19

Miniature Dual Axes Confocal Microscope for Real Time *In Vivo* Imaging

Wibool Piyawattanametha and Thomas D. Wang
[1]*Departmentsof Applied Physics, Biology, Electrical Engineering Microbiology & Immunology, Radiology, and Pediatrics James H. Clark Center (Bio-X), Stanford,*
[2]*NECTEC, Pathumthani,*
[3]*Faculty of Medicine, Chulalongkorn University, Prathumwan,*
[1]*CA*
[2,3]*Thailand*

1. Introduction

Today, disease interpretation of excised tissue is performed by analyzing biopsy specimens with a tabletop microscope [1]. While this method is effective, the process can be limited by sampling error, processing costs, and preparation time. In addition, the interpretive accuracy of the specimens can be affected by artefacts associated with tissue sectioning, paraffin embedding, and histochemical staining. Thus, a lot of effort has gone into the development of new methods that perform real time *in vivo* imaging with sub-cellular resolution.

Confocal microscopy is a powerful optical imaging method that can achieve sub-cellular resolution in real time. The technique of optical sectioning provides clear images from "optically thick" biological tissues that have previously been collected with large, tabletop instruments that occupy the size of a table [2, 3]. They can be used to collect either reflectance or fluorescence images to identify morphological or
molecular features of cells and tissues, respectively. Moreover, images in both modalities can be captured simultaneously with complete spatial registration. This approach uses a "pinhole" placed in between the objective lens and the detector to allow only the light that originates from within a tiny focal volume below the tissue surface to be collected. For miniature instruments, the core of an optical fiber is used as the "pinhole."

Recently, significant progress has been made in the development of endoscope-compatible confocal imaging instruments for visualizing inside the human body. This direction has been accelerated by the availability, variety and low cost of optical fibers, scanners, and light sources, in particular, semiconductor lasers. These methods are being developed for use in the clinic as well as in small animal imaging facilities. The addition of a miniature real-time, high resolution imaging instrument can help guide tissue biopsy and reduce pathology costs. However, these efforts are technically challenging because of the demanding performance requirements for small instrument size, high image resolution, deep tissue penetration depths, and fast frame rates.

The performance parameters for miniature *in vivo* confocal imaging instruments are governed by the specific application. An important goal is the early detection and image

Source: Advances in Solid State Circuits Technologies, Book edited by: Paul K. Chu,
ISBN 978-953-307-086-5, pp. 446, April 2010, INTECH, Croatia, downloaded from SCIYO.COM

guided therapy of disease in hollow organs, including colon, esophagus, lung, oropharynx, and cervix. Applications can also be found for better understanding of the molecular mechanisms of disease in small animals. In particular, localization of pre-malignant (dysplastic) lesions in the digestive tract can guide tissue biopsy for early detection and prevention of cancer. In addition, visualization of over expressed molecular targets in small animal models can lead to the discovery of new drugs.

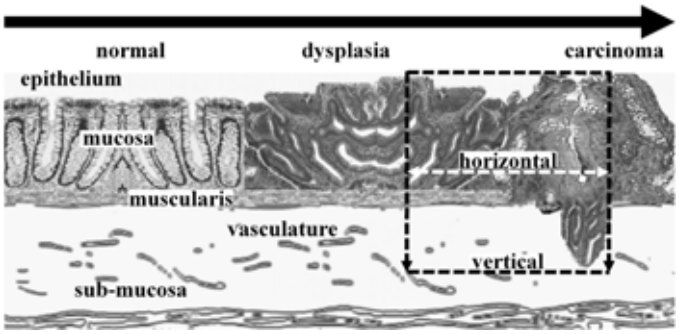

Fig. 1. Dysplasia represents a pre-malignant condition in the epithelium of hollow organs, such as the colon and esophagus. The dual axes confocal architecture has high dynamic range that is suitable for imaging in the vertical cross-sectional plane to visualize disease processes with greater tissue penetration depths.

As shown in Fig. 1, dysplasia originates in the epithelium and represents an important step in the transformation of normal mucosa to carcinoma. Dysplasia has a latency period of approximately 7 to 14 years before progressing onto cancer and offers a window of opportunity for evaluating patients by endoscopy who are at increased risk for developing cancer. The early detection and localization of dysplastic lesions can guide tissue resection and prevent future cancer progression. Dysplastic glands can be present from the mucosal surface down to the muscularis. Thus, an imaging depth of ~500 μm is sufficient to evaluate most early epithelial disease processes.

On reflectance imaging, sub-cellular resolution (typically <5 μm) is needed to identify nuclear features, such as nuclear-to-cytoplasm ratio. On fluorescence imaging, high contrast is needed to distinguish between the target and background. With both modalities, a fast imaging frame rate (>4 Hz) is necessary to avoid motion artefact.

2. Single axIs confocal architecture

A. Configuration of optics

Recent advances in the development of microlenses and miniature scanners have resulted in the development of fiber optic coupled instruments that are endoscope compatible with high resolution, including single [4-8], and multiple fiber [7, 9] strategies. Different methods of scanning are also being explored [10-14].

All of these endoscope compatible designs use a single axis design, where the pinhole (fiber) and objective are located along one main optical axis. A high NA objective is used to achieve sub-cellular resolution and maximum light collection, and the same objective is used for both the illumination and collection of light. In order to scale down the dimension of these

instruments for endoscope compatibility, the diameter of the objective must be reduced to ~5 mm or less. As a consequence, the working distance (WD) as well as the field-of-view (FOV) is also decreased, as shown by the progression of the 3 different objectives in Fig. 2. The tissue penetration depth also decreases, and is typically inadequate to assess the tissue down to the muscularis, which is located at a depth of ~500 μm and is an important landmark for defining the early presence of epithelial cancers.

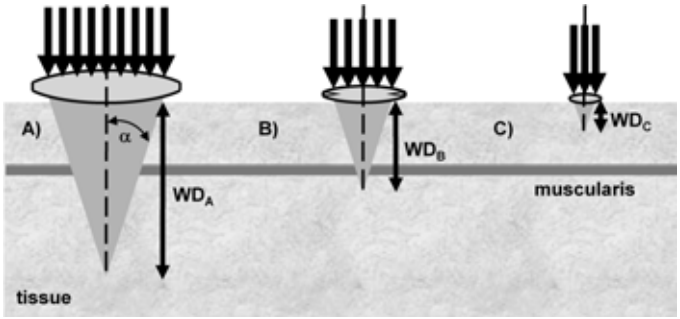

Fig. 2. For endoscope compatibility, the diameter of a single axis confocal microscope must be scaled down in size (A→B→C), resulting in a reduced working distance and limited tissue penetration depth.

B. Resolutions

For the conventional single axis architecture, the transverse, Δr_s, and axial, Δz_s, resolution between full-width-half-power (FWHP) points for uniform illumination of the lenses are defined by the following equations [3]:

$$\Delta r_s = \frac{0.37\lambda}{n\sin\alpha} \approx \frac{0.37\lambda}{n\alpha}; \quad \Delta z_s = \frac{0.89\lambda}{n(1-\cos\alpha)} \approx \frac{1.78\lambda}{n\alpha^2} \tag{1}$$

where λ is the wavelength, n is the refractive index of the medium, α is the maximum convergence half-angle of the beam, $NA = n\sin\alpha$ is the numerical aperture, and sinα≈ α for low NA lenses. Eq. (1) implies that the transverse and axial resolution varies as 1/NA and 1/NA², respectively. A resolution of less than 5 μm is adequate to identify sub-cellular structures that are important for medical and biological applications. To achieve this resolution in the axial dimension, the objective lens used requires a relatively large NA (>0.4). The optics can be reduced to the millimeter scale for *in vivo* imaging, but requires a sacrifice of resolution, FOV, or WD. Also, a high NA objective limits the available WD, and requires that the scanning mechanism be located in the pre-objective position, restricting the FOV and further increasing sensitivity to off-axis aberrations.

C. Commercial systems

Two endoscope compatible confocal imaging systems are commercially available for clinical use. The EC-3870K (Pentax Precision Instruments, Tokyo, Japan) has an integrated design where a confocal module (Optiscan Pty Ltd, Victoria, Australia) is built into the insertion tube of the endoscope, and results in an overall diameter of 12.8 mm, as shown in Fig. 3a [15]. This module uses the single axis optical configuration where a single mode optical fiber is aligned on-axis with an objective that has an NA ≈ 0.6. Scanning of the distal tip of the

optical fiber is performed mechanically by coupling the fiber to a tuning fork that vibrates at resonance. Axial scanning is performed with a shape memory alloy (nitinol) actuator that can translate the focal volume over a distance of 0 to 250 μm below the tissue surface. Excitation is provide at 488 nm (peak absorption of fluorescein) by a semi-conductor laser, and a transverse and axial resolution of 0.7 and 7 μm, respectively, has been achieved. The images are collected at a frame rate of either 0.8 or 1.6 Hz to achieve a FOV of either 1024x1024 or 1024x512 pixels, respectively. The dimension of the confocal instrument by itself is ~5 mm. When a suspicious lesion is identified, the confocal window located on the distial tip is placed into contact with the tissue to collect images. A separate instrument channel can be used to obtain pinch biopsies of tissue.

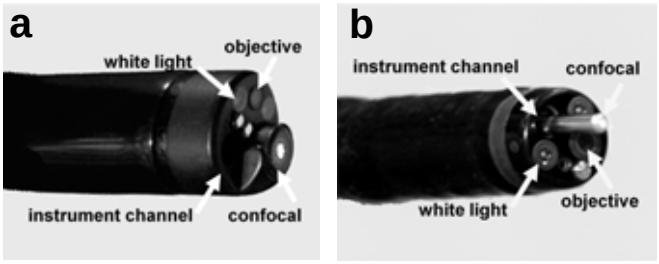

Fig. 3. a) The EC-3870K (Pentax) has a confocal module (Optiscan) integrated into the endoscope insertion tube. b) The Cellvizio® GI is a confocal miniprobe that passes through the instrument channel of the endoscope.

The Cellvizio® GI (Mauna Kea Technologies, Paris, France) uses a set of miniprobes that range in diameter from 1.5 to 2.5 mm, and passes through the standard instrument channel of medical endoscopes, as shown in Fig. 3b. This instrument moves independently of the endoscope, and its placement onto the tissue surface can be guided by the conventional white light image [8, 15]. This miniprobe consists of a fiber bundle with ~30,000 individual fibers that is aligned on-axis with an objective that has an NA ≈ 0.6. The core of each individual fiber acts as a collection pinhole to reject out-of-focus light. Scanning is performed at the proximal end of the bundle in the instrument control unit with a 4 kHz oscillating mirror for horizontal lines and a 12 Hz galvo mirror for frames. In this design, axial scanning cannot be performed. Instead, separate miniprobes that have different working distances are needed to optically section at different depths. Excitation is provided at 488 nm, and the transverse and axial resolution of these instruments ranges from 2.5 to 5 μm and 15 to 20 μm, respectively. Images are collected at a frame rate of 12 Hz with a FOV of either 600x500 μm^2 or 240x200 μm^2.

3. Dual axes confocal architecture

D. Configuration of optics

So far, the aforementioned miniaturization techniques in the previous section deploy a conventional single-axis confocal architecture that has the objective and optical fiber aligned along the same optical axis. In order to overcome some of these limitations for endoscope compatibility and *in vivo* imaging, we have developed the novel dual axes confocal configuration, shown in Fig. 4. We use two fibers oriented along separate optical axis of

different low NA objectives to spatially separate the light paths for illumination and collection [16, 17]. The region of overlap between the two beams (crossed at a half angle θ from the midline) defines the focal volume, hence the resolution, and can achieve sub-cellular dimensions. A very low probability exists for light scattered by tissue along the illumination path (blue cone) to enter the low NA collection objective (green cone), thus significant improvement in the dynamic range of detection can be achieved.

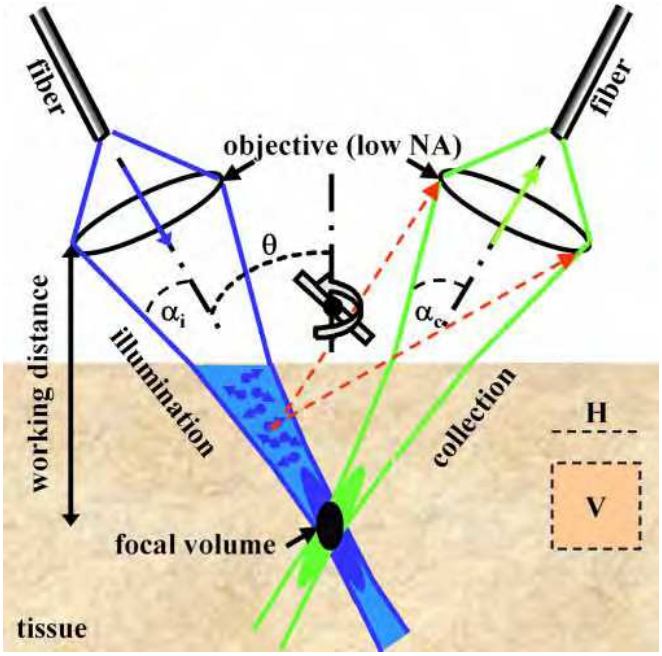

Fig. 4. Novel dual axes confocal architecture uses separate optical fibers and low NA lenses for off-axis light collection, achieving long working distance, high dynamic range, and scalability while preserving resolution.

Furthermore, the low NA objectives enable an increased working distance so that the scanning mirror can be placed on the distal (tissue) side of the lens (post-objective position), resulting in less sensitivity to off-axis aberrations [17]. In this configuration, the beams always pass through the low NA objectives on axis, resulting in a diffraction-limited focal volume that can then be scanned over a large FOV, limited by the performance of the scanner rather than by the optics. This design feature allows for the instrument to be scaled down in size to millimeter dimensions for compatibility with medical endoscopes without loss of performance.

We first develop the theory to explain the unique performance features of the dual axes confocal architecture by characterizing the point-spread function (PSF) and dynamic range. Then, we demonstrate the scaled down implementation of this configuration in miniature prototypes. Because of the challenges of packaging in such a small form factor, we first demonstrate a handheld (10 mm diameter) instrument and then an endoscope-compatible (5.5 mm diameter) prototype, using the same MEMS mirror and scanhead optics.

E. Definition of coordinates

The coordinates for the dual axes confocal configuration are shown in Fig. 5. The illumination (IO) and collection (CO) objectives represent separate low NA lenses. The maximum convergence half-angles of the illumination and collection beams are represented by α_i and α_c, respectively. The separate optical axes are defined to cross the z-axis (z_d) at an angle θ. The main lobe of the PSF of the illumination objective is represented by the light gray oval. This lobe has a narrow transverse but a wide axial dimension.

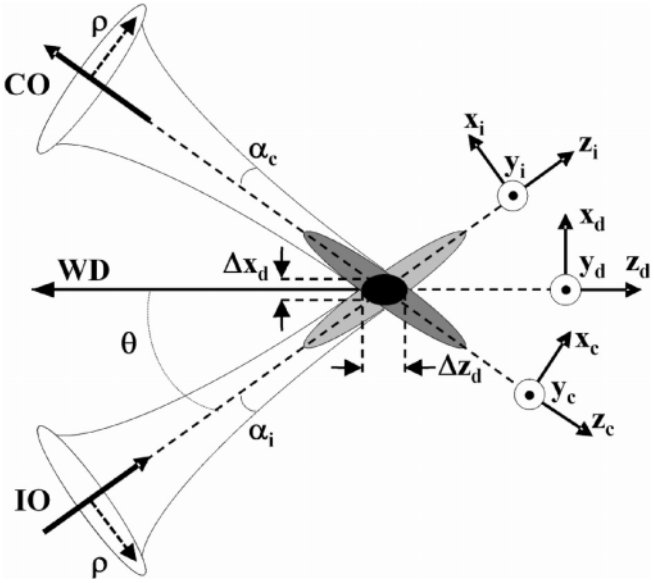

Fig. 5. Coordinates for dual axes confocal configuration

Similarly, the main lobe of the PSF of the collection objective is similar in shape but symmetrically reflected about z_d, as represented by the dark gray oval. For dual axes, the combined PSF is represented by the overlap of the two individual PSF's, represented by the black oval. This region is characterized by narrow transverse dimensions, Δx_d and Δy_d (out of the page), and by a significantly reduced axial dimension, Δz_d, which depends on the transverse rather than the axial dimension of the individual beams where they intersect.

F. Point spread function

The dual axes PSF can be derived using diffraction theory with paraxial approximations [18]. The coordinates for the illumination (x_i,y_i,z_i) and collection (x_c,y_c,z_c) beams are defined in terms of the coordinates of the main optical axis (x_d,y_d,z_d), and may be expressed as follows:

$$\begin{aligned} x_i &= x_d\cos\theta - z_d\sin\theta & x_c &= x_d\cos\theta + z_d\sin\theta \\ y_i &= y_d & y_c &= y_d \\ z_i &= x_d\sin\theta + z_d\cos\theta & z_c &= -x_d\sin\theta + z_d\cos\theta \end{aligned} \tag{2}$$

The maximum convergence half-angles of the focused illumination and collection beams in the sample media are represented as α_i and α_c, respectively. The angle at which the two

beams intersect the main optical axis is denoted as θ. A set of general dimensionless coordinates may be defined along the illumination and collection axes, as follows [19]:

$$\begin{aligned} u_i &= k_i n z_i \sin^2 \alpha_i & u_c &= k_c n z_c \sin^2 \alpha_c \\ v_i &= k_i n \sqrt{x_i^2 + y_i^2} \sin \alpha_i & v_c &= k_c n \sqrt{x_c^2 + y_c^2} \sin \alpha_c \end{aligned} \tag{3}$$

The wavenumbers for illumination and collection are defined as $k_i = 2\pi/\lambda_i$ and $k_c = 2\pi/\lambda_c$, respectively, where λ_i and λ_c are the wavelengths, and n is the index of refraction of the media.

The amplitude PSF describes the spatial distribution of the electric field of the focused beams. Diffraction theory may be used to show that the PSF of the illumination and collection beams is proportional to the Huygens-Fresnel integrals below [18]:

$$U_i(v_i, u_i) \propto \int_0^1 W_i(\rho) J_0(\rho v_i) e^{-j u_i \rho^2/2} \rho d\rho \tag{4}$$

$$U_c(v_c, u_c) \propto \int_0^1 W_c(\rho) J_0(\rho v_c) e^{-j u_c \rho^2/2} \rho d\rho \tag{5}$$

where J_0 is the Bessel function of order zero, and ρ is a normalized radial distance variable at the objective aperture. The weighting function, W(ρ), describes the truncation (apodization) of the beams. For uniform illumination, W(ρ) = 1. For Gaussian illumination, the objectives truncate the beams at the $1/e^2$ intensity, resulting in a weighting function of $W(\rho) = e^{-\rho^2}$. In practice, the beams are typically truncated so that 99% of the power is transmitted. For a Gaussian beam with a radius ($1/e^2$ intensity) given by w, an aperture with diameter πw passes ~99% of the power. In this case, the weighting function is given as follows:

$$W(\rho) = e^{-(\pi\rho/2)^2} \tag{6}$$

For the single axis configuration, the illumination and collection PSF's at the focal plane ($u_i = u_c = 0$) are identical functions of the radial distance ρ, and can both be given by U_s using the substitution $v = knr \sin\alpha$, as follows:

$$U_s(v) \propto \int_0^1 W_i(\rho) J_0(\rho v) \rho d\rho \tag{7}$$

The resulting signal at the detector V from a point source reflector in the media is proportional to the power received, and is given by the square of the product of the overlapping PSF's as follows:

$$V = A\left|U_i U_c\right|^2 \tag{8}$$

where *A* is a constant.

Similarly, since the depth of focus for each individual beam, described within the exponential term in the integral product of Eqs. (3) and (4), is much larger than that of the

transverse width, the exponential term may be neglected. As a result, the detector output v_d for the dual axes configuration for uniform illumination ($W = 1$), is given as follows:

$$V_d \propto \left(\frac{2J_1(v_i)}{v_i}\right)^2 \left(\frac{2J_1(v_c)}{v_c}\right)^2 \tag{9}$$

This expression can be combined with Eqs. (2) and (3) to derive the result for transverse and axial resolution with uniform illumination as follows [16]:

$$\Delta x_d = \frac{0.37\lambda}{n\alpha\cos\theta}; \quad \Delta y_d = \frac{0.37\lambda}{n\alpha}; \quad \Delta z_d = \frac{0.37\lambda}{n\alpha\sin\theta} \tag{10}$$

Note that for the dual axes configuration, the axial resolution is proportional to $1/NA$, where $NA = n\sin\alpha \approx n\alpha$, rather than $1/NA^2$, as is the case for the single axis design [3]. For example, with uniform illumination and the following parameters: $\alpha = 0.21$ radians, $\theta = 30$ degrees, $\lambda = 0.785$ μm and $n = 1.4$ for tissue, Eq. (10) reveals a result for the dual axes configuration of Δx_d = 1.1 μm, Δy_d = 1.0 μm, and Δz_d = 2 μm for the transverse and axial resolutions, respectively. Thus, sub-cellular resolution can be achieved in both the transverse and axial dimensions with the dual axes configuration using low NA optics but not with the single axis architecture.

For an endoscope-compatible instrument, delivery of the illumination and collection light is performed with use of optical fibers and is more appropriately modeled by a Gaussian rather than a uniform beam. With this apodization, the detector response for the dual axes configuration from a point source reflector in the media, given by Eq. (9), may be solved numerically as a function of transverse (x_d and y_d) and axial (z_d) dimensions. The integrals are calculated in Matlab, and use the weighting function with 99% transmission. In comparison, this model reveals a result of Δx_d = 2.4 μm, Δy_d = 2.1 μm, and Δz_d = 4.2 μm for the transverse and axial resolutions, respectively. Thus, the use of optical fibers, modeled by a Gaussian beam, produces results that are slightly worse but still comparable to that of uniform illumination [19].

G. Dynamic range

Differences in the dynamic range between the single and dual axes confocal configurations can also be illustrated with this model [18]. The calculated axial response for the single axis design with Gaussian illumination is shown by the dashed line in Fig. 6a, where optical parameters are used that achieve the same axial resolution (FWHM) of 4.2 μm. The result reveals that the main lobe falls off in the axial (z-axis) direction as $1/z^2$, and reaches a value of approximately -25 dB at a distance of 10 μm from the focal plane (z = 0). In addition, a number of side lobes can be appreciated.

In comparison, the response for the dual axes configuration, shown by the solid line in Fig. 6a, reveals that the main lobe rolls off in the axial (z-axis) direction as $\exp(-kz^2)$, and reaches a value of -60 dB at a distance of 10 μm from the focal plane ($z = 0$). Thus, off-axis illumination and collection of light in the dual axes architecture results in a significant improvement in dynamic range and in an exponential rejection of out-of-focus scattered light in comparison to that for single axis. This advantage allows for the dual axes configuration to collect images with deeper tissue penetration and with a vertical cross-section orientation. The transverse response with Gaussian illumination is shown in Fig. 6b.

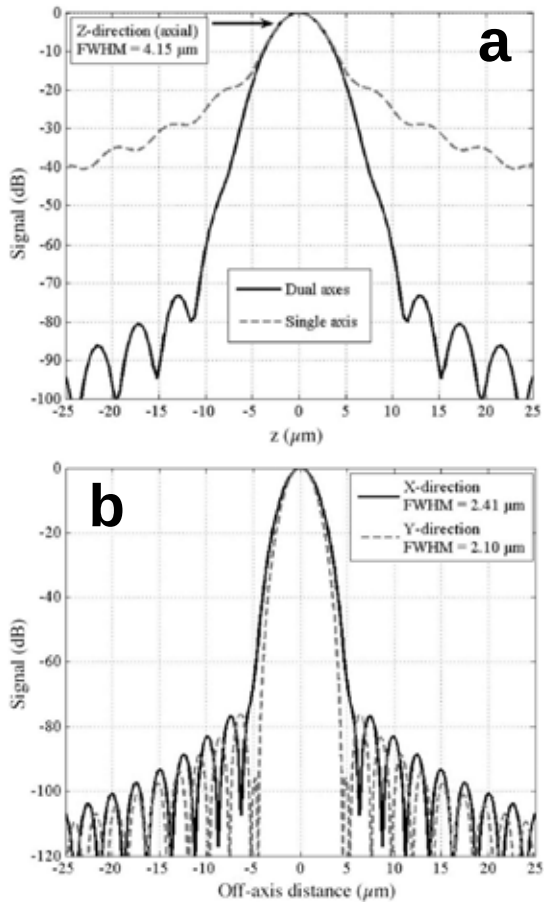

Fig. 6. Dynamic range of novel dual axes confocal architecture. a) The axial response of the single axis (dashed line) configuration falls off as $1/z^2$ and that for the dual axes (solid line) design falls off as $\exp(-kz^2)$, resulting in a significant improvement in dynamic range, allowing for vertical cross-sectional imaging to be performed. b) Transverse (X-Y direction) response.

H. Post-objective scanning

In confocal microscopes, scanning of the focal volume is necessary to create an image. In the single axis architecture, the high NA objectives used limit the working distance, thus the scan mirror is by convention placed on the pinhole (fiber) side of the objective, or in the pre-objective position, as shown in Fig. 7a. Scanning orients the beam at various angles to the optical axis and introduces off-axis aberrations that expand the focal volume. In addition, the FOV of pre-objective scanning systems is proportional to the scan angle and the focal length of the objective. The diameter of the objective limits the maximum scan angle, and as this dimension is reduced for endoscope compatibility, the focal length and FOV are also diminished.

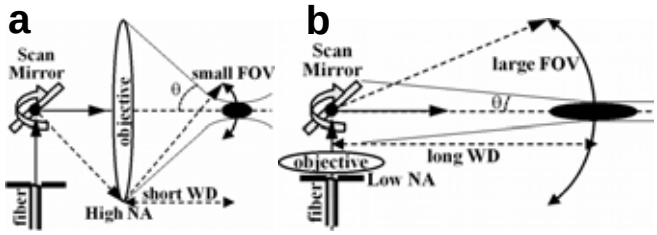

Fig. 7. a) For pre-objective scanning, illumination light is incident on the objective off-axis, resulting in more sensitivity to aberrations and limited FOV. b) With post-objective scanning, the incident light is on-axis, less sensitivity to aberrations, and large FOV. Post-objective scanning is made possible by the long WD produced by the low NA objectives used in the dual axes architecture.

In the dual axes configuration, the low NA objectives used creates a long working distance that allows for the scanner to be placed on the tissue side of the objective, or in the post-objective position [17]. This design feature is critical for scaling the size of the instrument down to millimeter dimensions for *in vivo* imaging applications without losing performance. As shown in Fig. 7b, the illumination light is always incident on-axis to the objective. In the post-objective location, the scan mirror can sweep a diffraction-limited focal volume over an arbitrarily large FOV, limited only by the maximum deflection angle of the mirror. Moreover, the scanner steers the illumination and collection beams together with the intersection of the two beams oriented at a constant angle θ and with the overlapping focal volume moving without changing shape along an arc-line. This property can be conceptualized by regarding the dual axes geometry as being equivalent to two separate beams produced from two circles in the outer annulus of a high NA lens containing a central obstruction (or a large central hole). A flat scan mirror deflects both beams equally, and thereby preserves the overlapping region without introducing aberrations to the beams.

I. Improved rejection of scattering

In the dual axes confocal architecture, the off-axis collection of light significantly reduces the deleterious effects of tissue scattering on the dynamic range of detection and allows for deeper ballistic photons to be resolved [20]. These features provide the unique capacity to collect vertical cross-sectional images in the plane perpendicular to the tissue surface. This is the preferred view of pathologists because differences from the normal patterns of tissue differentiation are revealed in the direction from the lumen to the sub-mucosa.

1. Optical configurations

The improvement in rejection of light scattered by tissue can be illustrated by comparing the dynamic range of detection between the single and dual axes optical configurations with equivalent axial resolution, as shown in Fig. 8a and 8b. The incident beams are modeled with a Gaussian profile because this is representative of light delivered through an optical fiber. For the single axis configuration, this beam is focused into the tissue by an ideal lens (L1). A mirror (M) is embedded in the tissue at the focal plane (parallel to the x-y plane) of the objective lens. In this scheme, the rays that reflect from the mirror pass back through the lens L1, deflect at an angle off the beam splitter, and are focused by an ideal lens (L2) on to a pinhole detector. For the dual axes set-up, the incident Gaussian beam is focused into the

tissue by an ideal lens (L3) with its axis oriented at an angle θ = 30° to the z-axis, and an ideal lens (L4) focuses the backscattered beam, with its axis z´ at an angle -30° to the z-axis, onto the pinhole detector. As before, a mirror (M) with its plane perpendicular to the z-axis and passing through the coincident focuses of the lenses is embedded in the tissue to reflect the incident light to the detector. In both configurations, the lens system has a magnification of 1 from the focal plane to the pinhole detector.

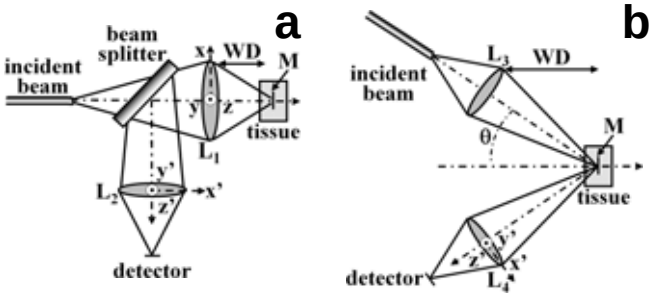

Fig. 8. a) single axis and b) dual axes optical configurations are used to evaluate the axial response at the detector.

In order to achieve an equivalent -3 dB axial resolution (FWHM), the NA's for the single and dual axes configurations are defined to be 0.58 and 0.21, respectively. From diffraction theory, discussed above, the theoretical transverse and axial resolutions for the PSF for dual axes at a wavelength λ = 633 nm with an average tissue refractive index of 1.4 and NA = 0.21 are found to be Δx = 1.16 μm, Δy = 1.00 μm, and Δz = 2.00 μm.3 The mirror is placed at a distance of 200 μm below the tissue surface in the focal plane of the objective lenses for both the single and dual axes configurations. This depth is representative of the imaging distance of interest in the epithelium of hollow organs. The calculations performed to analyze the effects of tissue scattering on light are based on Monte Carlo simulations using a non-sequential ray tracing program (ASAP® 2006 Breault Research Organization, Tucson, AZ). Three assumptions are made in this simulation study: 1) multiple scattering of an incoherent beam dominates over diffraction effects, 2) the non-scattering optical medium surrounding the lenses and the tissue (the scattering medium) is index matched to eliminate aberrations, and 3) absorption is not included to simplify this model and because there is much larger attenuation due to the scattering of ballistic photons.

2. Mie scattering analysis

We use Mie theory with the Henyey-Greenstein phase function *p(θ)* to model the angular dependence of tissue scattering, as follows [21, 22]:

$$p(\theta) = \frac{1}{4\pi} \frac{1-g^2}{(1+g^2-2g\cos\theta)^{3/2}} \tag{11}$$

where *g*, the anisotropy factor, is defined as

$$g = \langle \cos\theta \rangle = \int_0^{2\pi} \int_0^{\pi} \cos\theta \cdot p(\theta) \sin\theta d\theta d\varphi. \tag{12}$$

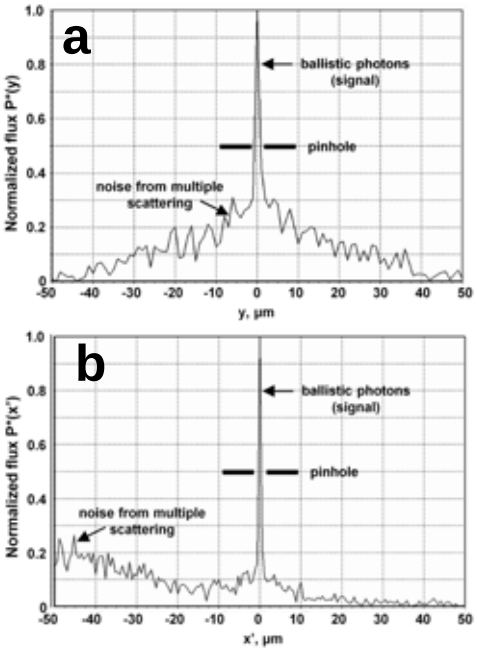

Fig. 9. Distributions of photon flux in tissue scattering model. The peak value of multiple scattered photons for A) single axis is co-located with the confocal pinhole while that for B) dual axes is separated by ~50 μm. As a consequence, the ballistic photons for dual axes result in a greater signal-to-noise ratio.

Given the average scatterer size, refractive index, and concentration, the attenuation coefficient μ_s and anisotropy g are determined and provided to the ASAP program as simulation parameters. For a tissue phantom composed of polystyrene spheres with a diameter of 0.48 μm, refractive index 1.59, and a concentration of 0.0394 spheres/μm^3 in water, the values g = 0.81 and μ_s = 5.0 mm^{-1} at λ = 633 nm are calculated from Mie theory [23]. For single axis, P(y′) is defined as the photon flux distribution along the y′-axis at the detector. The photon flux can be normalized by defining $P^*(y') = P(y')/P_{max}$, where P_{max} is the maximum flux. The normalized flux $P^*(y')$ consists of ballistic (signal) and multiple scattered (noise) photons, as shown in Fig. 9a [24]. The maximum flux for both the signal and noise components arrive at center of the detector. A confocal pinhole placed in front of the detector can filter out some but not all of this "noise," resulting in a reduced signal-to-noise ratio (SNR). For dual axes, the detector is angled off the optical axis by 30 deg. P(x′) is defined as the photon flux distribution along the x′-axis at the detector. The photon flux can be normalized by defining $P^*(x') = P(x')/P_{max}$, where P_{max} is the maximum flux at the detector. Fig. 9b shows that normalized photon flux distribution for dual axes also exhibits a ballistic and multiple scattered components. However, for dual axes, the peak flux of multiple scattered photons arrives ~50 μm lateral to the center of the detector where the ballistic photons arrive, a consequence of off-axis collection. Thus, there is much less "noise" for the confocal pinhole (diameter ~1 μm) to filter out, resulting in a higher SNR and dynamic range.

3. Improvement in dynamic range

An implication of this result is that the dual axes configuration has improved dynamic range compared to that of single axis. This difference can be quantified by determining the axial response at the detector. This can be done by calculating the photon flux f(Δz) as the mirror is displaced along the *z*-axis in the tissue. The flux is calculated using Monte-Carlo simulations in ASAP with the mirror at positioned in the range -10 μm < Δz < 10 μm with respect to the focal plane at $z = 0$, which is located at 200 μm below the tissue surface. The flux is then normalized according to $F(\Delta z) = f(\Delta z)/f(0)$. The axial response is shown in Fig. 10a for various pinhole diameters D, including 1, 2 and 3 μm, which correspond to typical fiber core dimensions. Note that for each pinhole diameter, the dual axes (DA) configuration has significantly better dynamic range than that of single axis (SA). Note that the introduction of tissue scattering results in a reduction of the dynamic range compared to that found in free space, as shown by Fig. 6a.

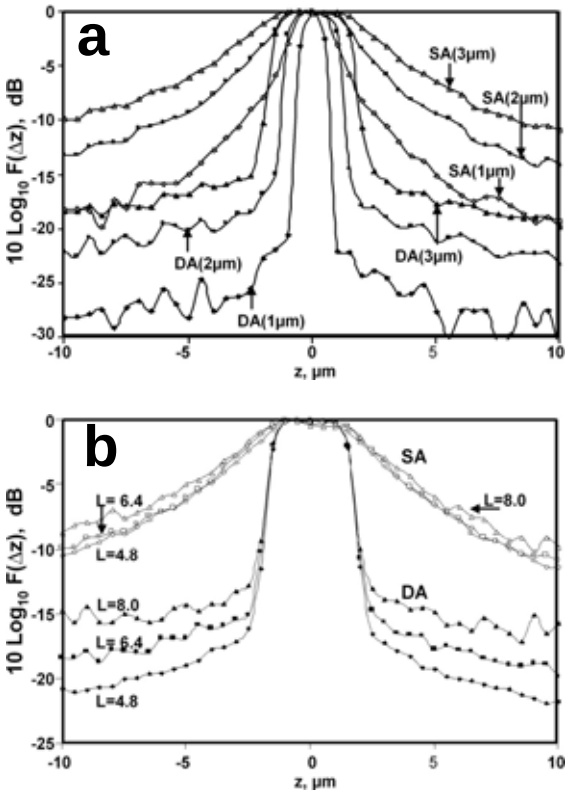

Fig. 10. Axial response for single and dual axes geometries. The dual axes (DA) configuration has a much greater dynamic range than that for single axis (SA) given different a) pinhole diameters (1, 2, and 3 μm) and b) optical lengths L (4.8, 6.4 and 8.0).

We can also determine the axial response of the detector for various optical lengths in tissue. This analysis reveals differences in the dynamic range between the single and dual axes

configuration for tissues with various scattering properties. The total optical length L is defined as twice the product of the scattering coefficient μ_s and the tissue depth t, or $L = 2\mu_s t$. The factor of two originates from the fact that the total path length is twice the tissue depth. The axial response is shown in Fig. 10b for various optical lengths L, including 4.8, 6.4, and 8.0. Note that for each optical length L, the dual axes (DA) configuration has significantly better dynamic range than that of single axis (SA). These values of L are typical parameters of gastrointestinal epithelium. At $\lambda = 633$ nm, μ_s is about 7 mm^{-1} for esophagus tissue [24] and about 20 mm^{-1} for normal colon mucosa [25]. The range of tissue depths spanned by L = 4.8 to 8 for esophagus and colon is 340 μm to 570 μm and 120 μm to 200 μm, respectively. In addition, these results shows that for single axis only minimal changes occur in the dynamic range with approximately a factor of 2 difference in optical thickness L, while for dual axes significant changes occur over this thickness range. Furthermore, scattering does not appear to alter the FWHM of the axial response for either single or dual axes over this range of lengths.

4. Geometric differences produced by off-axis detection

The superior axial response of the dual axes confocal architecture has a simple geometric explanation. When the mirror moves away from the focal plane by ±Δ, the centroid of the beam is steered away from the optical axis by $\pm 2\Delta\sin\theta$ from where the center of the pinhole is located [20]. Even taking into consideration diffraction and the broadening of the out-of-focus beam, the beam intensity decreases exponentially when $\Delta > D/2$ (for $\theta = 30°$). But in the single axis case, many of the photons scattered near the vicinity of the focal plane (±Δ) are collected by the detector through the pinhole. Thus, the spatial filtering effect by a pinhole for the single axis configuration is not as effective as that for dual axes. The implication of this effect for imaging deep in tissue is evident. In the single axis case, many of the multiple scattered photons that arrive from the same direction as that of the ballistic photons, starting from the surface to deep within the tissue, are collected by the detector despite the presence of a pinhole to filter the out-of-focus light. This explains why in Fig. 9a the single axis configuration has a large noise component alongside the ballistic component. Thus, the dual axes confocal architecture provides optical sectioning capability that is superior to that of the conventional single axis design in terms of SNR and dynamic range, and this result can be generalized to a range of relevant pinhole sizes. As a result, the dual axes architecture allows for imaging with greater tissue penetration depth, thus is capable of providing images in the vertical cross-section with high contrast. The implementation of the dual axes confocal configuration to an endoscope compatible instrument for collection of both reflectance and fluorescence has significant implications for *in vivo* imaging by providing both functional and structural information deep below the tissue surface.

4. Tabletop dual axes confocal imaging instruments

The dual axes confocal architecture was first implemented as a tabletop instrument using readily available optical components to demonstrate the proof of concept of off-axis illumination and collection with post-objective scanning. In particular, the primary advantages of the dual axes configuration including high dynamic range and deep tissue penetration are revealed by vertical cross-sectional images with either reflectance or fluorescence. The combination of these two imaging modes forms a powerful strategy for integrating structural with functional information.

The dual axes optical design incorporates a solid immersion lens (SIL) made from a fused-silica hemisphere at the interface where the two off-axis beams meet the tissue. This refractive element minimizes spherical aberrations that occur when light undergoes a step change in refractive index between two media. The curved surface of the SIL provides a normal interface for the two beams to cross the air-glass boundary. Fused silica is used because its index of refraction of $n = 1.45$ is closely matched to that of tissue. Note that as the beams are scanned away from their neutral positions, they will no longer be incident to the surface of the SIL and small aberrations will occur. Another feature of the SIL is that its curved surface increases the effective NA of the beams in the tissue by a factor of n, the index of refraction, and produces higher resolution and light collection efficiency. On the other hand, the SIL acts to reduce the scanning displacement of the beams in the tissue by a factor of $1/n$ so that larger deflections are needed to achieve the desired scan range.

J. Horizontal cross-sectional imaging instrument

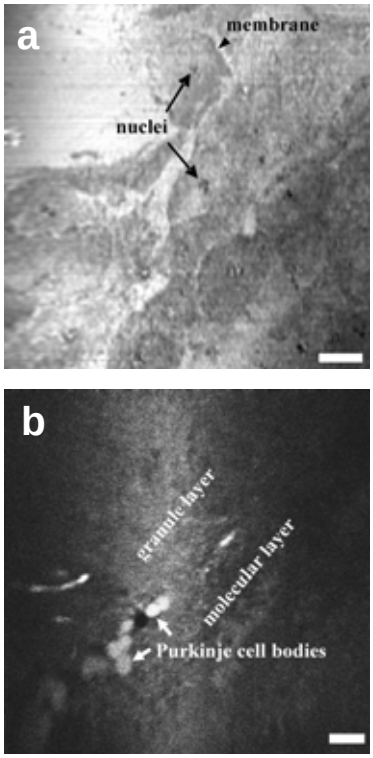

Fig. 11. Horizontal cross-sectional dual axes images ex vivo. a) Squamous esophageal mucosa collected at z = 0 μm with λ = 488 nm reveals sub-cellular features, including cell nuclei (arrows) and membrane (arrowhead), scale bar 20 μm. b) Normal colonic mucosa at z = 150 μm with λ = 1.3 μm illumination reveals circular crypts with colonocytes surround the lumen and lamina propria (LP) filling space in between the crypts, scale bar 50 μm.

Reflectance imaging takes advantage of subtle differences in the refractive indices of tissue micro-structures to generate contrast. The backscattered photons can provide plenty of

signal to overcome the low NA objectives used for light collection in the dual axes configuration. The first reflectance images were collected with a tabletop system that used a 488 nm semiconductor laser that delivered illumination into a single mode optical fiber that was focused by a set of collimating lenses with NA = 0.16 to a spot size with ~400 μW of power [16]. These parameters produced a transverse and axial resolution of 1.1 and 2.1 μm, respectively. The reflected light was collected by a complementary set of optics. The off-axis illumination and collection was performed at θ = 30° to the main optical axis. Reflectance images were collected in horizontal cross-sections of freshly excised specimens of esophagus *ex vivo*. As shown in Fig. 11a, the cell membrane and individual nuclei of squamous (normal) esophageal mucosa can be appreciated in the image collected at z = 0 μm, scale bar 20 μm.

Much greater image contrast can be achieved with fluorescence imaging where the use of optical reporters, such as GFP, and exogenous probes can reveal over expression of molecular targets. The same tabletop dual axes microscope was also used to collect fluorescence images with a long pass filter to block the excitation light and photomultiplier tube (PMT) for detection [26]. In Fig. 11b, a fluorescence image of the cerebellum of a transgenic mouse that constituitively expresses GFP under the control of a β-actin promoter at a depth of z = 30 μm is shown, scale bar 50 μm. Purkinje cell bodies (arrows) can be seen as large round structures aligned side by side in a row, separating the granule layer and the molecular layer.

K. Vertical cross-sectional imaging instrument

1. Reflection imaging mode

In order to collect vertical cross-sectional images, heterodyning can be used to provide a coherence gate that filters out illumination photons that are multiply-scattered and travel over longer optical paths within the tissue [17]. This approach is demonstrated with a fiber optic Mach-Zehnder interferometer, shown in Fig. 12a. A broadband near-infrared source produces light centered at λ = 1345 nm with a 3 dB bandwidth of 35 nm and a coherence length in tissue of ~50 μm. A fiber coupler directs ~99% of the power to the illumination path, which consists of a single mode optical fiber (SMF_1) with a collimating (CL_1) and focusing lens (FL_1) with NA = 0.186. The axes of illumination and collection are oriented at $\theta = 30^o$ to the midline. Light reflected from the tissue is collected by the second set of focusing (FL_2) and collimating (CL_2) lenses into another single mode fiber (SMF_2). The lens and fiber parameters are the same for both the illumination and collection beams. The fiber optic coupler directs ~1% of the source into a reference beam which is frequency shifted by an acousto-optic modulator at 55 MHz for heterodyne detection. An adjustable optical delay is used to increase the signal by matching the optical path length of the reference beam to that of the ballistic photons. An adjustable optical delay is used to increase the signal. In addition, a polarization controller consisting of two half-wave plates and a single quarter-wave plate is used to maximize the signal. The reference and collection beams are combined by a 50/50 coupler and the resulting heterodyne signal is detected by a balanced InGaAs detector (D_1, D_2) with a bandwidth of 80 MHz. The resulting electronic signal is then processed with a band pass filter (BPF) with a 3 MHz bandwidth centered at 55 MHz, then demodulated (DM), digitized by a frame grabber (FG), and displayed (D).

In this heterodyne detection scheme, the reference beam essentially provides amplification of the weak collection beam via coherent optical mixing, and enables the measurement of reflected light with a dynamic range larger than 70 dB. Post-objective scanning is performed with the scan mirror (SM) placed distal to the objective lenses. Reflectance images were

collected from fresh biopsy specimens taken from the squamo-columnar junction of subjects with Barrett's esophagus. Specimens with dimensions of ~3 mm were resected with jumbo biopsy forceps, and the mucosal surface was oriented normal to the z-axis. Vertical cross-sectional images were collected with depth of 1 mm. From Fig. 12b, squamous (normal) mucosa is present over the left half of the image with an intact epithelium (EP). The other structures of normal esophageal mucosa, including the muscularis mucosa (MM), sub-mucosa (SM), and muscularis propria (MP), can also be identified. Columnar mucosa consistent with intestinal metaplasia is seen over the right half of the image, and reveals the presence of pit epithelium (PE) [17]. These findings correlate with the tissue micro-structures seen on histology.

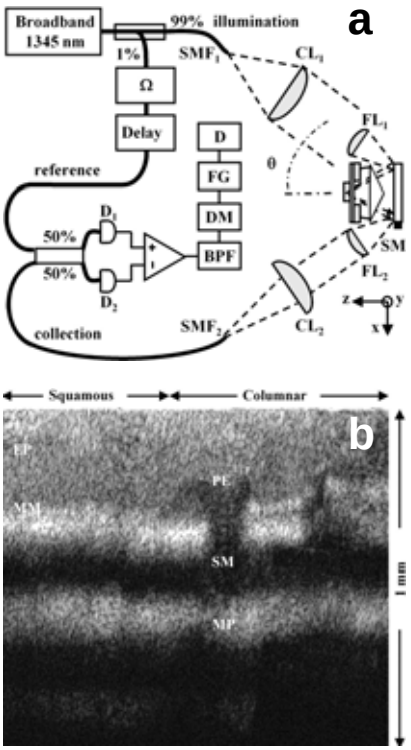

Fig. 12. Vertical cross-sectional dual axes confocal reflectance images ex vivo. a) Schematic of optical circuit for heterodyne detection, details discussed in the text. b) Reflectance image of squamo-columnar junction in esophagus with vertical depth of 1 mm. Squamous (normal) mucosa reveals epithelium (EP) and muscularis mucosa (MM) over left half. Columnar (intestinal metaplasia) mucosa shows pit epithelium (PE) over right half. Submucosa (SM) and musclaris propria (MP) are seen on both sides.

2. Fluorescence imaging mode

Fluorescence detection adds an entirely new dimension to the imaging capabilities of the dual axes architecture. Detection in this mode offers an opportunity to achieve much higher image contrast compared to that of reflectance and is sensitive to labeled molecular probes