Partial coherence in microscopy and a userguide for `microlith`

Shalin B. Mehta
www.mshalin.com

March 26, 2014

1 Introduction

`microlith` is an image simulation software. The unique feature of `microlith` is its ability to accurately simulate partially coherent microscopy methods such as dark-field, phase-contrast, and differential interference contrast. It also provides accurate simulation of fluorescence contrast and coherent imaging system. In this document, I review role of partial coherence in microscopy[1], and how to use `microlith` for simulating images produced by coherent, partially coherent, and incoherent microscopes.

A manuscript describing the algorithms, accuracy, and applications of `microlith` is under review. If you use `microlith` for work related to a scientific publication, please cite our preprint: Shalin B. Mehta and Rudolf Oldenbourg, "`microlith`: image simulation for biological phase microscopy," arxiv:1305.7149 [phys.opt].

`microlith` is a portmanteau of words microscopy and lithography, since it can be used for simulating images of lithography masks as well. The dictionary meaning of `microlith` is "a small stone tool that has been sufficiently worked to be distinguishable from the workshop waste or accident." This software has similarly received some testing and usage so as to provide interesting and reliable results. But, there are *very likely* bugs in the implementation. Please treat simulations as a complement to real experiments and theoretical analysis, instead of as a substitute. The `microlith` package is provided 'as is' under GPLv3 license without any implied warranties. I welcome feedback and collaborative input.

2 Partial coherence in imaging

A modern microscopist would reckon that quantitative imaging is said to furnish an indispensable instrument in the treatment of nearly every recondite question in modern biology (with apologies to Lord Kelvin, who reckoned that the Fourier's theorem is said to furnish an indispensable instrument in the treatment of nearly every recondite question in modern physics). In biological microscopy, fluorescence contrast is the most popular form of contrast. The reason for popularity of fluorescence is not only its molecular specificity, but the linear image formation model that applies to fluorescence imaging. Under several experimental conditions, the image produced by a fluorescence microscope is the linear convolution of the specimen's fluorophore density with the point spread function of the system. This linearity has allowed researchers to exert computational control over fluorescence imaging with methods such as deconvolution for 3D reconstruction, structured illumination for super-resolution, and adaptive optics for aberration compensation.

The linear image formation in fluorescence is the consequence of the incoherence among fluorophores - fluorophores do not emit in synchrony even if they are excited in synchrony. Therefore, the total detected intensity is linear sum of intensities contributed by individual fluorophores. The same applies to imaging of stars - two stars do not emit in synchrony. Therefore, advances in anstronomical imaging are easily tranferred to biological fluorescence microscopy.

Imaging of specimens that are not fluorescent or self-luminous can be classified in following categories based on the type of illumination employed:

[1] the review part of this document is adapted from the first chapter of my thesis[1]

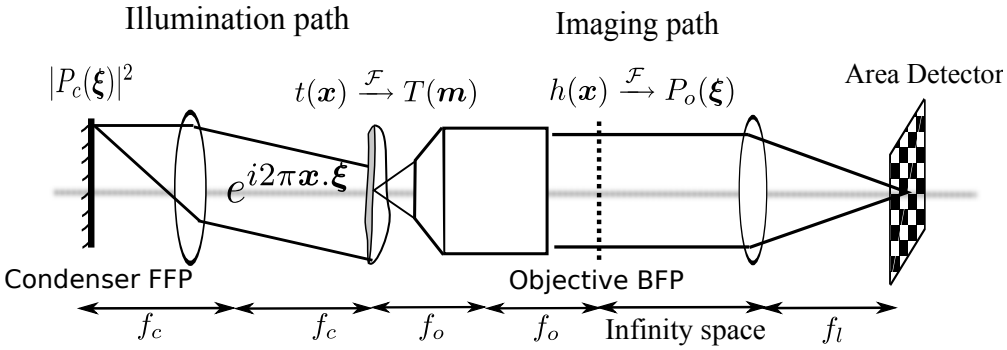


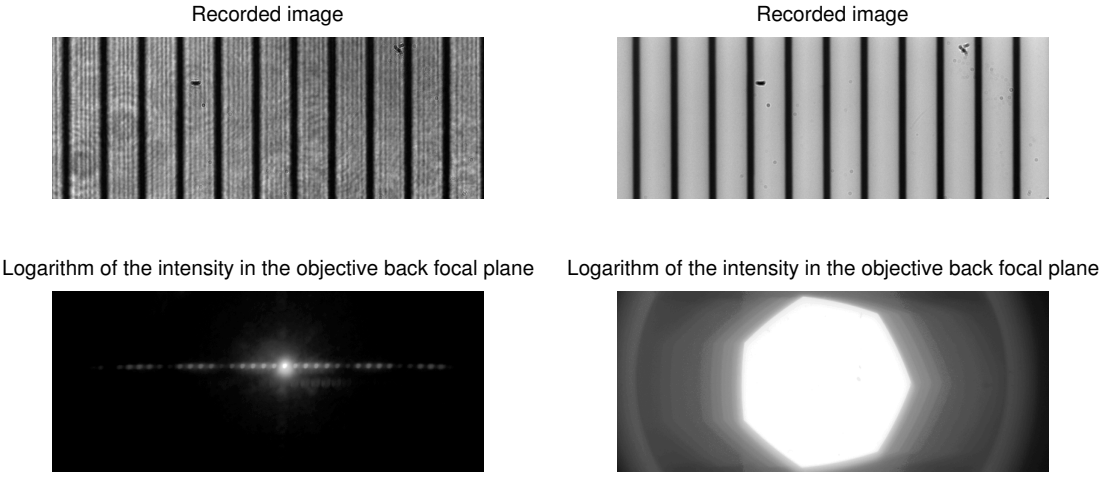
Figure 1: A general schematic of a partially coherent imaging system. The independent variables typeset in bold indicate two dimensional vectors. $t(\mathbf{x})$ is the specimen's transmission, $T(\mathbf{m})$ is the specimen's spectrum, $\mathrm{P_c}(\boldsymbol{\xi})$ is the pupil function of the condenser front focal plane (FFP), $\mathrm{P_o}(\boldsymbol{\xi})$ is the pupil function of the objective back focal plane (BFP), and $\mathcal{F}$ indicates Fourier transform. Each point of the condenser located at $\boldsymbol{\xi}$ illuminates the specimen with a plane-wave $e^{i2\pi \mathbf{x}.\boldsymbol{\xi}}$.

- **Spatially coherent imaging:** Spatially coherent imaging methods use illumination originating from a point source, such as, laser. After interaction with the specimen, the amplitude of spatial coherent light is linearly dependent on distribution of the absorption and optical path length (OPL) of the specimen. Coherent methods such as holography are designed to recover amplitude of the light from measured intensity after it has interacted with specimen – from which absorption, refractive index or thickness may be untangled if enough measurements are made. Although fast, single-view holography has poor 3D imaging performance. Specifically, the spatial-frequency coverage of single-view holography along the axial direction is inadequate to provide true 3D imaging [2]. One way to image 3D information is to introduce tomography, through either object rotation or illumination scanning [3, 4]. Coherent illumination is sensitive to imperfections in the light path, which exhibit themselves in the form of speckle noise and mottle. Moreover, the direct measurement of phase is wrapped and unwrapping procedure is sensitive to noise [5].

- **Spatially partially coherent imaging:** On the other hand, performance of phase imaging systems can be improved by employing *simultaneous* illumination from a large range of directions. Such an illumination can be achieved by using an incoherent source in conjunction with high-NA illumination optics, leading to partially coherent field at the specimen plane. Partially coherent methods, in contrast to coherent methods, produce an image which depends bi-linearly [2] on the specimen's transmission [6, 7, 8]. These methods can be designed to be sensitive to the gradient or curvature of the specimen phase, allowing retrieval of the phase via integration. Phase contrast, dark-field, and differential interference contrast are examples of partially coherent methods. Due to simultaneous illumination from several directions, a single partially coherent image provides better lateral resolution, axial resolution, and immunity to instrumental imperfections in comparison to a coherent image. Experimentally, partially coherent configurations are easier to setup as they use conventional sources such as halogen lamp, mercury-arc lamp, and light emitting diodes (LED). Thus, partially coherent imaging provides some experimental advantages of tomographic coherent imaging, without requiring several measurements. Nevertheless, the simultaneous illumination with several incoherent plane-waves breaks the linearity of image formation and makes it difficult to invert the imaging process to retrieve the specimen phase.

Figure 1 shows a general schematic of a partially coherent imaging system. The performance of the imaging system is described by the intensity of the illumination pupil $\mathrm{P_c}$ and the amplitude of the objective pupil $\mathrm{P_o}$. The so called coherence ratio $S = \mathrm{NA}_c/\mathrm{NA}_o$ determines 'how coherent'

[2]The term 'bi-linearity' implies that the image at a given point in the image space depends on integral over pairs of points in the specimen space.



(a) Nearly coherent illumination produced using a point aperture at the center of the condenser front focal plane

(b) Partially coherent illumination with a condenser aperture opened to nearly half of the objective aperture

Figure 2: Intensity in the image plane (top row) and in the objective back focal plane (bottom row) with (a) almost coherent and (b) partially coherent illumination. The specimen is a stage micrometer with 10μm spacing that is usually supplied with Zeiss microscopes. The spatially coherent illumination was created by placing a piece of paper with a hole at the center in the condenser turret. To avoid spectral blur of diffraction orders, the halogen lamp illumination was filtered through 546/10nm filter. I used 63X 1.4NA objective to capture the maximum possible number of diffraction orders. The experiment was carried out using Zeiss Axiovert 200M in collaboration with Rudolf Oldenbourg at Marine Biological Laboratory.

the image formation is and critically affects the imaging properties. For the spatially coherent imaging system, $S = 0$. As S increases, the coherence of the field that illuminates the specimen reduces.

Figure 2 illustrates the effect of illumination with images of a stage micrometer taken with nearly coherent and partially coherent illumination along with corresponding intensity distributions in the objective back focal plane (BFP). As seen from images of the micrometer, the partially coherent image is much better in terms of resolution, and signal to noise ratio (SNR). In particular, the coherent noise in the form of the fringes produced by the imperfections in the light-path are observed in the coherent image, but not in partially coherent image. The observable dynamic range of intensities in the back focal plane was much higher than what could be recorded by the available 12-bit camera. Therefore, imaging of back focal plane was carried out at several exposures – starting with the lowest exposure which properly imaged the brightest, i.e., un-diffracted order, and increasing the exposure by a factor of 2 sequentially. The images acquired at different exposures were combined into a single high dynamic range image. Note that increasing the exposure by a factor of 2 corresponds to acquiring additional 1-bit of dynamic information. The back focal plane images shown above are the logarithm of the high dynamic range data (24-bit) constructed from 12 low dynamic range images (each 12-bit).

It is instructive to compare the intensity distribution in the BFP between coherent and partially coherent illumination. In the coherent case, we clearly see the diffraction orders produced by the specimen. Note that the objective pupil $\mathrm{P_o}(\boldsymbol{\xi})$ filters the diffraction orders that contribute to the image, and hence acts as a low-pass filter. The intensity distribution in the BFP is nothing but the filtered power spectrum, $|T(\mathbf{m})\mathrm{P_o}(\mathbf{m})|^2$, of the specimen's transmission. Apart from the spectral peaks produced by the lines spaced apart at the distance of 10μm, we also observe a broader variation (which follows the sinc^2 pattern) due to finite width of the lines. The image amplitude (whose magnitude sqaure is the recorded intensity) is given by inverse Fourier transform of the filtered specimen spectrum, i.e., $I = |\mathcal{F}_m^{-1} T(\mathbf{m})\mathrm{P_o}(\mathbf{m})|^2$.

Unlike the coherent case, the diffraction orders overlap in the partially coherent case. The un-diffracted light (or zero spatial frequency of the specimen) produces the image of the condenser aperture at the center of the objective BFP, seen as the hexagonal shape in fig. 2b. Other spatial

frequencies ($\mathbf{m}$) of the specimen produce a shifted image of the condenser aperture, $\mathrm{P_c}(\boldsymbol{\xi}-\mathbf{m})$ in the BFP, each of which appears as a shifted hexagon in fig. 2b. Taking an alternate view, one may say each point of the hexagonal condenser aperture $\mathrm{P_c}(\boldsymbol{\xi})$ produces a shifted version of the specimen's spectrum $T(\mathbf{m}-\boldsymbol{\xi})$ in the objective BFP[3]. The inverse Fourier transform of the shifted and filtered specimen spectrum, $T(\mathbf{m}-\boldsymbol{\xi})\mathrm{P_o}(\mathbf{m})$, forms the image due to the condenser point $\boldsymbol{\xi}$. Since the image formation by a single condenser point is spatially coherent, the intensity at the image plane due to source point $\boldsymbol{\xi}$ can be written as,

$$I_{\boldsymbol{\xi}}(\mathbf{x}) = \left|\mathcal{F}^{-1}_{\mathbf{m}}[T(\mathbf{m}-\boldsymbol{\xi})\mathrm{P_o}(\mathbf{m})]\right|^2 \tag{1}$$

$$= \left|\int T(\mathbf{m}-\boldsymbol{\xi})\mathrm{P_o}(\mathbf{m})e^{2\pi i\mathbf{m}.\mathbf{x}}\mathrm{d}\mathbf{m}\right|^2$$

$$= \iint T(\mathbf{m_1}-\boldsymbol{\xi})T^*(\mathbf{m_2}-\boldsymbol{\xi})\mathrm{P_o}(\mathbf{m_1})\mathrm{P_o}^*(\mathbf{m_2})e^{2\pi i(\mathbf{m_1}-\mathbf{m_2}).\mathbf{x}}\mathrm{d}\mathbf{m_1}\mathrm{d}\mathbf{m_2} \tag{2}$$

In above equations, the amplitude spectrum of the image is linearly dependent on the amplitude spectrum of specimen. However, the intensity spectrum consists of *mixing* of pairs of spatial frequencies in the amplitude spectrum of the specimen. If above relationships are expressed in space domain (as shown in Ref. [9, Sec.21.3.1]), the intensity at each point in the image is seen to be dependent on pairs of points of the specimen amplitude. Such a dependence of the image intensity on amplitude of the specimen is called 'bi-linear' (i.e., linear in pairs) dependence.

In partially coherent imaging, mutual incoherence of the condenser points leads to the incoherent superposition of the intensity due to each point. Therefore, the total image intensity is given as,

$$I(\mathbf{x}) = \int |P_c(\boldsymbol{\xi})|^2 I_{\boldsymbol{\xi}}(\mathbf{x})\mathrm{d}\boldsymbol{\xi} \tag{3}$$

Due to superposition of intensities, the observed intensity is not linear in the specimen transmission. However, bi-linear dependence of the intensity on specimen transmission still holds as can be seen by substituting eq. 1 in eq. 3:

$$I(\mathbf{x}) = \int |P_c(\boldsymbol{\xi})|^2 \iint T(\mathbf{m_1}-\boldsymbol{\xi})T^*(\mathbf{m_2}-\boldsymbol{\xi})\mathrm{P_o}(\mathbf{m_1})\mathrm{P_o}^*(\mathbf{m_2})e^{2\pi i(\mathbf{m_1}-\mathbf{m_2}).\mathbf{x}}\mathrm{d}\mathbf{m_1}\mathrm{d}\mathbf{m_2}\mathrm{d}\boldsymbol{\xi} \tag{4}$$

The condenser points located far from the optical axes bring the information about the higher spatial frequencies of the specimen within the imaging pupil. Therefore, partially coherent imaging provides higher resolution (more information about the specimen) than does coherent illumination. In the above equations, the contributions of the specimen and the system are not separated. However, a change of variable $\mathbf{m}-\boldsymbol{\xi}=\mathbf{m'}$ separates the contribution of the specimen spectrum and the system pupils, leading to the notion of the transmission cross-coefficient (TCC) introduced by Hopkins [6].

When imaging is performed with polychromatic illumination, the recorded intensity is affected by following factors:

- Detector sensitivity
- Source spectrum
- Specimen dispersion (or lack thereof)
- Chromatic aberration in the imaging path

In the linear regime of light-matter interaction, image formation at each wavelength is necessarily incoherent with the image formation at another wavelength. Therefore, sum over source concept extends to the wavelength dimesion.

Due to experimental advantages of high resolution, high light throughput, high signal to noise ratio, and ease of setup, partially coherent methods such as DIC have been popular for wide range of biological applications in optical [10, 11] microscopy. Recently, X-ray microscopy [12, 13, 14] has also seen a surge of interest in partially coherent phase imaging with laboratory sources. In optical lithography, engineering of partially coherent illumination has long been pursued to push

[3]Phase-modulation of $e^{i2\pi\mathbf{x}.\boldsymbol{\xi}}$ imposed by the condenser point $\boldsymbol{\xi}$ at the specimen plane implies spectral shift in the BFP.

the finest feature size that can be etched [15, 16]. For all of these fields, it is attractive to develop an accurate image simulation tool. `microlith` provides such a tool for optical microscopy and lithography.

3 How to use `microlith`?

In the following, we assume that the user wants to simulate partially coherent imaging. For the simpler case of simulating incoherent (fluorescent) imaging, please see the help in the software. To obtain quick overview of the steps in simulation, please try out the testbenches included in the package.

3.1 Overview

The inputs to the `microlith` toolbox are, 1) complex transmission of the specimen $t(x)$, 2) complex transmission of the objective pupil that accounts for the aberrations and optical processing $\mathrm{P_o}(\boldsymbol{\xi})$, and 3) intensity distribution of the condenser pupil $\mathrm{P_c}(\boldsymbol{\xi})$. The specimen transmission is sampled over spatial grid measured in μm. The user can choose between one of the standard contrast configurations and specify parameters such as objective numerical aperture (NA_o), condenser numerical aperture (NA_c), wavelength (λ) and other contrast dependent parameters from which the matrices representing the objective and condenser pupil are computed by the toolbox. The user interacts with the toolbox in physical coordinates. Within the toolbox, all spatial variables are expressed in normalized units of λ/NA_o and all spatial-frequency variables are expressed in units of NA_o/λ. The user can specify arbitrary pupils in normalized co-ordinates for simulating custom contrast method or a lithography system. For lithography systems, $t(x)$ is the complex transmission of the mask whose aerial image can be computed with `microlith`.

Several MATLAB scripts prefixed `TestBench` are included in the distribution to illustrate how to use `microlith` and to point out how to setup accurate simulation. Also try `>> doc microlith` to get an overview of the class. Following are some important points to note:

- For each source point $\boldsymbol{\xi}$, `microlith` calculates the inverse Fourier transform shown in eq. 1 using FFT algorithm. To avoid spurious blurring among features near the edge of the simulation grid, it is important to ensure that the simulation grid extends beyond all light transmitting features by at least the width of the image of a point under given settings.

- Since in most real-life image simulation problems, the undiffracted light (corresponding to the DC term in the spectrum of the specimen) is a major contributer to the final image, it is crucial to setup the frequency grid such that DC term is sampled. Since the frequency grid is symmetric in positive and negative frequencies, this condition implies that the frequency grid is of odd length. In current implementation, size of the frequency grid is the same as the size of the spatial grid. Therefore, spatial grid sampled with odd length vectors provides more accurate results than even length vectors.

- The user interacts with `microlith` in physical coordinates (μm, real numerical aperture etc.), except when specifying a custom microscope for simulation. Within `microlith` simulation is performed in normalized optical coordinates. If you are specifying a custom microscope, it is important to know that the the optical units used in `microlith`are optical units defined by Born and Wolf [17] divided by 2π. In the units used by Born and Wolf, the first zero of airy disk occurs at $0.61 \times 2\pi = 3.83$ and the first on-axis minima occurs at normalized defocus of 4π. In the optical units used by `microlith`, the first zero of airy disk occurs at 0.61 and the axial minima at 2. These normalizations make it possible to accurately sample the minima of the point spread function on a coarser grid as compared to the optical coordinates used in Born and Wolf.

3.2 Installation

`microlith` is implemented as a MATLAB class. To install, just add the source-code directory and its sub-directories (that contains `@microlith` folder, other helper functions, and some testbenches) to your MATLAB search path. Several scripts that show usage of `microlith` are included, their filenames begin with `TestBench...`. These scripts can be modified to simulation

problem of interest to the user. The scripts are commented to point out how to ensure accuracy of simulated results and some interesting facts.

3.3 Step-wise guide to simulation with quasi-monochromatic illumination

Each function (or method) in `microlith` class has been documented with help. Type `>>help microlith` to obtain the usage information about the class. There are essentially three steps in simulating any image at given wavelength.

Specify the grid and allocate memory The main `microlith` function (called constructor) accepts discrete grid axes (along transverse and axial directions) as inputs and allocates the memory required for simulation. Type `>>help microlith/microlith` for detailed help.

Generate pupils of the imaging system based on specified parameters The `computesys` function accepts a string variable specifying the contrast method and a structure specifying the parameters to calculate the objective and condenser pupils that specify the imaging system.

Type `>>help microlith/computesys` for help regarding supported contrast methods and corresponding parameters.

Compute the image from specified transmission The `computeimage` function accepts a 2D matrix (representing specimen transmission) as an input and computes the final intensity image, which will be a 3D matrix if the grid axis along the axial direction is a vector. For `'coherent'` systems, image amplitude (rather than intensity) is returned. Otherwise, intensity image is returned. Type `>>help microlith/computeimage` for detailed help.

Compute 3D mutual coherence Another function to note is `computecoherence`, which computes 3D complex mutual coherence in the specimen volume from the condenser pupil by employing the idea of "generalized source" proposed by McCutchen [18]. Type `>>help microlith/computecoherence` for detailed help.

3.4 Polychromatic illumination

You can signal your intention to simulate polychromatic image formation by supplying a vector instead of a scalar for wavelength variable. The vector is assumed to contain wavelengths at which the image simulation should be performed. You can specify the detector sensitivity and source spectrum as a function of wavelength vector (both are normalized internally). One of the two approaches is taken for simulating polychromatic image.

1. **Specimen is not dispersive and chromatic aberration due to imaging system is not complex:** You can also specify chromatic aberrations of the imaging system by including `ChromaticShift` and `ChromaticScale` vectors as a function of the wavelength vector. The chromatic shift and scaling are easily assessed using multi-color sub-resolution beads. In this situation, the polychromatic image is simply a shifted (by ChromaticShift), scaled (due to wavelength and ChromaticScale), and weighted (due to detector sensitivity and source spectrum) sum of an image at central wavelength. This computation is relatively fast. This level of complexity should suffice when simulating images of cells with well-corrected optics.

2. **Specimen is dispersive or chromatic aberration is more complex:** To account for dispersion or wavelength dependent complex aberrations, you can specify specimen transmission and imaging pupils at each wavelength. In this case, the image simulation time extends by the number of wavelengths at which the simulation is performed.

References

[1] S. B. Mehta, "Phase-space perspective of partially coherent imaging systems: Applications to Biological Phase Microscopy," Ph.D. Thesis, National University of Singapore, Nov. 2010. http://scholarbank.nus.edu.sg/handle/10635/19182

[2] S. S. Kou and C. J. R. Sheppard, "Imaging in digital holographic microscopy," *Optics Express*, vol. 15, pp. 13640–13648, Oct. 2007.

[3] S. Vertu, J. J. Delaunay, I. Yamada, and O. Haeberlé, "Diffraction microtomography with sample rotation: influence of a missing apple core in the recorded frequency space," *Central European Journal of Physics*, vol. 7, no. 1, p. 22–31, 2009.

[4] S. S. Kou and C. J. R. Sheppard, "Image formation in holographic tomography: high-aperture imaging conditions," *Applied optics*, vol. 48, no. 34, pp. 168–175, 2009.

[5] D. C. Ghiglia and M. D. Pritt, *Two-dimensional phase unwrapping: theory, algorithms, and software*. Wiley New York, 1998.

[6] H. H. Hopkins, "On the diffraction theory of optical images," *Proc. R. Soc. Lond. A*, vol. 217, pp. 408–432, May 1953.

[7] C. J. R. Sheppard and A. Choudhury, "Image formation in the scanning microscope," *J. Mod. Opt.*, vol. 24, no. 10, pp. 1051–1073, 1977.

[8] T. Wilson and C. J. R. Sheppard, *Theory and Practice of Scanning Optical Microscope*. Academic Press, London, 1984.

[9] W. Singer, M. Totzeck, and H. Gross, *Handbook of Optical Systems, vol 2, Physical Image Formation*. Wiley, New York, 2005.

[10] S. Inoué and K. R. Spring, *Video microscopy : the fundamentals*. New York: Plenum Press, 2nd ed., 1997.

[11] S. Inoué, "Windows to dynamic fine structures, then and now.," *FASEB J*, vol. 13 Suppl 2, pp. S185–S190, Dec. 1999.

[12] O. von Hofsten, M. Bertilson, and U. Vogt, "Theoretical development of a high-resolutiondifferential-interference-contrast optic for x-raymicroscopy," *Opt. Express*, vol. 16, pp. 1132–1141, Jan. 2008.

[13] A. Olivo and R. Speller, "A coded-aperture technique allowing x-ray phase contrast imaging with conventional sources," *Applied Physics Letters*, vol. 91, p. 074106, 2007.

[14] R. A. Lewis, "Medical phase contrast x-ray imaging: current status and future prospects," *Physics in Medicine and Biology*, vol. 49, no. 16, pp. 3573–3583, 2004.

[15] F. M. Schellenberg, *Resolution Enhancement Techniques in Optical Lithography*, vol. MS 178 of *SPIE Milestone Series*. 2004.

[16] A. K.-K. Wong, *Resolution Enhancement Techniques in Optical Lithography*. SPIE Publications, Mar. 2001.

[17] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge: Cambridge University Press, 7th expanded ed ed., 1999.

[18] C. W. McCutchen, "Generalized source and the van cittert-zernike theorem: A study of the spatial coherence required for interferometry," *JOSA*, vol. 56, no. 6, pp. 727–732, 1966.