Figure 2C compares the standard deviation of x_0, v_x and x_c estimates (interpreted as the precision) with the CRLB limits of Figure 2A and shows that the two are very close for T values from 0.1 to 0.4 s. The window effect on the long T side is almost identical to that of Figure 2A, whereas the diverging behavior of the uncertainty in v_x and x_0 estimations on the short T side disappears to a large extent. This is attributed to the presence of bias in estimations of these parameters from the simulated data (see Figure 2B), as opposed to the unbiased estimator assumption of the CRLB analysis. Further discussion on the discrep-

ancy at the short T end is going to be provided following the next figure, where a more direct connection can be made with the Algorithm section. Estimation of x_c performs the best also in terms of precision, with <0.08 pixel (<7.1 nm) standard deviation for all $T \leq 0.4$ s. Again, precision of estimates of x_m that assume a stationary molecule decreases significantly for $T > 0.25$ s, where it becomes worse than both the x_c and x_0 precisions.

All of these observations are in agreement with the relatively abstract interpretations of the FIM. Analysis of both CRLB predictions and numerical simulation results brings us to the conclusion that it is more advantageous to describe position with the midframe values (x_c, y_c) instead of initial ones (x_0, y_0), and rely more on the position estimation, than on the velocity estimation. In fact, the whole parametrization in this study could have been done in the form

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \begin{bmatrix} v_x \\ v_y \end{bmatrix}(t - T/2) \quad t \in [0, T]; \quad (26)$$

however, we decided to keep our initial-point-based parametrization to attract more attention to this interesting result. The comparison of x_c with x_m clearly shows that the implicit assumption of molecule motion during image acquisition improves the accuracy and precision of the molecule localization. As expected, this tendency becomes more pronounced for higher integration times T, where the image distortion due to molecule motion increases.

As the next conclusion from Figure 2, we note that both accuracy and precision of estimation only slightly depend on T. This suggests that in a real experiment, the choice of a suitable T must be based on other criteria rather than the performance of the estimation algorithm. Working with a short exposure time would be beneficial for several reasons: First and most importantly, this would reinforce validity of the uniform in-frame motion assumption of our imaging model. Second, the Kalman filter would be updated more frequently. Third, the resulting trajectory would have a higher time resolution. It is true that the in-frame velocity estimation becomes worse for short T. However, our main goal here is to determine the position of the molecule as a function of time, and the in-frame velocity is just an auxiliary parameter that helps in achieving better resolution in position (see Fig. 2), and predicting where to place the estimation window in the next frame. If the molecule velocity on the time scale of T is needed, it can be determined more precisely from the difference of midframe positions of consecutive images in the recorded stack.

Based on these conclusions, we carry out a similar analysis for a range of velocities and fixed T. This represents the condition in a real experiment. Figure 3A shows the CRLB limits in x_c and v_x estimations as a function of v_x values from 0.5 to 15 pixel/frame for $T = 50$ ms ($v_x = 0.89$ to 26.67 μm/s), and a window of 15×15 pixels. λ_0 and λ_{bg} are again taken to be 15,000 and 300 photons/s. The maximum error within the studied velocity range turns out to be <1.5 pixel/frame (<2.67 μm/s) for v_x, and <0.5 pixel
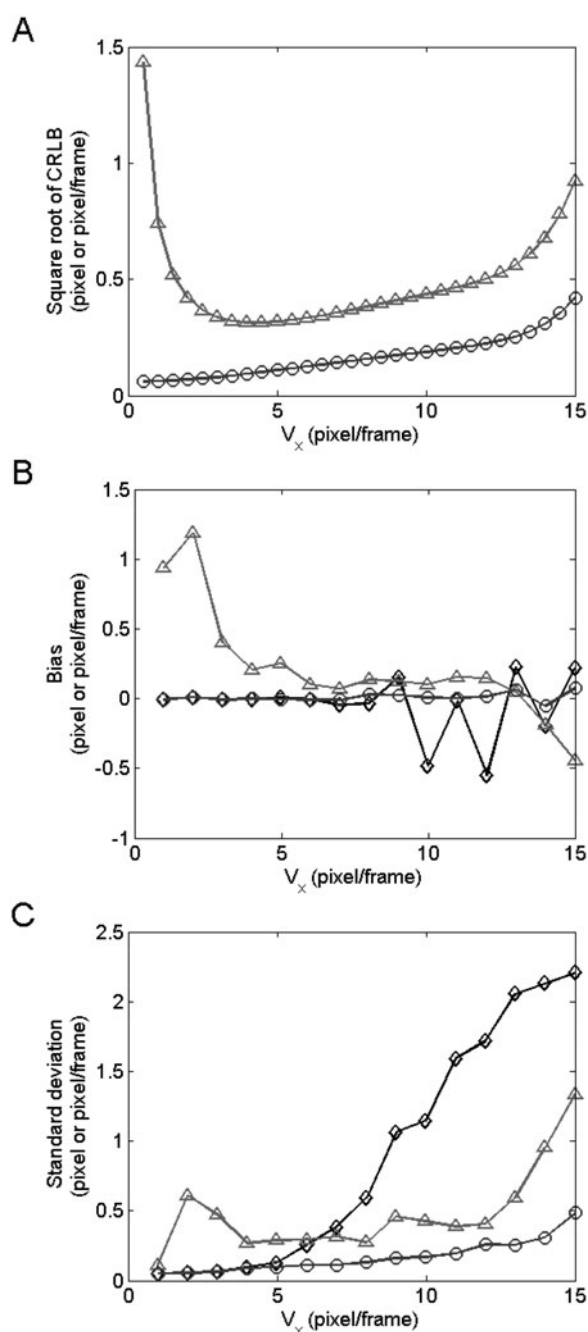

Figure 3. Performance of the proposed estimation algorithm as a function of v_x for parameters v_x (green triangles), and x_c (red circles), based on their (**A**) CRLB limits, (**B,C**) bias and standard deviation of estimates from simulated images. For comparison, corresponding characteristics of x_m obtained from same set of simulated images are also shown (black diamonds). T is kept fixed at 50 ms for all presented data.

(<44.4 nm) for x_c. The estimation algorithm performs fairly well for both parameters, except for an increase in the bias in v_x estimation for $v_x \leq 3$ pixel/frame, and the finite window effect in Figure 3C. For x_c estimation, the bias is <0.1 pixel (8.9 nm) for all v_x, whereas the precision is better than 0.2 pixel (17.8 nm) up to $v_x = 12$ pixel/frame. The negligible bias and the standard deviation comparable to

the CRLB limit indicate that, for $T = 50$ ms and typical λ_0 and λ_{bg} values, the midframe position (x_c, y_c) can be estimated as accurately and precisely as its theoretical limit for a wide range of molecule velocities. Similar to Figure 2, accuracy and precision of the estimates of x_m drop significantly with increasing v_x, underlining again the benefits of the newly proposed tracking algorithm over the traditional stationary-molecule schemes.

The growing behavior in the CRLB bound for the v_x parameter (see Fig. 3A) as velocity decreases (in $v_x < 3$ pixels/frame region) can be explained by the increasing condition number of the Fisher information matrix. As previously noted in the Algorithm section, FIM is singular for $v_x = 0$ and it is close to singular for low velocity values. The corresponding high values of CRLB implies that being unbiased is a too stringent condition to pose on the estimator for low velocities (Usman et al., 1993; Eldar, 2008). In fact, as shown in Figure 3B, the estimation algorithm automatically imposes a bias in this region to reduce standard deviation, which is shown in Figure 3C. For the practical purposes what matters is the actual mean square error between the true parameter and the estimated value, which can be written as

$$E[(v_x - \hat{v}_x)^2] = \text{bias}(\hat{v}_x)^2 + \text{variance}(\hat{v}_x). \qquad (27)$$

Based on Figures 3B and 3C, we can deduce that the variance-bias trade-off implied by equation (27) is adjusted by the estimation algorithm to avoid the potential high variance values for an unbiased estimator (reflected by the CRLB curve) at low velocities.

High-Level Analysis

So far, the discussions have focused on the performance of the in-frame (low-level) analysis of our tracking algorithm. Using these results, we now continue with the piezo-stage experiments, where the high-level Kalman filtering part is also used.

In these experiments, we compare the measured $(\mathbf{y}_k)$ and filtered $(\hat{\mathbf{x}}_{k|k})$ position and velocity data of a molecule with their actual values $(\mathbf{x}_k)$ obtained from the position sensor readings. We calculate the actual velocity from position sensor data at a 200 times lower rate (50 Hz), to discard the high-frequency noise in the position data. The in-frame estimations were done with 400 iterations for the first frame, and with 100 iterations afterward, where the starting values in a frame were taken as the results of the previous one (Kalman time updates for x_c, y_c, v_x, v_y; and in-frame estimates for λ_0, λ_{bg}). The operation of Kalman filtering is related to the covariance matrices Q and R. In forming Q, we consider the expected deviations from the uniform motion assumption and use approximate values reflecting such deviations. Working with higher process noise variance values tolerates more deviations from the assumed model, but it also results in less accurate estimates. In a real experiment, we can only enter approximate values for the process noise variance, but as shown in Wu et al. (2010), the result will not change much, unless an order of magnitude
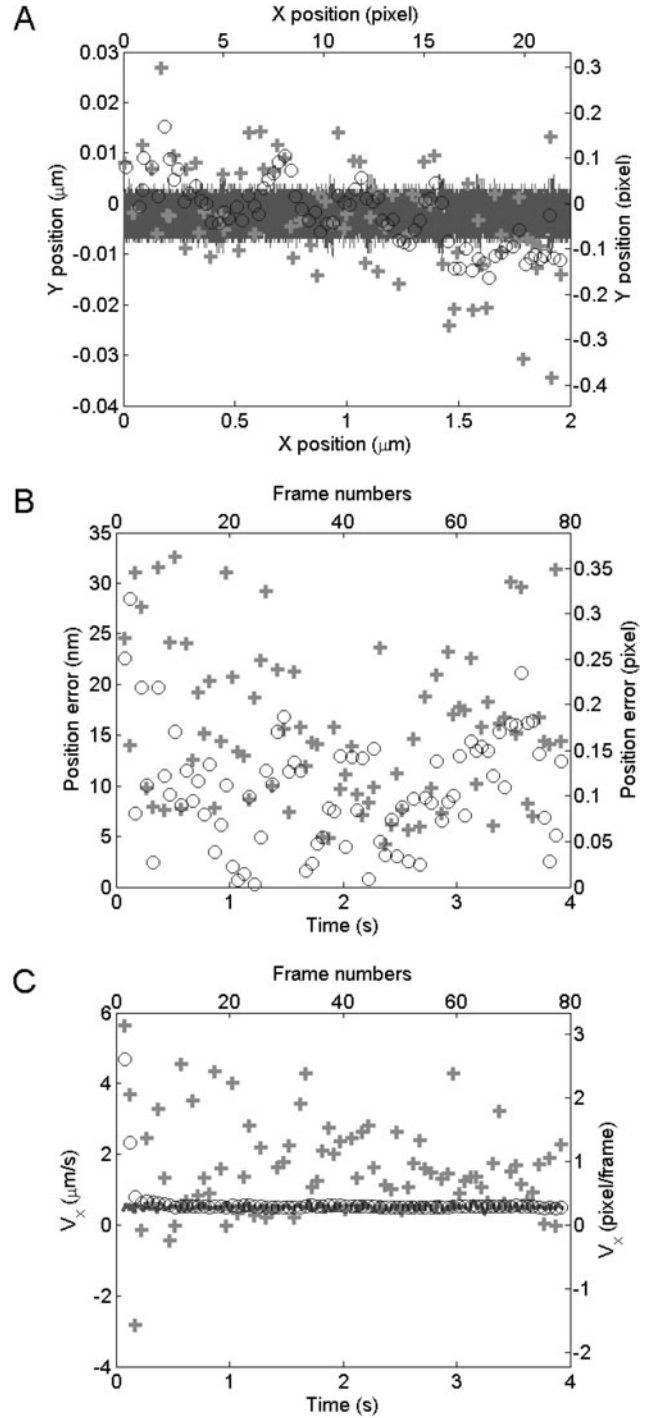

Figure 4. A representative result from the piezo-stage experiments, showing (**A**) the overall linear trajectory, (**B**) the measurement error, and (**C**) v_x data for the actual (red line), measured (green pluses), and Kalman-filtered (blue circles) values.

error is made. For the elements of R, we use the sum of squares of the typical biases and standard deviations in position and velocity estimation suggested by Figure 3 $[0.2^2 + 0^2 = 0.04 \text{ pixel}^2 = 316 \text{ nm}^2$ for position, and $0.2^2 + 0.4^2 = 0.2 \text{ (pixel/frame)}^2 = 0.63 \text{ } (\mu\text{m/s})^2$ for velocity].

Figure 4A shows the case for a linear motion along the x-axis. The movie was recorded with $T = 50$ ms and

analyzed using a window of 15 × 15 pixels. Owing to the perfect agreement between the assumed and actual motions, we used a small process noise variance of 1 nm^2 for position and (1 nm/0.05 s)^2 for velocity. In comparing position data, a constant (but unknown) offset (D_x, D_y) between the position of the analyzed molecule and the origin of the stage position sensor readings had to be eliminated to bring the two to the same origin. This was done by taking the average over k of the differences between the filtered and actual positions (i.e., $D_x = \langle \hat{\mathbf{x}}_{k|k}[1] - \mathbf{x}_k[1] \rangle_k$, and $D_y = \langle \hat{\mathbf{x}}_{k|k}[2] - \mathbf{x}_k[2] \rangle_k$), and subtracting it from the filtered/measured positions. After this necessary shift, we see that both the measured and filtered trajectories agree well with the actual one.

Figure 4B plots the position error against time (or frame numbers) for the measured and filtered trajectories. The position error is defined as $\sqrt{\Delta x^2 + \Delta y^2}$, where Δ denotes the difference of the measured/filtered values from the actual values at the midframe times. We see that both measured and filtered data reveal the actual motion within average errors of 0.17 pixel (15.54 nm) and 0.11 pixel (9.65 nm), respectively. Considering the difficulty in comparing the values from image analysis with the ones from the real system, this error likely involves other contributions from the shifting procedure, or from slight misalignments or astigmatism within the imaging setup. The level of accuracy obtained here is therefore quite satisfactory for a real experiment. We further see that the errors in measured and filtered trajectories do not differ significantly. This is explained by the in-frame position estimations' precision already being comparable to the fluctuations in the sensor reading. Since the sensor reading itself has a standard deviation of 0.03 pixel (2.6 nm) in the (stationary) y-axis (see Fig. 4A), error in the filtered data cannot be smaller than $\sqrt{2} \times 0.03 = 0.04$ pixel (3.6 nm).

The situation with the velocity measurement, however, is different. In this case, our in-frame estimations have a larger error than the real velocity fluctuations, and the effect of Kalman filtering becomes clear. This is shown in Figure 4C, where the v_x estimates of the same molecule are plotted together with their filtered and actual counterparts. We see that Kalman filtering settles quickly within first three frames and corrects the noisy estimations. A similar trend would be observed in filtered position estimations also, when working with images of a worse SNR.

In addition to the linear motion case presented above, we also wanted to test our algorithm for a more general type of motion. Figure 5 summarizes the results of tracking a molecule moving along a sinusoidal trajectory of $y = 2\sin((2\pi/2 \ \mu m)x)$ with the period of 1 s. As in the case of linear motion, the movie was recorded with $T = 50$ ms, analyzed using a window of 15 × 15 pixels, and the offset between the measured and actual position origin was corrected. A higher process noise variance [100 nm^2 for position, and 4 (μm/s)^2 for velocity] was used to compensate the difference between the actual and assumed motions. This resulted in less effective filtering as compared to the
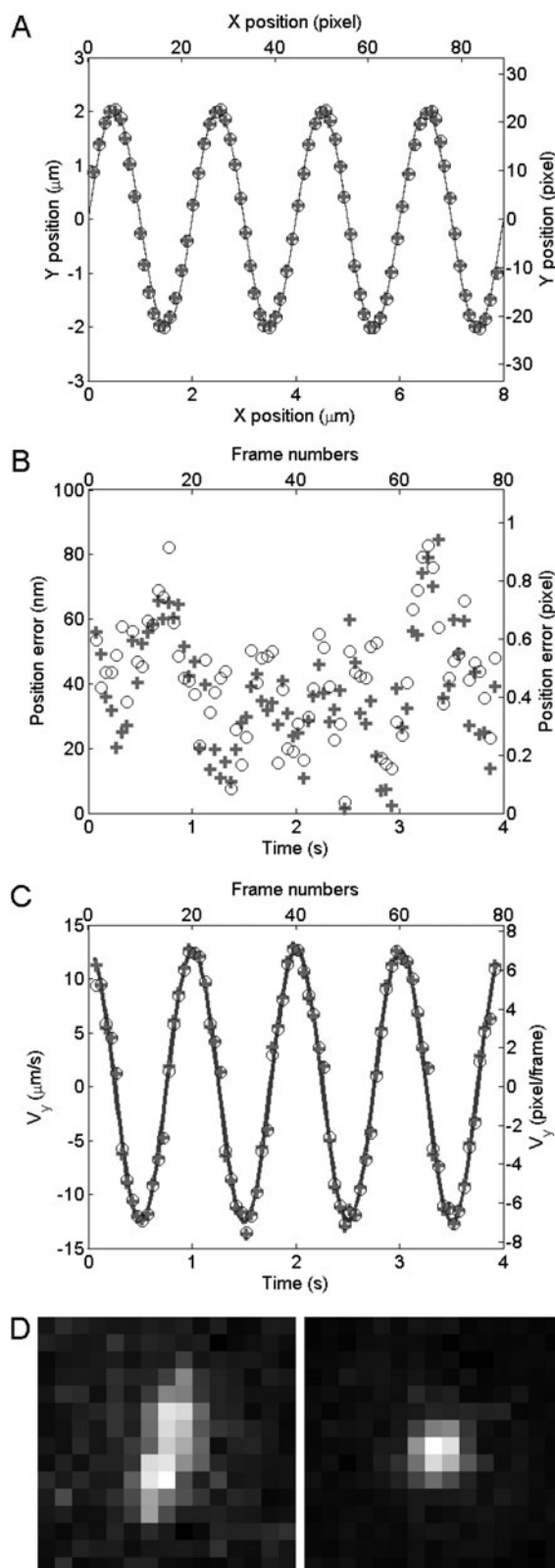

Figure 5. Result from the piezo-stage experiments with accelerated motion, showing (**A**) the overall sinusoidal trajectory, (**B**) the measurement error, and (**C**) v_y data for the actual (red line), measured (green pluses), and Kalman-filtered (blue circles) values. The two images in image **D** correspond to representative maximum-velocity (left) and maximum-acceleration (right) instants occurring at 19th and 25th frames, respectively.

linear motion experiment. In fact, the Kalman filtering was almost completely turned off, as evidenced by the coinciding filtered and measured values (see Fig. 5A for position and Fig. 5C for velocity). In other words, the whole tracking algorithm worked based on the raw in-frame estimations. Despite the absence of corrections from Kalman filtering, we see in Figure 5B that the error in position estimation still remains smaller than 1 pixel (90 nm) and has an average value of 0.42 pixel (37.43 nm). Although this error is higher than the linear motion case, it must be considered together with the level of velocity and acceleration present in this motion. The maximum values of the (time-varying) velocity and acceleration along y-direction were 7 pixel/frame (12.6 μm/s) and 2.2 pixel/frame^2 (79 μm/s^2), respectively. Two 15 $\times$ 15 pixel images from one of the maximum velocity and maximum acceleration instants are provided in Figure 5D to demonstrate the appreciable distortion. This illustrates the potential advantage of our algorithm in analyzing fast molecules. Furthermore, the algorithm performs sufficiently well even for accelerated motion that violates the assumptions of the used imaging model.

Conclusions

In this work, we introduced a new model of image formation for mobile incoherent point sources of light (e.g., single molecules) and developed an algorithm for retrieving the position and velocity of such light sources from a recorded stack of images. The core of the tracking algorithm is based on in-frame estimation of parameters of a moving light source that is further refined by Kalman filtering of the raw estimates. We illustrated that the new imaging model provides a better precision and accuracy of position estimates of a moving molecule, in comparison to the standard imaging schemes assuming a stationary light source.

We were able to track molecules under uniform motion with an average position error of 0.11 pixel (9.65 nm). We also applied our algorithm to accelerated motion, a case that does not match with the assumptions of our model, and still obtained reasonable success. We achieved tracking of molecules moving with a speed as fast as 7 pixels/frame (12.6 μm/s), and acceleration as high as 2.2 pixel/frame^2 (79 μm/s^2) within a mean error of 0.42 pixel (37.43 nm). These results can be improved by including in-frame acceleration into the proposed algorithm.

The Kalman filtering demonstrated here can also be improved by including λ_0 and/or the in-frame acceleration of the molecule to the state vector. This would further expand the applicability of the algorithm to more complex modes of motion. Similarly, updating the measurement error variances adaptively, depending on λ_0 or velocity, can be an effective remedy for blinking molecules.

The main area of application of the presented tracking procedure lies in studying directed molecular transport, for example in living cells, fluid flows within microfluidic systems, or optical potential landscapes. In principle, stochastic Brownian motion of a molecule can be accounted for either on the level of in-frame estimation by evaluating apparent broadening of the observed point spread function (Schuster et al., 2002), or in the Kalman filter by considering the influence of the random position fluctuations of the molecule on the measured trajectory variance (Wu et al., 2010). Even more general imaging models will become possible with the advent of imaging detectors that can deliver arrival times of individual photons. Although the presented imaging model is two-dimensional, our approach can be readily generalized to three dimensions with a suitable form of the point spread function.

Acknowledgments

References

Abraham, A.V., Ram, S., Chao, J., Ward, E.S. & Ober, R.J. (2009). Quantitative study of single molecule location estimation techniques. *Opt Express* **17**, 23352–23373.

Abraham, A.V., Ram, S., Chao, J., Ward, E.S. & Ober, R.J. (2010). Comparison of estimation algorithms in single-molecule localization. *Proceedings of SPIE*, vol. 7570, 757004-1.

Anthony, S.M. & Granick, S. (2009). Image analysis with rapid and accurate two-dimensional Gaussian fitting. *Langmuir* **25**, 8152–8160.

Cheezum, M.K., Walker, W.F. & Guilford, W.H. (2001). Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophys J* **81**, 2378–2388.

Eldar, Y.C. (2008). Rethinking biased estimation: Improving maximum likelihood and the Cramer-Rao bound. *Found Trends Signal Process* **1**, 305–449.

Evans, E. (2001). Probing the relation between force—lifetime—and chemistry in single molecular bonds. *Annu Rev Biophys Biomol Struct* **30**, 105–128.

Genovesio, A. & Olivo-Marin, J.-C. (2003). Tracking fluorescent spots in biological video microscopy. *Proceedings of SPIE*, vol. 4964, pp. 98–105.

Hamamatsu. (2009). EM-CCD Technical Note. Technical report. Hamamatsu Photonics K.K., Systems Division.

Ho, W. (2002). Single-molecule chemistry. *J Chem Phys* **117**, 11033–11061.

Hoffmann, A., Nettels, D., Clark, J., Borgia, A., Radford, S.E., Clarke, J. & Schuler, B. (2011). Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: Recurrence analysis of single particles (RASP). *Phys Chem Chem Phys* **13**, 1857–1871.

Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. & Ha, T. (2008). Advances in single-molecule fluorescence methods for molecular biology. *Annu Rev Biochem* **77**, 51–76.

Kailath, T., Sayed, A.H. & Hassibi, B. (2000). *Linear Estimation*. Upper Saddle River, NJ: Prentice Hall.

Kay, S.M. (1993). *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory*. Upper Saddle River, NJ: Pearson Education.

Kues, T., Peters, R. & Kubitscheck, U. (2001). Visualization and tracking of single protein molecules in the cell nucleus. *Biophys J* **80**, 2954–2967.

Moerner, W.E. (2007). New directions in single-molecule imaging and analysis. *Proc Nat Acad Sci USA* **104**, 12596–12602.

Ober, R.J., Ram, S. & Ward, E.S. (2004). Localization accuracy in single-molecule microscopy. *Biophys J* **86**, 1185–1200.

Pampaloni, F., Lattanzi, G., Jonas, A., Surrey, T., Frey, E. & Florin, E.L. (2006). Thermal fluctuations of grafted microtubules provide evidence of a length-dependent persistence length. *Proc Nat Acad Sci USA* **103**, 10248–10253.

Pralle, A., Keller, P., Florin, E.L., Simons, K. & Horber, J.K.H. (2000). Sphingolipid-cholesterol rafts diffuse as small entities in the plasma membrane of mammalian cells. *J Cell Biol* **148**, 997–1007.

Saxton, M.J. & Jacobson, K. (1997). Single-particle tracking: Applications to membrane dynamics. *Annu Rev Biophys Biomol Struct* **26**, 373–399.

Schmidt, T., Schutz, G.J., Gruber, H.J. & Schindler, H. (1996). Local stoichiometries determined by counting individual molecules. *Anal Chem* **68**, 4397–4401.

Schuster, J., Cichos, F. & Borczyskowski, C. (2002). Diffusion measurements by single-molecule spot-size analysis. *J Phys Chem A* **106**, 5403–5406.

Schutz, G.J., Sonnleitner, M., Hinterdorfer, P. & Schindler, H. (2000). Single molecule microscopy of biomembranes. *Mol Membr Biol* **17**, 17–29.

Smal, I., Niessen, W. & Meijering, E. (2008). A new detection scheme for multiple object tracking in florescence microscopy by joint probabilistic data association filtering. *Int Symp Biomed Imag: From Nano to Macro (ISBI-08)*, Paris, France. pp. 264–267.

Smith, C.S., Joseph, N., Rieger, B. & Lidke, K.A. (2010). Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nat Methods* **7**, 373–375.

Thompson, R.E., Larson, D.R. & Webb, W.W. (2002). Precise nanometer localization analysis for individual fluorescent probes. *Biophys J* **82**, 2775–2783.

Usman, M., Hero, A., Fessler, J. & Rogers, W. (1993). Bias-variance tradeoffs analysis using uniform CR bound for a SPECT system. *Nuclear Science Symposium and Medical Imaging Conference*, 1993 IEEE Conference Record. vol. 3, pp. 1463–1467.

Wong, Y., Lin, Z.P. & Ober, R.J. (2011). Limit of the accuracy of parameter estimation for moving single molecules imaged by fluorescence microscopy. *IEEE Trans Signal Process* **59**, 895–911.

Wu, P.H., Agarwal, A., Hess, H., Khargonekar, P.P. & Tseng, Y. (2010). Analysis of video-based microscopic particle trajectories using Kalman filtering. *Biophys J* **98**, 2822–2830.

Yorulmaz, M., Kiraz, A. & Demirel, A.L. (2009). Motion of single terrylene molecules in confined channels of poly(butadiene)-poly(ethylene oxide) diblock copolymer. *J Phys Chem B* **113**, 9640–9643.