Label-free *in vivo* pathology of human epithelia with a high-speed handheld dual-axis confocal microscope

Chengbo Yin

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Jonathan T. C. Liu, Chair

Eric J. Seibel

Milind Rajadhyaksha

Program Authorized to Offer Degree:

Mechanical Engineering Department

University of Washington

Abstract

Label-free *in vivo* pathology of human epithelia with a high-speed handheld dual-axis confocal microscope

Chengbo Yin

Chair of the Supervisory Committee:
Jonathan T. C. Liu
Mechanical Engineering Department

Oral cancer is one of the most common malignancies worldwide, with an approximate global prevalence of 300,000 individuals in 2012. Roughly 50,000 new cases and over 10,000 deaths are attributed to this disease each year in the United States. Dentists and head & neck physicians are trained to recognize aberrations in gross morphology that are indicative of malignancy, as well as to utilize a number of wide-field imaging techniques to improve the visualization of lesions, some of which can detect disease with high sensitivity but poor specificity. As a result of poor diagnostic specificity, the majority of suspected lesions, when biopsied and analyzed via histopathology, reveal benign conditions rather than premalignant or malignant lesions. In addition to the high costs and time associated with obtaining large numbers of unnecessary tissue samples for pathological analysis, the need for an invasive biopsy often results in patient discomfort, noncompliance, complications and/or diagnostic delays. Reflectance confocal microscopy can

potentially provide a real-time non-invasive method to triage and guide excisional biopsy. However, miniaturization is a challenge and previous *in vivo* systems have necessitated trade-offs between critical parameters such as device size, imaging speed (frame rate), imaging depth, resolution, and field-of-view. In this thesis project, a fully packaged miniature line-scanned reflectance microscope, which combines a dual-axis confocal architecture with MEMS-based beam scanning, was built to achieve an optimized clinical device for real-time micropathological detection of oral lesions. This handheld device, with a tip diameter of 12 mm, enables high-contrast sub-cellular imaging (lateral resolution: 1μm, axial resolution: 1.72 μm) deep within tissues of the oral cavity. Furthermore, high-speed (>20 Hz) optical sectioning was achieved, which greatly enhanced the clinical usability of this device by minimizing motion artifacts and enabling high-quality real-time video mosaicking. However, due to the size limitation of the miniature device (12-mm diameter tip), the field of view limits the capability of the device for imaging a large extent of tissue with high resolution. To extend the imaging field during the clinical use of such devices, we have developed a real-time video mosaicking method based on a previous post-processing mosaicking method, which allows users to efficiently survey larger areas of tissue. For the images that are 400 x 400 pixels in size, real-time mosaicking is possible at >30 frames/sec. Unlike other real-time mosaicking methods, our strategy can accommodate image rotations and deformations that often occur during clinical use of a handheld microscope. We perform a feasibility study to demonstrate that the use of real-time mosaicking enables significantly more efficient sampling of a desired imaging field when using a handheld dual-axis confocal microscope.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor Prof. Jonathan T.C. Liu for the continuous support of my Ph.D. study and related research. His guidance helped me in all the time of research and writing of this thesis. I appreciate the time he spent on me. He is a good mentor for me, since he wasn't only focusing on making me to be a Ph.D. Instead, he taught me lots of things that would benefit my whole life. A five-year Ph.D. program was a long and tough journey. Having Prof. Liu as my mentor was lucky for me.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Eric Seibel, Prof. Dan FU, and Dr. Milind Rajadhyaksha, for their insightful comments and encouragement from their expertise. They made me to understand I should always keep the big picture in my mind. This is also what Prof. Liu always emphasizes to me. In particular, I am grateful to Dr. Rajadhyaksha and his research group (Dr. Gary Peterson and Dr. Kivanc Kose) for supporting me on my research during our collaboration.

My sincere thanks also go to Dr. Matt Keller and Dr. Wenbo Wang, who provided me an opportunity to join their team as an intern. I really enjoyed the summer intern in 2019.

I want to thank all my labmates. Dr. Ye Chen and Dr. Soyoung Kang made me feel our lab is like a warm family. Thank Dr. Peter Wei for being a good research partner. In particular, I am grateful to Dr. Yu Wang, Dr. Adam Glaser and Dr. Steven Leigh. They are my colleagues, mentors and friends who I always can trust. I am lucky to have them.

Last but not the least, I would like to thank my family. My parents always trust me and support me to do what I believe is correct. They worked hard so that I could only focusing on getting my Ph.D. I have a family, they made me feel I am not the only child of my parents. In the end, I want

to sincerely apologize to my grandparents for not being able to accompany with them. I really miss you and I believe that we will meet again in one day.

Chapter 1. INTRODUCTION

1.1 CLINICAL PROBLEM

Oral cancer is one of the most common malignancies worldwide, with an approximate global prevalence of 300,000 individuals in 2012 [1]–[3]. In the USA, about 50,000 new cases are diagnosed each year and there are more than 10,000 deaths [2]–[4]. The lower lip vermilion, the tongue, and the floor of the mouth are the most common sites of oral cancer [5]. Outside of the US, in regions like Southeast Asia where oral hygiene is less advanced and there is heavier use of chewing tobacco and betel nuts, oral cancers can also be prevalent on the buccal mucosa and sides of the tongue [6]. Unfortunately, oral cancers are often detected at a late stage in which the disease has already metastasized. The 5-year survival rate for patients with squamous cell oral carcinoma, which represents 90% of oral cancer patients, is only 50% [7]–[11]. As with most cancers, the early and accurate diagnosis of oral cancers is critical for patient survival [7]–[16]. However, a major challenge for the clinical diagnosis of precancerous or early-stage oral lesions is differentiating between malignant lesions and a host of benign conditions under gross visual examination, such as lichen planus, lupus erythematosus, leukoedema, white sponge nevus, frictional keratosis, cheek/lip/tongue biting, contact lesions, and smoker's palate [17].

1.2 EXISTING SOLUTIONS FOR ORAL CANCER DETECTION

1.2.1 *Traditional histopathology approaches*

The microscopic evaluation of surgical and biopsy specimens with analog bright field microscopes, through a process known as histology, is currently regarded by the medical community as the “gold standard” for the diagnosis of diseases, including the diagnosis of

precancerous and cancerous lesions [18]. However, this core technology for anatomic pathology has certain limitations that can lead to poor inter-observer concordance and limited diagnostic accuracy for prognostication and prediction of treatment response. For example, the invasive physical resection of tissues is often not desired by patients, especially in cases where the tissues are of functional importance (e.g., brain), cosmetic value (e.g., skin), and in cases where there is a low probability of malignancy (e.g., biopsies of suspicious lesions in the oral cavity). In addition, dynamic information (e.g., blood flow) and physiological parameters (e.g., pH, oxygenation, electrolyte concentration, etc.) are often lost or altered during *ex-vivo* tissue processing, which can reduce diagnostic accuracy. Once tissues are excised, the standard procedure of sample preparation - involving fixation, dehydration, wax embedding, sectioning, mounting of tissue sections on glass slides, and staining - is labor-intensive, complex, and time-consuming, potentially resulting in treatment delays and process-induced errors. Moreover, the multi-step procedure is known to introduce artifacts (e.g., shrinkage and cracking due to dehydration) as well as sampling errors since only a small fraction of most tissue specimens are processed onto glass slides for imaging.

For detecting and diagnosing oral cancers, there are two major challenges that result from this reliance on traditional histopathology: 1) Due to the discomfort, trauma, risk-of-infection, time, and cost of obtaining (and processing) biopsy samples, the patient's oral cavity is undersampled such that precancerous lesions often remain undetected. 2) Due to the large number of benign conditions that appear suspicious under gross examination, the majority of biopsied tissues are normal. Furthermore, dentists and physicians are hesitant to biopsy suspicious lesions until they have progressed to a more-obvious (but late-stage) malignancy. Clearly, these two challenges exacerbate one another and serve to create an "ill-posed problem" for the early detection of oral cancers.

1.2.2 *Wide-field microscopy approaches*

In response to the two challenges mentioned above, various wide-field imaging methods have been developed and commercialized to improve visual examination. For example, toluidine blue staining (also known as tolonium chloride or TBlue staining) has been used for decades to assist with the identification of epithelial dysplasia. Published results indicate high sensitivity but low specificity as the dye also tends to stain benign conditions that are associated with inflammation [9], [11], [12], [14], [15]. A chemiluminescence technique (ViziLite, Zila Pharmaceuticals, Phoenix, AZ) was approved in 2002 in the USA for improving the identification, visualization, and monitoring of oral precancerous lesions. Published studies report that the ViziLite tool enhances intraoral visualization of white lesions (sensitivity as high as 100%) but is not able to discriminate between keratotic, inflammatory, malignant, or potentially malignant oral mucosal white lesions[7], [10], [13], [15](specificity reported as between 0 and 15%). A recent ViziLite study, in 2011, yielded similar results with a sensitivity of 77% and specificity of 28%[19]. Other techniques include a light emission technique (Microlux DL, AdDent, Danbury, CT) and a narrowband tissue autofluorescence imaging technique (VELscope, LED Dental Inc., White Rock, BC, Canada). Both techniques have been shown to improve the visualization of lesions but are unable to discriminate between benign and malignant lesions [7], [9], [11], [16], [20]. In short, while existing wide-field techniques for oral cancer detection improve the sensitivity of detection, their specificity is limited, further exacerbating the "ill-posed problem" described in the previous paragraph by necessitating more benign biopsies. In addition to the aforementioned commercialized techniques, reflectance and fluorescence imaging - of endogenous autofluorescence and exogenous fluorophores - are being investigated as alternatives for oral cancer detection [21]–[24]. While these approaches have shown some success in identifying

malignancy, early studies report that the ability to distinguish between precancer, cancer, benign lesions, inflammation, and infections, is challenging [21], [24]. Furthermore, Pavlova et al. [21] have reported that variations in stromal fluorescence and epithelial fluorescence are very different in normal and diseased tissue and can result in ambiguous results when averaged together during conventional wide-field fluorescence imaging, thereby implying the need for optical sectioning. A review in 2011 concluded: "there is insufficient evidence to demonstrate that its use [autofluorescence imaging] as an adjunct to conventional oral screening provides additional benefit to conventional oral cancer screening alone [25]."

1.2.3 *Confocal microscopy approaches*

The previously mentioned wide-field imaging techniques lack the resolution to directly visualize nuclear, cellular, and glandular morphology, which have been the cardinal features informing "gold-standard" histopathological diagnosis for decades. While wide-field imaging is ideal for rapid surveillance of an entire oral cavity, there is a need for a complementary imaging technique to triage the large number of "false-positive" results obtained by such widefield approaches. Therefore, we are proposing to develop a miniature confocal microscope to rapidly obtain images of glandular, cellular, and nuclear detail for diagnosing suspicious tissues *in vivo*, and to guide the acquisition of excisional biopsies. This optical-sectioning technology has the potential to both improve the early detection of oral cancers, as well as to significantly reduce the time, cost, and patient discomfort associated with the acquisition of large numbers of unnecessary biopsies. The proposed technology for *in vivo* point-of-care pathology [26], as an adjunct to wide-field imaging techniques, has the potential to be applied not only for early detection, but also for guiding tumor resection. The ability to discriminate between normal and tumor tissues in real time at the surgical margins represents a high-impact practical application of *in vivo* pathology. Example surgical

applications within the field of ENT oncology include oral mucosal lesions, thyroid nodules for follicular adenoma versus carcinoma, and parathyroid glands for normal versus adenoma. During oral (mandibulectomy) surgery, assessment of bone tissue may be possible, which is significant since bone cannot be processed for frozen pathology.

Therefore, there has been a long-standing interest in developing *in vivo* microscopes for real-time non-invasive microscopic examination of vital tissues. Such in situ evaluation can circumvent some of the drawbacks of *ex vivo* tissue processing and can provide immediate (and potentially more accurate) feedback to the clinicians, thus accelerating and improving diagnoses and treatments.

1.3 SUCCESS OF REFLECTANCE CONFOCAL MICROSCOPY FOR DERMATOLOGY

In major studies that involved blinded analysis and correlation to pathology in 350 patients, both melanocytic lesions (including melanomas) and nonmelanocytic lesions were detected with sensitivity of 92% and specificity of 70-84% [27], [28]. This is a significant advance over the specificity of 32-39% with the current standard of visual examination by dermoscopy [27], which is limited mainly to melanocytic lesions. Another study reported sensitivity of 97% and specificity of 98% [29]. Basal cell carcinomas, too, are detected with 97% specificity and 92% sensitivity [30]. Note that inflammatory infiltrates are often associated with basal cell carcinomas, and this detection accuracy was achieved in the presence of such conditions. The combination of optical sectioning, resolution and contrast allows cancers and pre-cancers to be differentiated from normal skin and from benign and inflammatory lesions. These prior successes give us confidence that a miniaturized reflectance confocal microscope would have a high impact for the detection and management of head and neck cancers.

1.4 CONFOCAL MICROSCOPY

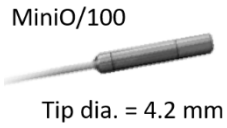
Over the past few decades, various portable research microscopes, and commercialized *in vivo* microscopy systems, have been developed to address a host of clinical needs [29], [31]–[36], [37], [38]–[57]. Many of these devices have been based on the technology of confocal microscopy [58], which provides cross-sectional images of intact specimens (i.e., optical sectioning) with high resolution and contrast (i.e., signal-to-background ratio, SBR).

There are three commercial handheld confocal microscopes currently on the market. The Cellvizio systems (Mauna Kea Technologies, Paris, France) are confocal microscopes for preclinical research and FDA-approved clinical applications. In these systems, laser scanning is performed on the proximal end of a coherent fiber bundle such that the flexible fiber probe can be extremely thin. The CellVizio systems output 12 frames per second, utilizing 488- and 660 -nm lasers as excitation for fluorescence imaging. Imaging parameters are probe dependent with 1.4 - 3.5 μm lateral resolution, 10 - 30 μm axial resolution, 0 - 70 μm imaging depth (fixed for each probe) and 240^2 - 600^2 μm^2 fields of view. The sizes of the probes range from 0.40- to 4.2-mm in terms of tip diameters. The Convivo system (Carl Zeiss AG, Oberkochen, Germany) operates at 488 nm. The system uses a miniature mechanical scanner at the tip of the device which can only output 0.75 frames per second. However, resolution is higher than the CellVizio system, with ~ 0.5 μm lateral resolution, ~ 4.75 μm axial resolution and field of view of ~ 475^2 μm^2. The diameter of the probe tip is 5 mm. The VivaScope 3000 (Caliber I.D. Inc., Andover, MA) is designed for reflectance imaging at 830 nm. This device can output more than 6 frames per second with ~ 1.25 μm lateral resolution, ~ 5 μm axial resolution, and a field of view of ~ 750^2 μm^2 field of view. Since this device uses galvo scanners and a large objective lens for imaging, the probe is too bulky for imaging the oral cavity. Note that the Cellvizio system and the Convivo systems are not

designed for reflectance imaging. In summary, each system has pros and cons, but there is a need for a system that combines the best in terms of resolution, imaging speed, and contrast. In terms of size, the device should be small enough to probe the more-exposed regions of the oral cavity.

In confocal microscopy, a spatial filter (e.g., a pinhole or a slit) is placed at a conjugate image plane so that only the signal originating from a localized focal volume within the tissue is efficiently transmitted onto a detector, while out-of-focus and multiply scattered photons (from tissue regions away from the focal volume of interest) are largely blocked by the pinhole (or slit). In order for a conventional confocal microscope (referred to as a single-axis confocal, SAC, microscope in this article) to achieve subcellular resolution while retaining adequate field-of-view (FOV) and working distance (WD), a bulky high numerical aperture (NA) objective lens, and a complex scanning mechanism, are typically needed. Even though a fiber bundle could help to shrink the size of the tip, the spatial resolution is limited when a coherent fiber bundle is used. A portable SAC microscope that images with high resolution and contrast is technically difficult to engineer but has been realized by several groups in academia and industry [31]–[41], [45]–[48]. As an alternative to conventional SAC microscopy, dual-axis confocal (DAC) microscopy relaxes, to some degree, the requirements for high-NA focusing in order to achieve efficient optical sectioning and provides certain benefits (as described later in this article) for the miniaturization of confocal microscopes down to the scale of several millimeters. In addition, DAC microscopy has demonstrated superior contrast (i.e., SBR) and imaging depth compared to SAC microscopy. In particular, with a line-scanned DAC microscope configuration (as described later in Chapters 3&4), more than 30 frames per second is possible, which significantly reduces the motion artifacts caused by handheld usage. All these advantages can be of clinical value for non-invasive real-time *in vivo* pathology.

Table 1.1. Specifications of commercial handheld confocal microscopes, and our custom device

	Cellvizio system (Mauna Kea Technologies)	Convivo system (Carl Zeiss AG)	VivaScope 3000 (Caliber I.D. Inc.)	Handheld LS-DAC microscope (our custom device)
Lateral resolution (μm)	1.4 – 3.5	~ 0.5	~ 1.25	~ 1.0
Axial resolution (μm)	10 - 30	~4.75	~5	~ 1.72
Field of view (μm^2)	240^2 - 600^2	~ 475^2	750^2	450^2
Frame rate (frames/sec)	12	0.75	≥ 6	≥ 20
Photo of the device				

Chapter 2. BACKGROUND

2.1 DAC ARCHITECTURE

Dual-axis confocal (DAC) microscopy, which was inspired by the concept of "theta" confocal microscopy , was developed in the early 2000s to address certain limitations of conventional SAC microscopes [62], [63]. DAC microscopy achieves optical sectioning using the same basic principle of confocal detection as conventional SAC microscopy, in which a spatial filter (i.e., a pinhole or slit) is used to reject out-of-focus light. The fundamental difference between SAC and DAC microscopy is how the focal volumes are generated and defined. In the SAC configuration, the illumination and detection optics share the same objective (**Figure 2.1 (a)**), which alone defines the focal volume of the system. However, in a DAC system (**Figure 2.1 (b)**), the illumination and collection beam paths (blue and green lines, respectively) do not overlap except at their foci. The

effective focal volume of a DAC system is therefore defined by the region where the two individual foci intersect (black oval).

The spatially separated illumination and collection path of the DAC architecture provides certain advantages. First, in a DAC configuration, the effective axial resolution (optical-sectioning thickness) of the system is proportional to 1/NA, rather than 1/ NA2 as in a SAC configuration, allowing a DAC system to achieve more effective optical sectioning using low-NA lenses (NA < 0.5), which are often lower in cost and more easily miniaturized. Second, the use of low-NA beams provides a longer WD such that it is possible to place a scanning mirror between the focusing optics and the sample [64]. Such "post-objective scanning" allows the beams to maintain their on-axis alignment through the focusing optics regardless of the angle of the scanning mirror. By eliminating the possibility of off-axis aberrations, simple inexpensive low-NA lenses, such as injection-molded aspheres, may be used rather than bulky compound objectives. Low-NA lenses are also more-easily scaled down to millimeter dimensions while maintaining a reasonable FOV and WD. Third, the DAC architecture has been shown to improve the imaging contrast (thus imaging depth) in highly scattering fresh tissues, mainly because the smaller acceptance cone of the low-NA collection beam, and the well-separated illumination and collection beam paths, are better for rejecting the multiply scattered photons [65], [66] that are the main contributors to the background for confocal microscopy within fresh tissues. Finally, the illumination and collection beam paths can be independently aligned to compensate for any chromatic effects due to the Stokes shift of fluorophores.

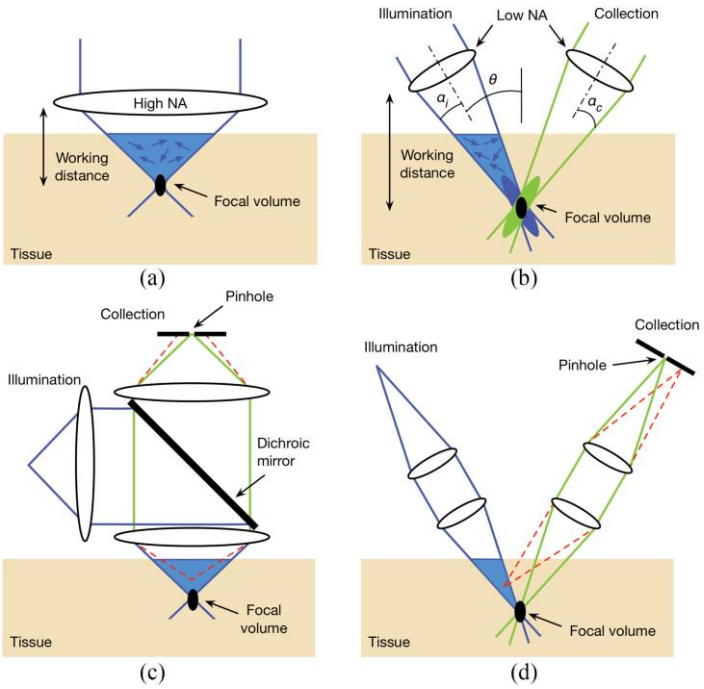


Figure 2.1. Comparison of the optical configurations for a conventional single-axis confocal (SAC) microscope and a dual-axis confocal (DAC) microscope. (a) In order to achieve a tight focal volume (black oval), a SAC microscope requires a high-NA objective lens. This results in a short working distance that makes miniaturization and beam scanning more difficult. (b)A DAC microscope uses low-NA off-axis illumination and collection beams, in which the focal volume is defined by the overlapping foci of the two beams. The use of low-NA beams allows for a longer working distance, which provides advantages for miniaturization and beam scanning. (c) In SAC microscopy, out-of-focus light (an example beam path is shown with the dashed red lines) is not completely rejected by the pinhole. (d) In DAC microscopy, out-of-focus light is directed away from the pinhole and is more optimally rejected, thereby improving the signal-to-background ratio.

The point spread function (PSF) derived from diffraction theory predicts the theoretical response of an imaging system under ideal conditions (i.e., the diffraction-limited performance). In a typical DAC system (**Figure 2.2**), two low-NA beams with focusing angles of α_i and α_c, respectively, intersect at a half-angle of θ. Using diffraction theory under the paraxial approximation [62], [67], the PSF of each beam can be calculated and is depicted by the blue and

green cigar-shaped ovals in **Figure 2.2 (b)**. The overall PSF of the DAC microscope is defined by the product (the black oval) of the intersecting PSFs of the illumination and collection beams. For truncated circular Gaussian beams, the amplitude PSF, U, which describes the spatial distribution of the electric-field amplitude of a beam, is proportional to the Huygens-Fresnel diffraction integral [68]:

$$U \propto \int_0^{\alpha} W(\rho) J_0(\rho) \rho d\rho \qquad (1)$$

where W is the weighting function that accounts for the beam truncation, J0 is the zero-order Bessel function, ρ is a normalized spatial variable that describes the distance from the optical axis of the beam (i.e., the beam radius). The upper limit of the integral, a, is determined by the size of the aperture. Here we assume the use of pure Gaussian beams that are not truncated by apertures (i.e., a = ∞). For the case of apodized (truncated) beams, a more-detailed treatment may be found in a previous publication [67].

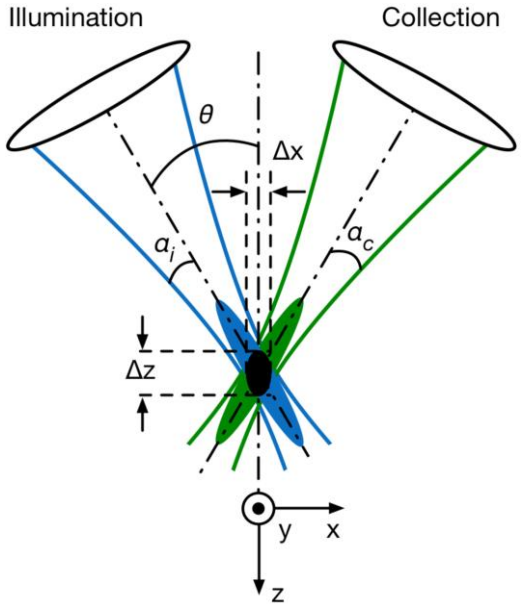


Figure 2.2. The DAC microscope architecture. Two low-NA beams (illumination and collection) with focusing angles of α_i and α_c, respectively, intersect at a half-angle of θ. The focal volume (black oval) of the system is defined by the product of the intersecting point spread functions

(PSFs) of the illumination (blue) and collection (green) beams. The dimensions of the focal volume (Δx, Δy, and Δz) correspond to the spatial (lateral and axial) resolutions of the system.

The response of the DAC system to a delta-function point object (power received at the detector), I_{DAC}, is proportional to the square of the product of the illumination and the collection amplitude PSFs, U_i and U_c:

$$I_{DAC} \propto |U_i \cdot U_c|^2 \quad (2)$$

The full-width at half-maximum (FWHM) extent of IDAC is often used to quantify the dimensions of the DAC focal volume and provides an approximation of the spatial resolution of the imaging system. By assuming that the illumination and collection wavelengths are identical ($\lambda_i = \lambda_c = \lambda$), and that the focusing NA of the illumination and collection beams are also identical ($\alpha_i = \alpha_c = \alpha$), the FWHM spatial resolution along each dimension can be calculated as [42]:

$$\Delta x = \frac{0.297\lambda}{n\alpha \cdot cos\theta}, \Delta y = \frac{0.297\lambda}{n\alpha}, \Delta z = \frac{0.297\lambda}{n\alpha \cdot sin\theta} \quad (3)$$

where n is the refractive index of the medium, and $0 < \theta < \pi/2$.

In comparison, the theoretical resolution derived from diffraction theory for a SAC microscope system with uniform illumination are [63], [69]–[73]:

$$\Delta x = \Delta y = \frac{0.4\lambda}{NA}, \Delta z = \frac{1.4\lambda}{NA^2} \quad (4)$$

These results indicate that the spatial resolution of a DAC system in the x and z directions (the plane of intersection of the DAC beams) are θ-dependent, and that the axial response, Δz, of a DAC system is inversely proportional to the NA of the lenses, rather than to the square of the NA as with a SAC microscope. Therefore, a DAC system is able to provide effective optical sectioning even when low-NA lenses are used, especially when NA < 0.5, as is the case for DAC microscopy.

In addition, as shown in **Figure 2.3 (a)**, the axial response of a DAC microscope will be steeper compared with a SAC microscope that has an identical FWHM axial resolution.

The PSF and axial response derived from diffraction theory cannot predict the effects of tissue scattering, which prevents ballistic photons (i.e., unscattered photons that are "diffraction limited" in terms of their trajectory and focusing abilities) from penetrating deeply within tissues. Therefore, Monte-Carlo ray tracing models have previously been developed to simulate the performance of various DAC microscope configurations in scattering media [65], [66], [74](**Figure 2.3 (b)**), including the role of θ and α on sectioning performance (contrast) [66], [74], [75]. The simulation results showed that both the imaging contrast (signal to background ratio, SBR) and axial resolution consistently improve as θ is increased, suggesting that the crossing angle of the two beams in a DAC system should be maximized when possible. In general, the contrast is more sensitive to the crossing angle whereas the resolution is sensitive to both θ and α (the NA of the beams) [75]. Maximizing both parameters typically gives the best performance, but also implies larger device sizes and/or shorter working distances, and potentially creates additional aberrations in the system.

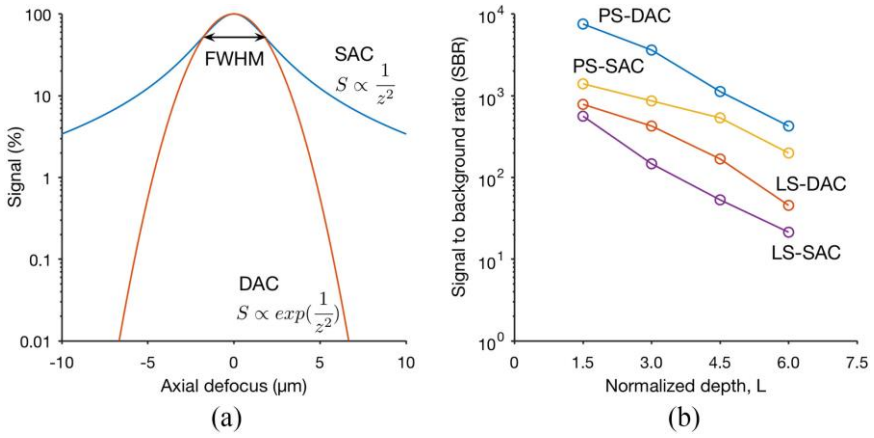


Figure 2.3. Quantitative comparison of the axial-sectioning response (a) and contrast (b) of typical SAC and DAC systems. (a) The theoretical axial response of a SAC and DAC microscope is shown, in which a point object is translated through the focus of the microscope in the z direction.

In this case, the SAC and DAC microscopes have equivalent axial resolutions (FWHM). The signal rolls off more quickly in a DAC system (red) than in a SAC system (blue), showing that the axial sectioning performance (rejection of out-of-focus light) is superior for the DAC configuration. (b) Monte-Carlo scattering simulations to compare the contrast (signal-to-background ratio, SBR) of various microscope configurations when imaging an in-focus reflective object within highly scattering biological tissues, as a function of depth. The results show that both the point-scanned (PS) and the line-scanned (LS) versions of DAC microscopy provide superior image contrast in scattering media when compared with their SAC microscopy counterparts. The normalized depth refers to the number of mean free paths that ballistic photons would travel in a round-trip perpendicular path from the tissue surface to the focal volume, i.e., $L = 2\mu_s d$, where μ_s is the scattering coefficient, and d is the imaging depth in the direction normal to the tissue surface. [74].

2.2 DAC Imaging Systems

2.2.1 *Basic Components*

1) Geometric Considerations: DAC microscope prototypes were initially developed as tabletop systems with primarily off-the-shelf optical components. For example, in several early systems [62], [64], [66], [67], [76]–[78], two low-NA (typically around 0.2) objective lenses were oriented at a half crossing angle of 30 deg. Note that the NA and the crossing angle have been varied in different DAC systems, and have often been constrained by pragmatic concerns such as working distance and device size, as well as the position and size of the scanning mechanism (e.g., galvanometric or MEMS scanning mirrors).

2) Post-Objective Scanning: As discussed previously, the long working distance of low-NA beams provides room for a scanning mirror to be placed after the focusing optics, allowing the system to avoid the off-axis aberrations that occur with pre-objective scanning. This scanning scheme allows for large FOVs to be achieved even when small and simple focusing optics are used. Note that the same scanning mirror may be used to steer both the illumination and collection

beams simultaneously (scanning the illumination beam while de-scanning the collection beam), which helps to ensure that the beams remain well-aligned [42], [64].

3) Hemispherical Solid Immersion Lens: The use of a hemispherical fused-silica solid immersion lens (SIL) has been a distinguishing feature of many previous DAC microscope prototypes [67], [76], as well as a related open-top light sheet microscope design described recently [79]. The SIL provides several advantages:

1. Minimizing off-axis aberrations: the curved surface of the hemisphere provides a normal interface for both the illumination and collection beams as they transition from air into the glass SIL, such that off-axis aberrations (e.g., coma and astigmatism) are minimized.
2. Minimizing spherical aberrations: the wave-front curvature of the focused beams is matched to the curved surface of the hemisphere, minimizing the spherical aberrations that would result from focusing a beam through a flat interface between two media with different refractive indices (e.g., from air to glass). Note that slight aberrations still occur while scanning the beams away from their ideal neutral positions.
3. Refractive index matching: since the refractive index of fused silica (n = 1.45) is similar to that of most biological tissues, aberrations are minimized as the beams travel across the interface between the tissue sample and the flat surface (distal surface) of the hemisphere.
4. Increasing the effective focusing NA: if a beam is being focused from air into a higher-index material through a flat interface, the NA is preserved due to Snell's law (where NA = n sinα). However, the curved interface of the hemisphere, when well-aligned, preserves the ray angles of the focused beams, thus increasing the NA by a factor n. On the other hand, the curved surface of the hemisphere also acts to de-magnify the scanning range of the beams, to first order, by 1/n. For example, in the case of fused silica (n = 1.45), a 250-

μm axial translation of the stage causes the focal volume to translate ~150 μm within the sample. Similarly, lateral translations are also de-magnified by roughly 1/n.

2.2.2 *DAC Microscopy Variations*

Early DAC microscopes utilized point-focused Gaussian illumination in conjunction with point-by-point confocal detection along a raster-scanned [57, p. 20078] or Lissajous-scanned trajectory [42]. While these systems demonstrated the advantages of the DAC architecture, their speed was often limited due to the use of point scanning. In addition to improving imaging speed, the ability to image deeper was also desired for certain *in vivo* imaging applications. To extend the potential of DAC microscopy, a wide range of DAC-microscopy variants have been explored. These approaches can be grouped into two main categories based on their aims: (1) improving the imaging speed by modifying either the scanning mechanism (e.g., with complex scanning mirrors [57], [80]) or the illumination pattern (e.g., line scanning [74], [76]); (2) improving the imaging depth (i.e., image contrast) by using different illumination sources and beam profiles (e.g., near infrared lasers [57], [64], Bessel beams [81]) or through advanced detection methods (e.g., temporal gating, lock-in detection [64], [67], [82]). Due to space constraints, this section highlights only the variants that have been developed within the past five years.

1) Line-Scanned Dual-Axis Confocal (LS-DAC) Microscopy: *In vivo* imaging with a handheld device is subject to motion artifacts induced by the subjects and users. For example, early miniaturized point-scanned (PS) DAC systems (**Figure 2.4 (a)**) had a limited frame rate of<5 Hz, which led to frequent motion artifacts (blurring and distortions) during *in vivo* use. Although it is possible for a PS-DAC system to acquire images at video rate [57], [70], the FOV is often limited or the scanning mechanisms are complicated, expensive, and difficult to scale down in size. As a result, a simpler line-scanned (LS) DAC microscopy (**Figure 2.4 (b)**) approach has been utilized

in recent years to improve the imaging speed, with a trade-off in image contrast (SBR) due to the loss of confocality along one dimension (along the focal line). In brief, LS-DAC microscopes illuminate a focal line in the sample, instead of a localized point. That focal line is imaged by the collection optics onto a linear-array detector. Since an entire line of pixels is imaged and acquired simultaneously, the focal line (or tissue) only needs to be scanned in one direction to generate a 2D image. In addition to reducing the cost and complexity of the scanning mechanism, line-by-line data collection also has the potential to improve the signal-to-noise ratio (SNR) of the images because of the longer pixel integration times.

The first LS-DAC microscope utilized a cylindrical lens in the illumination path to generate a long focal line along the y axis of the imaging plane [76]. On the detection side, instead of using a single-mode fiber as a pinhole, as was done in previous PSDAC systems, a digital slit (cropped region) was defined within a detector array to serve as a spatial filter for confocal detection. This configuration eliminated the need for a fast 2D scanning mirror to create an image, such that high-speed imaging could be achieved with a slow 1D galvanometric mirror [78], [83]. The frame rate was thus primarily limited by the sensitivity of the detector, rather than the mechanical scanning mechanism. It has been shown that LS-DAC and PS-DAC microscopy exhibit comparable imaging performance at shallow depths (<150 μm), but that the image quality (contrast) of LS-DAC microscopy deteriorates at deeper depths because of the reduced confocality of a line-scanned system [76].

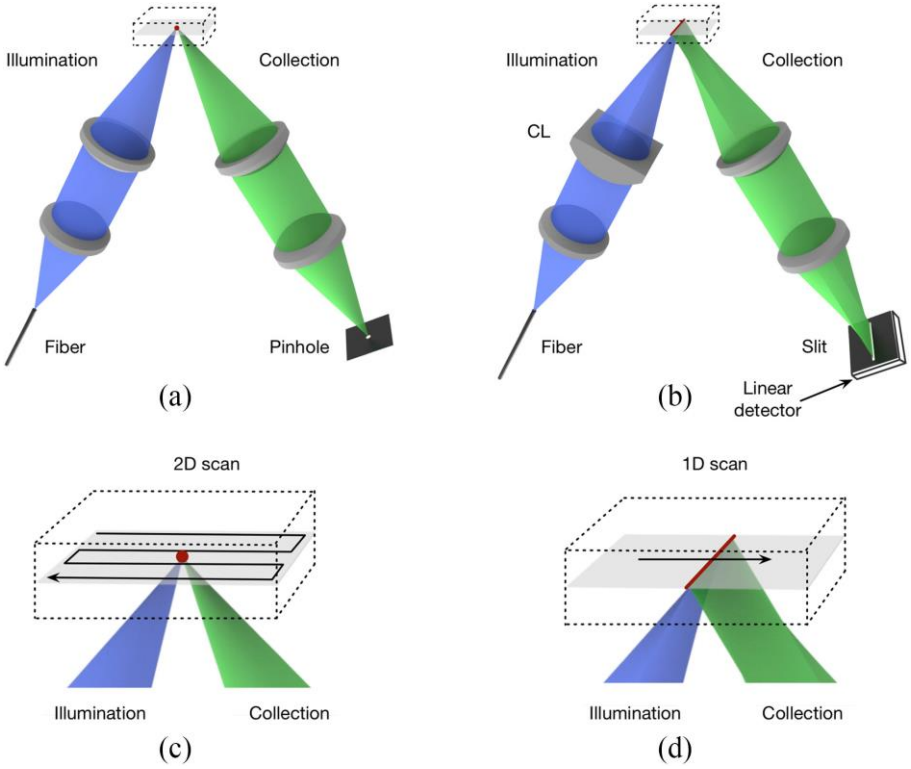


Figure 2.4. (a) In a point-scanned DAC system, light is tightly focused to a point within the sample, and a pinhole is used for confocal detection. To create an image, the point is scanned in two dimensions (e.g., following a raster- or Lissajous-scanned trajectory) and the image is generated point by point. (b) In the case of a line-scanned DAC system, the illumination objective lens is replaced with a cylindrical lens (CL) so that light is focused to a thin line within the sample, and a slit is used for confocal detection. The focal line only needs to be scanned in one dimension to create a 2D image. (c)–(d) Zoomed-in views of the scanning trajectories described in (a) and (b) are shown, respectively.

2) Sheet-Scanned Dual-Axis Confocal (SS-DAC) Microscopy: As discussed in the previous section, one of the drawbacks of LS-DAC microscopy is the deteriorated image contrast due to the loss of confocality along the focal line. Sheet-scanned (SS) DAC microscopy was developed to mitigate the reduced performance in LS-DAC microscopy by utilizing the additional spatial information provided by a 2D detector array [84]. The key principle of this approach is to partially mitigate the loss of confocality by utilizing the useful information contained in the slightly

"out-of-focus" regions near the focal line that are typically rejected by a physical slit in a line-scanned confocal microscope. In particular, if the crossing angle of a DAC microscope approaches 90 deg, then the collection arm essentially images a "light sheet" generated by the illumination arm, from which additional information is provided that can be used to perform 3D deconvolution.

The SS-DAC concept was first demonstrated with a scientific complementary metal-oxide-semiconductor (sCMOS) detector array to image an oblique light sheet (effectively, an angled light sheet). By scanning the angled light sheet, a thin 3D volume of data was acquired, which could then be used for deconvolution of the main LS-DAC image if the PSF of the system was known (either measured or simulated). This simple technique was shown to improve the spatial resolution and contrast of the LS-DAC system [84].

Note that SS-DAC is slightly different from the recently developed technology of light-sheet microscopy (LSM), also known as selective plane illumination microscopy (SPIM) [79], [85]. In a SS-DAC microscope, the illumination beam is focused with a moderate NA to generate a fairly localized focal line, whereas LSM typically utilizes a lower NA to generate a thicker light sheet with a long depth of focus (in order to image a larger 3D volume by scanning the light sheet in only one direction). With LS-DAC and SS-DAC microscopy, the goal is to generate a 2D image by scanning in one direction, rather than a 3D volume of data. However, with SS-DAC, some 3D information is captured (similar to LSM), but for the purposes of 3D deconvolution to improve one 2D image, rather than to obtain volumetric information. LSM generates more data (in 3D) but sacrifices contrast due to the use of a very low-NA illumination beam whereas LS-DAC and SS-DAC generate 2D images, but with higher contrast [86]. In general, DAC microscopy is more ideal for imaging highly scattering fresh tissues (including *in vivo*) whereas LSM is ideal for rapid 3D microscopy of relatively transparent model organisms and optically cleared ex vivo tissues.

3) Modulated-Alignment Dual-Axis (MAD) Confocal Microscopy: MAD confocal microscopy is a technology that combines the inherent strengths of focal-modulation microscopy [87] and PS-DAC microscopy, with the aim of improving image contrast and depth[82]. In a PS-DAC microscope, illumination and collection beams are spatially separated except at one single point (at the focus of the microscope). Optical sectioning with DAC microscopy relies on the precise alignment of the illumination-beam's focus and detection-beam's focus at the sub-micron level. For example, it has been shown that the confocal signal is reduced by an order of magnitude when the two beams are offset by only 1.4 times the beam radius (i.e., a distance on the order of a micrometer). This feature provides the opportunity to implement a "spatial overlap modulation" technique that was originally used in the context of nonlinear microscopy [88].

The first MAD confocal microscope utilized an acousto-optic deflector (AOD) in the illumination beam path to sinusoidally scan the illumination beam over a small range (+/− a few micrometers from the well-aligned condition) in the direction perpendicular to the plane defined by the dual-axis beams. This spatial modulation was performed at a frequency f, resulting in a modulated signal at a frequency of 2f, which could be detected and distinguished from the static (non-modulated) background signal using 2f lock-in detection. This strategy was shown to improve the image contrast (SBR) by ∼6 dB in scattering media in comparison to standard PS-DAC microscopy [82]. There are a few limitations to the MAD confocal microscopy approach that have been mitigated to some degree with recent improvements to the technology [89]. First, the diffracted light used as a spatially modulated illumination source in this system can vary in intensity over time due to the fact that the AOD diffraction efficiency typically varies with scanning angle, which can lead to a modulated background signal that competes with the MAD signal. Second, the acoustic wave within an AOD crystal has a limited propagation speed (∼3.63

mm/μs in the early prototypes), which limits the modulation rate and thus the maximum frame rate for MAD imaging. Finally, the MAD technique may be limited in tissues with refractive heterogeneities, which introduce aberrations and misalignments of the beams that can reduce the modulation depth of the MAD signals.

4) Bessel Dual-Axis Confocal (DAC) Microscopy: Bessel beams have been investigated as a means of improving deep-tissue microscopy in highly scattering and heterogeneous media [81], [90]–[93]. As discussed in the previous sections, DAC microscopy requires the precise intersection of two beams at their respective foci (micron scale). Therefore, the pointing accuracy of the beams and the quality of their foci are critical for optimal performance. A few recent studies have explored the adaptation of Bessel illumination for DAC microscopy and have shown that Bessel beams exhibit improved pointing accuracy and beam quality in samples with refractive heterogeneities, in comparison to conventional Gaussian beam [81], [93]. Consequently, spatial resolution is maintained more effectively with Bessel-DAC microscopy compared with standard Gaussian-DAC microscopy. One drawback of Bessel illumination is that the diffraction side lobes contain a significant amount of the beam energy and contribute to an out-of-focus background that reduces image contrast [81]. Various approaches have been proposed to mitigate this effect for other imaging modalities, such as through the use of two-photon excitation and structured illumination, etc. [90], [92], [94], [95] Similar strategies are still under investigation for DAC microscopy systems.

5) Divided-Pupil Systems: As an alternative to using two separated objectives as discussed in the previous sections, off-axis illumination and collection has also been achieved using a single high-NA lens with a “divided pupil”, i.e., using one half of the lens for illumination and the other half for collection [96]–[100]. It should also be noted that a few light-sheet microscopy variants –

for example, oblique plane microscopy [101] and swept confocally aligned planar excitation (SCAPE) microscopy [102] – have also utilized a similar configuration (off-axis illumination and collection beam paths that share one large objective lens) to achieve high-speed volumetric imaging.

2.3 PORTABLE DAC SYSTEMS

Perhaps the greatest constraint for the design of *in vivo* microscopes is size. Although the DAC architecture has many unique properties that significantly simplify its miniaturization, smaller form factors are typically associated with reduced performance, as well as increased design complexity and cost. Each clinical device is designed for a specific biomedical application and careful deliberation is necessary to arrive at the most optimal set of design trade-offs. This section surveys some of the miniature DAC systems that have been built to address clinical applications of the detection of head and neck cancers. Furthermore, since portable DAC systems can be used for the other scenarios, this survey also includes some existing portable DAC systems for the applications of gastrointestinal endoscopy, dermatopathology, and neurosurgery.

2.3.1 *Label-Free in Vivo Reflectance Microscopy of Skin and the Oral Cavity*

Imaging the nuclear morphology and tissue architecture down to the dermal-epidermal junction (~100 μm below the skin surface) is valuable for the diagnosis and treatment of basal cell carcinoma, one of the most common cancers of the skin and oral cavity. As an alternative to invasive biopsy and histopathology, label-free reflectance confocal microscopy can provide a powerful method to allow suspicious lesions to be non-invasively and rapidly examined in real time. These clinical applications require the imaging device to be able to achieve (1) a resolution of 5 μm or less to distinguish nuclear morphology, (2) an imaging depth of at least 100 μm to reach

the deeper layers of interest (dermal-epidermal junction), (3) a small imaging head that can fit within the oral cavity, and (4) a high frame rate to reduce motion artifacts during handheld use. A miniature reflectance LS-DAC was developed recently to address these clinical needs and design criteria. This device is similar to the fluorescence version that was discussed in the previous section but is more compact owing to the use of a small and inexpensive (<$100) linear detector array integrated within the device. Although not as sensitive as the detector used in the fluorescence LS-DAC device, the low-cost linear array in the reflectance LS-DAC device has sufficient sensitivity for reflectance imaging, which generates much more signal than fluorescence imaging.

The portable divided-pupil line-scanned confocal microscope developed at Memorial Sloan Kettering Cancer Center is an additional example of a clinical device for label-free *in vivo* reflectance imaging of human skin (**Figure 2.5**) [103]. In a divided-pupil system, the pupil of a high-NA (0.9, water-immersion) objective is divided into two halves, one for the illumination beam and the other half for the collection beam. An endoscopic relay lens was incorporated at the distal end so that the oral cavity could be accessed. With such reflectance-based devices, sub-cellular optical sectioning has been achieved at a depth of ~100 μm in human skin, while achieving a frame rate of 8 Hz over a FOV that is comparable to that of a standard 20X objective lens. Note that reflectance confocal microscopy has also been used as a complementary imaging modality to another common reflectance-mode optical imaging technique, optical coherent tomography (OCT). Reflectance confocal microscopy provides high resolution images at a limited imaging depth while OCT provides lower resolution images at a greater imaging depth [104].

This chapter provides a review of DAC microscope technology, an optical imaging modality that utilizes low-NA beams to achieve effective optical sectioning and superior image contrast in biological tissues. In contrast to the conventional confocal microscope invented by Minsky in 1957

[58], a DAC microscope utilizes spatially separated off-axis illumination and collection beam paths that only intersect at their foci – an optical architecture inspired by the initial works of Stelzer et al. [59], [61], and Webb et al. [63] in the 1990s – to improve imaging contrast/depth and to improve miniaturization. Since its first introduction in 2003 [64], DAC microscopy has been significantly improved in terms of imaging depth (>500 μm), speed (>30 fps), resolution (<1 μm laterally; <3 μm axially), and size (e.g., endoscope-compatible). Several portable DAC systems have been developed for a wide range of clinical applications such as for intraoperative guidance and for early disease detection, with some systems currently in the process of clinical testing. In addition to confocal microscopy (SAC and DAC), there are many other portable optical sectioning microscopy systems such as multiphoton microscopy [49]–[52], OCT [53], [105], structured illumination microscopy [54]–[56], etc. These technologies are also being investigated as non-invasive and real-time alternatives to conventional biopsy and histopathology and have been shown to be useful for certain clinical applications. Many of these systems have been discussed in a prior review [26]. Finally, it should be noted that the research and development of these miniature clinical devices has both facilitated and benefited from the advancement of miniature optical components such as MEMS scanners, ultra-small lenses, fiber-optics technologies, as well as compact detectors and other hardware for high-speed image acquisition and processing.

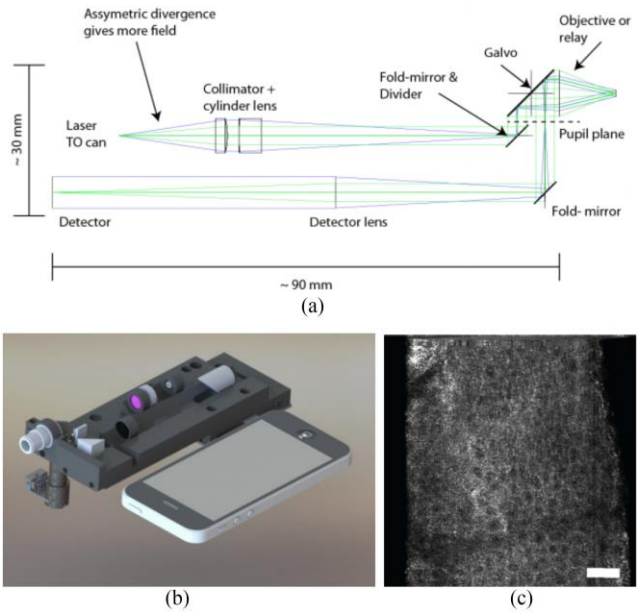


Figure 2.5. (a) Optical circuit of a compact line-scanned divided pupil confocal system. The pupil of the objective lens is physically divided into two halves, one to generate the illumination beam and the other for the collection path. A single galvanometric mirror is used to scan both beams to create an image. (b) A design rendering of the system is shown alongside a smartphone. (c) A label-free in vivo image is shown of human epidermis. [103].

2.3.2 *In Vivo Endoscopic Microscopy of Hollow Organs*

There is a clinical need for improved early detection and image-guided therapy of diseases in hollow organs such as the gastrointestinal (GI) tract. A critical requirement for such applications is that the device be small enough to fit within the instrument channel of a standard GI endoscope, which has a diameter of several millimeters. An endoscope-compatible DAC microscope with a diameter of 5.5-mm was developed at Stanford University for *in vivo* GI imaging at a frame rate of 5 Hz [106]. The device was deployed through the instrument channel of an endoscope and was used to image the colonic mucosa of patients after topical application of an FDA-approved contrast agent (sodium fluorescein) [106]. Facilitated by the advancement of MEMS technology, a number of variations upon this initial DAC endomicroscopy design have been developed with improved

performance for a broad range of clinical applications [33], [80], [107]–[109]. For example, a recent publication from the University of Michigan describes a DAC endomicroscope that incorporates a state-of-the-art 3D MEMS scanner that can alternate between a "tilting mode" (with actuation along the x and y axes) and a "pistoning mode" (with actuation along the y and z axes) to image either *en-face* optical sections or vertical optical sections in real time (**Figure 2.6**) [33], [110]. This specific topic of DAC endomicroscopy has been discussed in greater detail in a separate review [111].

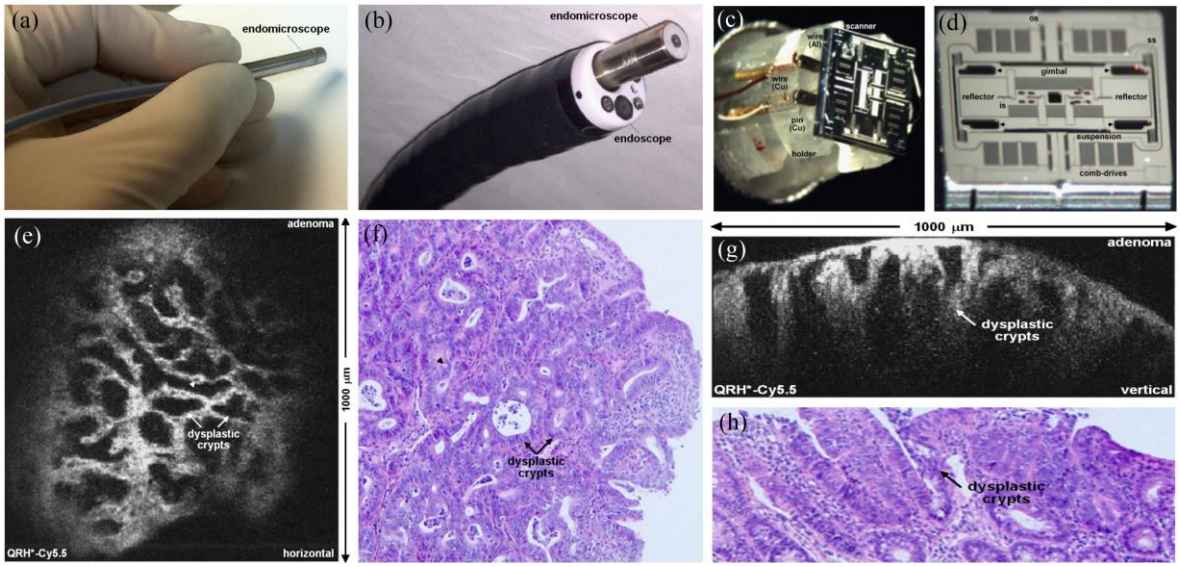


Figure 2.6. Photographs of (a) a miniature DAC endo-microscope that is fitted within (b) a clinical GI endoscope. The head of the imaging probe has a diameter of 5.5 µm. (c)–(d) Photograph and scanning electron micrograph (SEM) of a custom-developed tri-axial MEMS scanner that enables the user to switch between two orthogonal imaging planes (either the en face plane or the vertical plane) in real time. (e) An example of *en face* optical sectioning of mouse colon after intravenous injection of a Cy5.5-labeled peptide, showing the dysplastic crypts (arrows) and goblet cells (arrowheads). (f) A corresponding H&E-stained histology section is shown of the mouse colon. (g) An example of vertical optical sectioning of the same mouse colon, showing EGFR expression from a region of adenoma up to 430 µm below the surface. (h) A corresponding H&E-stained histology section is shown of the mouse colon. [107].

2.3.3 *Intraoperative Neurosurgical Guidance*

Surgical resection (i.e., debulking) is the first step in the treatment of many brain tumors such as gliomas, in which a greater extent of resection has been associated with improved outcomes. However, the complete resection of gliomas is challenging because the tumor at the margins is often indistinguishable from the surrounding normal brain, and there is no quantitative metric (e.g., tumor-cell density) by which to optimize the extent of resection for these diffuse tumors that infiltrate far beyond the radiologically defined margins. While only a small set of untargeted fluorescence contrast agents (e.g., FITC, ICG, etc.) are approved for clinical use, a new compound, 5-ALA, has recently been approved by the U.S. FDA for neurosurgical guidance, in which a fluorescent metabolic byproduct (PpIX) of the orally ingested agent (5-ALA) is used to highlight the bulk tumor regions and to improve the extent of resection. However, image contrast is often still ambiguous and weak near the diffuse margins of the tumor. It has been suggested that intraoperative microscopy, which can provide images that approach the gold-standard of histopathology, may have value for neurosurgical oncologists to maximize the extent of resection while minimizing neurological damage. In particular, portable optical-sectioning microscopes provide sufficient resolution to detect and potentially quantify the sparse and disseminated tumor-cell populations at the margins of diffuse gliomas. Such sparse cell populations are often not visualized by other imaging technologies (e.g., low-power surgical microscopes, MRI, CT) since they lack the spatial resolution to detect individual disseminated cells, even if such isolated cells are effectively labeled by a contrast agent. A pen-sized handheld DAC microscope with a distal diameter of 1.8 mm was described in 2010, in which images were obtained from the brains of living mice that were genetically engineered to develop medulloblastoma [42]. The device was able to achieve cellular resolution (4 μm laterally and 8 μm axially) at an imaging depth of up to

250 μm with the use of low- NA (~0.075) beams focused with a parabolic mirror. A biaxial MEMS scanning mirror was used for post-objective scanning of the focal volume over a FOV of approximately 0.4 mm by 0.4 mm. The MEMS mirror was axially translated with a piezoelectric actuator to adjust the imaging depth by up to 250 μm. In addition, a customized gradient-index (GRIN) "needle lens" located at the distal tip was used, with a diameter of just 1.8 mm. A major limitation of this prototype was its slow frame rate (~4 Hz) due to the point-by-point Lissajous scanning pattern that was used, which made the device vulnerable to motion artifacts. In addition, the spatial resolution was not ideal. A handheld LS-DAC microscope was developed (**Figure 2.7**) recently with significantly improved frame rates and spatial resolution [29]. As discussed in Sec. III-B, the LS-DAC architecture significantly simplifies the requirements of the scanning mechanism for high-speed imaging. A robust commercial MEMS mirror was used to scan the focal line in one dimension to create *en face* images at video rate (>16 fps). The improved resolution was achieved by using a custom-developed 1:3 demagnifying relay objective at the distal end of the device. The relay lens effectively increased the NA of the beams as well as the crossing angles, at the cost of reducing the FOV. The system was able to achieve a lateral resolution of 1.1 μm with 2.0-μm axial resolution (optical sectioning thickness) over a FOV of roughly 350 μm by 350 μm. This high-speed LS-DAC microscope enabled depth-resolved imaging of red blood cells trafficking within the capillaries of a living mouse, as well as high-contrast imaging of ex vivo tissues stained with FDA-approved fluorophores (e.g., methylene blue). A trade-off of using line scanning in miniature devices is that confocal detection must be achieved with a digital line detector that is directly integrated within the device, instead of using a fiber-coupled point detector (in the case of a miniature PS-DAC microscope).

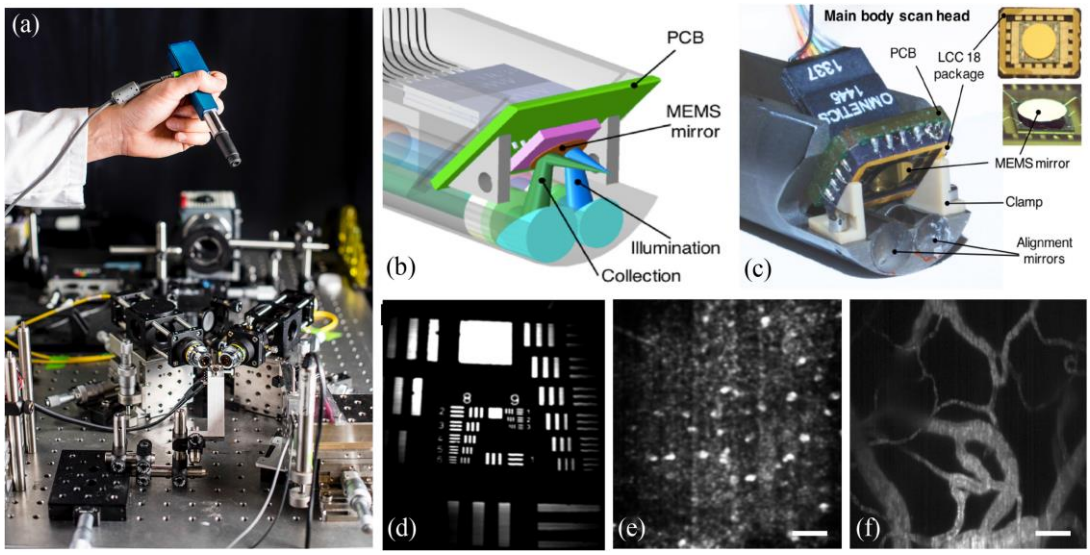


Figure 2.7. A handheld video-rate LS-DAC microscope for intraoperative guidance. (a) A miniature LS-DAC is held above a large tabletop LS-DAC prototype [photograph courtesy of Dennis Wise at the University of Washington]. (b)–(c) A design rendering and photograph, respectively, of the scan head of the handheld LS-DAC device. A single MEMS mirror is used to scan both beams in one dimension to rapidly generate 2D images. (d) An image of a 1951 USAF resolution target, showing the ability to resolve objects as small as ~1 μm. (e) A label-free in vivo image of hyper-reflective nuclei in human oral buccal mucosa. (f) A fluorescence in vivo image (maximum intensity projection from a depth range of 50 to 100 μm) of the vasculature of a mouse ear after retro-orbital injection of FITC-dextran. All scale bars represent 50 μm.

Chapter 3. PROOF-OF-CONCEPT PROTOTYPE OF MINIATURE *IN VIVO* MEMS-BASED LINE-SCANNED DUAL-AXIS CONFOCAL MICROSCOPE

3.1 MINIATURE MICROSCOPE DESIGN

3.1.1 *Microscope modules*

The miniature LS-DAC microscope developed in this study consists of three major modules (**Figure 3.1**): (1) a main body housing the optics that are unique for the illumination beam (blue) and collection beam (green), the MEMS scanning mirror and alignment mirrors; (2) a relay objective lens with a lens cap that provides 3x magnification from the focal plane of the microscope within tissue to the back focal plane that is scanned by the main-body optics; and (3) a detector array module.

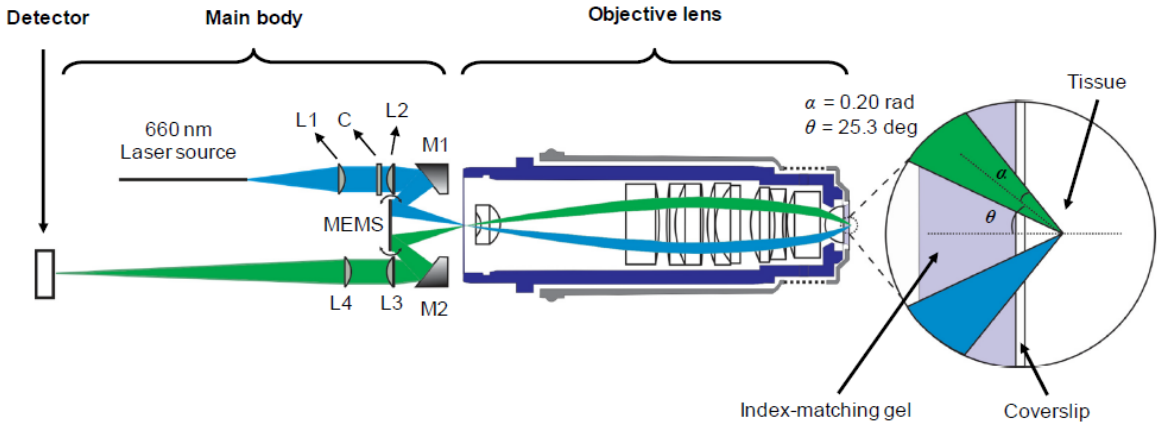


Figure 3.1. Optical circuit of the miniature LS-DAC microscope. The illumination beam path is colored blue whereas the collection beam path is colored green. The main body houses the MEMS scanning mirror, the mirrors (M1 and M2) used to align the dual-axis beams such that they intersect at the back focal plane of the objective lens, and the primary optics (L1, L2, L3 & L4) that focus the dual-axis beams at the back focal plane. The custom objective lens relays the beams from the back focal plane (at the left side of the objective) to the front focal plane in tissue (right side) with

3x de-magnification. The focusing angle of the beams in tissue, α, and crossing angle, θ, enable high-contrast optical sectioning with a resolution of 1 – 2 μm in the lateral and axial dimensions, respectively.

Compared to a previous tabletop LS-DAC system [83], the only major differences for this miniature system are the incorporation of a miniature MEMS scanner (Mirrorcle Technologies Inc.), rather than a large galvanometric scanner, as well as the design and utilization of a custom 3x relay objective (**Figure 3.1**). The 3x objective lens not only boosts the crossing angle (θ) and NA of the dual-axis beams (α) but is also designed to provide effective index matching of these high-NA beams into fresh tissue specimens with minimal aberrations. As a result, the spatial resolution of the miniature system surpasses that of the previous tabletop LS-DAC system (see sections 2.5 and 3.2 for details).

3.1.2 *Illumination focusing module*

A single-mode optical fiber (SM670) is used to couple laser radiation at a wavelength of 660 nm into the illumination path of the main body (Gaussian beams are assumed throughout this work). An illumination fiber module, assembled by GRINTECH GmbH in Jena, Germany, consists of a series of three doublet achromat lenses packaged within a stainless-steel cylindrical tube with an inner diameter of 3.0 mm and an outer diameter of 3.2 mm as shown in **Figure 3.2 (a)**. Lenses L1 and L2 are spherical achromats that were purchased from Edmund Optics in Barrington, NJ (catalog numbers 45262 and 45090, respectively) and then reduced in diameter (3.0-mm diameter) by BMV Optical in Ottawa, Ontario, Canada. Lens C is a custom cylindrical achromat lens, fabricated by BMV Optical, that is based on the lens prescription of a spherical achromat from Edmund Optics (catalog number 45420). Without the cylindrical lens, C, the focusing module would generate a focused spot. The addition of the cylindrical lens converts the illumination fiber

module into an anamorphic focusing system, which causes one axis of the illumination beam to focus first, thereby generating a focal line as desired for LS-DAC microscopy. This optical design is similar to our previous large-scale tabletop LS-DAC prototypes [76], [84]. The focal lengths of lenses L1 and L2 are 9 mm and 12 mm, respectively. This causes the NA of the illumination beam to be reduced from the 0.12 1/e2 NA emitted from the single-mode fiber to a focusing NA of 0.09 at the back focal plane of the 3x objective. The focal length of lens C was selected to provide a focal line with a full-width at half maximum (FWHM) length of 1.2 mm, which corresponds to a 0.4- mm long focal line within tissue (after the 3x objective).

3.1.3 *Main body design*

As shown in **Figure 3.2 (c)** and **(d)**, the main body features two cylindrical channels, one for the illumination fiber module (described in the previous section), and one for the collection-side optics. The collection-side channel holds a pair of doublet achromat lenses, L3 and L4, purchased from Edmund Optics (catalog numbers 63692 and 45345, respectively) and then reduced in diameter (3.2-mm diameter) by BMV Optical. The focal lengths of L3 and L4 are 12 mm and 60 mm, respectively. Therefore, the lenses serve to magnify the image of the focal line onto the linear detector array. As described later, this 5x magnification, coupled with the 3x magnification of the objective lens, provides sufficient magnification (15x) to allow the detector array (with 6.5-μm pixel spacing) to fully sample the focal line (lateral resolution of ~1 μm in tissue) according to the Nyquist sampling criterion. Both the illumination and collection beams are designed to focus and intersect at the back focal plane of the objective lens at a half crossing angle, θ, of 11 deg. The alignment of the beams is achieved by rotating and translating (axially) a pair of 45-deg mirrors that fit within the cylindrical channels in the main body (**Figure 3.2b–3.2d**). The alignment mirrors are adjusted to deflect the illumination and collection beam paths at 11 deg with respect to the x

axis and are then permanently fixed in place. The MEMS mirror is positioned above the beams to re-direct them in the axial (z) direction, with a half crossing angle of 11 deg with respect to the z axis.

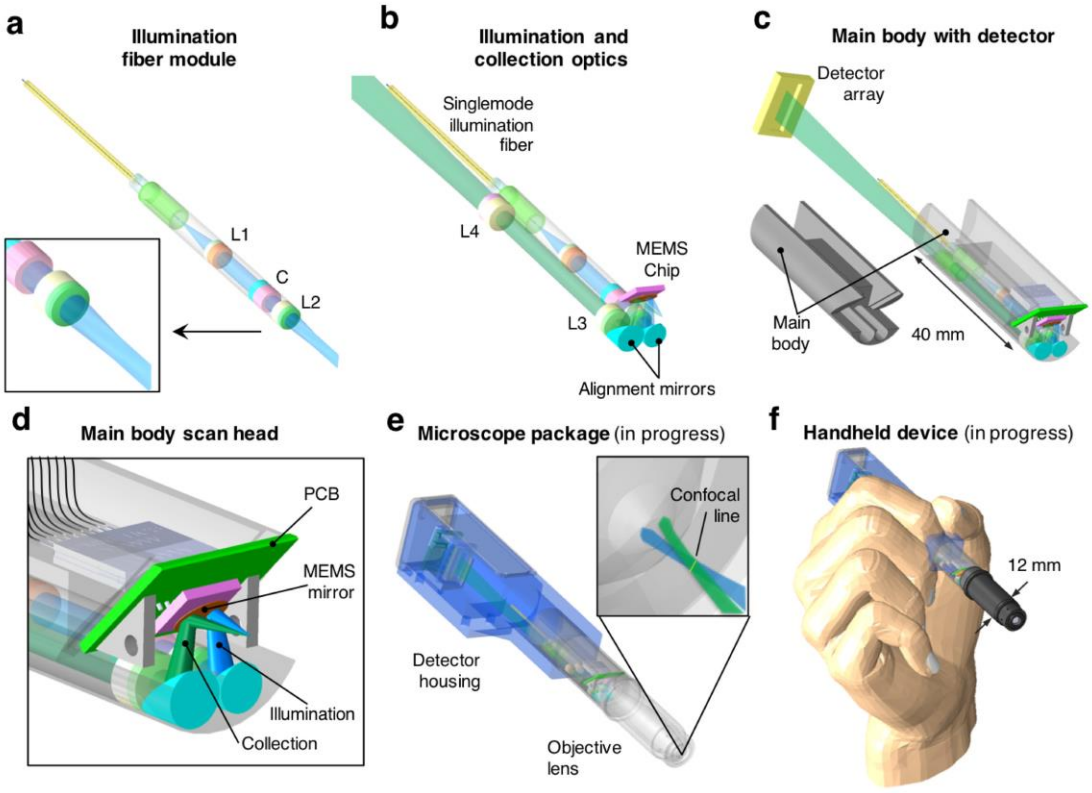


Figure 3.2. (a) Line-focusing illumination fiber module. (b) Illumination and collection optics within the main body. (c) Main body with detector. (d) Main body scan head. (e) Microscope package (was in progress). (f) Handheld device (was in progress).

3.1.4 *MEMS scanning mirror*

The MEMS scanning mirror utilized in our device was purchased from Mirrorcle Technologies in Richmond, CA. The MEMS mirror surface has a circular profile, with a diameter of 3.8 mm, and is actuated in two dimensions (tip / tilt) with electrostatic comb drives. In the LS-DAC device, only one dimension of scanning is necessary to create a two-dimensional image (in a line-by-line fashion). However, the extra dimension of scanning may be useful for correcting certain image

distortions that may be present in our current LS-DAC system. The resonant frequency of the MEMS mirror is in the ~200 Hz range. This would be too slow for a point-scanned system but is more than adequate for a line-scanned system in which a maximum frame rate of 200 frames/sec would theoretically be possible if photon counts are sufficiently high, as may be the case for reflectance microscopy. The square MEMS scanner chip measures 4.3 by 4.3 mm, and is glued and wire-bonded by the manufacturer (Mirrorcle Tech. Richmond, CA) within a rectangular surface-mount LCC18 package that measures 8.89 by 7.24 mm. The LCC18 package is soldered onto a PCB chip (measuring 10.16 by 8.64 mm) that was custom designed by our group and fabricated by Advanced Circuits Inc. The PCB chip provides a mounting surface for the MEMS package, and also routes the electrical connections from the LCC18 MEMS package to a row of conductive vias located on the backside of the PCB. Note that a total of 8 electrical connections, plus a ground connection, are required to fully actuate the two-dimensional MEMS mirror. The 8 high-voltage MEMS connections are supplied by wires that are connected to the backside of the PCB through a two-part miniature 9-pin connector ("Nanominiature" line) from Omnetics Connector Corp., Minneapolis, MN (part numbers A79000- 001 and A79001-001, respectively) as shown in **Figure 3.2 (d)**.

3.1.5 *Objective lens and optical ray tracing*

The purpose of introducing the objective lens into the miniature system is to increase the focusing NA of the illumination and collection beams located at the back focal plane of the objective, as well as to increase the crossing angle of the beams, which improves the resolution of the miniature system [42], [67]. As mentioned previously, the 1/e2 NA of the Gaussian illumination beam at the back focal plane of the objective (in air, where $n_1 = 1$) is equal to $n1\sin\alpha 1 = 0.09$. The objective

lens magnifies the NA at the back focal plane by a factor of 3 and the sample is assumed to have an index of approximately $n_2 = 1.34$. Therefore,

$$\alpha_2 = \sin^{-1}\left(3 \cdot \frac{n_1}{n_2} sin\alpha_1\right) = 0.202 \quad (5)$$

Likewise, since θ1 = 11 deg, θ2 at the front focal plane in tissue is 25.3 deg. According to diffraction-theory calculations [42], [67], in which the illumination and collection beams are assumed to be symmetric and identical (approximately true for our device), these parameters should result in the following theoretical spatial resolutions (in tissue with an index n = 1.34):

$$\begin{aligned} \Delta x &= \frac{0.446\lambda}{n\left(\frac{\pi}{2} \cdot \alpha\right) cos\theta} = 0.77\ \mu m, \\ \Delta y &= \frac{0.446\lambda}{n\left(\frac{\pi}{2} \cdot \alpha\right)} = 0.70\ \mu m, \\ \Delta z &= \frac{0.446\lambda}{n\left(\frac{\pi}{2} \cdot \alpha\right) sin\theta} = 1.64\ \mu m. \end{aligned} \quad (6)$$

A ray-trace analysis of the miniature system was performed in ZEMAX and is shown in **Figure 3.3**. The spot diagram from the illumination path (**Figure 3.3 (b)**) shows that the focal line within tissue does not exhibit significant aberrations, with a root mean square (RMS) spread of < 1 μm across the line. As shown in Eq. (6), this spread is within the diffraction-limited FWHM width of the microscope. When the MEMS mirror is tilted by 4.5 deg, the line is translated by 200 μm in the x direction. Again, the spot diagram shows an RMS spread of < 1 μm across the line, which is excellent for the edge of the field of view. Note that due to the geometry of the beam-scanning method, the focal line tilts slightly as it scans in the x direction. However, this artifact is easily removed in software during image reconstruction and display. For the collection path, we analyzed the optical performance by setting up three ideal point sources at various locations along the focal line in tissue and carried out ray-trace simulations from these ideal point sources to the detector plane. The spot diagrams at the detector plane, shown in **Figure 3.3 (d)**, demonstrate that

the RMS spread is significantly less than the diffraction-limited FWHM spot size of ~15 μm that is expected at the detector. This corresponds to a 1-μm spot size within the tissue, due to the 15x magnification provided by the 3x objective plus the 5x collection optics (section 3.1.3). The custom objective was designed and fabricated in collaboration with Photon Gear Inc. (Ontario, NY).

3.1.6 *Detector array*

A miniature linear-detector array module, shown in **Figure 3.2 (c)**, is currently being developed to fully package our microscope as a handheld device for clinical use. For preliminary assessment of the optical design and to evaluate the performance of the microscope, a sCMOS camera (Hamamatsu Orca Flash 4.0) with a two-dimensional (2D) detector was used to mimic the one-dimensional (1D) linear detector that will be incorporated in the final design. A thin rectangular region of interest within the sCMOS array serves as a digital slit, as described previously for a large-scale tabletop LS-DAC microscope system [76], [83], [84]. In short, the sCMOS camera allows rapid acquisition of a 2048 × 8-pixel region at the center of the camera. Since the pixel spacing is 6.5 μm, we bin the center 3 rows of pixels to create a digital slit of thickness 6.5 × 3 = 19.5 μm. This approximately matches the diffraction-limited FWHM spot size of ~15 μm that is expected at the detector (see section 3.1.5). Exposure times of 125 μs (a line-acquisition rate of 8 kHz) are utilized to collect 500 lines per frame at an imaging rate of 16 frames/sec. In summary, our device images a field of view (FOV) of 300 × 300 μm with a sampling density of 500 × 600 pixels (x by y).

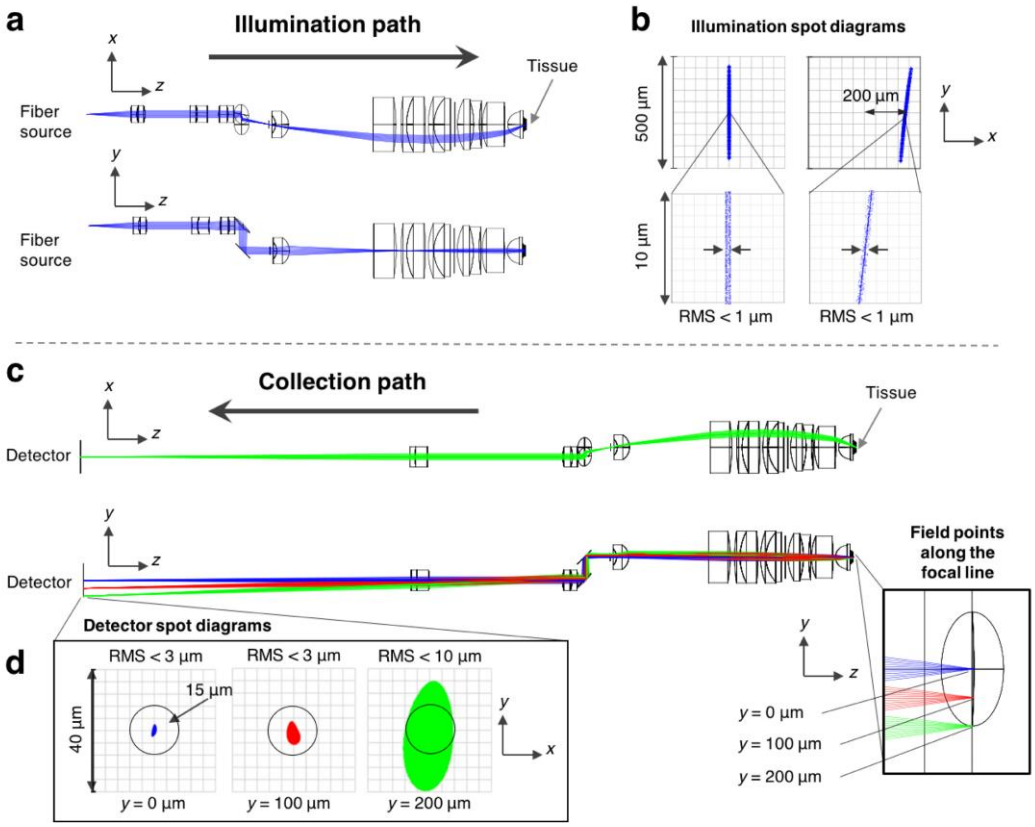


Figure 3.3. Ray-trace simulations. (a) Illumination beam path shown as two orthogonal views (x-z plane and y-z plane). (b) Illumination spot diagrams for the MEMS mirror in its neutral position (left) and for the MEMS mirror tilted by 4.5 deg such that the focal line is offset by 200 µm in the x direction at the sample. The spot diagrams indicate that aberrations are minimal since the spread in the rays is < 1 µm RMS at both the center of the field of view (neutral position, left) as well as at the edge of the field of view (x ~200 µm, right). (c) Collection beam path shown as two orthogonal views (x-z plane and y-z plane). (d) Collection spot diagrams for photons originating from three positions along the focal line at the sample (blue: y = 0 µm, red: y = 100 µm, and green: y = 200 µm).

3.2 RESULTS

3.2.1 *Machined and fabricated components*

As shown in **Figure 3.4 (a)**, an illumination fiber module forms a subassembly with the main body (section 3.1.2). The main body with optical components is shown in **Figure 3.4 (b).** The illumination fiber module (**Figure 3.4 (a)**) is inserted into a cylindrical illumination channel and is secured with set screws (**Figure 3.4 (b)**). The collection lenses, L3 and L4, as well as the alignment mirrors, M1 and M2, are secured within their respective channels using UV-curing glue. **Figure 3.4 (c)** shows the scan head of the main body (CAD model shown in **Figure 3.2 (d)**). The clamps support the MEMS chip at a 45-deg angle and at the correct axial position to ensure proper alignment of the device. The main body and the objective lens (fabricated by Photon Gear Inc., Ontario, NY) are shown in **Figure 3.4 (d)**. In the eventual handheld device, an outer tube will surround and hold these components together (**Figure 3.4 (e) and (f)**).

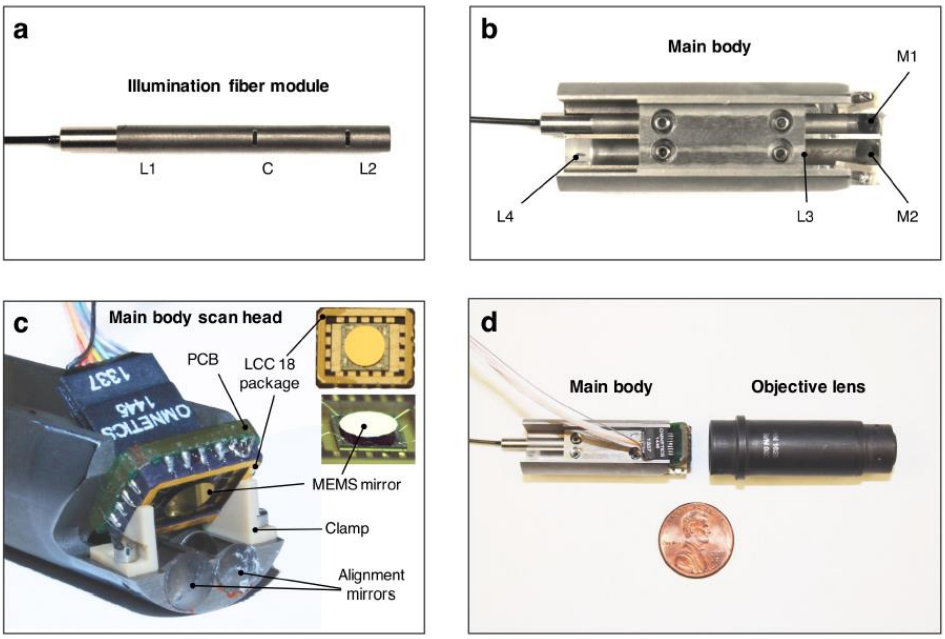


Figure 3.4. Photographs of microscope components. (a) Illumination fiber module housing two spherical doublet achromats, L1 and L2, and a cylindrical doublet achromat, C. (b) Main body with optical components installed, including the two alignment mirrors, M1 and M2, and the

collection lenses (spherical doublets), L3 and L4. (c) Main body scan head and detailed photos of the MEMS chip. The MEMS chip is wire-bonded into an LCC18 package, which is soldered onto a custom PCB. The MEMS PCB module is held in place by two triangular clamps. (d) Main body and objective lens.

3.2.2 *Reflectance-based characterization of device performance*

The axial response of the miniature system to a flat reflective surface was measured. **Figure 3.5(a)** and **(b)** are plots of the axial response on a linear and log scale, respectively. The FWHM optical-sectioning thickness at the center of the field of view (FOV) is measured to be ~2.0 μm. **Figure 3.5 (b)** shows that the background signal, in the absence of a mirror, is below 0.1% of the maximum intensity from a mirror at the focus. The presence of diffraction sidelobes is due, in part, to slight clipping of the Gaussian beams by apertures, such as the 3.0-mm diameter illumination lenses. The edge response of the microscope to a chrome knife edge on glass is plotted in **Figure 3.5 (c)**, showing a 10% to 90% transition width in the x direction (X10-90) of 1.1 μm. The corresponding transition width in the y direction is similar, as expected based on Eq. (6), and is measured to be Y10-90 = 1.0 μm. The image of a reflective 1951 USAF resolution test chart, shown in **Figure 3.5 (d)**, shows the ability of the microscope to resolve features at the micron scale. Note that there is vignetting at the edges of the FOV due to slight field curvature introduced by the scanning MEMS mirror. However, this field curvature (< 10-μm deviation in focal depth over the entire FOV) is not a significant issue when imaging thick three-dimensional tissues. As expected, the resolution is slightly degraded at the edges of the FOV (+/− 150 μm from the center). The worst FWHM resolution values measured at the edges of the FOV were approximately 3.1 μm in the axial dimension (section thickness) and approximately 1.3 μm in the lateral dimensions.

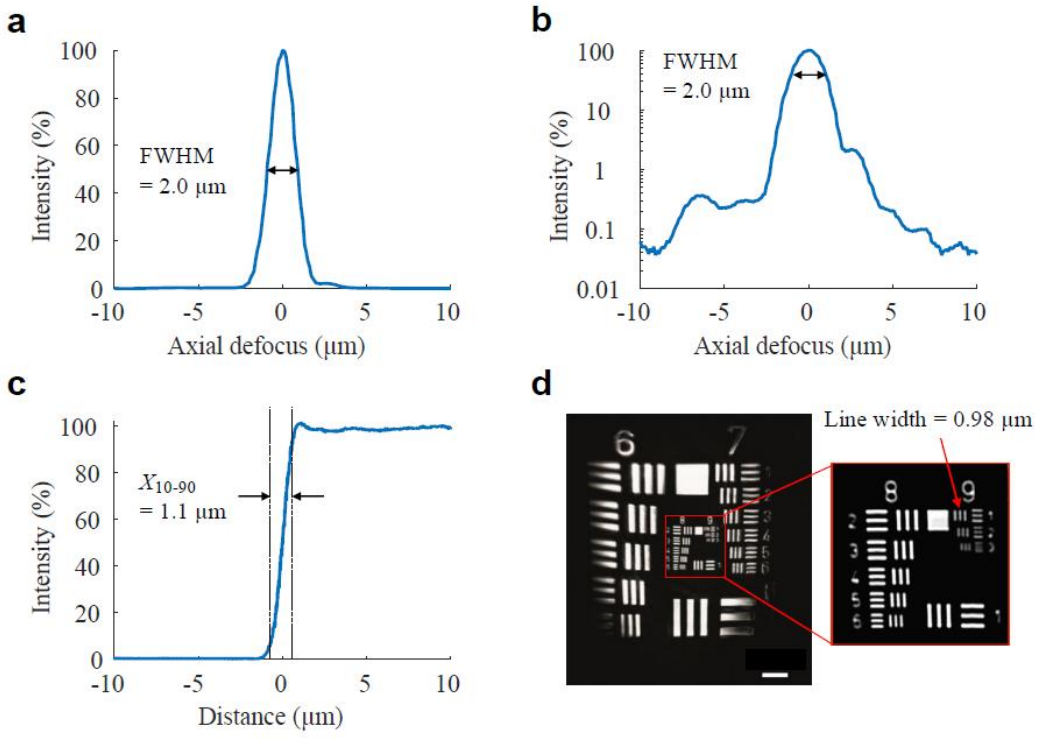


Figure 3.5. (a) Axial response to a flat mirror, plotted on a linear scale, showing a FWHM optical sectioning thickness of 2.0 µm. (b) Axial response to a flat mirror, plotted on a log scale, showing > 30 dB of dynamic range (1000-fold attenuation in signal) as the mirror is translated away from the focal plane. (c) Edge response to a chrome knife edge on a glass substrate. (d) Image of a reflective USAF bar target. The scale bar represents 20 µm.

3.2.3 *Tissue images*

To demonstrate the ability of our device to acquire fluorescence images, we imaged fresh mouse tissues and blood flow within the capillaries of a mouse ear. Fresh mouse tissues were stained in 1% methylene blue for 10 - 30 min, and then rinsed in PBS to remove excess methylene blue from the tissue surface. Methylene blue is known to preferentially label cell nuclei, as is apparent from the kidney images shown in **Figure 3.6 (c)** and **(e)**. Corresponding H&E pathology images, from the same approximate tissue regions imaged by the miniature LS-DAC microscope device, are also shown in **Figure 3.6**.

To demonstrate the ability to visualize dynamic processes at 16 frames/sec, we imaged red blood cells trafficking within the capillaries of a mouse ear. The blood plasma was labeled via retro-orbital injection of a high-molecular-weight fluorescent dextran (Cy5.5-dextran, M.W. = 500 kDa, 2.5 mg per injection in a 150-μL volume, Nanocs Inc., DX500-S5-1). The intravenous injection of the fluorescent dextran causes the red blood cells to appear as dark shadows against the bright fluorescence from the plasma within the capillaries. **Figure 3.7** shows a maximum-intensity projection (in the axial or depth direction) of the vasculature over a range of depths from 120 to 170 μm. The colormap in **Figure 3.7** encodes for depth, with shallower vessels colored blue and deeper vessels colored red.

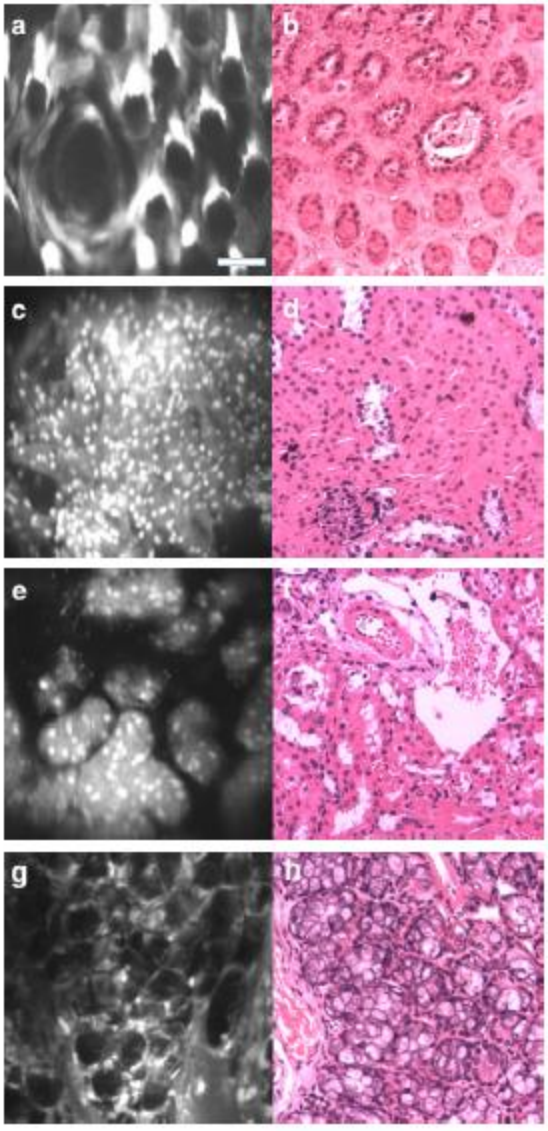


Figure 3.6. (a) Mouse tongue image at a depth of ~50 μm. (b) Histologic section (H&E staining) of corresponding tissue. (c) Mouse kidney image at a depth of ~100 μm. (d) Histologic section (H&E staining) of corresponding tissue. (e) Mouse kidney image at a depth of ~50 μm. (f) Histologic section (H&E staining) of corresponding tissue. (g) Mouse colon image at a depth of ~70 μm. (h) Histologic section (H&E staining) of corresponding tissue. The scale bar represents 50 μm.

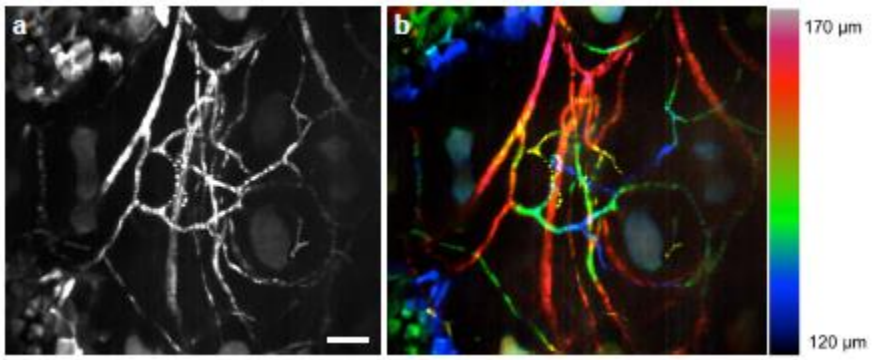


Figure 3.7. (a) Depth projection of vasculature in a mouse ear, imaged between 120- and 170-µm deep. The trafficking of blood cells can be observed within the vessels at an in vivo imaging speed of 16 frames/sec. The scale bar represents 50 µm.

3.3 DISCUSSION

Previous miniature point-scanned dual-axis confocal (PS-DAC) microscopes have had modest axial resolutions (optical sectioning) in the 5 – 10 µm range, limited frame rates of up to 4 frames/sec and have relied on sophisticated MEMS scanners designed to operate at frequencies in the kHz range [42], [112]. The slow frame rates of such devices made them susceptible to motion-induced image artifacts, and the fast 2D MEMS scanners were challenging to develop and optimize for use in a reliable clinical device. Therefore, in this study, we developed a miniature line-scanned dual-axis confocal (LS-DAC) microscope that utilizes a robust commercial MEMS mirror to achieve a high frame rate of >16 frames/sec (and potentially much higher). While line-scanned confocal microscopes sacrifice confocality in one dimension and are therefore less efficient at rejecting out-of-focus and multiply scattered background light compared to their point-scanned counterparts, this study demonstrates that a miniature LS-DAC microscope is capable of high-speed imaging within fluorescently labeled fresh tissues (*ex vivo* and *in vivo*) with excellent sensitivity and resolution.

A major goal of this study was to demonstrate the ability to miniaturize a LS-DAC microscope without sacrificing performance compared to a table-top LS-DAC prototype that was recently developed for high-speed optical sectioning [66], [83]. The miniature system described here exhibits a lateral and axial resolution of 1 – 2 μm, which is slightly superior to that of our previous tabletop systems. However, as shown in the axial response plot in **Figure 3.5 (b)**, the miniature system exhibits diffraction noise that slightly deteriorates the contrast (SBR) of the microscope in comparison to the tabletop LS-DAC devices. We believe that these diffraction sidelobes are a result of the beam clipping that is difficult to avoid due to the necessity for miniature optical components in our design – a necessary trade-off. However, at the shallow imaging depths (~100 – 200 μm) at which the LS-DAC technology is designed to image best [76], [84], this slight loss of contrast is not a major factor and our results demonstrate the ability to image fluorescently labeled tissues with high contrast and resolution. Furthermore, we demonstrate the ability image fluorescent vasculature with sufficient sensitivity and contrast to visualize the trafficking of blood cells at 16 frames/sec.

In the next chapter, we developed a fully package the main body optics, 3x objective lens, and linear detector array into a handheld package for clinical use, including for oral-cancer detection and for guiding brain-tumor resection procedures. We also developed a sterilizable lens cap to provide an effective interface between the distal tip of the microscope and the patients' tissues, as well as to enable passive or active axial translation of the focal plane by clinical end users for imaging at various depths.

In Chapter 5, intraoperative and point-of-care diagnostic use of our device was facilitated by implementing real-time mosaicking algorithms, as evidenced by recent studies that explored the use of such algorithms for microscopy and scanning fiber endoscopes [113]–[116]. Recent

advances in surgical robotics and telesurgery also motivate the integration of the LSDAC system with robotic devices [116], [117]. Additional studies have demonstrated the feasibility of co-registering images from wide-field-of-view (FOV) devices (e.g. MRI and CT) with limited-FOV devices (e.g. miniature microscopes) to facilitate device localization in traditional and robot-assisted surgeries [118]. This work represents a first step towards realizing real-time *in vivo* pathology for a variety of diagnostic and therapeutic applications and will pave the way for several first-in-human feasibility studies in the near future.

Chapter 4. LABEL-FREE *IN VIVO* PATHOLOGY OF HUMAN EPITHELIA WITH A HIGH-SPEED HANDHELD DUAL-AXIS CONFOCAL MICROSCOPE

In this chapter, a fully packaged handheld LS-DAC microscope based on the design described in chapter 3 was built for label-free *in vivo* pathology of human epithelia. As mentioned in the previous chapters, handheld devices for the early detection of skin and oral malignancies, and/or surgical guidance of a variety of anatomical sites, should ideally acquire images at a high frame rate in order to minimize motion artifacts during clinical use on patients. While most previous DAC microscopes have utilized point scanning, in which an image is constructed by scanning a localized focal volume in two dimensions for 2D imaging, the DAC microscope described in this study utilizes line scanning (LS) in order to achieve a high frame rate. In a line-scanned confocal microscope, the illumination beam is focused to a line within the specimen, and a detection slit is used in front of a linear detector array, instead of a pinhole, to reject out-of-focus light. While a line-scanned system sacrifices one dimension of confocality (along the focal line), simulations and experiments have demonstrated that a LS-DAC microscope is capable of achieving adequate contrast (SBR) when imaging near tissue surfaces (approximately 100-μm deep) in comparison to a point-scanned DAC microscope [74], [76].

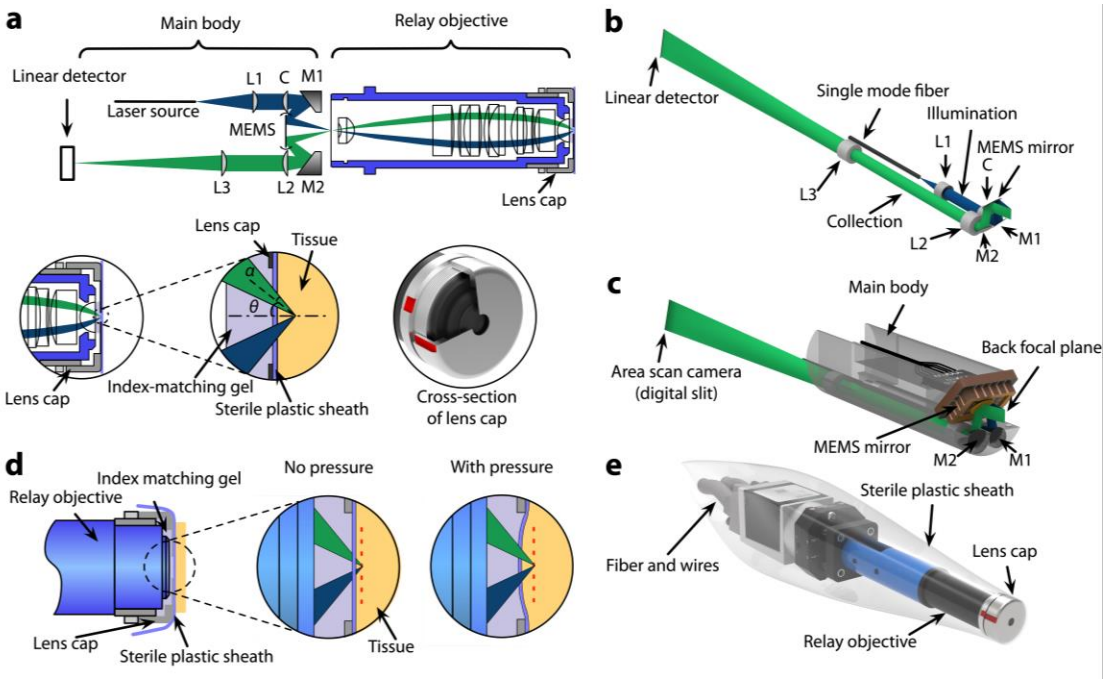
In comparison to the earlier proof-of-concept prototype that was not fully packaged for clinical use [29], a number of technical improvements are reported here: (1) A optimized illumination module has been fabricated to improve the imaging resolution and contrast. (2) A portable detector has been incorporated into a fully packaged handheld device to enable portable clinical use. (3) A sterile lens cap has been designed to enable pressure-sensitive adjustment of the

optical-sectioning depth by the user during imaging. Collectively, these technical advances have allowed us to obtain first-in-human reflectance images of skin and oral mucosa.

4.1 MECHANICAL AND OPTICAL DESIGN OF THE FULLY PACKAGED HANDHELD MICROSCOPE

The handheld LS-DAC microscope developed in this study consists of three major modules (**Figure 4.1** (a)): (1) a main body that houses the optics for the illumination (blue) and collection (green) beams, a MEMS scanning mirror, and two alignment mirrors; (2) a custom relay objective lens with a lens cap that provides 3x magnification; and (3) a portable linear detector array. A single-mode optical fiber (SM670) is used to couple laser radiation at a wavelength of 660 nm into the illumination path of the main body (Gaussian beams are assumed throughout this work). A newly optimized illumination fiber module, assembled by GRINTECH GmbH (Jena, Germany), consists of two doublet achromat lenses packaged within a stainless steel cylindrical tube with an inner diameter of 3.0 mm and an outer diameter of 3.2 mm. Lens L1 is a spherical achromat from Edmund Optics (Barrington, NJ, catalog number 45262) that has been reduced in diameter (3.0-mm diameter) by BMV Optical (Ottawa, Ontario, Canada). Lens C is a custom cylindrical lens, fabricated by BMV Optical, that is based on the lens prescription of a spherical achromat from Edmund Optics (catalog number 45090). Compared to the fiber module utilized in a previous prototype [29], which utilized three lenses, the new two-lens fiber module is simpler to assemble and exhibits reduced diffraction sidelobes in the profile of the illumination focal line due to reduced clipping of the Gaussian beam (see Results). The MEMS scanning mirror utilized in our device is from Mirrorcle Technologies in Richmond, CA (shown in **Figure 4.1** (d)). The MEMS chip is packaged into an LCC18 package (measuring 8.89 by 7.24 mm) and soldered on to a custom-designed PCB chip (measuring 10.16 by 8.64 mm) manufactured by Advanced Circuits

Inc. The detector in **Figure 4.1 (c)** is a 2D detector array (Basler ace acA2000-340km) with 2048 by 1088 pixels, in which the size of each pixel measures 5.5 by 5.5 µm. To utilize this detector as a linear array, a region of interest of 4 rows by 2048 pixels was binned to generate a 1 x 2048 output. A novel lens cap was designed to provide a means for adjusting the imaging depth (shown in **Figure 4.1 (c)**). The distal face of the lens cap provides a flat surface that comes into contact with the tissue during imaging. A 3.5-mm diameter hole at the center of the distal face of the lens cap provides optical access and is covered with a sterile plastic film that wraps around the entire device to maintain sterility. As the user adjusts the pressure of the device against the tissue, the tissue deflects slightly into the hole of the lens cap, which in turn allows the microscope to image



more deeply.

Figure 4.1. (a) Optical circuit of the handheld LS-DAC microscope. The blue and green colors indicate the illumination and collection beams, respectively. The main body houses the MEMS scanning mirror. The mirrors (M1 and M2) are used to align the dual-axis beams such that they intersect at the back focal plane of the custom relay objective. The objective relays the beams from the back focal plane (at the left side of the objective) to the front focal plane in tissue (right side) with 3x magnification. The focusing angle of the beams in tissue, α, and crossing angle, θ, enable

high-contrast optical sectioning with micron-scale resolution (see text). The lower right inset shows a cross-sectional view of the lens cap. (b) Illumination and collection optics within the main body. (c) A design rendering of the precision-machined main body, within which the alignment mirrors and optical components are mounted. (d) A cross-sectional view of the distal end of the objective with a lens cap installed. When the pressure of the lens cap against the tissue is adjusted, the tissue curves slightly into the hole at the tip of the lens cap, which changes the depth at which the focal plane is located beneath the tissue surface. A sterile but flexible plastic sheath forms a seal around the lens cap. (e) A 3D rendering of the fully packaged device.

4.2 ELECTRICAL AND CONTROL SYSTEM OF THE HANDHELD MICROSCOPE

The imaging system is controlled by a LabVIEW program that runs on a standard PC (shown in **Figure 4.2 (a)**). The MEMS controller amplifies the voltage signals from the USB port on the PC to scan the MEMS mirror (triangular waveform). A FPGA-based frame grabber (NI PCIe-1473R) collects video data from the Basler detector and stitches the lines to form 2D images. When the LabVIEW program is started, the initial data acquisition is triggered by the next available HSYNC signal from the detector, and the scanning of the MEMS mirror is triggered through software. Since these two tasks are triggered by two unsynchronized trigger sources, a MATLAB script is embedded in the LabVIEW program to provide software-based synchronization to prevent image drifting.

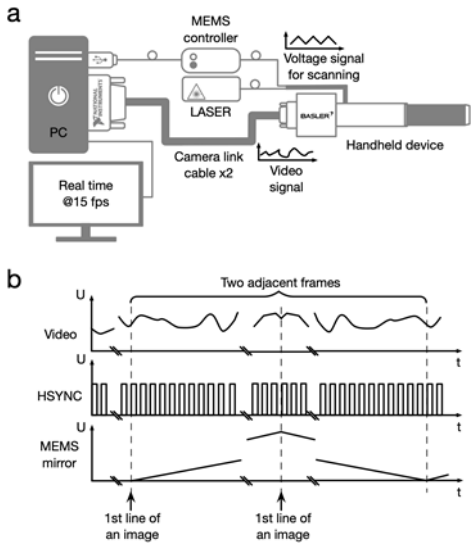


Figure 4.2. (a) Diagram of electronic components and connections. A LabVIEW program controls the scanning of the MEMS mirror and acquires video signals from the detector for real-time display and data storage. A MEMS controller connected to the PC via a USB cable outputs amplified voltage signals to scan the MEMS mirror. An FPGA-based frame grabber board in the PC collects video signals from the detector. (b) Timing diagram of the system.

4.3 RESULTS

The image of a reflective 1951 USAF resolution test chart, shown in **Figure 4.3 (a)**, shows the ability of the microscope to resolve features at the micron scale. **Figures 4.3 (b)** and **(c)** are plots of the axial response to a flat mirror on a linear and log scale, respectively. The FWHM optical-sectioning thickness at the center of the field of view (FOV) is measured to be ~1.72 μm.

Figure 4.3 (c) shows that the background signal in the axial-response plot, as the mirror is translated away from the focal plane, is ~0.01% of the maximum signal from a mirror located at the focal plane. This is a significant improvement from our initial miniature device [29], which leads to improved contrast and imaging depth in tissues. Note that there is vignetting at the edges of the FOV due to slight field curvature introduced by the scanning MEMS mirror.

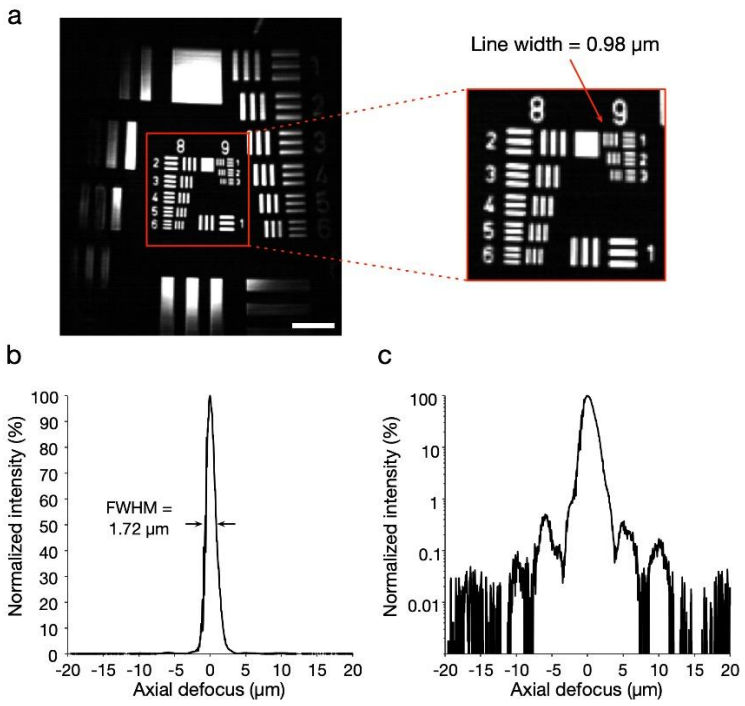


Figure 4.3. (a) Diagram of electronic components and connections. A LabVIEW program controls the scanning of the MEMS mirror and acquires video signals from the detector for real-time display and data storage. A MEMS controller connected to the PC via a USB cable outputs amplified voltage signals to scan the MEMS mirror. A FPGA-based frame grabber board in the PC collects video signals from the detector. (b) Timing diagram of the system.

To demonstrate the ability of our device to acquire label-free reflectance images *in vivo*, we imaged human facial skin and oral mucosa from healthy volunteers under IRB approval and patient consent. All images are acquired in real time at a frame rate of 15 frames/sec. For direct comparison to imaging with point-scanning, we also imaged with a handheld point-scanned single-axis confocal microscope, the VivaScope 3000 (Caliber I.D. Inc., Andover, MA), which has a frame rate of 7 frames/sec. **Figure 4.4 (a)** and **(e)** show photographs of the two devices as a size comparison. The images in **Figure 4.4 (b)-4.4 (c)** show distinct morphological features in the facial skin such as the stratum spinosum (red arrow) around a hair follicle (green arrow) in **Figure 4.4 (b)** and the epidermis (red arrow) and dermal papillae (green arrow) at the dermal-epidermal

junction in **Figure 4.4 (c)**. **Figure 4.4 (d)** shows distinct hyper-reflective nuclei in the squamous cells of the buccal mucosa. **Figure 4.4 (f) - 4.4 (h)** show reflectance images of similar features obtained at the same skin and oral mucosa sites with the point-scanned SAC microscope. Visual comparison shows that optical sectioning and resolution is preserved for the LS-DAC approach and is comparable to that of PS-SAC approach down to the basal cell layer (~50-150 μm depth), which confirms our earlier modeling and experimental measurements [74], [76], [119].

4.4 DISCUSSION

Compared to the images collected by the PS-SAC (VivaScope 3000) device, speckle noise is more apparent in the images collected by the handheld LS-DAC microscope. This is, in part, due to a narrower confocal slit, which preserves resolution and thin sectioning, but at the trade-off of higher speckle contrast [120]. Future devices can mitigate speckle noise, if desired, by increasing the physical slit width and/or altering the magnification of the collection optics, with the attendant tradeoffs described in the literature [120]. Furthermore, instead of using the coherent laser as the illumination of the device, an incoherent light source, such as LED, can potentially be the light source of the device to fundamentally avoid the appearance of the speckle noise [121].

As described in previous publications, the signal-to-background ratio (SBR) of a point-scanned confocal architecture is superior to a line-scanned architecture due to loss of confocality along the focal line of a line-scanned device [74], [75], [86]. However, the dual-axis configuration acts to mitigate this deterioration in SBR, somewhat, because of the fact that the illumination and collection beams are spatially separated, except where they intersect at their respective foci [74], [119]. This is seen in the image comparisons between the LS-DAC device and the point-scanned single-axis confocal (PS-SAC) VivaScope device, in which the PS-SAC device exhibits slightly improved contrast (SBR) at deeper depths, such as at the dermal-epidermal junction in **Figure 4.4**

(b) and **(g)**. Since the LS-SAC device is smaller than the VivaScope device, diffraction noise (sidelobes) from lens apertures may be more severe and could be an additional source of minor degradations in contrast (SBR). This is a consequence of the fact that large beam diameters are desired to achieve relatively large NAs (high resolution) with a long working distance but are severely constrained in a miniature system in which all optical components (including lens apertures) are necessarily small. However, at shallow depths, image quality is comparable between the LS-DAC and PS-SAC approaches, as expected and as previously shown with tabletop systems [76].

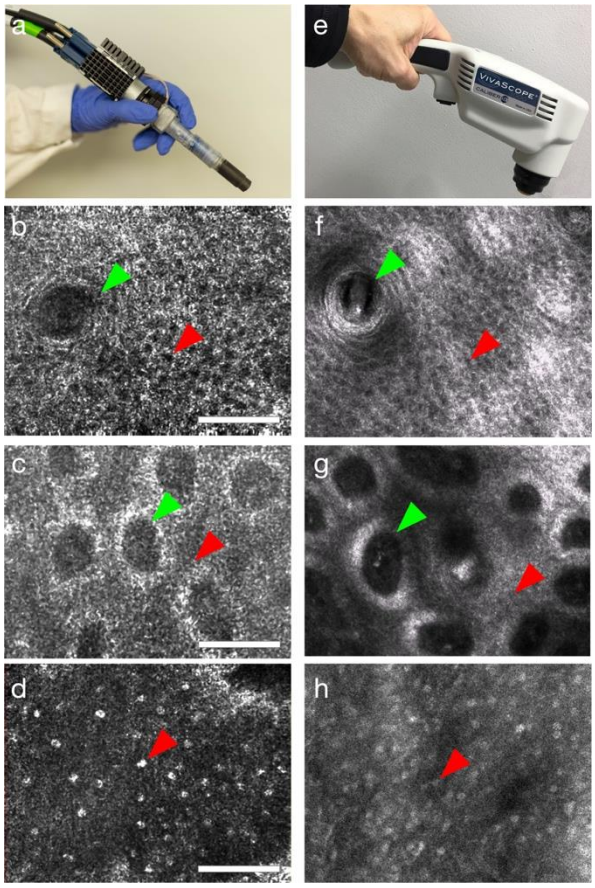


Figure 4.4. (a) Photograph of the handheld LS-DAC microscope. (b) Human facial skin image (stratum spinosum around a hair follicle). Scale bar: 100µm. (c) Human facial skin image (stratum basalis and dermal-epidermal junction). Scale bar: 100µm. (d) Image of human oral mucosa (oral epithelial nuclei). Scale bar: 100µm. (e) Photograph of a handheld PS-SAC (VivaScope 3000)

microscope. (f-h) Label-free images of corresponding tissues collected by the PS-SAC microscope.

In summary, a fully packaged handheld LS-DAC microscope was developed with pressure-sensitive depth control (via a novel lens cap design), which is the first device of its kind to be used for *in vivo* imaging of human skin and oral mucosa. Compared to previous miniature point-scanned dual-axis confocal (PS-DAC) microscopes (4 frames/sec), this handheld LS-DAC microscope has a much higher frame rate (15 frames/sec), with reduced motion artifacts. The axial and lateral resolution (FWHM of 1.7 and 1.1 μm, respectively) were also improved [106], [112]. As mentioned previously, maximizing the frame rate of a handheld device is critical for minimizing motion artifacts during handheld use, and also for enabling effective video mosaicking [122]. Note that for the reflectance-based LS-DAC device developed in this study, the imaging speed is limited not by SNR (low photon counts) but rather by the limited readout rate of the detector used in the device. In the future, higher imaging speeds of >30 frames/sec should be possible with higher-speed linear array detectors. As mentioned in Chapter 2, the lower lip and the tongue are the two most common sites for oral cancer. Even though our device does not have a flexible probe to reach all locations within the oral cavity, the small diameter of the 3x relay lens allow the users to image most areas of the lower lip, the tongue and the buccal mucosa (inner cheek).

The LS-DAC design architecture has allowed us to develop a system with a miniature form factor that enables clinical use to image skin and oral mucosa, and potentially other exposed tissues, such as during surgical resection procedures. Despite a slight reduction in contrast, the improved imaging speed in the thesis shows that the LS-DAC approach is an attractive design choice.

Chapter 5. IMPLEMENTATION OF REAL-TIME VIDEO MOSAICKING FOR IMAGE GUIDANCE

As mention in Chapter 1, in the case of oral cancer detection, dentists and physicians often can visually identify suspicious lesions but are hesitant to biopsy them since the majority of such lesions are benign. Our noninvasive handheld microscope enables “image-guided biopsy” by providing label-free in vivo images that approximate the gold standard of invasive histopathology [123], [124]. However, due to the small FOV of most handheld and endoscopic microscopes, it is difficult for users to know if they have adequately sampled/imaged the lesion when imaging a large extent of tissue is desired. As shown in **Figure 5.1 (a)**, real-time video mosaicking would guide users as they seek to interrogate the lesion and would allow them to determine if they have successfully sampled the suspicious region. Not only for oral cancer detection, real-time video mosaicking is also a useful tool which enables surgical guidance, where the goal is to minimize positive margins. **Figure 5.1 (b)** shows that after removing the majority of a tumor, the boundary of the wound (i.e. the surgical margin) may still contain residual tumor which necessitates additional resection procedures [125], [126]. Rather than randomly sampling small FOVs with a handheld microscope, in which certain tissue regions may be redundantly imaged while other regions are missed, real-time mosaicking can help surgeons to efficiently survey as much of the surgical margin as possible in order to locate and excise residual malignancies.

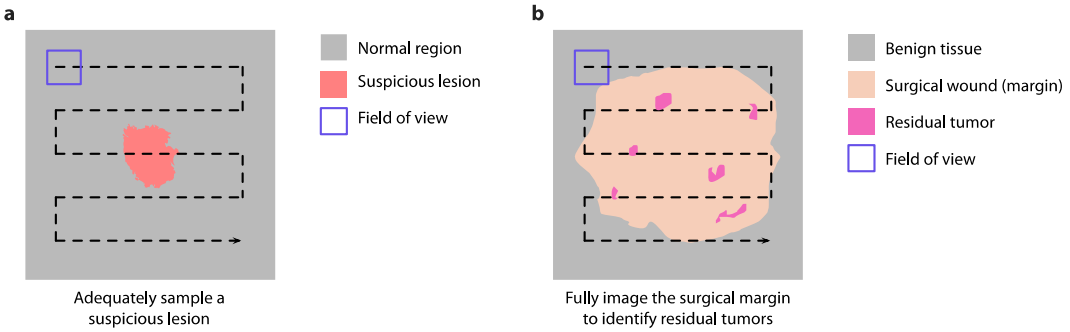


Figure 5.1. Two scenarios that could benefit from real-time video mosaicking over a large field of view. (a) For biopsy guidance, adequately sampling a suspicious lesion can be difficult with a handheld device that has a limited FOV (sub-millimeter). (b) For surgical guidance, imaging larger areas at the surgical margins improves a surgeon's ability to identify residual disease for continued resection.

The combination of an automated scanning stage with post-processed image mosaicking is the most common solution for extending the FOV of a microscopic system [127]–[136]. However, the incorporation of a bulky robotic scanning mechanism is often not ideal for handheld or endoscopic microscopy applications. Post-processed mosaicking of overlapping image frames collected with a handheld device (with no robotic positioning) is possible for extending the FOV but lacks of real-time position tracking [115], [122], [128], [137], [138]. Therefore, it is challenging for a user to efficiently sample a large area of tissue, and large regions of tissue will invariably be missed, or redundantly sampled. Some real-time image mosaicking methods have been developed for stitching adjacent overlapping images, but can only handle image translations without rotations or deformations [115], [137], [138]. Since image rotations and deformations are common during handheld in vivo imaging of irregular tissue surfaces, we developed a real-time mosaicking method that could handle image translations, rotations and deformations.

We modified a previous slow post-processing mosaicking method [122] so that real-time mosaicking is possible at >30 frames/sec with images that are 400 x 400 pixels in size, as for

example collected with a custom-developed handheld dual-axis confocal microscope [29], [123], [139]. The primary goal of our mosaicking method is to guide clinical users as they seek to image a large tissue area both fully and efficiently, which is useful for accurately interrogating heterogenous tissues. It also allows the users to locate and image small lesions when using a relatively large handheld device that blocks a user's visual access to these small lesions

5.1 METHODS

5.1.1 *Real-time image registration*

Our real-time video mosaicking method reads raw video frames from the hard drive saved by the LabVIEW vi mentioned in Chapter 4, performs real-time feature-based image registration, and then achieves image stitching (**Figure 5.2**). These image frames are generated at 20 - 30 frames/sec in this study (400 pixels by 400 pixels per frame). The C/C++ application that we developed, which is based on the Qt platform, performs sequential image registration on adjacent image frames. For every two adjacent frames, the image registration algorithm generates a homography matrix to describe the two-dimensional spatial relationship between the two adjacent frames.

As shown in **Figure 5.2 (b)**, the image-registration procedure consists of four sub-routines: 1) feature extraction, 2) feature matching, 3) filtered random sample consensus (RANSAC) to generate a homography matrix, and 4) image overlay. For feature extraction, the "speeded-up robust features (SURF)" algorithm is utilized [140]. Compared to the method used in a previous post-processing mosaicking method for feature detection and description, called "scale-invariant feature transform (SIFT)," the SURF algorithm is significantly faster [140], [141]. Despite this increase in speed, the registration speed of SURF, running on a CPU (Intel, i7-8700), still cannot match our desired real-time frame rate (up to 30 frames/sec). Therefore, a parallel-computing version of SURF (called from the OpenCV library) running on a standard GPU (Nvidia, GTX

1080) was implemented so that > 30 frames can be registered in one second. Further improvements to image registration speed can be achieved with the use of more-powerful CPUs and GPUs in the future.

Once key features are identified/generated for each image frame, the key features for every pair of adjacent frames are stored in the GPU memory, such that a feature-matching procedure can be used to identify common features shared between adjacent frames. For this, we use the "fast library for approximate nearest neighbors (FLANN)" method [142], which is also a parallel computing algorithm called from OpenCV, which runs on the GPU. FLANN rapidly finds the most similar features in adjacent frames, and links them together as matched pairs.

Based on the matched pairs generated by FLANN, the "random sample consensus (RANSAC)" algorithm generates a homography matrix to describe the spatial relationship between two adjacent frames (which accounts for rotations, translations, and deformations). To improve the accuracy of image registration, a method was developed to reduce matching errors and converge upon an optimal homography matrix. Here we assumed that the frame rate of our handheld imaging device would be sufficiently high such that rotational transformations would be relatively small between adjacent frames (e.g. < 5 deg). Under this low-angle assumption, the majority of matching feature pairs should have approximately the same distance. As shown in **Figure 5.2 (b)**, the matched pairs of features corresponding to the most common spatial distance are extracted in our "filtered RANSAC" procedure so that the proportion of incorrect matches is reduced prior to computation of a final homography matrix. This filter helps to increase the accuracy of image registration, especially for low-contrast or noisy datasets.

Image-registration failures can occur when there are severe motion artifacts within single frames or when there is insufficient overlap between two adjacent images (~80% or less). A set of

thresholds are used to check if there are sufficient matching features between frames (~30 correct matches or more), or if the resulting translation and rotation coordinates are within a reasonable range (< 50 microns, < 5 deg). These thresholds are empirical and can be adjusted for different imaging devices and imaging conditions. If image registration fails, the mosaicking algorithm pauses until a new frame can be correctly registered to the last stitched frame. Our image mosaicking algorithm is not designed to stitch images obtained at different tissue depths. However, at high frame rates, there is often sufficient similarity between adjacent frames such that mosaicking can be performed while the imaging depth is slowly changed. Note that this implies that our mosaicking algorithm is not globally optimized and that positioning errors will inevitably increase as a function of distance traveled (errors are ~1% of the total distance traveled, based on experimental observations). Once the RANSAC procedure is done, images are overlaid based on the homography matrices with no averaging or blending between frames (**Figure 5.3**). Since the goal of our method is to provide a real-time guide to the user for efficient sampling of large imaging fields, the image quality of the real-time mosaic is not a priority. If a high-quality mosaicked image is desired, post-processing routines can be used in which images are seamlessly blended together (**Figure 5.4**).

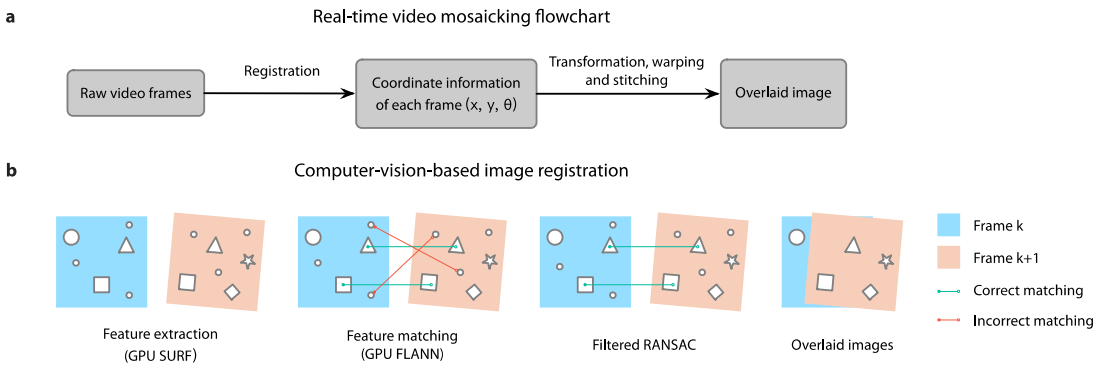


Figure 5.2. Real-time video mosaicking flowchart and image-registration sub-routines. (a) Main steps of the real-time video mosaicking workflow. (b) Computer-vision-based image registration method, which accommodates image translations, rotations and deformations.

As the microscope is translated slowly across a tissue surface, a LabVIEW-based user interface is for users to visualize raw frames (20 – 30 frames/sec in this study) and a pop-up window shows the growing mosaicked image in real time.

5.1.2 *Potential solution for generating high-quality mosaicked images*

The primary purpose of this real-time mosaicking method is not to generate high-quality images for diagnostic purposes, but rather to guide users to efficiently survey a large area of the tissue or to locate a desired sub-region of the tissue. Therefore, we did not include any image blending functions in our method. However, since the users need to review the mosaicked image, here we provide a potential solution for image blending by post processing.

As mentioned previously, the overlapping region between two adjacent frames must be large enough so that the SURF algorithm is able to recognize sufficient number of key features for matching. However, when mosaicking the images, large overlapping regions redundantly cover the same area and lead more seams. In other words, stitching the images with small overlapping regions can help to reduce the number of seams and save computational resources. Therefore, our

solution is to choose a subset of the raw frames based on the coordinates saved by the real-time mosaicking application. By this way, each pair of adjacent frames for stitching will have ~10% overlap. For the finial stitching, blending the images by averaging or using a graph-cut based method will work. In the following Result section, we used an ImageJ plugin to linearly average the images for generating the mosaicked image.

5.2 *RESULTS*

A pilot study was performed to demonstrate that the use of real-time mosaicking enables significantly more efficient spatial coverage when performing handheld imaging with a miniature dual-axis confocal microscope (20 frames/sec, FOV = 350 x 350 μm). For this simple study, users were tasked with the challenge of comprehensively imaging a 3-by-2-mm area of fluorescently labeled (acridine orange) mouse kidney tissue within 2 min. This task was performed with and without the assistance of video mosaicking and was repeated ten times for each condition. To maximize scanning efficiency and to minimize redundancy, users were advised to follow a raster-scanning pattern when moving the handheld device across the tissue. However, without any real-time feedback, scanning paths were difficult to control, where users could not determine how far they had traveled in any direction, therefore leading them to image too far or not far enough. With real-time video mosaicking, users could see how far they had traveled with the device (i.e. a distance scale bar was provided in the LabVIEW-based graphical user interface). Users could therefore more-easily scan the tissue with minimal redundancy and high efficiency. The ultimate performance metric was “coverage percentage,” which refers to the percentage of the 3 x 2 mm tissue region that was sampled with the device within the allotted 2 min time frame.

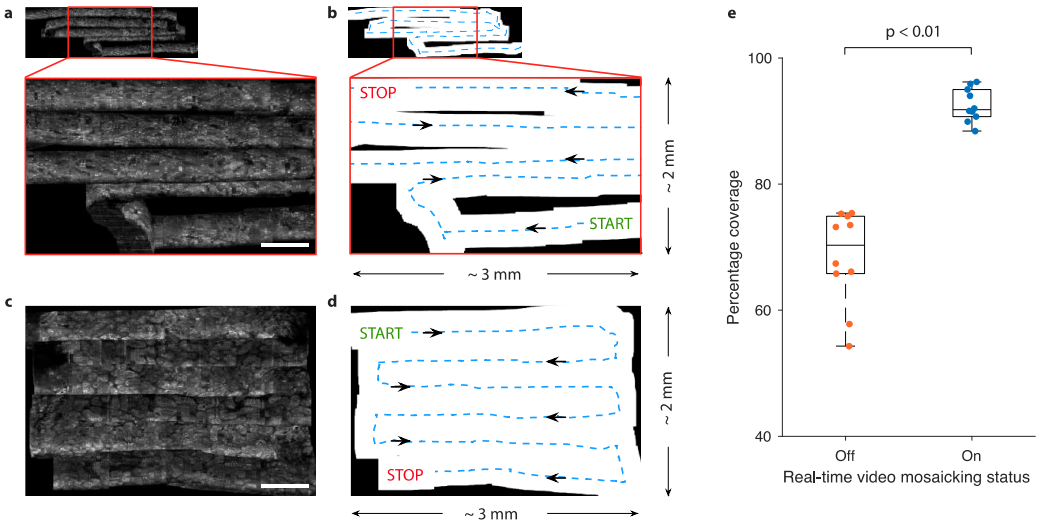


Figure 5.3. Users were tasked with imaging a 3 x 2 mm tissue region within a 2 min time frame. (a) A stitched image collected without real-time image mosaicking. The red box indicates the targeted imaging region (3 x 2 mm). Scale bar: 500 µm (b) The white areas indicate the imaged regions, and the dashed lines depict the trajectory of the device over time. (c) A stitched image collected with real-time image mosaicking. Scale bar: 500 µm (d) The white areas indicate the imaged regions, and the dashed lines are the trajectory of the device. (e) A boxplot of the coverage percentage for a 3 x 2 mm tissue region imaged both with and without real-time mosaicking guidance (~2-min imaging duration).

Figure 5.3 (a) shows one example of a stitched image collected without the guidance of real-time image mosaicking. Since users were unable to know how far they scanned, they often scanned further than needed. **Figure 5.3 (b)** shows that significant areas within the targeted 3 x 2 mm region were missed (black areas). The dashed lines in **Figure 5.3 (b)** show the approximate trajectory of the handheld device over time. As a comparison, **Figure 5.3 (c)** shows one example of a stitched image collected with real-time image mosaicking. **Figure 5.3 (d)** shows that most of the targeted 3 x 2 mm region was imaged (the white areas). The dashed trajectory indicates that the user could optimize their motions to follow an efficient raster-scanning pattern with minimal redundancy or overshooting. A box plot of the percentage coverage, both with and without real-time mosaicking,

is shown in **Figure 5.3 (e).** The median coverage rate with real-time video mosaicking is close to 95%, with all imaging attempts yielding a coverage percentage of at least 88%. As a comparison, with no real-time guidance, the median coverage percentage was ~70% with a large variation in coverage rates ranging from 55% to 75%. These results show that users can efficiently and fully image a large area of tissue with the assistance of a real-time video mosaicking method.

Finally, to show that the data collected from our study can be used to generate a high-quality stitched image (with post processing), an ImageJ plugin was used. **Figure 5.4 (a)** shows a post-processed image of a mouse kidney specimen used in our pilot study (same dataset as shown in **Figures 5.3 (c) & 5.3 (d)).**

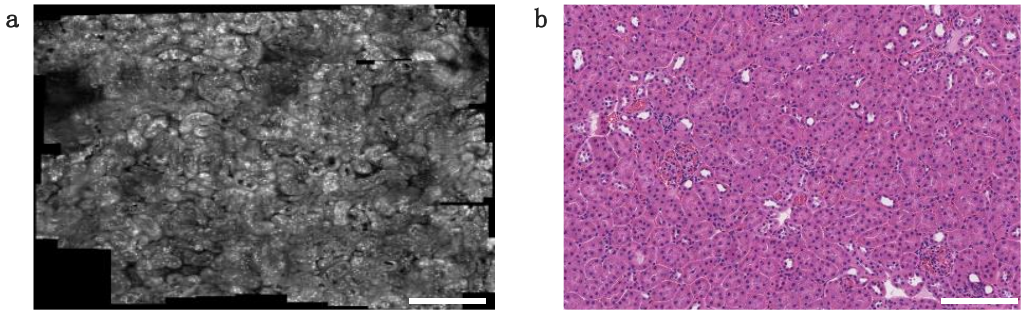


Figure 5.4. (a) Post-processed image mosaic of fluorescently labeled (acridine orange) mouse kidney imaged with a handheld line-scanned dual-axis confocal microscope. (b) A corresponding H&E histology image of mouse kidney. Scale bar: 500 µm

5.3 DISCUSSION

In summary, we implemented a feature-based video mosaicking method to guide users of handheld microscopes as they seek to survey large areas of tissue efficiently and comprehensively. To further improve these methods, higher frame rates and/or larger fields of view would be helpful to ensure high levels of overlap between frames even if the device is translated more quickly across the tissue surface. However, higher frame rates and larger images will demand faster algorithms

and CPUs/GPUs for real-time computation. Currently, feature-based mosaicking methods will always fail if users move the device too rapidly, or if tissue contact is lost. In the future, adding positioning sensors with sub-millimeter accuracy could provide a means for guiding users back to a precise tissue location for resumed mosaicking. Such positioning sensors could also correct for spatial errors that accumulate over time with feature-based mosaicking. Such a strategy could be combined with global mosaicking methods [115], [129], [136], in which image frames are registered not only with adjacent frames in time, but also with adjacent frames in space (acquired in the past) to continuously correct for accumulated spatial errors and tissue deformations. Nonetheless, we have shown through a pilot study that even the most rudimentary real-time mosaicking method can have a large impact on the usability and efficiency of a clinical handheld microscope for in vivo disease detection and surgical guidance.

Chapter 6. CONCLUSION AND FUTURE WORK

Histopathology is the clinical gold-standard for cancer diagnosis. However, it is invasive, destructive of tissues, time-consuming of slide preparation and scanning, and limited in terms of sampling extent [29], [123]. For some point-of-care clinical applications, the ability to obtain noninvasive real-time pathology data could have significant advantages compared to traditional histopathology [26], [119], [143], [144]. As we mentioned in Chapter 1, noninvasive real-time pathology could be of value for disease screening applications, in which invasive biopsy is often not justified in terms of cost and risk [123], [124]. In addition, for surgical guidance (mentioned in Chapter 5), the time and cost of rapid frozen section analysis prevent surgeons from examining more than a trace amount of tissue [127]. Especially for critical organs such as the brain, noninvasive assessment of tissues would be much preferred [125], [126]. Therefore, this thesis described a handheld MEMS-based line-scanned dual-axis confocal microscope with real-time video mosaicking to deliver point-of-care pathology to the patients. The frame rate of the device is up to 30 frames/sec (400 x 400 pixels in size), in which mitigates the motion artifact for handheld usage. The device provides subcellular level microscopic images with 1-μm lateral resolution and 1.72-μm axial resolution.

In Chapter 3, we demonstrated a proof-of-concept prototype to verify the performance of the optical and mechanical design. Optical simulation, mechanical design and alignment of the system were the three main tasks for building the prototype. Results showed that the lateral and axial resolutions of the device are 1.1 μm and 2 μm, respectively. Due to the limitation of the frame rate of the detector that we were using on the proof-of-concept system, the frame rate of the device was limited to 15 frames/sec. We compared fluorescence images of different mouse tissues with their corresponding H&E images. We also took a video of blood flowing in a live mouse ear to show

that our device can non-invasively image deep inside the tissue with high resolution. In addition, it proves that the frame rate was high enough to capture the blood flowing. The Z-projection image showed that the optical sectioning capability provides microscopic information at different depths in the tissue.

In Chapter 4, a fully packaged handheld device was developed. The frame rate of the device was increased to 20 frames/sec (>30 frames/sec is possible) by using a faster detector. As we mentioned in Chapter 3, due to the assembling method of the old fiber module, the axial response of the device had a sidelobe, which worsened the SBR. To fix this issue, we utilized a new design of the fiber module. Eventually, the axial response of the fully packaged device showed the SBR was significantly improved. The axial resolution was also improved to 1.72 μm. Results showed reflectance in-vivo images and videos of human skin and oral mucosa. The video also showed that the lens cap allowed the users to adjust the imaging depth by adjusting the pressure they push the device against the tissue surface. Compared to a commercial handheld single-axial confocal microscope, our device has smaller size and much faster frame rate.

In Chapter 5, a real-time video mosaicking method was implemented to the handheld device for imaging guidance. The primary goal of developing this tool was to enlarge the limited FOV of the device for clinical applications, since imaging large extent of tissue is always desired. By modifying a previous post-processing method, the real-time video mosaicking method could stitch images with transformation, rotation and deformation. The current stitching speed is around 30 frames/sec. Upgrading the hardware of the PC can increase it. Results of the percentage coverage study showed that, with the assistance of the real-time video mosaicking method, the users could adequately sample a larger suspicious lesion or fully image the surgical margin to identify residual tumors.

6.1 FUTURE WORK

There is some additional work to further improve the performance of the device for point-of-care clinical applications. This section presents some of these directions.

For a handheld microscope, a higher frame rate could allow the users to scan the device across the surface of the tissue faster with minimal motion artifacts. For example, in the current device, the framerate of 30 FPS realistically limits the physical movement of the device to ~2 mm/sec. While this framerate is clinically compatible, there are still scenarios where the device may move faster. Overall, a higher frame rate would improve a clinician's efficiency in terms of imaging and mosaicking a given region of tissue, which would reduce the patients' discomfort since it could shorten the imaging time. As mentioned in Chapter 5, higher frames rates can avoid the failure of image registration caused by rapid motion of the device. However, there are several challenges that we must meet. 1) A higher frame rate requires a faster detector. When considering faster detectors, the size limitation of a handheld device may ultimately be an issue since high speed cameras can be large and bulky. 2) A higher framerate means a shorter exposure time. As a result, we may need to increase the illumination power to make sure the camera has enough signal (the signal-to-noise ratio). 3) A higher frame rate means a larger data size. Especially for real-time video mosaicking, faster algorithms and more powerful CPUs and GPUs may be necessary.

Enlarging the FOV of the device will allow a single raw frame to contain more key features for matching. It helps the image registration to be more robust and accurate. When the FOV is enlarged, even though the overlapping percentage between two adjacent frames is small, the algorithms still can find enough key features for image registration so that the users can move the handheld device faster. In addition, clinicians can survey the area of tissue even quicker and reduce the patients' discomfort. The challenge of enlarging the FOV is that we need to redesign the lens

and the optical circuit of the device. It may be hard to keep the size of the device small while increasing the size of the lens for a larger FOV.

Motorized Z-scanning would help users to easily change the depth of imaging. Currently, a lens cap is used to manually adjust the imaging depth over a range of ~0.2 mm. A motorized lens cap and/or motorized optical components could allow users to easily adjust the imaging depth during imaging. In addition, "controlled" Z-scanning could help to maintain a constant imaging depth for enhanced mosaicking performance. For all solutions, the biggest challenge is maintaining a small size for a handheld device.

Embedding a motion sensor to the device could help with real-time video mosaicking. As we mentioned in Chapter 5, the motion tracking of the device is currently based on the image registration between two adjacent frames. Once rapid device motion of >2.1 mm/sec occurs, the motion tracking fails since the two adjacent images may not sufficiently overlap. A motion sensor could potentially complement our computer-vison based mosaicking method and allow users to return to an approximate location at which computer-vision-based mosaicking can resume.

For the clinical side, large clinical studies are necessary to assess the sensitivity and specificity of our device for detecting various malignancies. It would also be beneficial to use the device to study additional organs. In this thesis, we only explore tissues like human skin and human oral mucosa. However, the image quality of the device may vary between tissue types, and further studies are needed to validate the performance beyond the tissues explored in this thesis.

In summary, the handheld MEMS-based dual-axis confocal microscope that we built enables high-speed in-vivo sub-cellular microscopic imaging for point-of-care clinical applications. This thesis presents the work towards realizing real-time in vivo pathology for a variety of diagnostic and therapeutic applications.

BIBLIOGRAPHY

[1] J. Ferlay *et al.*, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer*, vol. 136, no. 5, pp. E359–E386, Mar. 2015, doi: 10.1002/ijc.29210.

[2] D. M. Saman, "A review of the epidemiology of oral and pharyngeal carcinoma: update," *Head Neck Oncol.*, vol. 4, p. 1, Jan. 2012, doi: 10.1186/1758-3284-4-1.

[3] S. Warnakulasuriya, "Global epidemiology of oral and oropharyngeal cancer," *Oral Oncol.*, vol. 45, no. 4, pp. 309–316, 2009.

[4] J. Ferlay, H.-R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin, "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008," *Int. J. Cancer*, vol. 127, no. 12, pp. 2893–2917, Jun. 2010, doi: 10.1002/ijc.25516.

[5] J. J. Sciubba, "Oral Cancer," *Am. J. Clin. Dermatol.*, vol. 2, no. 4, pp. 239–251, Aug. 2001, doi: 10.2165/00128071-200102040-00005.

[6] N. Gupta *et al.*, "Changing Trends in oral cancer - a global scenario," *Nepal J. Epidemiol.*, vol. 6, no. 4, pp. 613–619, Dec. 2016, doi: 10.3126/nje.v6i4.17255.

[7] A. T. Eriksson, M. M. Corcuera, J. C. Trapero, J. C. Sánchez, and A. B. Martínez, "Analysis of new diagnostic methods in suspicious lesions of the oral mucosa," *Med. Oral Patol. Oral Cir. Bucal Ed Inglesa*, vol. 14, no. 5, p. 1, 2009.

[8] O. Kujan, A.-M. Glenny, J. Duxbury, N. Thakker, and P. Sloan, "Evaluation of screening strategies for improving oral cancer mortality: a Cochrane systematic review," *J. Dent. Educ.*, vol. 69, no. 2, pp. 255–265, 2005.

[9] M. W. Lingen, J. R. Kalmar, T. Karrison, and P. M. Speight, "Critical evaluation of diagnostic aids for the detection of oral cancer," *Oral Oncol.*, vol. 44, no. 1, pp. 10–22, 2008.

[10] E. S. Oh and D. M. Laskin, "Efficacy of the ViziLite system in the identification of oral lesions," *J. Oral Maxillofac. Surg.*, vol. 65, no. 3, pp. 424–426, 2007.

[11] N. L. Rhodus, "Oral cancer and precancer: improving outcomes," *Compend. Contin. Educ. Dent. Jamesburg NJ 1995*, vol. 30, no. 8, pp. 486–488, 490–494, 496-498 passim; quiz 504, 520, Oct. 2009.

[12] J. B. Epstein and P. Güneri, "The adjunctive role of toluidine blue in detection of oral premalignant and malignant lesions," *Curr. Opin. Otolaryngol. Head Neck Surg.*, vol. 17, no. 2, pp. 79–87, 2009.

[13] C. S. Farah and M. J. McCullough, "A pilot case control study on the efficacy of acetic acid wash and chemiluminescent illumination (ViziLite™) in the visualisation of oral mucosal white lesions," *Oral Oncol.*, vol. 43, no. 8, pp. 820–824, 2007.

[14] J. B. Epstein, S. Silverman, J. D. Epstein, S. A. Lonky, and M. A. Bride, "Analysis of oral lesion biopsies identified and evaluated by visual examination, chemiluminescence and toluidine blue," *Oral Oncol.*, vol. 44, no. 6, pp. 538–544, 2008.

[15] L. L. Patton, J. B. Epstein, and A. R. Kerr, "Adjunctive techniques for oral cancer examination and lesion diagnosis: a systematic review of the literature," *J. Am. Dent. Assoc.*, vol. 139, no. 7, pp. 896–905, 2008.

[16] B. Balevi, "Evidence-based decision making: should the general dentist adopt the use of the VELscope for routine screening for oral cancer?," *J. Can. Dent. Assoc.*, vol. 73, no. 7, p. 603, 2007.

[17] J. Reibel, “Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics,” *Crit. Rev. Oral Biol. Med.*, vol. 14, no. 1, pp. 47–62, 2003.
[18] K. P. Schepman and I. van der Waal, “A proposal for a classification and staging system for oral leukoplakia: a preliminary study,” *Eur. J. Cancer. B. Oral Oncol.*, vol. 31, no. 6, pp. 396–398, Nov. 1995, doi: 10.1016/0964-1955(95)00032-1.
[19] K. H. Awan, P. R. Morgan, and S. Warnakulasuriya, “Utility of chemiluminescence (ViziLite™) in the detection of oral potentially malignant disorders and benign keratoses,” *J. Oral Pathol. Med.*, vol. 40, no. 7, pp. 541–544, Aug. 2011, doi: 10.1111/j.1600-0714.2011.01048.x.
[20] L. McIntosh, M. J. McCullough, and C. S. Farah, “The assessment of diffused light illumination and acetic acid rinse (Microlux/DL™) in the visualisation of oral mucosal lesions,” *Oral Oncol.*, vol. 45, no. 12, pp. e227–e231, Dec. 2009, doi: 10.1016/j.oraloncology.2009.08.001.
[21] I. Pavlova, M. Williams, A. El-Naggar, R. Richards-Kortum, and A. Gillenwater, “Understanding the Biological Basis of Autofluorescence Imaging for Oral Cancer Detection: High-Resolution Fluorescence Microscopy in Viable Tissue,” *Clin. Cancer Res.*, vol. 14, no. 8, pp. 2396–2404, Apr. 2008, doi: 10.1158/1078-0432.CCR-07-1609.
[22] Ching-Fen Jiang, Chih-Yu Wang, and Chun-Ping Chiang, “Comparative study of protoporphyrin IX fluorescence image enhancement methods to improve an optical imaging system for oral cancer detection,” vol. 16, pp. 076006-16–9, 2011.
[23] R. Paczona, S. Temam, F. Janot, P. Marandas, and B. Luboinski, “Autofluorescence videoendoscopy for photodiagnosis of head and neck squamous cell carcinoma,” *Eur. Arch. Otorhinolaryngol.*, vol. 260, no. 10, pp. 544–548, Nov. 2003, doi: 10.1007/s00405-003-0635-6.
[24] Darren M. Roblyer *et al.*, “Comparison of multispectral wide-field optical imaging modalities to maximize image contrast for objective discrimination of oral neoplasia,” vol. 15, pp. 066017-15–9, 2010.
[25] P. López-Jornet and T. D. la Mano-Espinosa, “The efficacy of direct tissue fluorescence visualization in screening for oral premalignant lesions in general practice: an update,” *Int. J. Dent. Hyg.*, vol. 9, no. 2, pp. 97–100, May 2011, doi: 10.1111/j.1601-5037.2010.00446.x.
[26] J. T. Liu, N. O. Loewke, M. J. Mandella, R. M. Levenson, J. M. Crawford, and C. H. Contag, “Point-of-care pathology with miniature microscopes,” *Anal. Cell. Pathol.*, vol. 34, no. 3, pp. 81–98, 2011.
[27] P. Guitera, G. Pellacani, C. Longo, S. Seidenari, M. Avramidis, and S. W. Menzies, “In Vivo Reflectance Confocal Microscopy Enhances Secondary Evaluation of Melanocytic Lesions,” *J. Invest. Dermatol.*, vol. 129, no. 1, pp. 131–138, Jan. 2009, doi: 10.1038/jid.2008.193.
[28] G. Pellacani, P. Guitera, C. Longo, M. Avramidis, S. Seidenari, and S. Menzies, “The Impact of In Vivo Reflectance Confocal Microscopy for the Diagnostic Accuracy of Melanoma and Equivocal Melanocytic Lesions,” *J. Invest. Dermatol.*, vol. 127, no. 12, pp. 2759–2765, Dec. 2007, doi: 10.1038/sj.jid.5700993.
[29] C. Yin *et al.*, “Miniature in vivo MEMS-based line-scanned dual-axis confocal microscope for point-of-care pathology,” *Biomed. Opt. Express*, vol. 7, no. 2, pp. 251–263, 2016.

[30] S. Nori *et al.*, “Sensitivity and specificity of reflectance-mode confocal microscopy for in vivo diagnosis of basal cell carcinoma: A multicenter study,” *J. Am. Acad. Dermatol.*, vol. 51, no. 6, pp. 923–930, Dec. 2004, doi: 10.1016/j.jaad.2004.06.028.
[31] K. B. Sung *et al.*, “Near real time in vivo fibre optic confocal microscopy: sub-cellular structure resolved,” *J. Microsc.*, vol. 207, no. 2, pp. 137–145, 2002.
[32] H.-J. Shin, M. C. Pierce, D. Lee, H. Ra, O. Solgaard, and R. Richards-Kortum, “Fiber-optic confocal microscope using a MEMS scanner and miniature objective lens,” *Opt. Express*, vol. 15, no. 15, pp. 9113–9122, 2007.
[33] L. Liu, E. Wang, X. Zhang, W. Liang, X. Li, and H. Xie, “MEMS-based 3D confocal scanning microendoscope using MEMS scanners for both lateral and axial scan,” *Sens. Actuators Phys.*, vol. 215, pp. 89–95, Aug. 2014, doi: 10.1016/j.sna.2013.09.035.
[34] K. Carlson *et al.*, “In vivo fiber-optic confocal reflectance microscope with an injection-molded plastic miniature objective lens,” *Appl. Opt.*, vol. 44, no. 10, pp. 1792–1797, 2005.
[35] C. Olsovsky *et al.*, “Handheld tunable focus confocal microscope utilizing a double-clad fiber coupler for <italic>in vivo</italic> imaging of oral epithelium,” *J. Biomed. Opt.*, vol. 22, no. 5, p. 056008, May 2017, doi: 10.1117/1.JBO.22.5.056008.
[36] R. Richards-Kortum, C. L. Smithpeter, B. S. Bowman, and M. R. Descour, “Fiber-optic confocal imaging apparatus and methods of use,” US6370422B1, 09-Apr-2002.
[37] F. Berier, S. Bourriaux, M. Genet, B. Viellerobe, A. Loiseau, and B. Abrat, “Miniaturized focusing optical head in particular for endoscope,” US7221824B2, 22-May-2007.
[38] M. Kanai, “Condensing optical system, confocal optical system, and scanning confocal endoscope,” US7338439B2, 04-Mar-2008.
[39] K. Kumar *et al.*, “Handheld histology-equivalent sectioning laser-scanning confocal optical microscope for interventional imaging,” *Biomed. Microdevices*, vol. 12, no. 2, pp. 223–233, Apr. 2010, doi: 10.1007/s10544-009-9377-6.
[40] A. A. Tanbakuchi, A. R. Rouse, J. A. Udovich, K. D. Hatch, and A. F. Gmitro, “Clinical confocal microlaparoscope for real-time in vivo optical biopsies,” *J. Biomed. Opt.*, vol. 14, no. 4, p. 044030, Aug. 2009, doi: 10.1117/1.3207139.
[41] T. D. Wang *et al.*, “Functional imaging of colonic mucosa with a fibered confocal microscope for real-time in vivo pathology,” *Clin. Gastroenterol. Hepatol.*, vol. 5, no. 11, pp. 1300–1305, 2007.
[42] J. T. Liu *et al.*, “Micromirror-scanned dual-axis confocal microscope utilizing a gradient-index relay lens for image guidance during brain surgery,” *J. Biomed. Opt.*, vol. 15, no. 2, pp. 026029–026029, 2010.
[43] E. J. Seibel, C. M. Brown, J. A. Dominitz, and M. B. Kimmey, “Scanning Single Fiber Endoscopy: A New Platform Technology for Integrated Laser Imaging, Diagnosis, and Future Therapies,” *Gastrointest. Endosc. Clin. N. Am.*, vol. 18, no. 3, pp. 467–478, Jul. 2008, doi: 10.1016/j.giec.2008.05.001.
[44] C. M. Lee, C. J. Engelbrecht, T. D. Soper, F. Helmchen, and E. J. Seibel, “Scanning fiber endoscopy with highly flexible, 1 mm catheterscopes for wide-field, full-color imaging,” *J. Biophotonics*, vol. 3, no. 5–6, pp. 385–407, Jun. 2010, doi: 10.1002/jbio.200900087.
[45] R. Kiesslich *et al.*, “Confocal laser endoscopy for diagnosing intraepithelial neoplasias and colorectal cancer in vivo,” *Gastroenterology*, vol. 127, no. 3, pp. 706–713, Sep. 2004.
[46] A. L. Polglase, W. J. McLaren, S. A. Skinner, R. Kiesslich, M. F. Neurath, and P. M. Delaney, “A fluorescence confocal endomicroscope for in vivo microscopy of the upper- and the lower-GI tract,” *Gastrointest. Endosc.*, vol. 62, no. 5, pp. 686–695, 2005.

[47] K. Kumar, K. Hoshino, and X. Zhang, “Handheld subcellular-resolution single-fiber confocal microscope using high-reflectivity two-axis vertical combdrive silicon microscanner,” *Biomed. Microdevices*, vol. 10, no. 5, p. 653, May 2008, doi: 10.1007/s10544-008-9176-5.
[48] A. A. Tanbakuchi, J. A. Udovich, A. R. Rouse, K. D. Hatch, and A. F. Gmitro, “In vivo imaging of ovarian tissue using a novel confocal microlaparoscope,” *Am. J. Obstet. Gynecol.*, vol. 202, no. 1, pp. 90.e1-90.e9, Jan. 2010, doi: 10.1016/j.ajog.2009.07.027.
[49] C. L. Hoy *et al.*, “Miniaturized probe for femtosecond laser microsurgery and two-photon imaging,” *Opt. Express*, vol. 16, no. 13, pp. 9996–10005, Jun. 2008, doi: 10.1364/OE.16.009996.
[50] M. T. Myaing, D. J. MacDonald, and X. Li, “Fiber-optic scanning two-photon fluorescence endoscope,” *Opt. Lett.*, vol. 31, no. 8, pp. 1076–1078, Apr. 2006, doi: 10.1364/OL.31.001076.
[51] K. Murari, Y. Zhang, S. Li, Y. Chen, M.-J. Li, and X. Li, “Compensation-free, all-fiber-optic, two-photon endomicroscopy at 1.55 μm,” *Opt. Lett.*, vol. 36, no. 7, pp. 1299–1301, Apr. 2011, doi: 10.1364/OL.36.001299.
[52] F. Helmchen, M. S. Fee, D. W. Tank, and W. Denk, “A Miniature Head-Mounted Two-Photon Microscope: High-Resolution Brain Imaging in Freely Moving Animals,” *Neuron*, vol. 31, no. 6, pp. 903–912, Sep. 2001, doi: 10.1016/S0896-6273(01)00421-4.
[53] X. Li, C. Chudoba, T. Ko, C. Pitris, and J. G. Fujimoto, “Imaging needle for optical coherence tomography,” *Opt. Lett.*, vol. 25, no. 20, pp. 1520–1522, Oct. 2000, doi: 10.1364/OL.25.001520.
[54] N. Bozinovic, C. Ventalon, T. Ford, and J. Mertz, “Fluorescence endomicroscopy with structured illumination,” *Opt. Express*, vol. 16, no. 11, pp. 8016–8025, May 2008, doi: 10.1364/OE.16.008016.
[55] T. N. Ford, D. Lim, and J. Mertz, “Fast optically sectioned fluorescence HiLo endomicroscopy,” *J. Biomed. Opt.*, vol. 17, no. 2, p. 021105, Mar. 2012, doi: 10.1117/1.JBO.17.2.021105.
[56] M. R. Kyrish *et al.*, “Needle-based fluorescence endomicroscopy via structured illumination with a plastic, achromatic objective,” *J. Biomed. Opt.*, vol. 18, no. 9, p. 096003, Sep. 2013, doi: 10.1117/1.JBO.18.9.096003.
[57] J. T. Liu *et al.*, “Miniature near-infrared dual-axes confocal microscope utilizing a two-dimensional microelectromechanical systems scanner,” *Opt. Lett.*, vol. 32, no. 3, pp. 256–258, 2007.
[58] M. Marvin, “Microscopy apparatus,” US3013467A, 19-Dec-1961.
[59] E. H. K. Stelzer and S. Lindek, “Fundamental reduction of the observation volume in far-field light microscopy by detection orthogonal to the illumination axis: confocal theta microscopy,” *Opt. Commun.*, vol. 111, no. 5, pp. 536–547, Oct. 1994, doi: 10.1016/0030-4018(94)90533-9.
[60] S. Lindek and E. H. K. Stelzer, “Confocal theta microscopy and 4Pi-confocal theta microscopy,” in *Three-Dimensional Microscopy: Image Acquisition and Processing*, 1994, vol. 2184, pp. 188–195, doi: 10.1117/12.172093.
[61] S. Lindek and E. H. K. Stelzer, “Resolution improvement by nonconfocal theta microscopy,” *Opt. Lett.*, vol. 24, no. 21, pp. 1505–1507, Nov. 1999, doi: 10.1364/OL.24.001505.

[62] T. D. Wang, M. J. Mandella, C. H. Contag, and G. S. Kino, “Dual-axis confocal microscope for high-resolution in vivo imaging,” *Opt. Lett.*, vol. 28, no. 6, pp. 414–416, Mar. 2003, doi: 10.1364/OL.28.000414.
[63] R. H. Webb and F. Rogomentich, “Confocal microscope with large field and working distance,” *Appl. Opt.*, vol. 38, no. 22, pp. 4870–4875, Aug. 1999, doi: 10.1364/AO.38.004870.
[64] T. D. Wang, C. H. Contag, M. J. Mandella, N. Chan, and G. S. Kino, “Dual-axes confocal microscopy with post-objective scanning and low-coherence heterodyne detection,” *Opt. Lett.*, vol. 28, no. 20, pp. 1915–1917, Oct. 2003, doi: 10.1364/OL.28.001915.
[65] L. K. Wong, M. J. Mandella, G. S. Kino, and T. D. Wang, “Improved rejection of multiply scattered photons in confocal microscopy using dual-axes architecture,” *Opt. Lett.*, vol. 32, no. 12, pp. 1674–1676, Jun. 2007, doi: 10.1364/OL.32.001674.
[66] J. T. Liu, “Efficient rejection of scattered light enables deep optical sectioning in turbid media with low-numerical-aperture optics in a dual-axis confocal architecture.,” *J Biomed Opt*, vol. 13, p. 034020, 2008.
[67] J. T. Liu *et al.*, “Dual-axes confocal reflectance microscope for distinguishing colonic neoplasia,” *J. Biomed. Opt.*, vol. 11, no. 5, pp. 054019–054019, 2006.
[68] M. Born *et al.*, “Principles of Optics by Max Born,” *Cambridge Core*, Oct-1999. [Online]. Available: /core/books/principles-of-optics/D12868B8AE26B83D6D3C2193E94FFC32. [Accessed: 24-Nov-2018].
[69] J. Pawley, *Handbook of biological confocal microscopy*. Springer Science & Business Media, 2010.
[70] M. Rajadhyaksha, R. R. Anderson, and R. H. Webb, “Video-rate confocal scanning laser microscope for imaging human tissues in vivo,” *Appl. Opt.*, vol. 38, no. 10, pp. 2105–2115, 1999.
[71] T. R. Corle, C.-H. Chou, and G. S. Kino, “Depth response of confocal optical microscopes,” *Opt. Lett.*, vol. 11, no. 12, pp. 770–772, Dec. 1986, doi: 10.1364/OL.11.000770.
[72] C. J. R. Sheppard and T. Wilson, “Depth of field in the scanning microscope,” *Opt. Lett.*, vol. 3, no. 3, pp. 115–117, Sep. 1978, doi: 10.1364/OL.3.000115.
[73] R. H. Webb, “Confocal optical microscopy,” *Rep. Prog. Phys.*, vol. 59, no. 3, p. 427, 1996, doi: 10.1088/0034-4885/59/3/003.
[74] Y. Chen, D. Wang, and J. T. Liu, “Assessing the tissue-imaging performance of confocal microscope architectures via Monte Carlo simulations,” *Opt. Lett.*, vol. 37, no. 21, pp. 4495–4497, 2012.
[75] Y. Chen and J. T. C. Liu, “Optimizing the performance of dual-axis confocal microscopes via Monte-Carlo scattering simulations and diffraction theory,” *J. Biomed. Opt.*, vol. 18, no. 6, p. 066006, Jun. 2013, doi: 10.1117/1.JBO.18.6.066006.
[76] D. Wang, Y. Chen, Y. Wang, and J. T. C. Liu, “Comparison of line-scanned and point-scanned dual-axis confocal microscope performance,” *Opt. Lett.*, vol. 38, no. 24, pp. 5280–5283, 2013.
[77] D. Meza, D. Wang, Y. Wang, S. Borwege, N. Sanai, and J. T. C. Liu, “Comparing high-resolution microscopy techniques for potential intraoperative use in guiding low-grade glioma resections,” *Lasers Surg. Med.*, vol. 47, no. 4, pp. 289–295, Apr. 2015, doi: 10.1002/lsm.22347.

[78] L. Wei, Y. Chen, C. Yin, S. Borwege, N. Sanai, and J. T. C. Liu, “Optical-sectioning microscopy of protoporphyrin IX fluorescence in human gliomas: standardization and quantitative comparison with histology,” *J. Biomed. Opt.*, vol. 22, no. 4, p. 46005, 01 2017, doi: 10.1117/1.JBO.22.4.046005.
[79] A. K. Glaser *et al.*, “Light-sheet microscopy for slide-free non-destructive pathology of large clinical specimens,” *Nat. Biomed. Eng.*, vol. 1, no. 7, p. 0084, Jul. 2017, doi: 10.1038/s41551-017-0084.
[80] H. Ra *et al., Two-Dimensional MEMS Scanner for Dual-Axes Confocal Microscopy*. .
[81] Y. Chen, A. Glaser, and J. T. C. Liu, “Bessel-beam illumination in dual-axis confocal microscopy mitigates resolution degradation caused by refractive heterogeneities,” *J. Biophotonics*, vol. 10, no. 1, pp. 68–74, 2017, doi: 10.1002/jbio.201600196.
[82] S. Y. Leigh, Y. Chen, and J. T. C. Liu, “Modulated-alignment dual-axis (MAD) confocal microscopy for deep optical sectioning in tissues,” *Biomed. Opt. Express*, vol. 5, no. 6, pp. 1709–1720, Jun. 2014, doi: 10.1364/BOE.5.001709.
[83] Y. Chen *et al.*, “Video-rate in vivo fluorescence imaging with a line-scanned dual-axis confocal microscope,” *J. Biomed. Opt.*, vol. 20, no. 10, pp. 106011–106011, 2015.
[84] D. Wang, D. Meza, Y. Wang, L. Gao, and J. T. Liu, “Sheet-scanned dual-axis confocal microscopy using Richardson–Lucy deconvolution,” *Opt. Lett.*, vol. 39, no. 18, pp. 5431–5434, 2014.
[85] J. Huisken, “Optical sectioning deep inside live embryos by selective plane illumination microscopy,” *Science*, vol. 305, pp. 1007–1009, 2004.
[86] A. K. Glaser, Y. Wang, and J. T. Liu, “Assessing the imaging performance of light sheet microscopies in highly scattering tissues,” *Biomed Opt Express*, vol. 7, pp. 454–466, 2016.
[87] N. Chen, C.-H. Wong, and C. J. R. Sheppard, “Focal modulation microscopy,” *Opt. Express*, vol. 16, no. 23, pp. 18764–18769, Nov. 2008, doi: 10.1364/OE.16.018764.
[88] K. Isobe *et al.*, “Background-free deep imaging by spatial overlap modulation nonlinear optical microscopy,” *Biomed. Opt. Express*, vol. 3, no. 7, pp. 1594–1608, Jul. 2012, doi: 10.1364/BOE.3.001594.
[89] S. Y. Leigh, Y. Chen, and J. T. C. Liu, “Modulated-Alignment Dual-Axis (MAD) Confocal Microscopy Optimized for Speed and Contrast,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 10, pp. 2119–2124, Oct. 2016, doi: 10.1109/TBME.2015.2511581.
[90] T. A. Planchon *et al.*, “Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination,” *Nat. Methods*, vol. 8, no. 5, pp. 417–423, May 2011, doi: 10.1038/nmeth.1586.
[91] F. O. Fahrbach and A. Rohrbach, “Propagation stability of self-reconstructing Bessel beams enables contrast-enhanced imaging in thick media.,” *Nat Commun*, vol. 3, p. 632, 2012.
[92] D. Lorenser, C. C. Singe, A. Curatolo, and D. D. Sampson, “Energy-efficient low-Fresnel-number Bessel beams and their application in optical coherence tomography,” *Opt. Lett.*, vol. 39, no. 3, pp. 548–551, Feb. 2014, doi: 10.1364/OL.39.000548.
[93] Y. Chen and J. T. C. Liu, “Characterizing the beam steering and distortion of Gaussian and Bessel beams focused in tissues with microscopic heterogeneities,” *Biomed. Opt. Express*, vol. 6, no. 4, pp. 1318–1330, Mar. 2015, doi: 10.1364/BOE.6.001318.
[94] F. O. Fahrbach, V. Gurchenkov, K. Alessandri, P. Nassoy, and A. Rohrbach, “Light-sheet microscopy in thick media using scanned Bessel beams and two-photon fluorescence

excitation," *Opt. Express*, vol. 21, no. 11, pp. 13824–13839, Jun. 2013, doi: 10.1364/OE.21.013824.
[95] M. Zhao *et al.*, "Cellular imaging of deep organ using two-photon Bessel light-sheet nonlinear structured illumination microscopy," *Biomed. Opt. Express*, vol. 5, no. 5, pp. 1296–1308, May 2014, doi: 10.1364/BOE.5.001296.
[96] D. S. Gareau, S. Abeytunge, and M. Rajadhyaksha, "Line-Scanning Reflectance Confocal Microscopy of Human Skin: Comparison of Full-pupil and Divided-pupil Configurations," *Opt. Lett.*, vol. 34, no. 20, pp. 3235–3237, Oct. 2009.
[97] C. J. Koester, "Scanning mirror microscope with optical sectioning characteristics: applications in ophthalmology," *Appl. Opt.*, vol. 19, no. 11, pp. 1749–1757, Jun. 1980, doi: 10.1364/AO.19.001749.
[98] Y. G. Patel, M. Rajadhyaksha, and C. A. DiMarzio, "Optimization of pupil design for point-scanning and line-scanning confocal microscopy," *Biomed. Opt. Express*, vol. 2, no. 8, pp. 2231–2242, Jul. 2011, doi: 10.1364/BOE.2.002231.
[99] P. J. Dwyer, C. A. DiMarzio, J. M. Zavislan, W. J. Fox, and M. Rajadhyaksha, "Confocal reflectance theta line scanning microscope for imaging human skin in vivo," *Opt. Lett.*, vol. 31, no. 7, pp. 942–944, 2006.
[100] P. J. Dwyer, C. A. DiMarzio, and M. Rajadhyaksha, "Confocal theta line-scanning microscope for imaging human tissues," *Appl. Opt.*, vol. 46, no. 10, pp. 1843–1851, Apr. 2007, doi: 10.1364/AO.46.001843.
[101] S. Kumar, D. Wilding, M. B. Sikkel, A. R. Lyon, K. T. MacLeod, and C. Dunsby, "High-speed 2D and 3D fluorescence microscopy of cardiac myocytes," *Opt. Express*, vol. 19, no. 15, pp. 13839–13847, Jul. 2011, doi: 10.1364/OE.19.013839.
[102] M. B. Bouchard *et al.*, "Swept confocally-aligned planar excitation (SCAPE) microscopy for high-speed volumetric imaging of behaving organisms," *Nat. Photonics*, vol. 9, no. 2, pp. 113–119, Feb. 2015, doi: 10.1038/nphoton.2014.323.
[103] C. Glazowski, G. Peterson, and M. Rajadhyaksha, "Compact divided-pupil line-scanning confocal microscope for investigation of human tissues," in *Photonic Therapeutics and Diagnostics IX*, 2013, vol. 8565, p. 856523, doi: 10.1117/12.2012630.
[104] N. Iftimia *et al.*, "Combined reflectance confocal microscopy/optical coherence tomography imaging for skin burn assessment," *Biomed. Opt. Express*, vol. 4, no. 5, pp. 680–695, May 2013, doi: 10.1364/BOE.4.000680.
[105] X. Liu, M. J. Cobb, Y. Chen, M. B. Kimmey, and X. Li, "Rapid-scanning forward-imaging miniature endoscope for real-time optical coherence tomography," *Opt. Lett.*, vol. 29, no. 15, pp. 1763–1765, Aug. 2004, doi: 10.1364/OL.29.001763.
[106] W. Piyawattanametha *et al.*, "In vivo near-infrared dual-axis confocal microendoscopy in the human lower gastrointestinal tract," *J. Biomed. Opt.*, vol. 17, no. 2, pp. 0211021–0211024, 2012.
[107] G. Li *et al.*, "Visualizing Epithelial Expression in Vertical and Horizontal Planes With Dual Axes Confocal Endomicroscope Using Compact Distal Scanner," *IEEE Trans. Med. Imaging*, vol. 36, no. 7, pp. 1482–1490, Jul. 2017, doi: 10.1109/TMI.2017.2673022.
[108] Z. Qiu *et al.*, "Targeted vertical cross-sectional imaging with handheld near-infrared dual axes confocal fluorescence endomicroscope," *Biomed. Opt. Express*, vol. 4, no. 2, pp. 322–330, Feb. 2013, doi: 10.1364/BOE.4.000322.

[109] Z. Gao *et al.*, "In vivo near-infrared imaging of ErbB2 expressing breast tumors with dual-axes confocal endomicroscopy using a targeted peptide," *Sci. Rep.*, vol. 7, no. 1, p. 14404, Oct. 2017, doi: 10.1038/s41598-017-13735-z.
[110] H. Li *et al.*, "Integrated monolithic 3D MEMS scanner for switchable real time vertical/horizontal cross-sectional imaging," *Opt. Express*, vol. 24, no. 3, pp. 2145–2155, Feb. 2016, doi: 10.1364/OE.24.002145.
[111] W. Piyawattanametha and T. D. Wang, "MEMS-Based Dual-Axes Confocal Microendoscopy," *IEEE J. Sel. Top. Quantum Electron.*, vol. 16, no. 4, pp. 804–814, Jul. 2010, doi: 10.1109/JSTQE.2009.2032785.
[112] S. Y. Leigh and J. T. Liu, "Multi-color miniature dual-axis confocal microscope for point-of-care pathology," *Opt. Lett.*, vol. 37, no. 12, pp. 2430–2432, 2012.
[113] E. S. Flores *et al.*, "Intraoperative imaging during Mohs surgery with reflectance confocal microscopy: initial clinical experience," *J. Biomed. Opt.*, vol. 20, no. 6, pp. 061103–061103, 2015.
[114] V. Becker, T. Vercauteren, C. H. von Weyhern, C. Prinz, R. M. Schmid, and A. Meining, "High-resolution miniprobe-based confocal microscopy in combination with video mosaicing (with video)," *Gastrointest. Endosc.*, vol. 66, no. 5, pp. 1001–1007, 2007.
[115] K. E. Loewke *et al.*, "In vivo micro-image mosaicing," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 159–171, 2011.
[116] Y. Gong, T. D. Soper, V. W. Hou, D. Hu, B. Hannaford, and E. J. Seibel, "Mapping surgical fields by moving a laser-scanning multimodal scope attached to a robot arm," in *Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*, 2014, vol. 9036, p. 90362S.
[117] M. J. Lum *et al.*, "The RAVEN: Design and validation of a telesurgery system," *Int. J. Robot. Res.*, vol. 28, no. 9, pp. 1183–1197, 2009.
[118] A. E. Sonmez, W. M. Spees, A. Özcan, Z. Deng, A. G. Webb, and N. V. Tsekos, "Robot-assisted mechanical scanning and co-registration of Magnetic Resonance Imaging and light-induced fluorescence," in *2012 4th IEEE RAS EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2012, pp. 775–780, doi: 10.1109/BioRob.2012.6290699.
[119] L. Wei, C. Yin, and J. T. C. Liu, "Dual-Axis Confocal Microscopy for Point-of-Care Pathology," *IEEE J. Sel. Top. Quantum Electron.*, vol. 25, no. 1, pp. 1–10, Jan. 2019, doi: 10.1109/JSTQE.2018.2854572.
[120] C. Glazowski and M. Rajadhyaksha, "Optimal detection pinhole for lowering speckle noise while maintaining adequate optical sectioning in confocal reflectance microscopes," *J. Biomed. Opt.*, vol. 17, no. 8, pp. 0850011–0850015, 2012.
[121] A. Vakili, D. Xiong, M. Rajadhyaksha, and C. A. DiMarzio, "High brightness LED in confocal microscopy," in *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXII*, 2015, vol. 9330, p. 933006, doi: 10.1117/12.2078191.
[122] K. Kose, M. Cordova, M. Duffy, E. S. Flores, D. H. Brooks, and M. Rajadhyaksha, "Video-mosaicing of reflectance confocal images for rapid examination of large areas of skin in vivo," *Br. J. Dermatol.*, vol. 171, no. 5, p. 1239, 2014.
[123] C. Yin, L. Wei, S. Abeytunge, G. Peterson, M. Rajadhyaksha, and J. T. C. Liu, "Label-free *in vivo* pathology of human epithelia with a high-speed handheld dual-axis confocal microscope," *J. Biomed. Opt.*, vol. 24, no. 3, p. 030501, Mar. 2019, doi: 10.1117/1.JBO.24.3.030501.

[124] G. Peterson, D. K. Zanoni, M. Ardigo, J. C. Migliacci, S. G. Patel, and M. Rajadhyaksha, "Feasibility of a Video-Mosaicking Approach to Extend the Field-of-View For Reflectance Confocal Microscopy in the Oral Cavity In Vivo," *Lasers Surg. Med.*, vol. 51, no. 5, pp. 439–451, 2019, doi: 10.1002/lsm.23090.
[125] J. T. C. Liu, D. Meza, and N. Sanai, "Trends in Fluorescence Image-guided Surgery for Gliomas," *Neurosurgery*, vol. 75, no. 1, pp. 61–71, Jul. 2014, doi: 10.1227/NEU.0000000000000344.
[126] L. Wei, Y. Fujita, N. Sanai, and J. T. C. Liu, "Toward Quantitative Neurosurgical Guidance With High-Resolution Microscopy of 5-Aminolevulinic Acid-Induced Protoporphyrin IX," *Front. Oncol.*, vol. 9, p. 592, 2019, doi: 10.3389/fonc.2019.00592.
[127] D. S. Gareau, "Confocal mosaicing microscopy in Mohs skin excisions: feasibility of rapid surgical pathology," *J Biomed Opt*, vol. 13, p. 054001, 2008.
[128] F. Piccinini, A. Bevilacqua, and E. Lucarelli, "Automated image mosaics by non-automated light microscopes: the MicroMos software tool," *J. Microsc.*, vol. 252, no. 3, pp. 226–250, 2013, doi: 10.1111/jmi.12084.
[129] S. K. Chow *et al.*, "Automated microscopy system for mosaic acquisition and processing," *J. Microsc.*, vol. 222, no. 2, pp. 76–84, 2006, doi: 10.1111/j.1365-2818.2006.01577.x.
[130] M. A. Saldua, C. A. Olsovsky, E. S. Callaway, R. S. Chapkin, and K. C. Maitland, "Imaging inflammation in mouse colon using a rapid stage-scanning confocal fluorescence microscope," *J. Biomed. Opt.*, vol. 17, no. 1, Jan. 2012, doi: 10.1117/1.JBO.17.1.016006.
[131] T. Yang *et al.*, "Rapid imaging of large tissues using high-resolution stage-scanning microscopy," *Biomed. Opt. Express*, vol. 6, no. 5, pp. 1867–1875, Apr. 2015, doi: 10.1364/BOE.6.001867.
[132] L. Gao, "Extend the field of view of selective plan illumination microscopy by tiling the excitation light sheet," *Opt. Express*, vol. 23, no. 5, pp. 6102–6111, Mar. 2015, doi: 10.1364/OE.23.006102.
[133] Y. K. Tao *et al.*, "Assessment of breast pathologies using nonlinear microscopy," *Proc. Natl. Acad. Sci.*, vol. 111, no. 43, pp. 15304–15309, Oct. 2014, doi: 10.1073/pnas.1416955111.
[134] M. G. Giacomelli, Y. Sheikine, H. Vardeh, J. L. Connolly, and J. G. Fujimoto, "Rapid imaging of surgical breast excisions using direct temporal sampling two photon fluorescent lifetime imaging," *Biomed. Opt. Express*, vol. 6, no. 11, pp. 4317–4325, Oct. 2015, doi: 10.1364/BOE.6.004317.
[135] S. Abeytunge, Y. Li, B. Larson, R. Toledo-Crow, and M. Rajadhyaksha, "Rapid confocal imaging of large areas of excised tissue with strip mosaicing," *J. Biomed. Opt.*, vol. 16, no. 5, May 2011, doi: 10.1117/1.3582335.
[136] S. Preibisch and P. Tomancak, "Globally optimal stitching of tiled 3D microscopic image acquisitions.," *Bioinformatics*, vol. 25, pp. 1463–1465, 2009.
[137] N. Bedard, T. Quang, K. Schmeler, R. Richards-Kortum, and T. S. Tkaczyk, "Real-time video mosaicing with a high-resolution microendoscope," *Biomed. Opt. Express*, vol. 3, no. 10, pp. 2428–2435, Sep. 2012, doi: 10.1364/BOE.3.002428.
[138] J. L. Schroeder, M. Bakalar, T. J. Pohida, and R. S. Balaban, "Rapid Overlapping-Volume Acquisition and Reconstruction (ROVAR): Automated 3D Tiling for High-Resolution, Large Field-of-View Optical Microscopy," *J. Microsc.*, vol. 243, no. 1, pp. 103–110, Jul. 2011, doi: 10.1111/j.1365-2818.2011.03490.x.

[139] L. Wei, C. Yin, Y. Fujita, N. Sanai, and J. T. C. Liu, “Handheld line-scanned dual-axis confocal microscope with pistoned MEMS actuation for flat-field fluorescence imaging,” *Opt. Lett.*, vol. 44, no. 3, pp. 671–674, Feb. 2019, doi: 10.1364/OL.44.000671.
[140] H. Bay, T. Tuytelaars, and L. Van Gool, “SURF: Speeded Up Robust Features,” in *Computer Vision – ECCV 2006*, 2006, pp. 404–417.
[141] D. G. Lowe, “Object recognition from local scale-invariant features,” in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 1999, vol. 2, pp. 1150–1157 vol.2, doi: 10.1109/ICCV.1999.790410.
[142] M. Muja and D. G. Lowe, “Fast approximate nearest neighbors with automatic algorithm configuration.,” *VISAPP 1*, vol. 2, no. 331–340, p. 2, 2009.
[143] M. Rajadhyaksha, M. Grossman, D. Esterowitz, R. H. Webb, and R. Rox Anderson, “In Vivo Confocal Scanning Laser Microscopy of Human Skin: Melanin Provides Strong Contrast,” *J. Invest. Dermatol.*, vol. 104, no. 6, pp. 946–952, Jun. 1995, doi: 10.1111/1523-1747.ep12606215.
[144] M. Rajadhyaksha, S. González, J. M. Zavislan, R. Rox Anderson, and R. H. Webb, “In Vivo Confocal Scanning Laser Microscopy of Human Skin II: Advances in Instrumentation and Comparison With Histology11The authors have declared conflict of interest.,” *J. Invest. Dermatol.*, vol. 113, no. 3, pp. 293–303, Sep. 1999, doi: 10.1046/j.1523-1747.1999.00690.x.

APPENDIX A

Alignment procedures of the handheld device.

Step 1:

Level and fix the main body on a translation stage by using a V-block.

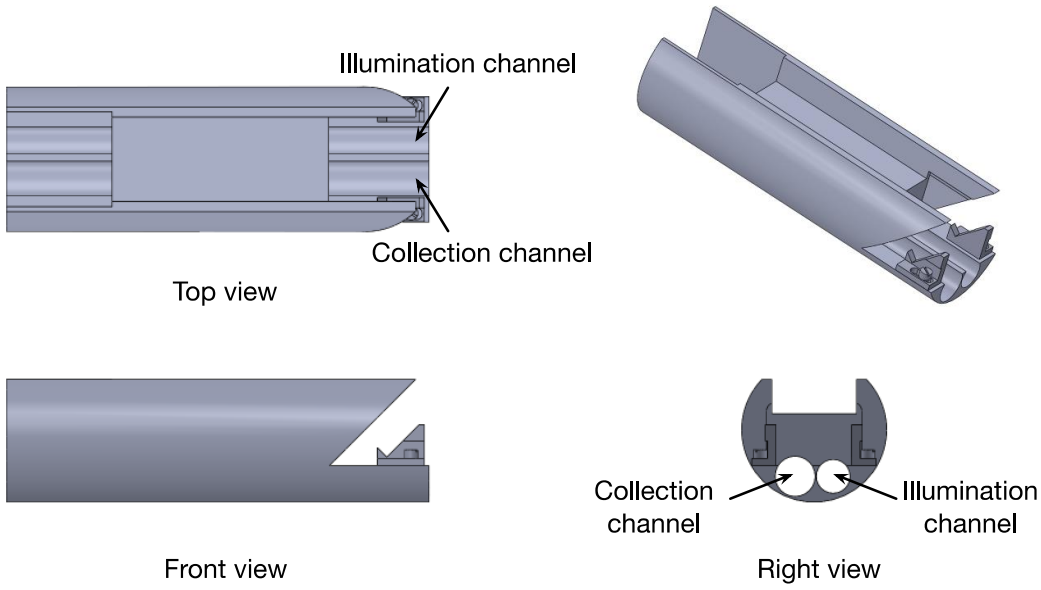


Figure A.1. Three standard views of the main body. There are two channels in the main body. One is for illumination and the other one is for collection.

Step 2:

Send a collimated beam through the center of the illumination channel. Make sure the beam is parallel to the channel.

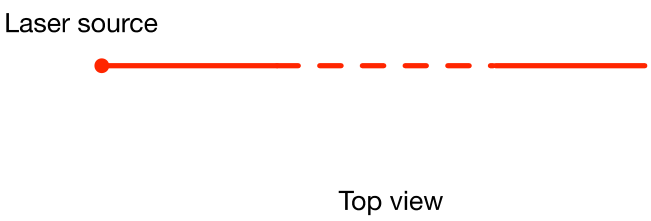


Figure A.2. A collimated beam travels through the illumination channel.

Step 3:

a) Insert the 3.2-mm-diameter prism into the front end of the illumination channel. Make sure the beam is reflected up vertically.

b) Rotate the prism by 11 deg (shown in fig. 6.3).

c) Secure the position of the 1st prism by UV glue.

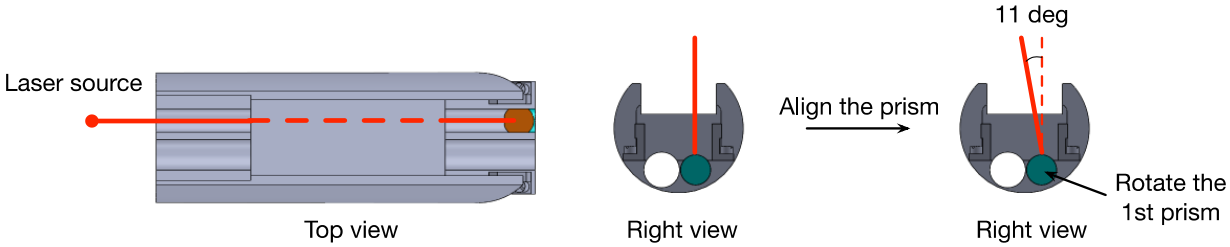


Figure A.3. The alignment procedure for aligning the first prism.

Step 4:

Replace the collimated beam by inserting the fiber module into the illumination channel. The tip of the fiber module should stay very close to the prism. Do not push the fiber module hard against the prism.

Figure A.4. Insert the fiber module into the illumination channel.

Step 5:

a) Insert the L2 lens into the collection channel and then send a collimated beam through the center of the collection channel.

b) Insert the 3.8-mm-diameter prism into the front end of the collection channel.

c) Use a 100x objective with a digital detector to observe the focused line (illumination) and the focused spot (collection).

d) Rotate the fiber module in the illumination channel to make the line-focused beam parallel to the illumination channel (the line-focused beam should be vertical in the focal plane after mounting the MEMS mirror).

e) Adjust the position of L2 and the 2nd prism along the collection channel and rotate the 2nd prism until we see the focused spot is located at the center of the focused line.

f) Make sure both beams travels in the XY plane after reflected by the prisms.

g) Fix the lens and the prism by UV glue.

h) Wait at least three hours or one night if possible and check the alignment again by using the 100x objective before starting the next step.

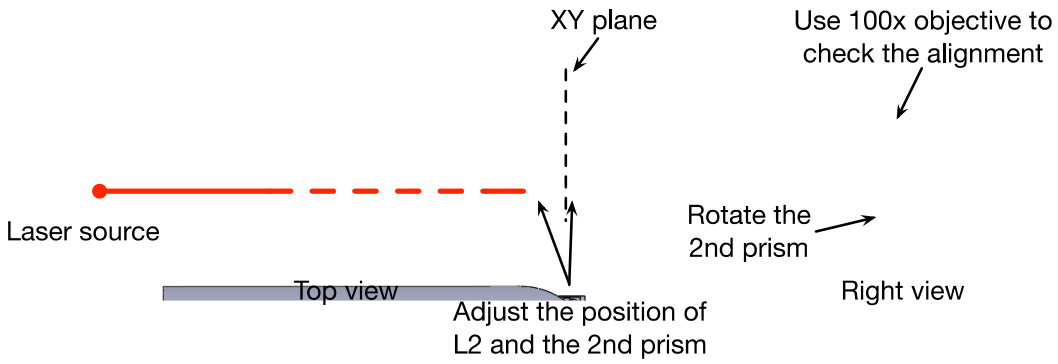


Figure A.5. The alignment procedure of aligning L2 and the 2nd prism.

Step 6:

Remove the collimated beam and glue L3 at the back end of the collection channel.

Figure A.6. Insert L3 into the collection channel.

Step 7:

a) Mount the MEMS mirror (shown in fig. 6.7) on the main body.

b) Make sure the beam is hitting the center of the MEMS mirror.

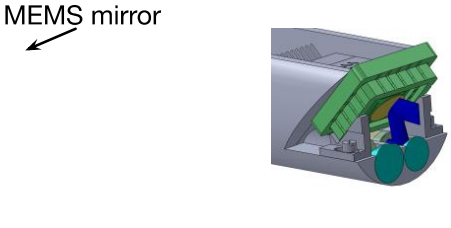


Figure A.7. Mount the MEMS on the main body.

Step 8:

a) Remove the main body from the V-block.

b) Mount the positioner (SSP05 - 16 mm cage plate positioner from Thorlabs Inc.) or the housing on a translation stage.

c) Connect the detector and the positioner by using adaptors (check the adaptor type based on detector's datasheet).

d) Insert the main body into the 1st housing (remember to connect the wires of the MEMS mirror before inserting).

e) Connect positioner and the 1st housing by three 4-mm rods for 16-mm cage system (Thorlabs Inc).

f) Place a USAF target at the focal plane.

g) Move the detector in X and Y directions and rotate the detector around Z axis by using the positioner to make sure the line-shape collection beam fulfills the linear detector or the center of the 2D detector.

h) Adjust the distance between the detector the main body until getting the best lateral resolution (should almost see the smallest element of group 7).

i) Secure the positioner, rods and the main body by set crews or glue.

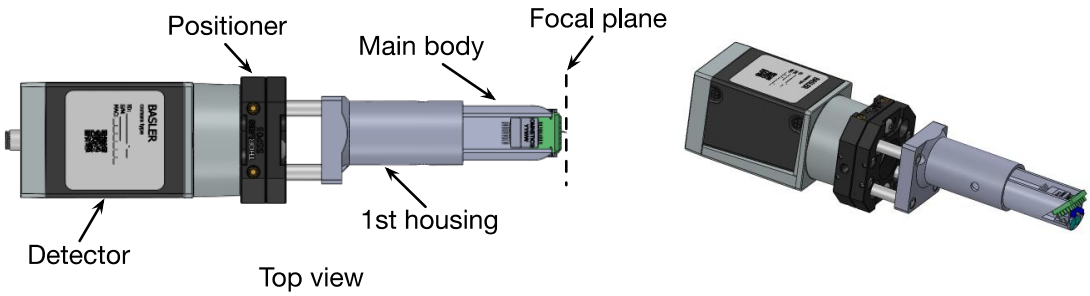


Figure A.8. Alignment procedure of aligning the detector.

Step 9:

1) Thread the lock ring and the 2nd housing onto the 1st housing.

2) Glue the 3rd housing and the 3x relay together and make sure to push the flange of the 3x relay against the 3rd housing.

3) Thread another lock ring onto the 3rd housing.

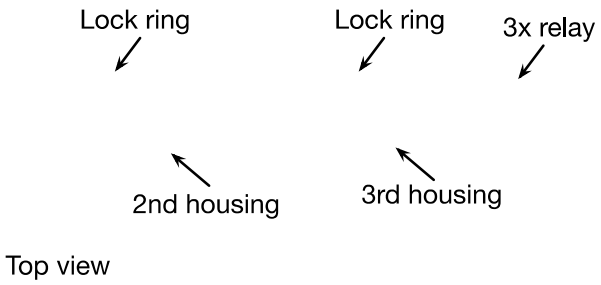


Figure A.9. Prepare the 3x relay and the rest of the housing parts.

Step 10:

1) Connect the 3rd housing with the 3x relay by using the 2nd housing.

2) Place a USAF target at the focal plane near the 3x relay. Adjust the position between the main body and the 3x relay to see the 1st or the 2nd element of the group 9 on the USAF target.

3) Do axial scanning to double-check the axial response (check the previous papers to see the ideal results: axial resolution <=1.8 μm, background < 0.1%).

Figure A.10. Alignment procedure for aligning the 3x relay.

Step 11:

a) Adjust the position of the 2nd housing to evenly hold the 1st and the 3rd housing.

b) Tight the lock rings to secure the position of the housing parts.

c) We can wrap tapes on the device to avoid any part loosing.

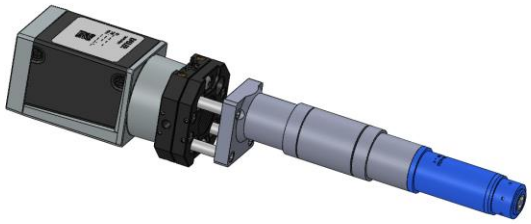


Figure A.11. Fully packaged device after alignment.