A Framework for Generating Realistic Synthetic Sequences of Dynamic Confocal Microscopy Images

William T. E. Pitkeathly
wtp628@bham.ac.uk
PSIBS DTC
University of Birmingham
Edgbaston, UK

Seyed Hamid Rezatofighi
hamid.rezatofighi@anu.edu.au
College of Engineering and Computer Science
The Australian National University,
ACT 0200, Australia

Joshua Z. Rappoport
j.rappoport@bham.ac.uk
School of Biosciences
University of Birmingham
Edgbaston, UK

Ela Claridge
E.Claridge@cs.bham.ac.uk
School of Computer Science
University of Birmingham
Edgbaston, UK

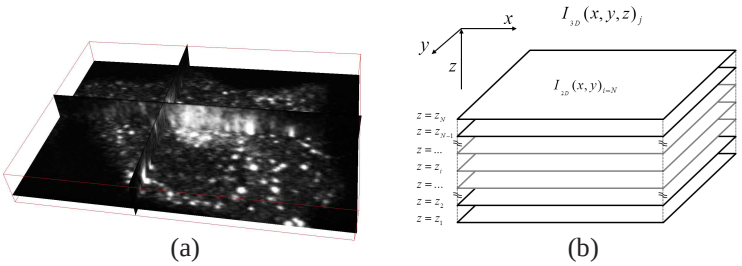
Abstract

In recent years many automated methods for detection and tracking of sub cellular structures in live cell fluorescence microscopy have been proposed. Because dependable ground truth from real data sets is difficult to obtain, most algorithms are tested on synthetic data where the ground truth is known. Differences between real and synthetic data sets can lead to imprecise judgement about an algorithm's performance. In this paper we present a method for generating realistic synthetic sequences of live cell confocal microscopy images that simulate the image formation as well as modelling the motion of dynamic structures during image acquisition using valid dynamic models. Sequences generated using this framework realistically reproduces the complexities existing in real confocal microscopy sequences.

1 Introduction

Confocal microscopy is a fluorescence microscopy technique used for imaging sub cellular structures in three dimensions (3D). Its optical sectioning capability gives confocal microscopes a much higher resolution along the axial (z) direction than conventional fluorescence microscopy, as well as a slightly superior lateral (x-y) resolution [10]. High speed confocal microscopy is particularly well suited for imaging intracellular traffic, such as vesicle dynamics as part of the endocytic pathway. Such studies involve acquiring large amounts of 3D data (tens to thousands of images) which typically contain ~ 10 to $\sim 10^3$ dynamic fluorescent vesicles which appear as spot like features. Analysis of such large amounts of

data via manual inspection is a painstaking and subjective process which has motivated the development of automated analysis techniques such as [5, 8]. Because the ground truth of the trajectories of features in real images is not available, the performance of automated detection and tracking techniques is quantitatively evaluated using synthetic data. The mathematical models for the dynamics of vesicles motion have been fairly well established and are easily simulated [2, 8]. However, current image simulations used for validation are based on simplistic assumptions about the imaging system and the sample features. For example, most simulations use fixed-shape Gaussian distributions to represent the particles of interest, use either a constant background or simple background structures, and have a spatially constant signal to noise ratio (SNR). In recent work [7] a frame work for generating synthetic



(a) (b)

Figure 1: (a) Orthogonal views of a single 3D confocal stack showing fluorescently tagged vesicles. (b) The schematic of a confocal stack array.

sequences of total internal reflection fluorescence (TIRF) microscopy has been presented. These synthetic TIRF images are very realistic because of the following advances: the image formation process of a TIRF microscope is simulated; valid dynamic models for vesicle motion are used; shape deformation of vesicles in motion is modelled; and spatio-temporal varying background extracted from real TIRF image sequences is used. In this paper we use the same concepts for modelling the image formation process of a confocal microscope. In particular, the microscope image formation model presented properly incorporates motion artefacts caused by vesicle motion during image acquisition, which are seen in real microscopy images and are not accounted for in other simulations.

2 Methods

2.1 Confocal Microscopy

Whereas conventional fluorescence microscopes illuminate the whole sample at one time, confocal microscopes use highly focussed laser light to locally illuminate the sample in order to minimise the illumination volume. Any light emitted from the sample and back towards the objective lens is passed through a pinhole aperture in the back focal plane of the optical system before falling onto a photosensitive detector [10]. The purpose of the pinhole is to limit the observation volume to the small region at the focal point and thus to prevent light from out-of-focus planes from reaching the detector. This permits what is known as ‘optical sectioning’ of a sample. It is this capability that gives confocal microscopes the advantage over conventional epi-fluorescence microscopes; because both the illumination volume and the observation volume are confined to the focal point of the objective lens which results in an overall better resolution, especially in the z-direction.

2.2 Confocal Measurement Model

The image formation process for confocal microscopy can be described mathematically by a convolution of a function describing the object being imaged, f, with a function describing the point spread function (PSF) of the system, h, and then distorting the image with the appropriate noise model. In 3D live cell confocal microscopy, a 3D space is sampled through series of parallel 2D rectangular x-y planes, at a set number of equally spaced intervals in the z direction. Usually, when describing the image formation of 3D microscopy systems, the function describing the object, f, is assumed to be a static three-dimensional $f = f(x,y,z)$ function with no time component [1, 11]. In this paper where we are modelling dynamic sub-cellular motion, the image formation model needs to also consider the evolution of the function during the time that the image is being acquired, hence $f = f(x,y,z,t)$. The point spread function of the objective lens, $h_{obj}(u,v)$, is modelled using an analytical expression for the diffraction pattern of light through a circular pupil with a perfect aberration free lens [3]. The point spread function (PSF) for confocal microscopes is approximately equal to the point spread function of the objective lens squared $h_{conf}(u,v) = \left|h_{obj}(u,v)\right|^2$, this is due to the fact that both the illumination and observation volumes are reduced to a diffraction limited sized spot [10].

$$h_{conf}(u,v) = \left|h_{obj}(u,v)\right|^2 = \left(\left|2\int_0^1 P(\rho)J_0(\rho v)\exp(iu\rho^2/2)\rho d\rho\right|^2\right)^2 \tag{1}$$

where: $u = 2\pi NA^2 z/\lambda$; $v = 2\pi NAr/\lambda$; $r = \sqrt{x^2+y^2}$; $P(\rho)$ is a pupil function with an aperture radius R, $J_0(.)$ is a first order Bessel function; $\rho = r/R$; NA is the numerical aperture of the lens; λ is wavelength excitation light used. This function is isotropic in the $x-y$ plane and anisotropic in the $y-z$ and $x-z$ planes giving it a characteristic 'bobbin' shape along the z-axis. Typically the physical resolution of a confocal microscope is, $r_{xy} \simeq 200nm$, and r_z = 0.6 to 2 μm, depending on the physical resolution of the diffractive optics of the particular microscope, as well as the diameter of the pin hole (which can be varied on most systems) [10].

Recorded images represent discretized point intensity measurements of the sample space in digitized array form. A complete data set for a live cell imaging experiment consists of J, 3D image stacks, $I_{3D}(x,y,z)_j$, $j = \{1,2,..,J\}$. A 3D confocal image stack, consists of a set of, N, 2D image slices $I_{2D}(x,y)_i$ slices corresponding to different cross sectional z planes through the sample, z_i, $i = \{1,2,..,N\}$. A single 2D slice, $I_{2D}(x,y)_i$, is represented by a 2D digital array of size n by m pixels. Often, in an attempt to increase the signal to noise ratio (SNR) of a 2D image slice, several scans of the same cross section of a sample are made in quick succession and the average of the point intensity measurements are used for pixel values, this is referred to as 'slice averaging'.

Because the function, $f = f(x,y,z,t)$, is constantly changing over time, an image of a 2D slice formally corresponds to the summation of, n_s, of four dimensional (4D) convolutions of the function space centred at z_i where n_s is the number of complete scans of the laser across the 2D cross section of the sample. The image formation model therefore becomes:

$$I_{2D}(x,y)_i\,|_{z=z_i} = \frac{1}{n_s}\sum_{s=1}^{n_s}\int_t^{t+t_e}\int\int\int f(x',y',z',t)\cdot h(x-x',y-y',z_i-z')\,dx'dy'dz'dt \tag{2}$$

We also denote the following relevant temporal variables: the instant before the first slice of stack, j, is acquired as T^j_{stack}; the time taken to complete a laser scan across a 2D x-y slice as t_e; the time taken to move the sample along the optical axis between adjacent image slices

(from z_i to z_{i+1}) as t_s; the time taken to move the sample from the z position of the last slice of the stack, z_N, to the position of the first slice in the next image stack z_1, as t_f. Therefore the instant before acquiring, slice, i, in stack j is $t_i^j = T_{stack}^j + (i-1)\cdot(t_e \cdot n_s + t_s)$.

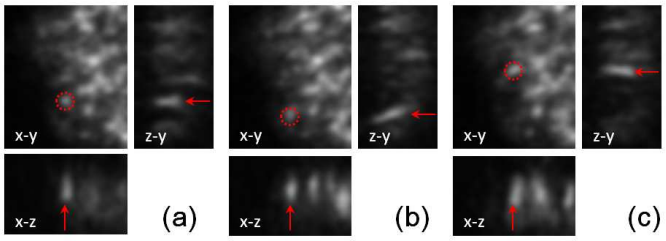


Figure 2: Examples of motion artefacts during image acquisition of a sub resolution vesicle on 3 consecutive time frames of a confocal microscope. For each time point the $x-y$ image shows an average projection of the images slices in the z-direction. (a) The vesicle of interest stays relatively stationary during the acquisition of the image stack, therefore its appearance in the $x-z$, and $z-y$ views, is the classical 'bobbin' shape of the confocal point spread function. In (b) the same vesicle moves significantly in the y direction during the acquisition of adjacent image slices; as a result the of the 'bobbin' appearance is sheared in the z-y plane. In (c) the vesicle stays relatively stationary again, and shows a similar appearance as in (a).

2.3 Dynamic Models

Movement of vesicles within cells is either due to diffusion in the cytosol, or along microtubules via motor proteins. The non-linear stochastic motion of vesicles can be simulated using a mixture of two linear dynamic models: random walk, nearly constant velocity with small accelerations [8]. Switching between these models resembles the tethering and docking, and linear motion of vesicles as they are trafficked within the cell [2]. The state of particle k is defined by the vector $x_t^k = [x, v_x, y, v_y, z, v_z]^T$, which describes the particles position and velocity in each dimension at time t. The particles state changes over time according the linear Gaussian model: $x_t^k = F x_{t-1}^k + \mathcal{N}(0,Q)$, where: $F = diag[F_i, F_i, F_i]$ is the dynamic model, and $\mathcal{N}(0,Q)$ is a zero mean Gaussian white noise process with covariance $Q = diag[Q_i, Q_i, Q_i]$, $i = \{1,2\}$, $i = 1$ for random walk, and $i = 2$ for constant velocity; Where q_1 and q_2 are constants which control the noise levels.

$$F_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q_1 = q_1 \begin{pmatrix} T^2 & 0 \\ 0 & T^2 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}, \quad Q_2 = q_2 \begin{pmatrix} \frac{T^3}{3} & \frac{T^2}{2} \\ \frac{T^2}{2} & T \end{pmatrix} \tag{3}$$

2.4 Image Simulation

In order to produce a sequence of simulated images first a set of the trajectories of vesicles are generated using the two dynamic models defined above. The temporal sampling factor is chosen as the smallest temporal variable of the system, t_e or t_s, as defined above. Since the typical slice scan speed, t_e, for high speed confocal microscopes is $\sim 10^{-2}$ to 10^{-3} seconds, and the maximum velocity of vesicles is $\sim 1\mu ms^{-1}$, we can assume that the that the sample is approximately stationary during the time a single slice scan is performed, therefore equation 2 changes to:

$$I_{2D}(x,y)_i|_{z=z_i} = \frac{1}{n_s}\sum_{s=1}^{n_s}\int\int\int f(x',y',z')|_{t=t_i^j+(s-1)t_e} \cdot h(x-x', y-y', z_i-z')\,dx'dy'dz' \tag{4}$$

The vesicles positions are plotted in at the time point of each slice scan in a 3D array $f(x',y',z')|_{t=t_i^j+(s-1)t_e}$, and then convolved as in equation 2.4. f is updated with the new particles positions for every slice scan. The trajectories of the simulated vesicles are confined to the region defined by an estimated background of a cell, which has been extracted from a real image sequence. Background structures can be estimated using the MPHD method [6]. This method is based on greyscale image reconstruction from mathematical morphology; it identifies spot like peaks and 'cuts' them off at the background level [6]. The result of the convolution is added the background image to form the final simulated image.

Fluorescence microscopy images are corrupted by a mixture of Poisson and Gaussian noise. The main source of noise is photon counting noise, and is governed by a Poisson distribution $\mathscr{P}(\cdot)$. An additional source of noise is read noise, which is an additive Gaussian process, $\mathscr{N}(\mu,\sigma^2)$, with mean μ and standard deviation σ. $\alpha > 0$ is the detector gain. A noisy image is thus represented as:

$$I(x,y,z,t)_n = \alpha\mathscr{P}(I(x,y,z,t)) + \mathscr{N}(\mu,\sigma^2) \tag{5}$$

3 Results

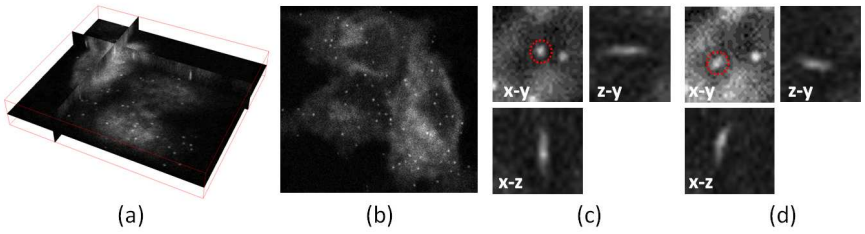


Figure 3: A simulated mage: (a) Orthogonal views of a single 3D stack of a simulated image with a background extracted from a real confocal image. (b) A single 2D slice from the 3D stack. (c) and (d) demonstrate motion artefacts from 2 consecutive image stacks (x-y views are average z-projections) (c) The highlighted vesicle is relatively stationary. (d) Shows the next time frame where the simulated vesicle moving at $\sim 1\mu s^{-1}$. The motion artefacts resemble those in Figure 2.

A sequence of images was produced using the proposed method with the parameters based on those from a real sequence acquired using a high-speed resonance scanner confocal microscope (see Figure 3). The PSF was generated using Equation 1 for a lens with numerical aperture (NA) of 1.49, using an excitation wavelength of $520nm$. Vesicles were simulated as sub resolution 3D ellipses with a diameter of $\sim 20nm$. The lateral (x-y) and axial (z) pixel resolution is $0.2\mu m$. The number of z-slices, $z_N = 20$, with $n_s = 2$ scans per slice. The temporal variables (in seconds): $t_e = 1/60$, $t_s = 1/60$ and $t_f = 1/60$. For vesicle motion dynamics: the temporal sampling factor $T = t_e$, $q_1 = 0.9$ and $q_2 = 0.7$. The maximum possible vesicle velocity was set to $\sim 1\mu s^{-1}$. Vesicles were allowed to switch between dynamics as in [7]. The background used was extracted from a sequence of real images using the MPHD method as in [6] using 5 frame temporal averaging.

4 Discussion

This paper has presented an accurate model for the image formation process of dynamic confocal images. Because the state of each particle, x_t^k, is known for the duration of any

sequences, they can be used to determine the detection and tracking accuracy of automated detection and tracking algorithms. For quantitatively evaluating the performance of a detection algorithm, metrics such as: true positive rate (TPR), false positive rate (FPR) and others as presented in [9] can be used. For tracking accuracy the root mean squared (RMS) error between any tracks produced by a tracking algorithm and those of the known trajectories of each particle can be calculated. RMS can either be calculated using only the particles position, or the state vectors directly if tracking is perfromed in state space (like in most probabilistic algorithms). An additional/alternative similarity measure for detection/tracking accuracy is the Jaccard similarity index [4].

The resulting simulations provide more realistic ground truth for validation of particle detection and tracking than has been previously proposed. In future work we intend extend the framework to model the dynamics of larger structures such as endosomes.

References

[1] David A Agard. Optical sectioning microscopy: cellular architecture in three dimensions. *Annual review of biophysics and bioengineering*, 13(1):191–219, 1984.

[2] J.G. Burchfield, J.A. Lopez, K. Mele, P. Vallotton, and W.E. Hughes. Exocytotic vesicle behaviour assessed by total internal reflection fluorescence microscopy. *Traffic*, 11:429–439, 2010.

[3] V.F. Canales, P.J. Valle, J.E. Oti, and M.P. Cagigal. Variable resolution with pupil masks. *Opt. Commun.*, 257(2):247–254, 2006.

[4] Nicolas Chenouard. *Advances in probabilistic particle tracking for biological imaging*. PhD thesis, PhD thesis, T él écom ParisTech, 2010.

[5] Auguste Genovesio, Tim Liedl, Valentina Emiliani, Wolfgang J Parak, Maité Coppey-Moisan, and J-C Olivo-Marin. Multiple particle tracking in 3-d+ t microscopy: Method and application to the tracking of endocytosed quantum dots. *Image Processing, IEEE Transactions on*, 15(5): 1062–1070, 2006.

[6] Seyed Hamid Rezatofighi, Richard Hartley, and William E Hughes. A new approach for spot detection in total internal reflection fluorescence microscopy. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 860–863, 2012.

[7] Seyed Hamid Rezatofighi, William TE Pitkeathly, Stephen Gould, Richard Hartley, Katarina Mele, William E Hughes, and James G Burchfield. A framework for generating realistic synthetic sequences of total internal reflection fluorescence microscopy images. In *Proc. IEEE Int. Symp. Biomed. Imag.*, pages 161–164, 2013.

[8] I. Smal, E. Meijering, K. Draegestein, N. Galjart, I. Grigoriev, A. Akhmanova, ME Van Royen, AB Houtsmuller, and W. Niessen. Multiple object tracking in molecular bioimaging by rao-blackwellized marginal particle filtering. *Med. Image Anal.*, 12(6):764–777, 2008.

[9] I. Smal, M. Loog, W. Niessen, and E. Meijering. Quantitative comparison of spot detection methods in fluorescence microscopy. *IEEE Trans. Med. Imaging*, 29(2):282–301, 2010.

[10] Robert H Webb. Confocal optical microscopy. *Reports on Progress in Physics*, 59(3):427, 1999.

[11] Tony Wilson. Confocal microscopy. *Academic Press: London, etc*, 426:1–64, 1990.