3-D VOLUMETRIC OPTICAL COHERENCE TOMOGRAPHY IMAGING
AND IMAGE ANALYSIS OF BARRETT'S ESOPHAGUS

by
WEI KANG

Submitted in partial fulfillment of the requirements
For the degree of Doctor of Philosophy

Dissertation Adviser: Dr. Andrew M. Rollins

Department of Biomedical Engineering
CASE WESTERN RESERVE UNIVERSITY

May, 2011
We hereby approve the thesis/dissertation of

Wei Kang

candidate for the _____ Ph.D ________________ degree *

(signed)_______ Andrew Rollins

(time and grade)

David Wilson

Amitabh Chak

Kenneth Singer

(date) March 4th, 2011

*We also certify that written approval has been obtained for any proprietary material contained therein.
Table of Contents

LIST OF TABLES .............................................................................................................. 7
LIST OF FIGURES ............................................................................................................. 8
ACKNOWLEDGEMENTS ................................................................................................. 14
LIST OF ABBREVIATIONS ............................................................................................ 15
ABSTRACT ...................................................................................................................... 17
CHAPTER 1. BACKGROUND AND SIGNIFICANCE .................................................... 19
  1.1 Barrett’s esophagus (BE) ..................................................................................... 19
  1.2 Current BE surveillance ...................................................................................... 20
  1.3 Optical coherence tomography (OCT) ................................................................. 22
    1.3.1 TD-OCT ........................................................................................................ 23
    1.3.2 Mathematical treatment ............................................................................... 27
    1.3.3 Spectral-domain OCT .................................................................................. 32
  1.4 Overview of EOCT for BE .................................................................................... 37
  1.5 Competing endoscopic imaging modalities ....................................................... 41
CHAPTER 2. ENDOSCOPIC OCT .................................................................................. 45
  2.1 Introduction .......................................................................................................... 45
  2.2 Methods ............................................................................................................... 46
    2.2.1 EOCT System Design .................................................................................. 46
    2.2.2 Spectrometer ................................................................................................ 49
    2.2.3 Balloon catheter probe ................................................................................ 49
List of Tables

**Table 1.1** Diagnostic performance of EOCT for esophageal disorder............. 39

**Table 4.1** Abbreviations of texture analysis measures................................. 103

**Table 5.1** Patient enrollment......................................................................... 124
List of Figures

Fig. 1.1 The appearance of esophageal mucosa as the progression from BE to EAC. (Source: http://pathology2.jhu.edu/beweb/cancer.cfm)............................. 20

Fig. 1.2 TD-OCT with a Michelson-Morley interferometry configuration. Light interferes only when the optical path lengths of the two arms are matched to within the coherence length of the broadband light source. For example, light reflected from layer a in the tissue only interferes with the mirror in position a′ and generates interferometric signal aa′. Depth is therefore discriminated.............. 25

Fig. 1.3 (a) A back-scattering profile or A-scan plotted as a function of in-depth position. (b) By stacking the A-scans rotationally, a cross-sectional esophageal OCT images is constructed. The dash arrow indicates the position of (a). (c) A 3-D OCT volume can be reconstructed by stacking the cross-sectional images along the longitudinal direction........................................................................... 26

Fig. 1.4 Illustration of the interferometric signal generated in Fig. 1.2. Position Za, Zb and Zc are the locations of zero path-length mis-match of the layers a, b and c in the sample. The red curves show the Gaussian profile of the signal, and l_c is the FWHM, also known as the in-depth or axial resolution (described below)... 29

Fig. 1.5 Diagram of SD-OCT system. A spectrometer is utilized to record the interferometric fringes. The fringes then undergoes several steps, including DC subtraction, re-sampling, dispersion compensation, inverse Fourier transform, and logarithmic compression to produce an image. The mirror in the reference arm is stationary. The sample is assumed to have 3 layers. The dash arrow indicates the position of zero path-length mismatch........................................... 33

Fig. 1.6 (a) Fiber-optic catheter probe that could be inserted through standard endoscopes. (b) The point scanning scheme of the short working distance probe. An EMR cap was utilized to stabilize the endoscope. The probe was inserted to obtained a cross-sectional image of a small region............................................ 38

Fig. 1.7 Endoscopic TD-OCT images from the study in dysplasia diagnosis [90]. (a) Low grade dysplasia (area marked by sunburst). (b) HGD (area marked by sunburst). Two criteria for diagnosing dysplasia were selected: reduced light scattering and loss of tissue architecture. Solid lines are each 2 mm in length.. 40
Fig. 2.1 (a) System diagram of the double-balloon-based EOCT system. Within the catheter, the beam can be located between two balloons for imaging without balloon-tissue contact (labeled beam position 1) or in a balloon for imaging with contact (labeled beam position 2). The inset photo (top right) shows the double-balloon catheter (compared with a dime) inserted through the GI endoscope and inflated. (b) The entire system was fit into a portable rack.

Fig. 2.2 (a) Two perpendicular views of the fiber-optic probe design of a GRIN lens with a polished cylindrical surface. The cylindrical surface compensates the astigmatism induced by the outer sheath. (b) An illustration when the cylindrical surface is poorly polished and asymmetric with respect to the optical axis.

Fig. 2.3 (a) The mechanical model of the second generation of probe design. A single-mode fiber is fixed in a ferrule followed by a glass spacer to expand the beam before entering a GRIN lens. A separate cylinder rod lens is attached after the GRIN lens to correct the astigmatism. (b) An assembled fiber-optic probe without metal housing compared to a dime.

Fig. 2.4 A photograph of the rotary-joint-pullback unit. The rotor of the fiber-optic rotary joint (FORJ) is driven by a motor for rotational scanning. The jig is fixed on a translational stage and driven by a linear actuator for longitudinal pulling.

Fig. 2.5 (a) Three dimensional drawing the pullback unit. The motor drives the gears and the pulleys, resulting in pulling the catheter probe. (b) The pullback unit is attached to the accessory channel on the endoscope.

Fig. 2.6 A sliding tubular device prevents air leakage while pulling the balloon. The rotary-joint-pullback unit pulls the inner tube sliding along a stationary outer tube. A three-way valve is used for connection to the air pump and regulator.

Fig. 2.7 The beam profile of the first generation probe measured at the focus in 2-D (a) and 3-D (b) views. The bar in (a) indicates 100 µm. The horizontal and vertical FWHM's are 22 µm and 31 µm, respectively. The side-lobes of the spot indicates that the beam is significantly degraded and no long Gaussian. (c) An EOCT image obtained with the first generation fiber-optic probe from porcine esophagus ex vivo.

Fig. 2.8. (a) & (c) Zemax simulations of the beam at the designed working distance without and with astigmatism correction, respectively. The designed spot size is 30 mm. (b) & (d) Beam shapes without and with astigmatism correction, respectively, as measured by a beam analyzer. The measured spot size is 33 mm in both horizontal and vertical directions. (Bars: 100 µm)
Fig. 2.9 One frame excerpted from a video recording demonstrating the insertion of the rotating fiber probe and inflation of the balloon in a porcine esophagus in vivo. The diameters of the probe and the inflated balloon are 2.4 mm and 18 mm, respectively.

Fig. 2.10 (a) A representative cross-sectional image of a porcine esophagus with the double-balloon imaging scheme. The layered structure that can be observed includes the squamous epithelium (SE), lamina propria (LP), muscularis mucosa (MM), submucosa (SM) and muscularis propria (MP). (b) One frame excerpted from a movie of 17 mm long section of porcine esophagus obtained with the double-balloon imaging scheme. (c) A representative cross-sectional image of the single-balloon imaging scheme. (d) 3D reconstruction of 10 mm long segment with single-balloon imaging scheme.

Fig. 2.11 (a) A 3-D reconstruction of Fig. 2.10(b) with true aspect ratio. (b) An enlarged view of the layered structure in a small region of interest. (c) A longitudinal view of the box in (a). (SE: squamous epithelium; LP: lamina propria, MM: muscularis mucosa, SM: submucosa; MP: muscularis propria.)

Fig. 3.1 (a) The helical imaging scheme of the side-view probe. (b) The motion artifacts are caused by the relative motion between the tissue and the sheath during two successive images. (Rad: radial; Rot: rotational; L: longitudinal) (c) The radial-time plane of 3D reconstruction shows the radial displacement. (d) The en face reconstruction by taking pixels at 300 μm deep of (c). (Bar: 1 mm)

Fig. 3.2 A representative raw image before the polar coordinate transformation. The vertical and horizontal directions in the raw image are radial and rotational in the esophagus, respectively. EP: epithelium; LP: lamina propria; MM: muscularis mucosa; SM: submucosa; MP: muscularis propria. Bars: 1 mm

Fig. 3.3 A tubular phantom with 8 longitudinal slots evenly distributed on the circumference of the wall.

Fig. 3.4 Flow chart of the motion artifact correction algorithm

Fig. 3.5 The en face views of the tubular phantom (in section 2.1) obtained without (a) and with (b) the mechanical motion artifact suppression. (Rot: rotational)

Fig. 3.6 Evaluation of automatic registration using 2-D sequences recorded in 3 swine. The mean normalized standard deviation (MNSD) evaluated how the pixels varied between subsequent frames in the sequence. MNSD of the
registered sequences was calculated as a function of the number of ROIs used for LBM. The MNSD scale begins at 0.28, which was the reference baseline MNSD attributable to image noise only, with no sample motion......................... 82

**Fig. 3.7** (a) The *en face* view of the muscularis mucosa layer in the excised esophagus from a 2-D time sequence of images. (b) The *en face* view of the in vivo 2-D sequence where physiological motion artifact is prominent. (c) The registered *en face* view of (b). (d) A radial-longitudinal view from the original 3-D sequence *in vivo*. (e) The registered view of (d). The arrow indicates the position of (g). (f) The *en face* view from the original 3-D sequence. Arrows indicate the position of (d). (g) The registered view of (f). Arrows indicate the position of (e). (h) The enlarged view of the square in (f). (i) The enlarged view of the square in (g). (Rad: radial; L: longitudinal; Rot: rotational. Bars: 1mm)................................. 85

**Fig. 3.8** (a) The *en face* view of the Barrett's esophagus with high grade dysplasia before motion artifact correction. The regions numbered 1, 2 and 3 are regions where Barrett's glands significantly distorted by motion artifacts. (b) The *en face* view after motion artifact correction. The oval shape of gland are mostly recovered in regions of corresponding numbers. (c) An cross-sectional images of the Barrett's esophagus. The balloon (on top of the tissue) was not large enough to expand the esophagus, leaving a gap between the tissue and the balloon. The gap is the reason why there is a dark region in the *en face* view in (a). (Rad: radial; Rot: rotational; L: longitudinal. Bars:1mm)................................. 89

**Fig. 3.9** (a) A 3-D reconstruction of the double-balloon sequence. Each image was then transform into the circular anatomic view without aligning the surface to a circle. The reconstruction was sectioned along the longitudinal direction to expose the mucosal wall. (b) A 3-D reconstruction of the single-balloon sequence............................................................................................................ 92

**Fig. 3.10** (a) *En face* view of sequence acquired by the double-balloon imaging scheme. (b) *En face* view after the registration algorithm applied. The algorithm was not able to correct the motion artifact sufficiently. (L: longitudinal. Rot: rotational.)................................................................. 93

**Fig. 4.1** (a) The catheter probe without a balloon fails to always focus the beam into the tissue. (b) The balloon-based catheter probe consistently maintains the focus on the mucosa. Dash circles indicate the working distance of the probe............................................................................................................... 94

**Fig. 4.2** Illustration of the backscattering profile. The tissue surface was considered the zero depth. The image intensity in the tissue region ($0 < z < z_i$)
decreased linearly, and became flat noise in the background region. The penetration depth was defined to be $z_i$.

**Fig. 4.3** (a) The region of 3 by 3 pixels used to calculate the texture parameters. $g_i$ and $g_i'$ are the intensities of the center-symmetric pair. (b) An example when each center-symmetric pair has different values. SAC equals -1 in this case. (c) & (d) Examples when each center-symmetric pair has identical values. SAC equals 1 in both cases.

**Fig. 4.4** (a) After the trajectory was generated, it was decomposed into radial (in black) and rotational (in red) components (first 100 images) (b) The amplitude spectra of (a). The frequency components at 0.22, 0.44 and 0.68 Hz were related to respiration. Those at 1.32, 2.62, 3.94 Hz were related to hearting beating. (c) The motion magnitudes at six pressure levels. There were no significant difference in mean magnitudes at six pressure levels ($p>0.05$) except the rotational component of respiration.

**Fig. 4.5** (a) Esophageal mucosal wall imaged with the single balloon. (b) Mucosal wall imaged with the double balloon. The layered structure and blood vessels in the double-balloon images are readily identified.

**Fig. 4.6** (a) The mean backscattering profile is plotted in blue. The corresponding images from which the profile was generated is shown at the bottom. The fitting line segments is plotted in red. (b) All penetration depths calculated from seven swine imaged with both single-balloon and double-balloon designs. Two-way ANOVA showed that the penetration depth has no significant difference ($p=0.03$).

**Fig. 4.7** Six texture parameters calculated from seven swine imaged by two balloon schemes. Two-way ANOVA showed that all parameters exhibit significant difference between the two balloon designs ($p<0.001$) except WVAR ($p=0.59$).

**Fig. 4.8** (a) The divot imaged with the single-balloon scheme. The epithelium is not visible at the divot due to the lost of tissue. (b) The divot imaged with the double-balloon scheme. (c) 3-D reconstruction of the divot imaged with the single-balloon scheme. The strong backscattering in the *en face* view is probably due to the blood from the muscularis mucosa at the divot.

**Fig. 5.1** (a) Normal esophageal image from patient 10. The typical normal esophagus has a thick and homogenous layer of epithelium with consistent thickness. The layered architecture can be clearly visualized, especially the boundary between epithelium and lamina propria. (b) Non-dysplastic Barrett's
esophageal image. the epithelium becomes thinner. The surface is no longer smooth but exhibits the finger-like topology highly assembling intestinal metaplasia. The layered architecture is less obvious but visible. (SE: squamous epithelium; LP: lamina propria, MM: muscularis mucosa, SM: submucosa; MP: muscularis propria. Bars: 1mm in two directions. Rad: radial. Rot: rotational.)

Fig. 5.2 (a) A cross-sectional image of human BE (possibly HGD), where the layered structure cannot be identified. (b) A cross-sectional image of normal esophagus, where the epithelium (EP), lamina propria (LP) and muscularis mucosa (MM) are clearly identified. (c) An enlarged view of the dash square in (a). (d) An enlarged view of the dash square in (d). (Bars: 1 mm in two directions.)

Fig. 5.3 (a) A cross-sectional image of the normal esophagus with a smooth surface. The balloon was not in fully contact with the tissue. (b) A cross-sectional image of BE with finger-like pattern on the surface. (c) 3-D reconstruction of (b) suggests that the finger-like pattern may be a cross-sectional view of the longitudinal folding (indicated by arrows). (Bars: 1mm in two directions. L: longitudinal. Rot: rotational.)

Fig. 5.4 (a) An en face view of Fig. 5.3(c) showing the glandular structure. (b) The corresponding cross-sectional view of the red arrows in (a). Yellow arrows indicate the corresponding Barret's glands found in the two views. (Bars: 1mm to two directions. L: longitudinal. Rot: rotational.)
Acknowledgements

Here I put a final touch on the work towards my PhD degree. I feel very fortunate to have worked on a dissertation for which I have both the ability and enthusiasm. I would like to acknowledge all the people who have motivated, mentored, collaborated with and supported me through all these years.

To my parents who always want me to stay as close to them as possible, but still encourage me to find my own future in a foreign country.

To my fiancée who has been there for me during the most difficult time, and been the happiest person for my success.

To my advisor Dr. Andrew Rollins for teaching me to be a creative, cautious, and humble scientist.

To my fellow lab mates for inspiring discussions and plenty of laughter.

To all my collaborators for the team work.

To all the friends I’ve made while at Case, thanks for the great times.
List of Abbreviations

ANOVA – analysis of variance
BE – Barrett’s esophagus
BVAR – Between-pair variance
CSAC – center-symmetric auto-correlation
EAC – esophageal adenocarcinoma
EOCT – endoscopic optical coherence tomography
EUS – endoscopic ultrasound
FD-OCT – Fourier-domain optical coherence tomography
FORJ – fiber-optic rotary joint
FWHM – full width at half maximum
GEJ – gastro-esophageal junction
GERD – gastroesophageal reflux disease
GRIN - gradient-index
GI – gastrointestinal
HGD – high-grade dysplasia
IRB – Institutional Review Board
IVUS – intravascular ultrasound
L – longitudinal
LBM – local block matching
LP – lamina propria
MM – muscularis mucosa
MNSD – mean normalized standard deviation
MP – muscularis propria
NBI – narrow-band imaging
NURD – non-uniform rotational distortion
OCT – optical coherence tomography
Rad – radial
RFA – radio-frequency ablation
ROI – region of interest
Rot – rotational
SAC – Normalized SCOV
SCOV – Gray level texture covariance
SD-OCT – spectral-domain optical coherence tomography
SE – squamous epithelium
SIM – specialized intestinal metaplasia
SM – submucosa
SNR – signal-to-noise ratio
SVR – Variance ratio
SS-OCT – swept-source optical coherence tomography
TD-OCT – time-domain optical coherence tomography
UHCMC – University Hospitals Case Medical Center
VAR – Local variance
WVVAR – Within-pair variance
Barrett's esophagus (BE) surveillance remains challenging, because even histopathology, the gold standard, is subject to sampling error. The esophageal mucosal area involved in BE can be 20 cm$^2$ or more. 3-D volumetric Imaging technology with high diagnostic accuracy may potentially guide and assist the standard histopathology, eliminate the sampling error and improve surveillance efficiency. It has been shown that endoscopic optical coherence tomography (EOCT) of a small mucosal area can obtain interpretable images of gastrointestinal mucosal microstructure, differentiate mucosal types and detect dysplasia in Barrett's esophagus. The realization of 3-D EOCT allows for further exploring the potential to fulfill the unmet need of comprehensive surveillance. The dissertation presents the step-by-step work from initially building the 3-D EOCT system for esophageal imaging, to eventually conducting the clinical trial. First, a system based on a spectral-domain OCT configuration will be described. The sample arm with the rotary-joint-pullback unit, double-balloon-based catheter
and miniature fiber-optic probe is the main hardware innovation allowing for 3-D imaging. Second, an automated motion artifact correction algorithm will be described. The algorithm successfully reveals the otherwise distorted microstructure in the esophageal mucosa, such as microvasculature network and the layered structure. The feasibility of 3-D imaging and motion artifact correction algorithm will be demonstrated in swine in vivo. Third, the balloon designs will be discussed in terms of safety and diagnostic efficiency for clinical trial. It will be shown that a low pressure level is sufficient for motion artifact suppression, and therefore reduce the risk of perforation. Images appearance is significantly influenced by balloon pressure/contact, which establishes the need to image the mucosa with the double-balloon design. Finally, the first cases in clinical trial of BE patients will be reported. The feasibility of 3-D EOCT imaging system are demonstrated. Images features such as layered structure, surface morphology and glandular structure are observed in BE patients. The 3-D EOCT system provides a platform allowing for comprehensive imaging in high quality, which can potential answer critical questions about what role OCT can play in dysplasia diagnosis during BE screening/surveillance.
Chapter 1. Background and significance

1.1 Barrett’s esophagus (BE)

BE and associated esophageal adenocarcinoma (EAC) have emerged as a major health care problem. The death rate of esophageal cancer ranked the seventh highest among all cancers in male since 2005 [1]. The estimated deaths for both genders in 2011 are 14,500, roughly half of whom have EAC. In recent decades, the incidence of EAC has increased more rapidly than that of any other malignant neoplasm in the United States [2] and Western Europe [3]. It is now the most common type of esophageal cancer in these countries. The reason(s) for this are not fully identified but the increase appears to be related in large part to a corresponding increase in the incidence of BE, a condition which increases the risk of EAC for at least thirty-fold [4-10].

BE is an acquired condition in the distal esophageal epithelium that has been damaged by gastroesophageal reflux disease (GERD) [11]. It develops through the process of metaplasia [12], in which the normal stratified squamous epithelium is replaced with a specialized columnar epithelium. The diagnosis of BE requires endoscopic examination to recognize a columnar type mucosa, and histopathologic evaluation of biopsy specimen to confirm specialized intestinal metaplasia (SIM) [6]. The mucosal area involved in BE can be 20 cm$^2$ or more in long segments [13-14]. BE is thought to undergo a long-term sequence of transformation from the non-dysplastic columnar Barrett’s epithelium, to BE
indefinite for dysplasia, to low-grade dysplasia, to high-grade dysplasia (HGD), and finally to invasive carcinoma (Fig. 1.1) [15-19]. This makes early diagnosis and early management of EAC a possibility.

Fig. 1.1 The appearance of esophageal mucosa as the progression from BE to EAC. (Source: http://pathology2.jhu.edu/beweb/cancer.cfm)

1.2 Current BE surveillance

Surveillance of BE patients is recommended for the purpose of diagnosing dysplasia or early carcinoma. Guidelines for surveillance of patients for BE suggest that endoscopy be performed in patients with a long-standing history of symptomatic GERD, particularly patients over the age of 50 years [4]. In most cases, BE is discovered during the evaluation of GERD symptoms such as heartburn, regurgitation, and dysphagia [6, 11, 20-21]. Once BE is confirmed histopathologically, patients are recommended to undergo periodic surveillance to diagnose dysplasia. Surveillance is practiced by a majority of endoscopists in the United States [22-23]. The rationale of surveillance is the poor 5 year survival
of EAC of 13% [24], and assumes that surveillance can reliably diagnosis dysplasia in BE [11]. There is considerable debate regarding the efficacy of surveillance endoscopy because of the lack of randomized trials to support its value. However, several studies do suggest the survival advantage of endoscopic surveillance [25-31].

Currently, BE surveillance relies on endoscopic visualization and biopsy. During the endoscopy, the patient is sedated and a videoendoscope is inserted transorally into the esophagus. The gastroenterologist then inspects the esophagus, stomach, and gastro-esophageal junction (GEJ) to detect dysplasia. Biopsy is followed according to the standard, so-called “Seattle” protocol (four-quadrant biopsies obtained with a large, jumbo forceps at 2 cm intervals within the Barrett’s mucosa), with or without the addition of vital staining techniques. The biopsy specimens are then sent for histopathologic evaluation.

The surveillance programs are less than ideal. Histopathology is currently the gold standard to diagnose dysplasia, but sampling errors are frequent and dysplasia as well as early stage cancers are often missed [32-33]. This random sampling error, as well as low sampling yield, pathology-associated costs, and delay in diagnosis among other limitations, have generated increasing interests in the development of optical techniques as a complementary means of enhancing the diagnosis of dysplasia [34-35]. Clearly, a surveillance method that comprehensively surveys the entire Barrett’s segment while reliably diagnosing dysplasia would reduce or eliminate sampling error, improve surveillance efficacy and thereby improve patient survival.
1.3 Optical coherence tomography (OCT)

OCT is an emerging, non-invasive imaging technique that provides microscopic cross-sectional images of biological samples [36]. It is often described analogous to ultrasound, because OCT detects the photons back-scattered by the sample as ultrasound detects the echoes. Unlike other optical imaging technologies, OCT has an axial (in-depth) resolution determined by the bandwidth of the light source. Depth discrimination in OCT is decoupled from the optics and accomplished by the coherence gate of the broadband light. Low coherence and interference allow OCT to have multiple unique features for a promising medical imaging modality. For example, it fills a nice niche that helps to bridge the gap between confocal microscopy (high resolution but low imaging depth) and ultrasound (low resolution but increased depth). Its imaging speed (up to 5 million in-depth scans per second [37]) and sensitivity (~110 dB) surpasses other contemporary optical imaging modalities. With fiber based implementations, miniature probing optics and Doppler flow measurement, OCT has recently made a significant impact in clinical ophthalmology [38] and cardiology [39]. Its potentials in gastroenterology [40], dermatology [41], dentistry [42], pulmonary [43], and gynecology [44] continue to be extensively investigated.

In this section, conventional time-domain OCT (TD-OCT) and its mathematical treatment will be briefly presented, followed by an introduction to spectral-domain OCT (SD-OCT), the second generation technology. Important imaging
parameters will be formulated and discussed for the introduction to endoscopic OCT (EOCT) in section 1.4

1.3.1 TD-OCT

OCT discriminates photons reflected from different depths in the sample, whose transit time are too fast for direct detection. In order to do that, a TD-OCT system employs a broadband light source and an interferometry such as the Michelson-Morley interferometry configuration depicted in Fig. 1.2. A low-coherent beam of light from a broadband light source is split into one arm with tissue sample (sample arm), and the other with a scanning optical delay line (reference arm). Light reflected from both arms is recombined and the interference is detected. Depth is discriminated, or gated, by low coherence interference when the optical path lengths of the sample and reference arms are matched to within the coherence length of the light source. As shown in Fig. 1.2, light reflected from surface “a” in the tissue propagates the same optical path length as light reflected from the mirror position “a’”, resulting in the interference signal “aa’”. When the difference in the optical path lengths of a and a’ exceeds the coherence length, no interference is generated. Therefore, a light source with a short coherence length is used to illuminate the sample in order to provide high axial (in-depth) resolution. A depth profile of back-scattering is called an A-scan. Fig. 1.3(a) shows a back-scattering profile from swine esophagus in the in-depth or radial direction. The term “time-domain” describes the fact that the A-scan is acquired as a function of time. A two dimensional image (B-scan) is built up by
stacking many A-scans while scanning the probe beam laterally across the sample (Fig. 1.3(b)). In the case of esophagus, the lateral direction is actually rotational. The lateral resolution of an OCT system is determined by the illumination and detection optics, which are usually designed to match the lateral resolution to the axial resolution. A three dimensional image is reconstructed by stacking the B-scans along the third dimension (the longitudinal direction in Fig. 1.3(c)). The A-scan rate is determined by the repetition rate of the scanning optical delay line, which is usually below 10k Hz [45-47]. The repetition rate is the imaging speed limiting factor for TD-OCT. 3D imaging of large dimensions, such as the esophagus, is time-consuming and impractical for in vivo clinical applications if TD-OCT is used.
Fig. 1.2 TD-OCT with a Michelson-Morley interferometry configuration. Light interferes only when the optical path lengths of the two arms are matched to within the coherence length of the broadband light source. For example, light reflected from layer a in the tissue only interferes with the mirror in position a' and generates interferometric signal aa'. Depth is therefore discriminated.
Fig. 1.3 (a) A back-scattering profile or A-scan plotted as a function of in-depth position. (b) By stacking the A-scans rotationally, a cross-sectional esophageal OCT images is constructed. The dash arrow indicates the position of (a). (c) A 3-D OCT volume can be reconstructed by stacking the cross-sectional images along the longitudinal direction.
1.3.2 Mathematical treatment

OCT and its major technical specifications will be described in math. The math is adapted from Izatt et al [48] and Fercher et al [49]. Light will be treated as scalar wave that is ergodic and stationary [50]. Polarization and field quantization will be ignored in this treatment.

A monochromatic plane wave from the light source with constant electric field amplitude $2A_i$ is incident on the interferometer. The electric field can be expressed in scalar form as $E_i = 2A_i \cos(k_0 z - \omega_0 t)$. When the wave is evenly split and recombined at the detector, the superposition field becomes

$$E_d = E_s + E_r$$
$$= \sqrt{R_s}A_i \cos(2k_0 l_s - \omega_0 t) + A_i \cos(2k_0 l_r - \omega_0 t)$$

where $d$, $s$ and $r$ indicate the detector, sample and reference. $R_s$ is the sample power reflectivity. $l_s$ and $l_r$ are the one-way optical path length of the sample and reference arms. The optical path length from the interferometer to the detector is ignored. The detector current is proportional to the time-averaged intensity of the superposition field.

$$i_d \propto \langle |E_d|^2 \rangle \propto \frac{1}{2} (1 + R_s) A_i^2 + \sqrt{R_s} A_i^2 \cos(2k_0 (l_r - l_s))$$

where $2\Delta l = 2(l_r - l_s)$ is the round trip path length mismatch and $\langle |E_d|^2 \rangle$ the time-averaged intensity. The first term of the summation on the right is a constant DC signal, whereas the second term represents the interference. The DC component is always eliminated either by balance detection or post-acquisition processing. The interferometric term ($i_d$) carries the depth information, given by

$$i_d(\Delta l) \propto \sqrt{R_s} A_i^2 \cos(2k_0 \Delta l) \quad (1.1)$$
Coherence is then added to the model. A broad-band light source with a Gaussian power spectral density is usually assumed [51] as

\[ S_{ii}(k) \propto \frac{2\sqrt{\ln 2}}{\Delta k \sqrt{\pi}} \exp \left\{ -\left( \frac{k - k_0}{\frac{\Delta k}{2\sqrt{\ln 2}}} \right)^2 \right\} \]

where \( k_0 \) is the center wavenumber and \( \Delta k \) the full width at half maximum (FWHM) spectral bandwidth expressed in wavenumber. Since the bandwidth of the light source is commonly expressed in wavelength, \( \Delta k \) is given by

\[ \Delta k = \frac{2\pi\Delta\lambda}{\lambda_0^2} \]

where \( \lambda_0 \) is the source center wavelength, and \( \Delta\lambda \) the FWHM bandwidth in wavelength. The interferometric current in this polychromatic case is the integral of equation (1.1) over the source spectrum, given by

\[ \tilde{i}_d(\Delta l) \propto \sqrt{R_s} A_i^2 \int_{-\infty}^{+\infty} S_{ii}(k) \cos(2k\Delta l) \, dk \]

\[ \propto \sqrt{R_s} A_i^2 \exp \left\{ -\left( \frac{\Delta l}{\frac{l_c}{2\sqrt{\ln 2}}} \right)^2 \right\} \cos(2k_0\Delta l) \quad (1.2) \]

\[ l_c = \frac{4(\ln 2)}{\Delta k} = \frac{2(\ln 2)\lambda_0^2}{\pi\Delta\lambda} \quad (1.3) \]

It can be seen from equation (1.2) that the polychromatic interferometric signal is Gaussian distributed as a function of the path length mismatch \( \Delta l \), and modulated sinusoidally. When there are several layers in the sample, each layer generates a modulated Gaussian signal as the scanning delay line sweeps.
(shown in Fig. 1.4). Depth is therefore discriminated. The sinusoid is usually filtered and the envelops are displayed usually in logarithmic scale.

Fig. 1.4 Illustration of the interferometric signal generated in Fig. 1.2. Position $Z_a$, $Z_b$, and $Z_c$ are the locations of zero path-length mis-match of the layers a, b and c in the sample. The red curves show the Gaussian profile of the signal, and $l_c$ is the FWHM, also known as the in-depth or axial resolution (described below)

The important parameters of an OCT system include the resolution, wavelength, penetration depth, sensitivity, imaging speed, etc. These parameters are inter-correlated. The trade-off limits the performance of OCT.

**Resolution**

The spatial resolving power of OCT is usually the most important parameter to determine whether the sample can be imaged. The axial resolution is usually defined as the FWHM of the interferometric envelope as shown in equation (1.3). A typical axial resolution for current light sources ranges from 5 μm to 10 μm. It should be noted that the resolution of an imaging system is typically defined as
the smallest distance of two objects that can be distinguished. However, no measurement technique is adopted to estimate the axial resolution in such a manner in OCT.

The lateral resolution is determined by the optical system. Assuming the probing beam to be Gaussian, the lateral resolution is defined as the FWHM diameter $\Delta d_{FWHM}$ of the probing beam amplitude distribution, given by

$$\Delta d_{FWHM} = 2\sqrt{\ln 2} w_0 = 2\sqrt{\ln 2} \frac{\bar{\lambda}^2}{\pi \sin^{-1} NA} \approx 2\sqrt{\ln 2} \frac{\bar{\lambda}^2}{\pi NA}$$

where $w_0$ is the width of the Gaussian beam waist, $\bar{\lambda}$ the mean wavelength, $NA$ the numerical aperture of the objective lens. $NA$ is usually very small, so that $\sin^{-1} NA \approx NA$. The lateral resolution of current OCT system ranges from 5 $\mu$m to 30 $\mu$m depending on the applications.

**Wavelength**

Light absorption and scattering determine the choice of wavelength of the light source, and limit the penetration depth of OCT imaging. Theoretical treatments [52-53] and investigations of tissue optical property [54-55] have suggested that the optimal penetration depth occurs near 1,300 nm and near 1,650 nm. The former wavelength has commonly been used because of the commercial availability.

**Penetration depth**

The penetration depth describes the depth from sample surface within which OCT is able to generate imaging contrast. The depth is roughly within the single
scattering regime. It has been shown that multiple scattered photons do not carry the image contrast [49]. The penetration depth may also be affected if an objective lens of high NA is used. In this case, the beam converges and diverges rapidly within the single scattering regime, where only the beam waist has the best lateral resolving power. The penetration depth of OCT is limited to 1-3 mm in dense tissue at 1,300 nm center wavelength,

**Signal-to-noise ratio (SNR)**

SNR determines the quality of the image. In a practical imaging situation, the dominating noise source is shot noise, arising from the random arrival of photons onto the photodetector. The shot-noise-limited SNR is given by [56]

\[
SNR_{TDOCT} = \frac{\rho R_s S_{TDOCT}}{2eB_{TDOCT}} \quad (1.4)
\]

where \(\rho\) is the photodetector responsivity, \(R_s\) the sample reflectivity, \(S_{TDOCT}\) the source power, \(e\) the electronic charge, \(B_{TDOCT}\) the noise equivalent bandwidth of the system [57], \(\Delta k_{FWHM}\) the FWHM bandwidth, \(z_{max}\) the imaging range, and \(\Delta t\) the A-scan acquisition time. SNR can be higher than 110dB for images of good quality.

**Imaging speed**

Imaging speed determines the temporal resolving power of OCT imaging. It is usually measured by the number of A-scans acquired per second. The maximal speed for TD-OCT is limited by the mechanical repetition rate of the reference
arm, which is about 10k A-scan/second. Because SNR is proportional to the A-scan acquisition time, the practical imaging speed is even lower.

### 1.3.3 Spectral-domain OCT

SD-OCT is one of the two equivalent Fourier-domain OCT (FD-OCT) technologies, and utilizes a spectrometer [58-61]. The other FD-OCT technology utilizes swept light source. It is therefore given the term Swept-source OCT (SS-OCT, or Optical Fourier-domain Imaging which is preferred by some researchers) [61]. SS-OCT was not utilized in this work and will not be discussed. The main advantage of FD-OCT is that the detection sensitivity (measured by SNR) is more than 100 folds higher than TD-OCT, thus allowing for rapid scanning.

A typical SD-OCT diagram is shown in Fig. 1.5. Compared with the time-domain setting in Fig. 1.2, the technical improvement is mainly utilizing the spectrometer as the detector. The interference spectrum is therefore recorded. There is no mechanical scanning part in the reference arm, where the mirror is stationary. Assuming the layers in the sample have path-length mismatches $\Delta l_a$, $\Delta l_b$, and $\Delta l_c$ with respect to the reference arm, each layer generates an interferometric signal given by equation (1.2)

$$i_d(\Delta l_m) \sim \sqrt{R_s(\Delta l_m) \exp \left( - \left( \frac{\Delta l_m}{l_c} \right)^2 \right) \cos(2\Delta l_m k)}, \quad (1.5)$$

$m = a, b, c$

where $R_s(\Delta l_m)$ is the reflectivity of the layer $m$, and wavenumber $k$ is now a variable because of the spectrometer. The term $A_i^2$ is omitted. When the light is
diffracted by the grating, the interference spectrum is a sinusoid recorded as a function of wavenumber $k$, called fringes. The spectrometer performs a Fourier transform to the interferometric signal from the spatial space ($\Delta l_m$) to the wavenumber space ($k$), with only the real part of the transform recorded ($e^{-i2\Delta l_m k} = \cos(2\Delta l_m k) - j \sin(2\Delta l_m k)$).

Fig. 1.5 Diagram of SD-OCT system. A spectrometer is utilized to record the interferometric fringes. The fringes then undergoes several steps, including DC subtraction, re-sampling, dispersion compensation, inverse Fourier transform, and logarithmic compression to produce an image. The mirror in the reference arm is stationary. The sample is assumed to have 3 layers. The dash arrow indicates the position of zero path-length mismatch.

Several steps are taken to produce an image, including DC subtraction, re-sampling, dispersion compensation, inverse Fourier transform from $k$ to $\Delta l_m$, and logarithmic compression. DC subtraction is very important because the spectrometer records the interferometric spectrum as well as the source spectrum. The source spectrum does not contribute to the image formation. Re-
sampling is required because the interferometric spectrum is distributed linearly in wavelength. This spectrum has to be converted to be linear in wavenumber, so that the inverse Fourier transform becomes a function in space as in equation (1.5). Dispersion is caused by the different refractive indices when light propagates in different material. It degrades the axial resolution of the image. The compensation is realized by expanding $k$ around the center wavenumber with Taylor series to adjust the phase of the OCT data. The second and third order coefficients are determined by optimizing the axial resolution. Finally, inverse Fourier transform is applied, and the image is log compressed and displayed.

Compared to TD-OCT, SD-OCT has the advantages in both SNR and imaging speed, which eventually allows for clinical application in real time. There are also three important limitations for this second generation OCT technology, which are complex conjugate ambiguity, sensitivity falloff and limited imaging range. Resolution and wavelength remain the same in SD-OCT.

**Signal-to-noise ratio (SNR)**

The SNR advantage of SD-OCT lies in the fact that the light is now distributed among the $N$ detectors in the spectrometer. The overall signal remains the same power because the signals recorded by the individual photo-detectors are coherent. On the other hand, the noise among the detector is independent. Assuming the source power $S_{SD\text{OCT}}$ and the noise equivalent bandwidth $B_{SD\text{OCT}}$ are the same as those in TD-OCT, the shot-noise-limited SNR in this case is given by [56]
There is a factor of \(N/2\) improvement. The factor \(N\) attributes to the fact that the noise is distributed among the \(N\) detectors. The factor 2 will be explained below.

**Imaging speed**

The SNR improvement of SD-OCT is not obvious compared to TD-OCT because the TD-OCT is almost always optimized for high SNR by acquiring images slowly. Alternatively, we can acquire A-scans much faster with SD-OCT while maintaining equivalent SNR. Since

\[
\beta_{SDOCT} \approx \frac{2 \Delta k_{FWHM} z_{max}}{(\pi \Delta t)}
\]

the A-scan acquisition time \(\Delta t\) can be reduced by \(N/2\) times theoretically. A typical SD-OCT system acquires more than 40k A-scan per second.

**Complex conjugate ambiguity**

This ambiguity is caused by the fact that only the real part of the spectrum is recorded. The cosine term in equation (1.5) is an even function of \(\Delta l\). Therefore, the inverse Fourier transform does not distinguish between \(\Delta l\) and \(-\Delta l\). This can be mitigated by moving the sample to one side of the zero path-length mismatch. Various approaches have been proposed to remove this ambiguity [62-72], but none of them has been examined for real time application by now.

**Sensitivity falloff**

Sensitivity falloff is the decrease in SNR as a reflector in the sample arm moved away from the zero path-length mismatch. It arises from the finite spacing of photodetector compared to the spot size of objective lens before the
photodetector. The photodetector, which is supposed to capture the entire spot, truncate it instead. It is equivalent convoluting a signal with a rectangular window. There is also a small amount of crosstalk between neighboring wavelengths which cannot be removed. This phenomenon can be modeled by $\delta k$, the spectral resolution of the spectrometer [60, 73]. The 6dB $SNR$ falloff point is defined as the position of the reflector where $SNR$ is 6dB lower than the zero path-length mismatch, given by

$$6dB \text{ falloff range} = \frac{2 \ln 2}{\delta k}$$

Recently, a linear-in-wavenumber spectrometer design has been proposed to significantly improve the falloff in SD-OCT [74-75].

**Imaging range**

Imaging range is the maximal $\Delta l$ in equation (1.5) which can still be resolved by the spectrometer. The detector array can be understood as sampling. $\Delta l$ determines the spatial frequency of the fringes. The spatial frequency should be less than half of the sampling frequency to avoid aliasing. The imaging range is given by

$$Imaging \text{ range} = \frac{N\pi}{2\Delta k}$$

where $N$ is the number of detector in the array and $\Delta k$ the wavenumber range that the spectrometer covers. The imaging range in a typical SD-OCT system is larger than the penetration depth. Therefore it is usually not a limiting factor for OCT imaging, unless the position of the sample surface changes dramatically.
1.4 Overview of EOCT for BE

The potential for OCT in BE diagnosis is apparent. Cross-sectional, or better yet, 3-D images in high resolution with detailed microstructure may provide "optical biopsy", which may guide and assist the current standard histopathology. In order to fulfill the potential of BE diagnosis, the technology should satisfy several criteria, 1) high accuracy and biopsy concordance, 2) easy interpretation with low interobserver variability, 3) large field-of-view imaging, 4) high-speed image acquisition and 5) cost effectiveness [76]. The first four aspects have undergone extensive investigations since the first esophageal imaging published in 1997 [77-78], while the last criterion has not been thoroughly investigated by far.

The first generation of EOCT systems employed the time-domain configuration together with fiber-optic catheter probes which could be inserted through standard endoscopes (Fig. 1.6(a)). The working distance of the OCT probes were about 1mm, which significantly limited the field of view. Only a cross-sectional image of a small esophageal region was acquired, similar to the point biopsy (Fig. 1.6(b)). The imaging speed was too slow to acquire volumetric images of the esophagus in vivo.
Fig. 1.6 (a) Fiber-optic catheter probe that could be inserted through standard endoscopes. (b) The point scanning scheme of the short working distance probe. An EMR cap was utilized to stabilize the endoscope. The probe was inserted to obtained a cross-sectional image of a small region.

In order to answer the question on diagnostic accuracy using TD-OCT, studies were first conducted to show that EOCT can obtain interpretable images in normal, premalignant and malignant esophagus [77-87]. The success attributed to both the high resolution, and the distinct back-scattering property of different tissues in the esophageal mucosa. The penetration depth was sufficient compared to the thickness of mucosa, where BE might occur. Several clinical trials were then carried out to investigate the feasibility of OCT to detect SIM for the purpose of BE screening [88-89], diagnose dysplasia for BE surveillance [90-93] and identify subsquamous Barrett's epithelium [94] or glandular structure [38]. Among these targets, dysplasia diagnosis is very important aspect because the current BE management depends on the degree of dysplasia. The diagnostic criteria for detecting abnormality in the epithelium were mainly based on the
glandular or layered architecture, surface architecture, reflectivity of epithelium and penetration depth. A few studies have shown that OCT may potentially identify SIM, differentiate non-dysplastic from dysplastic tissue (Fig. 1.7), and differentiate high-grade dysplasia/carcinoma from low-grade dysplasia. The results are summarized in Chart 1.1.

Table 1.1 Diagnostic performance of EOCT for esophageal disorder

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIM [89]</td>
<td>81%</td>
<td>57% - 66%</td>
</tr>
<tr>
<td>Dysplasia [90]</td>
<td>68%</td>
<td>82%</td>
</tr>
<tr>
<td>HGD/carcinoma [91]</td>
<td>83%</td>
<td>75%</td>
</tr>
</tbody>
</table>

These clinical trials concluded that OCT could identify SIM with an accuracy similar to that of endoscopy, and be used to target biopsies to areas of BE with a higher probability for the presence of dysplasia. One study reported a significant interobserver variability when diagnosing dysplasia [90]. Computer-aided diagnosis algorithms were then developed to provide a quantitative measurement [92-93]. The sensitivity of 78% - 82% and the specificity of 74% - 90% could potentially reduce diagnostic variability. Although these successes shed light on the feasibility of clinical application, TD-OCT technology did not have a significant advantage over the standard histopathology.
Fig. 1.7. Endoscopic TD-OCT images from the study in dysplasia diagnosis [90]. (a) Low grade dysplasia (area marked by sunburst). (b) HGD (area marked by sunburst). Two criteria for diagnosing dysplasia were selected: reduced light scattering and loss of tissue architecture. Solid lines are each 2 mm in length.

We and other have recently developed the second generation of EOCT for volumetric esophageal imaging [40, 95-99]. Researchers have realized that such 3-D EOCT may eliminate sampling error in dysplasia diagnosis, given the capability of comprehensive visualization. The new technology benefits from the ultra-high imaging speed of FD-OCT technology. SD-OCT and SS-OCT have been demonstrated the equivalent performance of imaging centimeters of esophageal segments in minutes. Miniature catheter probe with more than 10 mm working distance is another key technology which allows for scanning a lumen with a large diameter. Various probing optical designs have been reported [97-98, 100]. Balloon is an important component of the probe. They are also utilized to dilate the otherwise collapsing esophageal wall and center the catheter
probe in the lumen. A double-balloon design has been proposed in order to investigate the influence of balloon pressure/contact on tissue appearance [101].

The feasibility of this second-generation technology has been demonstrated in porcine esophagus in vivo [96, 101]. On-going clinical trials have been carried out by us and other groups [40], which hopefully will determine the clinical significance of EOCT. OCT has also been investigated as a comprehensive imaging tool to detect residual Barrett's epithelium after radio-frequency ablation (RFA) [102]. It may benefit the RFA procedure by reducing the recurrence of BE and the need for secondary ablation, and potentially further increase the complete eradication rate of intestinal metaplasia.

In conclusion, OCT has been demonstrated to detect abnormality in BE and can potentially eliminate sampling error during Barrett's screening/surveillance. Technological improvement and clinical trials are the two main aspects of current research for the ultimate goal to determine the role of OCT in BE diagnosis and management.

1.5 Competing endoscopic imaging modalities

Before detailed presentation of esophageal OCT, it is worthy of briefly reviewing the competing endoscopic imaging modalities in order to better understand the advantage and limitation of OCT. These technologies include endoscopic ultrasound (EUS), chromoendoscopy, narrow-band imaging (NBI), fluorescence endoscopy, Raman spectroscopy and confocal endoscopy.

EUS
EUS is a clinically well-established esophageal imaging modality. By detecting the echoes reflecting from the layer structure, EUS produces images with important anatomical correlation with the esophageal wall \cite{103}. EUS plays a central role in staging carcinoma \cite{104}, detecting lymph node metastasis \cite{105}, and differentiating mucosal versus submucosal tissue \cite{106}. EUS with high frequency can resolve the layers in the mucosa, but does not usually detect dysplasia \cite{107}. Therefore, it is usually not considered a potential tool for precancerous esophageal screening or surveillance.

**Chromoendoscopy**

Chromoendoscopy applies a chromogenic stain (i.e., optical absorbers of specific colors that bind to either normal or abnormal tissue) to the mucosa, and is most useful when combined with high magnification endoscopy \cite{108}. These stains include Lugol's solution, crystal violet \cite{109}, acetic acid \cite{110}, etc. Chromoendoscopy of the upper GI tract does not improve the detectability of dysplasia and "has failed to establish its role in diagnosing Barrett's oesophagus and Barrett's associated neoplasia so far" \cite{111}.

**NBI**

Rather than displaying a white-balanced endoscopic image, NBI adjusts the ratio of red, green and blue channels to enhance the contrast in mucosal morphology and micro-vasculature. The morphology and vascular pattern are altered in the transition from intestinal metaplasia to dysplasia \cite{112}. NBI has very high sensitivity of 94% - 100% and specificity of 74% - 99% in detecting HGD, but cannot different low grade dysplasia from non-dysplasia \cite{113-114}.
Also, it is believed that the morphologic/microvascular changes are not universally present in all cases [114].

Fluorescence endoscopy

Fluorescence endoscopy in GI tract is mainly measuring the fluorescence generated by endogenous tissue molecules (autofluorescence) or exogenous agents. The contrast is due to the difference in the biochemical composition and content of fluorophores among normal, dysplastic, and neoplastic tissues. Autofluorescence endoscopy has been possible to visualize dysplasia that were missed on white light endoscopy [112], but the overall result has not been encouraging. A randomized, crossover study comparing autofluorescence spectroscopy and standard white-light videoendoscopy has showed equivalent sensitivity and a lower specificity with fluorescence imaging [115]. A significant problem is the high level of noisy autofluorescence in the non-dysplastic BE mucosa. Fluorescence agents such as aminolevulinic acid [116] or Qdot [117] have also been investigated. The problems were either high noise background or possible toxicity.

Raman spectroscopy

Raman spectroscopy probes the biochemistry of the tissue by measuring the inelastically scattered light, which reveals the molecular bonds. An ex vivo study in BE biopsy has shown 88% sensitivity and 92% specificity for HGD using pathology as a standard [118]. The main challenge is the low intensity of the Raman signal, which is several orders of magnitude lower than (auto)fluorescence. To the best of the author’s knowledge, study on Raman
spectroscopy in an *in vivo* setting has not been reported yet. The reason may be the lack of suitable fiber optic probes which filter the fluorescence from the Raman signal [112].

Confocal microscopy

Confocal microscopy utilizes a pinhole in a standard microscopic optical setting to reject the out-of-plane photons for a higher resolution. This technology provides higher resolution than magnification endoscopy and allows for depth scanning. Keisslich *et al.* obtained 90% diagnostic accuracy for detection of neoplasia in BE patients [119]. However, faster imaging speed, greater penetration depth and wilder field of view may be required for the purpose of BE screening/surveillance.

This dissertation describes the work on 3D SD-OCT for BE imaging. The rest of the dissertation will be organized as follows. The hardware aspects of the system, especially the design and assembly of the sample arm, will be described in chapter 2. Motion artifact removal during 3-D reconstruction, the software aspect, will follow in chapter 3. Chapter 4 will discuss the choice of balloon design. Clinical trial in patients undergoing BE surveillance will be summarized in chapter 5, followed by the future plan in chapter 6.
Chapter 2. Endoscopic OCT

2.1 Introduction

Recent EOCT work has focused on increasing the image acquisition speed and the field of view. These improvements have enabled comprehensive imaging of segments of the esophagus in a clinically feasible time period (a few minutes). The key technologies enabling this improvement are FD-OCT [58-61] and balloon catheters [40, 95-98, 100].

Previously reported balloon catheters for EOCT have been delivered to the esophagus using a guidewire or an overtube [96-97]. After delivery, the balloon was inflated to center the probe within an expanded lumen and held stationary while the probe was rotated and pulled back in a helical scan. However, the pressure on the mucosa caused by direct balloon-tissue contact may affect the appearance of the tissue in OCT images. It has been previously shown that catheter pressure on colonic mucosa significantly alters the tissue appearance [120]. This raises the question of whether balloon pressure or contact affects OCT image features and dysplasia diagnosis in BE. To answer this question, a balloon-catheter-based EOCT platform must acquire comprehensive images with and without balloon-tissue contact. The platform must generate high quality OCT images rapidly in order to image segments of BE in patients. Also, the balloon catheter should be deployable through the accessory channel of the endoscope for user-friendly clinical procedures, endoscopic guidance of balloon positioning,
correlation of endoscopic and OCT images, and documentation of biopsy and OCT imaging sites [76].

In this chapter, I will describe an EOCT platform for in vivo esophageal imaging that meets the requirements listed above. The system features a unique double-balloon catheter design. By scanning the beam between the two balloons, images without balloon-tissue contact can be obtained. Also the balloon design is capable of the conventional imaging through the balloon, and enables direct comparison of tissue appearance with and without balloon-tissue contact. This system furthermore represents the first demonstration of a balloon EOCT catheter deployable through the endoscope. In addition, this EOCT system is based on a spectral-domain OCT (SD-OCT) engine and shows comparable imaging performance to previously demonstrated esophageal EOCT systems which have all been based on swept light source OCT engines [40, 95-97, 100]. The system was validated by high quality, comprehensive imaging of swine esophagus in vivo, both with and without balloon-tissue contact. In the method section, credits are given to the corresponding personnel for building the components to which I did not have significant contribution.

2.2. Methods

2.2.1 EOCT System Design

The EOCT system diagram is shown in Fig. 2.1(a). It employs a 13.5 mW, broadband SLD light source (centered at 1310 nm) with a FWHM bandwidth of 75 nm (IPSDM13xx, InPhenix, Inc., Livermore, CA). The measured axial
resolution in air is 13 μm. The light is delivered through a circulator then split by a 90/10 fiber coupler. 90% of the light goes to the sample arm, which consists of a rotary-joint-pullback unit and a balloon catheter. 90% of the light collected from the sample arm is directed back through the circulator to the spectrometer for high power efficiency. The spectral interference fringes are detected by a custom-designed, linear-in-wavenumber ($k$) spectrometer, which will be described below. Finally, the detected spectrum is read by a frame grabber (NI PCIe-1427; National Instruments Corporation, Austin, TX) whose line acquisition is synchronized to the rotation of the fiber probe. The image acquisition and display is coordinated by a custom software program written in C++ (developed by Dr. Yinsheng Pan). The entire system was assembled and fit into a portable rack, so that it can be moved the animal or clinical lab easily (Fig. 2.1(b)).
Fig. 2.1 (a) System diagram of the double-balloon-based EOCT system. Within the catheter, the beam can be located between two balloons for imaging without balloon-tissue contact (labeled beam position 1) or in a balloon for imaging with contact (labeled beam position 2). The inset photo (top right) shows the double-balloon catheter (compared with a dime) inserted through the GI endoscope and inflated. (b) The entire system was fit into a portable rack.
2.2.2 Spectrometer

The spectrometer enables one to see fringes on the spectrum and relate those fringes to a particular distance of a reflector in the tissue. Here, a novel linear-in-\(k\) spectrometer was designed (by Dr. Zhilin Hu) to operate with a high-speed InGaAs digital line-scan camera (SU-LDH 1024; SUI Goodrich, Princeton, NJ), adapted from a previous linear-\(k\) spectrometer [74]. The new line scan camera has 1024 pixels, with a pixel width of 25 \(\mu\)m. The spectrometer covers a bandwidth of 115 nm centered at 1310 nm. The imaging range is 4.4 mm. The optical resolution of the spectrometer is 13 \(\mu\)m. The -6dB sensitivity falloff range is 1.7 mm.

2.2.3 Balloon catheter probe

The endoscopic balloon catheter is a key component of the EOCT system and consists of three main components, a fiber-optic probe, a rotary-joint-pullback unit and balloons.

Fiber-optic probe

The fiber-optic probe is used to guide and focus the light onto the esophageal mucosa and to collect the image-bearing backscattered light. The first generation of fiber-optic design (by Dr. Zhilin Hu) used a customized gradient-index (GRIN) lens with a polished cylindrical surface. Fig. 2.2(a) illustrates the design from two perpendicular views to show the cylindrical surface. The beam from the optical fiber is directly focused by the GRIN lens. The cylindrical surface is necessary to
compensate for the astigmatism induced by the outer sheath, which is equivalent to a concave lens. A mirror reflects the beam 80 degree with respect to the optical axis of the lens, enabling side-viewing. There are two manufacturing problems with this design. First, it requires a long GRIN lens so that the surface can be polished. The GRIN lens has to be custom-made. Second, the axis of the cylinder has to accurately intercept the optical axis of the GRIN lens to avoid aberration. Fig. 2.2(b) illustrates a poorly polished surface which is asymmetric with respect to the optical axis. The manufacturing tolerance is not guaranteed.

Fig. 2.2 (a) Two perpendicular views of the fiber-optic probe design of a GRIN lens with a polished cylindrical surface. The cylindrical surface compensates the astigmatism induced by the outer sheath. (b) An illustration when the cylindrical surface is poorly polished and asymmetric with respect to the optical axis.

The idea to improve the optical design is to use as few custom-made components as possible, or at least to make the customized process as simple
as possible. Fig. 2.3(a) shows the mechanical model of the second (current) generation of probe design (SolidWorks, Dassault Systèmes SolidWorks Corp., Concord, MA). A single-mode fiber (SMF28) is fixed in a ferrule followed by a glass spacer to expand the beam before entering a GRIN lens with a pitch of 0.295. A separate cylinder rod lens with a focal length of 12 mm is attached after the GRIN lens to correct the astigmatism. All optical components have polished surfaces angled at 8 degrees to prevent back reflection. The optical components are on-shelf products except the spacer. The length of the spacer determines the beam angle entering the GRIN lens, and therefore the working distance of the lens group. 3.15mm and 2.95mm spacers are chosen to build working distances of 9mm for porcine esophagus, and 12mm for human esophagus, respectively. The lens group and ferrule are fit into a glass tube and then glued into a metal housing. A gold mirror is glued at the tip of the metal housing to deflect the beam by 80 degrees. The overall numerical aperture of the probe is 0.024. The total outer diameter of the fiber probe, including the metal housing, is 1.7 mm. Fig. 2.3 shows the size of an assembled probe without the metal housing.
Fig. 2.3 (a) The mechanical model of the second generation of probe design. A single-mode fiber is fixed in a ferrule followed by a glass spacer to expand the beam before entering a GRIN lens. A separate cylinder rod lens is attached after the GRIN lens to correct the astigmatism. (b) An assembled fiber-optic probe without metal housing compared to a dime.

**Rotary-joint-pullback unit**

The rotary-joint-pullback unit generates a helical scanning pattern enabling volumetric imaging. The rotating unit enables the fiber probe to scan the esophagus rotationally (Fig. 2.4). A miniature fiber-optic rotary joint (FORJ, RFCX-131-28, Princetel, Inc., Pennington, NJ) couples the light from the coupler (Fig. 2.1(a)) to the proximal end of the rotating probe. The stator of the fiber-optic rotary joint is fixed to a jig (designed by Dr. Hui Wang), while the rotor is supported and centered by a ball bearing. The rotor is driven by a motor (AKM11B, Kollmorgen, Radford, VA) via a belt.
Fig. 2.4 A photograph of the rotary-joint-pullback unit. The rotor of the fiber-optic rotary joint (FORJ) is driven by a motor for rotational scanning. The jig is fixed on a translational stage and driven by a linear actuator for longitudinal pulling.

The pullback unit pulls the catheter probe longitudinally. The first generation of pullback unit (designed by Dr. Zhilin Hu) pulls both the probe and the balloon(s). If the friction between the balloon and esophagus is small, the balloon slides easily within the lumen. Three dimensional drawing is shown in Fig. 2.5(a). The adapter connects and fixes the pull back system to the endoscope as shown in Fig. 2.5(b). Four pulling pulleys coated with silicon rubber for the resistance enhancement hold and pull the entire catheter. A screw controls the force between the pulleys to tighten the catheter. The pulleys are driven by a step
motor via gear sets. Computer software controls the motor to rotate and pull the probe proximally at a speed of 2 mm/s. The advantage of this design is that the longitudinal imaging length is arbitrary. However, there are several problems. First, the material of the outer sheath is Teflon. Pulling the sheath is not efficient because of the low coefficient of static friction. Second, inflating the balloon increases the friction between balloon and esophagus. The balloon does not necessarily slide against the esophagus, but rather stretches it. Last, sterilization of the gears and pulleys is complicated. This pullback design is not used frequently in the in vivo study.

![Diagram of pullback unit](image)

Fig 2.5 (a) Three dimensional drawing the pullback unit. The motor drives the gears and the pulleys, resulting in pulling the catheter probe. (b) The pullback unit is attached to the accessory channel on the endoscope.

The second generation of pullback unit pulls the fiber-optic probe within the outer sheath. In Fig. 2.4, The rotary joint unit is mounted on a sliding translational
stage (PT1, Thorlabs Inc., Newton, NJ). The stage is controlled by a linear actuator (T-LA28A, Zaber Technologies Inc., Vancouver, British Columbia, Canada). Therefore, the linear actuator controls the longitudinal movement of the probe. Because the probe and the sheath slide against each other, a sliding tubular device has been built and attached near the translational stage to avoid air-leakage (Fig. 2.6). The device consisted of two bronze tubes. The inner tube is attached to the rotary joint and the outer tube is attached to the outer sheath whose distal ends support the balloons (described below). A large Nylon balloon is used to prevent air-leakage between the tubes. A pressure meter is used to monitor the net balloon pressure.

![Fig 2.6 A sliding tubular device prevents air leakage while pulling the balloon.](image)

The rotary-joint-pullback unit pulls the inner tube sliding along a stationary outer tube. A three-way valve is used for connection to the air pump and regulator.

A 2m flexible tube encloses the optical fiber and transfers the torque and the force to the fiber probe. A transparent outer sheath is used to enclose the flexible tube while supporting the balloons. In order to be deployable through the GI endoscope, the diameter of the balloon catheter is restricted by the diameter of
the accessory channel. The outer sheath has a diameter of 2.4 mm, which passes through the 3.6mm diameter channel with the balloon deflated.

**Balloons**

The inflated balloons center the fiber probe in the esophageal lumen so that the beam from the fiber probe maintains focusing on the esophageal mucosa. The double-balloon design allows two imaging schemes to be employed in imaging the esophagus (with and without balloon-tissue contact). Fig. 2.1(a) has illustrated the double-balloon design. The two necks of the balloons are attached to the outer sheath using biocompatible epoxy. For imaging with no balloon-tissue contact, the images are obtained in the gap between the two balloons (labeled beam position 1). The fiber probe can also be placed within the distal balloon (labeled beam position 2) to acquire images through the balloon in the conventional way. Urethane low durometer balloons (18002501AC, Advanced Polymers Inc., Salem, NH) with a diameter of 18 mm are utilized. The length of the gap is restricted to approximately 15 mm to limit the extent of lumen contraction between the balloons, and keep the tissue within the axial imaging range. Holes are made in the outer sheath within the balloons so that they are inflated and deflated from the proximal end of the catheter. In order to estimate the position of the balloons in the esophagus, black Teflon heat-shrink bands are attached to the proximal side of the double-balloon on the outer sheath. These bands are 5cm apart and visible under endoscope to indicate the length of
catheter projection. The bands have a wall thickness of 0.2 mm so that they do not affect the insertion of the catheter.

The air pressure controlling device is connected to Teflon outer sheath by using a three-way valve (Fig. 2.6). An aquarium air pump (Rena Air 400, PlanetRena, Charlotte, NC) continuously pumps air to an air regulator (T700, ControlAir, Inc., Amherst, NH). The air is then directed to the three-way valve. By turning the knob of the air regulator, the pressure can be controlled from 0 to 130 mmHg. The diameter of the balloon is inflated to 18 mm for swine esophageal imaging (9 mm working distance probe), and 25 mm for human esophageal imaging (12 mm working distance probe). The variation of the pressure in the balloon is less than 5 mmHg.

2.2.4 Imaging protocol

The fiber probe is rotated at 10 revolutions per second and pulled back at 2 cm per minute. The incident light on the tissue is approximately 1 mW (6 mJ/cm², about 1% of the American National Standard Institute (ANSI) limit for skin exposure). The line scan camera acquires data at the full acquisition potential of 47k A-scans per second, which generates 95MB of raw data per second. Each frame consists of 4,700 A-scans. The A-scan pitch is 11 µm at a 9 mm working distance and the frame pitch is 33 µm. 1 cm to 3 cm segments of esophagus are image.

The animal imaging protocol has been approved by Institutional Animal Care and Use Committee at Case Western Reserve University. The animal models
were swine (20-40 kg) sedated with Telazol (8-12 mg/kg), intubated, ventilated and maintained under anesthesia with Isoflurane (1.5~2%) for the duration of the procedure. Heart rate and oxygen saturation were monitored.

The endoscope was operated by gastroenterologists with experience in endoscopic surveillance of BE. An initial endoscopic inspection was performed with the OCT catheter hidden in the accessory channel. When a site was identified for imaging, the distal balloon was projected out of the accessory channel so that the gap between the balloons was visible. The fiber probe was placed under endoscopic guidance such that the probe beam passed between the two balloons. The endoscope was then pulled back 5 cm from the region of interest for deploying the proximal balloon, and both balloons were inflated. Volumetric images were obtained in the region of interest without balloon-tissue contact. Then the two balloons were deflated and the proximal balloon was pulled back into the accessory channel. The distal balloon and the probe beam were placed at the same region of interest under endoscopic guidance and the tissue was imaged with the balloon in contact with the tissue.

2.3 Results

2.3.1 Fiber-optic probe

Fig. 2.7 shows an example of the beam profile of the first generation fiber-optic probes measured at the designed focus by a beam analyzer. Significant side lobes are observed as a result of aberration. The horizontal and vertical FWHM's are 22 μm and 31 μm, respectively. The ellipticity is 0.38. Fig. 2.7(c) is one EOCT
image obtained with this probe design from swine esophagus \textit{ex vivo}. The \textit{SNR} and penetration depth are low and no speckle can be observed. The reason is probably due to the fact that the probing beam is not Gaussian. The incident light can no longer be treated as plane wave near focus. The coherent gate is significantly distorted. It is also observed that there are multiple artificial rings, which are probably due to the internal reflection of the optical components in the probe. This design did not meet the requirement to obtain images of high quality, and was therefore discarded.
Fig 2.7 The beam profile of the first generation probe measured at the focus in 2-D (a) and 3-D (b) views. The bar in (a) indicates 100 µm. The horizontal and vertical FWHM's are 22 µm and 31 µm, respectively. The side-lobes of the spot indicates that the beam is significantly degraded and no long Gaussian. (c) An EOCT image obtained with the first generation fiber-optic probe from porcine esophagus ex vivo.

Fig. 2.8 shows the results of the second generation fiber-optic probe design. Figs. 2.8 (a) and (c) are optical simulations (Zemax, ZEMAX Development Corp., Bellevue, WA) of the beam at a 9mm working distance without and with astigmatism correction by the cylinder lens, respectively. The corrected spot size is designed to be 30µm FWHM. Figs. 2.8 (b) and (d) show the beam at a 9 mm
working distance before and after the astigmatism correction, respectively, as measured by a beam analyzer. The corrected spot has an ellipticity of 0.98 with a FWHM diameter of 33µm.

![Images of beam simulations and measurements](image)

Fig. 2.8. (a) & (c) Zemax simulations of the beam at the designed working distance without and with astigmatism correction, respectively. The designed spot size is 30 µm. (b) & (d) Beam shapes without and with astigmatism correction, respectively, as measured by a beam analyzer. The measured spot size is 33 µm in both horizontal and vertical directions. (Bars: 100 µm)

2.3.2 *Endoscopically guided balloon deployment*

Balloon deployment is accurately controlled endoscopically during the imaging procedure. Fig. 2.9 is a single-frame excerpt from a video recording the operation
of a balloon catheter in a porcine esophagus *in vivo*. The balloon is clearly visible so that its position can be adjusted. The entire balloon catheter is smoothly inserted through and pulled back in the accessory channel of the GI endoscope. The time required for balloon deployment is less than half a minute.

Fig. 2.9 One frame excerpted from a video recording demonstrating the insertion of the rotating fiber probe and inflation of the balloon in a porcine esophagus *in vivo*. The diameters of the probe and the inflated balloon are 2.4 mm and 18 mm, respectively.

### 2.3.3 Swine esophageal imaging *in vivo*

Comprehensive porcine esophageal images *in vivo* were obtained with both double and single balloon imaging schemes as shown in Fig. 2.10. The diameters of the imaged esophagi are around 18 mm, but in order to visualize the tissue structure, the images are displayed with a diameter of about 3 mm. The scale bars indicate 1 mm in the radial direction. Fig. 2.10(a) is a representative cross-sectional image of the porcine esophagus without balloon-tissue contact in the anatomic view. The imaging depth in the tissue exceeds 1 mm and the tissue layers and structures can be clearly observed, including the
squamous epithelium, lamina propria, muscularis mucosa, submucosa, muscularis propria, and blood vessels in the muscularis mucosa. Fig. 2.10(b) shows a frame excerpted from a movie displaying volumetric visualization of a 17 mm long segment obtained within 50 seconds (approximately 5 GB of raw data). It was observed that motion artifact caused distortion in the volumetric images along the longitudinal dimension. The distortion is corrected by segmenting the air-tissue interface and aligning the interface to a cylinder. Figs. 2.10(c) & (d) show a representative anatomic view and the corresponding 3D reconstruction obtained with the single-balloon scheme (10 mm segment obtained within 30 seconds, approximately 3 GB of raw data). The typical layers of the esophagus can also be clearly recognized. The image signal is slightly stronger and more uniform than the non-contact image, but some structures (e.g. blood vessels) are less apparent. The radial cropping magnifies the variance in the axial range of the tissue, and therefore the lumen appears more elliptical and/or eccentric in Figs. 2.10(a) and (c) than it would appear if the images were displayed with true aspect ratio.
Fig. 2.10 (a) A representative cross-sectional image of a porcine esophagus with the double-balloon imaging scheme. The layered structure that can be observed includes the squamous epithelium (SE), lamina propria (LP), muscularis mucosa (MM), submucosa (SM) and muscularis propria (MP). (b) One frame excerpted from a movie of 17 mm long section of porcine esophagus obtained with the double-balloon imaging scheme. (c) A representative cross-sectional image of the single-balloon imaging scheme. (d) 3D reconstruction of 10mm long segment with single-balloon imaging scheme.
Fig. 2.11(a) shows a 3-D reconstruction of Fig. 2.10(b) with true aspect ratio. Image details are difficult to visualize. The square frame indicates the position of Fig. 2.10(a). Fig. 2.11(b) magnifies a small region of interest in the double-balloon image where the layered structure can be cleared identified. This image also shows the typical speckle pattern when the tissue is imaged at the focus of the probing beam. Fig. 2.11(c) shows a longitudinal view of the box in Fig. 2.11(a). The tissue structure, including the blood vessel, can clearly be visualized. The rich image details and the comprehensive imaging capability are the major advantage of 3-D OCT over 2-D OCT.

Fig. 2.11 (a) A 3-D reconstruction of Fig. 2.10(b) with true aspect ratio. (b) An enlarged view of the layered structure in a small region of interest. (c) A longitudinal view of the box in (a). (SE: squamous epithelium; LP: lamina propria, MM: muscularis mucosa, SM: submucosa; MP: muscularis propria.)
2.4. Discussion

The described endoscope-compatible double-balloon EOCT catheter allows imaging of the esophageal mucosa with and without direct balloon-tissue contact, which will allow for studying the influence of pressure on detection of dysplasia in BE. The two imaging schemes generate volumetric images of the esophagus rapidly without exchanging balloons. Deployment of the catheter through the endoscope is important because the two-scheme protocol is only possible with endoscopic visualization. Other EOCT catheters, without balloons, were commonly designed for deployment through the endoscope [38, 45, 81-85, 87]. Endoscope-guided deployment of an EOCT catheter bears advantages such as visual guidance and placement of the catheter, documentation of the procedure, and correlation of OCT and endoscopic views of the same tissue sites. It will also facilitate biopsy-correlation studies and, importantly, minimize the time added to the procedure when EOCT is employed in the clinic. Endoscopic deployment has not been demonstrated in previous reports of balloon-EOCT imaging. After investigating the consequences of balloon-tissue contact in BE surveillance, it will be necessary to design a probe capable of imaging long segments of esophagus with optimum contact and pressure. In the demonstration presented here, the balloons are inflated to a larger diameter during double-balloon imaging than single-balloon imaging because the tissue within the gap has a smaller diameter than that supported by the balloons. Alternatively, if it is important to maintain the balloons at a fixed diameter for a
study, a shorter working-distance probe can be rapidly exchanged while the balloons remain in place in the esophagus.

The instrument reported here represents the first demonstration of a spectral-domain OCT system for gastrointestinal endoscopy. The usability of our SD-OCT system and image quality are comparable to previously demonstrated endoscopic swept-source OCT (SS-OCT) [40, 95-97]. While SD-OCT is the conventional technology in retinal OCT imaging (which is the dominant clinical application of OCT imaging) [121], SS-OCT has been employed for all previous implementations of FD-OCT for endoscopy. There are several reasons for using SS-OCT. First, because of commercially available components, SD-OCT is more readily implemented than SS-OCT in the 830 nm range commonly used for retinal imaging. At 1300 nm, commonly used in endoscopy, high quality components have been more readily available and more cost effective for SS-OCT than for SD-OCT. Rapidly scanning tunable lasers for SS-OCT were developed before fast line-scan cameras were available for SD-OCT at 1300 nm [95]. Furthermore, balloon-based EOCT requires a relatively long axial imaging range, to accommodate the axial position of the tissue which varies with radial position and moves due to physiological movement such as peristalsis, respiration, etc. SS-OCT more readily accomplishes long axial range because it generally benefits from less fall-off [122], can more easily resolve complex-conjugate ambiguity [123-125], and is not limited by a fixed number of spectral samples, as is SD-OCT. The endoscopic SD-OCT system reported in this dissertation makes use of an InGaAs line-scan camera with 1024 pixels and a
readout rate of 47,000 lines per second. This enabled SD-OCT that is not only fast enough for clinical pull-back procedures, but also has sufficient axial range (4.2 mm) to accommodate the variance of tissue position experienced in porcine esophagus in vivo. InGaAs line-scan cameras appropriate for SD-OCT with more pixels and faster read-out will become available in the near future, further improving the feasibility of SD-OCT for endoscopic applications. The fall-off is improved by use of the linear-in-wavenumber spectrometer, and the probe optics is corrected to create a nearly Gaussian beam, so that high quality images are readily obtained. From these results and observations, it is concluded that SD-OCT is a feasible alternative to SS-OCT for endoscopic imaging. The SD-OCT configuration will also allow for incorporating the ultra-broadband light source into the EOCT system for potential improvement of the axial resolution (~5 μm) [126].

Ultrahigh resolution OCT reduces speckle size and enables visualization of finer morphological features in esophageal images, as previously reported in a time-domain configuration [38, 127], and may potentially benefit BE diagnosis. However, for a fixed-sized detector array, accommodating a broader bandwidth would trade-off with a shorter axial imaging range.

Some preliminary observations of the effects of balloon-tissue contact are apparent in Figs. 2.10(a) and (c). The natural mucosal surface morphology is apparent in the double-balloon image, whereas the mucosal surface is compressed and smoothed by the balloon in the single-balloon image. The balloon-flattened surface results in more stable illumination and therefore more uniform image brightness. However, a study has shown that surface morphology
may be useful for detecting dysplasia in BE [93]. Detailed structure in the muscularis mucosa, especially blood vessels, is more clearly observed in the double-balloon images. The imaging depth in the single-balloon images is greater, with multiple layers of muscularis propria visible, which is consistent with previous observations [30]. While it was expected that the double-balloon imaging scheme resulted in higher variability of the axial position of the tissue as a function of radial position, the difference is small, as seen in Fig. 2.10. However, the movement of the tissue due to physiological movement was greater with double-balloon imaging compared to single-balloon imaging. These observations suggest the need to further investigate the advantages and disadvantages of balloon-tissue contact and pressure, how tissue features in BE are altered, and how the changes affect detection of dysplasia in BE. Studies in porcine esophagus on the influence of balloon will be further presented in chapter 4.

2.5. Conclusion

In conclusion, I have described an \textit{in vivo} esophageal SD-OCT imaging platform that generates high quality volumetric images with double-balloon and single-balloon imaging schemes and endoscope-deployable catheters. I have also demonstrated imaging of porcine esophagus using both schemes \textit{in vivo} and presented preliminary observations. Both animal feasibility study and human clinical trial will be performed with this platform.
Chapter 3. Motion artifact correction

3.1 Introduction

3-D EOCT allows for interrogation of a 20 cm² mucosal area or more, and therefore avoids the sampling error inherent in standard clinical biopsy-based examinations [13-14]. If the volumetric image sets are precisely reconstructed, 3-D visualization and analysis may allow one to detect the onset of BE more accurately and efficiently than standard tissue biopsy.

As described in chapter 2, the balloon serves to stabilize the side-viewing probe and suppress motion artifacts while the probe is rotated rapidly and pulled back slowly, scanning the OCT beam on the esophageal lumen while recording a helical 3D image set (Fig. 3.1(a)). However, even with the balloon, significant image distortion due to motion artifacts has been commonly observed [40, 95-98, 100, 128-129]. Fig. 3.1(b) illustrates the three components of motion artifact present in helical esophageal scanning. The beam rotates one revolution with slight pull-back to the position indicated by the dashed beam to generate one image. In the time between two successive images, the tissue experiences a relative displacement due to motion which causes image distortion, or motion artifact. The radial component of the motion artifact (in the direction of the probe beam) and the rotational component (in the direction of the probe rotation) create distortion detectable within individual frames and between successive frames. The longitudinal component of the artifact (in the direction of the pull-back) is not
clearly differentiated without detecting additional information, such as two-dimensional Doppler shift [129]. In this chapter, motion artifacts refer to the radial and rotational components if not otherwise specified. Fig. 3.1(c) illustrates an example of radial motion artifact by displaying a stack of cross-sectional OCT images obtained from a swine esophagus in vivo without pull-back of the probe. The third axis is time (i.e. successive frames) and the time-radial plane illustrates how the reconstruction is substantially distorted radially due to motion of the tissue. Fig. 3.1(d) is an en face plane extracted from the same volume within the layer of the muscularis mucosa (approximately 300 μm from the mucosal surface). In this view, rotational distortion is clearly observed. Therefore, in order to faithfully represent the imaged tissue, an image registration algorithm is needed for 3D reconstruction of the EOCT images [40, 95-96, 98]. Such an algorithm will also facilitate the quantification of the size and number of the subsquamous Barrett’s glands (approximately 100 μm in diameter) for possible treatment planning and assessment [94, 102, 127].
Fig. 3.1 (a) The helical imaging scheme of the side-view probe. (b) The motion artifacts are caused by the relative motion between the tissue and the sheath during two successive images. (Rad: radial; Rot: rotational; L: longitudinal) (c) The radial-time plane of 3D reconstruction shows the radial displacement. (d) The *en face* reconstruction by taking pixels at 300 $\mu$m deep of (c). (Bar: 1mm)

Post-acquisition registration technologies have been developed for correcting motion artifact in esophageal OCT images with limited success. The lumen surface has been aligned to a cylinder to compensate radial distortion [40, 95-96, 98], but this did not compensate rotational artifact. This alignment also requires lumen segmentation, which is time-consuming if not implemented automatically. A Doppler interferometer combined with the OCT system estimated the radial and longitudinal motion artifact in a rotating phantom [129]. It was not readily
practical for in vivo imaging because a second optical fiber limited the probe rotation. The registration algorithms utilized for removing motion artifact in intravascular ultrasound (IVUS) [130-133] and intravascular OCT imaging [134] may provide insights into effective algorithms for esophageal OCT. Groups have registered both surface contours [130] and tissue structures [133] to correct for motion. Esophageal OCT images consist of similar contours and image structures. However, a global translational and/or rotational transformation did not effectively suppress the common non-uniform rotational distortion (NURD) [134]. Applying registration algorithms to smaller regions of interest (ROI) could improve the results by enabling deformable transformations. Using smaller ROIs for both IVUS and intravascular OCT, researchers have assumed highly correlated speckle patterns between successive frames [131, 134]. However, speckle patterns between successive esophageal OCT images are not correlated because the lumen is relatively large and displacement between successive frames is significant. A practical approach for esophageal OCT imaging may compare the structural similarity between ROIs of moderate sizes, e.g. the local block matching (LBM) algorithm [132], and then apply local transformations to correct for rotational motion. The ROI size should be optimized so that it is large enough to contain structural features for correlation, but also small enough to correct for NURD.

This chapter describes a fully automated registration algorithm to suppress motion artifacts. First, the algorithm estimates the mucosal surface by detecting the strong radial gradient. After the images are radially aligned, LBM detects the
rotational artifact. Different size ROIs are chosen for comparison. Esophageal OCT images acquired from a swine model in vivo with the single balloon imaging scheme are reconstructed for evaluation. The limitation of applying the algorithm to double-balloon images will be discussed in section 3.4.

3.2 Method

3.2.1 Spectral-domain endoscopic OCT System

The spectral-domain endoscopic OCT system with the single balloon imaging scheme described in chapter 2 was utilized for imaging. A representative raw image before transformation into circular anatomic view (as in Fig. 2.10(b)) is shown in Fig. 3.2. Each vertical line is one A-scan, and the image is simply constructed by arranging all A-scans horizontally. The vertical and horizontal directions in the raw image are radial and rotational in the esophagus, respectively. The upper surface is the air-balloon-tissue interface, which is about 9 mm away from the probe. All the motion artifact correction will be done in the raw image.
Inconsistencies in probe rotation speed cause mechanical motion artifact and have been previously reported [96]. The inconsistency was compensated by recording the orientation (angle) of the A-scans. The A-scans were triggered at a constant frequency and later interpolated to generate A-scans evenly spaced rotationally. In the SD-EOCT system, A-scan data is acquired at a constant angular increment. 4,000 triggers per revolution are provided by the servomotor drive (S200; Kollmorgen, Radford, VA) in the rotary-joint-pullback unit to the line scan camera in the spectrometer. A tubular phantom was imaged to illustrate the mechanical motion artifact and its suppression. The phantom has 8 longitudinal slots evenly distributed on the circumference of the wall, as shown in Fig. 3.3. The phantom was imaged without and with the mechanical motion artifact suppression for comparison.
3.2.2 Imaging protocol

The swine animal model was utilized under the same protocol described in chapter 2. When acquiring images, the balloon was inserted approximately 45 cm from the mouth. It was inflated and maintained at a constant pressure of 30 mmHg. First, a cross-section of the same longitudinal position was imaged repeatedly without pulling the probe. Images were acquired for 25 seconds with ventilation (2-D sequences *in vivo*). The ventilation was then paused for 25 seconds to obtained images without motion artifact caused by respiration. After resuming the ventilation, 3-D volumetric images were obtained by pulling the probe for 1 cm (3-D sequences *in vivo*). The probe was then pushed forward 1 cm to the initial position. The process was repeated for 50, 70, 90, 110 and 130 mmHg. 3 swine were imaged in total. A cross-section of one excised swine esophagus was also imaged repeatedly without pulling the probe to show the mechanical motion artifact existed.
3.2.3 Image registration to suppress motion artifacts

The flow chart of the image registration is shown in Fig. 3.4. First, the balloon-tissue interface was automatically detected so that the images would be aligned radially. The gradient of each pixel was defined as the difference between the mean intensities within 200 \( \mu \text{m} \) below and above the current pixel. 200 \( \mu \text{m} \) was chosen because it was the approximate mean thickness of the epithelium. This definition helped identify the balloon-tissue interface, and also avoided local but strong gradients due to the image speckle. The detection of the interface was then formulated as a global optimization problem, which sought a contiguous curve with only one pixel from each A-scan. The summation of the gradient value on the curve was a global maximum among all the possible curves. The optimization was achieved by dynamic programming [135-136]. A one-dimensional low-pass filter was also applied to smooth the curve. The detected interface was aligned to a straight line by adjusting each A-scan along the radial direction.
LBM was applied to estimate the rotational motion. LBM matched ROIs in each image to the previous image by cross-correlation. In this work, each image was divided radially into $N$ ROIs with an equal number of A-scans. $N$ varied from 1 to 20 for comparison. The structural similarity between the current ROI and its reference was defined by the cross-correlation coefficient. The coefficient was calculated by shifting the current ROI in the rotational direction with respect to its reference. The shift that yielded the maximal coefficient was considered to be the rotational displacement of the middle A-scan in the ROI. After motion estimation by LBM, the displacements of all the A-scans between two middle A-scans were linearly interpolated. As a result, A-scans were assigned new unevenly-spaced rotational coordinates. The registered image was then reconstructed by bilinear interpolation at even rotational spacing. LBM does not distinguish tissue motion...
from actual variation in the longitudinal anatomical structure of the tissue, such as blood vessel branches. This phenomenon has been commonly studied in computer vision [137]. LBM simply translates the current ROI rotationally so that it appears as similar to the reference ROI as possible. This might generate small errors compared to the actual anatomy, which would accumulate as LBM is applied hundreds of times, and distort the longitudinal anatomical structure. A simple solution was implemented based on the assumption that anatomical changes occur more slowly than physiological motion. A spatial filter was applied to remove the slowly varying apparent rotational motion. The filter was a 5th order Butterworth high pass filter with the cutoff frequency at the ventilation rate (13 breaths/min = 0.022 frame⁻¹).

3.2.4 Evaluation of surface detection and overall registration

The surface segmentation algorithm was evaluated with 3-D sequences in vivo. Two images were selected from each pressure level in each swine. The lumen dimensions in these 36 images differed significantly, enabling the robustness of the segmentation to be examined. A blinded reader (Zhao Wang) manually segmented the contour of the tissue surface. The manually segmented contour was considered the gold standard. The segmentation error was defined as the distance between the manually segmented pixel and the automatically segmented pixel in the same A-scan. The mean and standard deviation of the errors of all A-scans in 36 images were calculated as evaluation metrics.
The overall alignment was evaluated on the 2-D sequences in vivo. Images to be evaluated were down-sampled 4 times so that random speckle did not affect the structural comparison. Images were then stacked along the time axis as shown in Fig. 3.1(c). The mean normalized standard deviation (MNSD) evaluated how the pixels varied along the time axis. The MNSD was defined as [134]:

$$\left\langle \sigma \right\rangle = \frac{1}{N_{tis}} \sum_{(r,c) \text{ in seq} \text{time}} \frac{\sigma(p_{rc})}{\left\langle p_{rc} \right\rangle},$$

where $p_{rc}$ is the pixel intensity at the $r$ row and $c$ column. $\left\langle p_{rc} \right\rangle$ and $\sigma(p_{rc})$ indicate the mean and the standard deviation at the position $(r, c)$ along the sequence. In order to avoid introducing variability from background noise into the measurement, MNSD was only calculated for pixels containing image signal in all images in the stack. $N_{tis}$ was the number of pixels included in the calculation. Higher MSND indicated less similarity along the time axis, and therefore more motion artifact. In order to quantify the motion artifact suppression, MNSD was calculated in the original and corrected sequences.

3.3 Results

3.3.1 Demonstration of mechanical motion artifact suppression

The tubular phantom was imaged without and with the mechanical motion artifact suppression method proposed in section 3.2.1. The images were obtained from the same cross-section without pulling the probe. Therefore, the images should be ideally identical along the time axis. Fig. 3.5(a) and (b) show the en face view before and after mechanical motion artifact correction. The dark region was the
Clearly, the mechanical motion artifact significantly distorted the images and was successfully suppressed by the constant-angle triggering mechanism.

Fig 3.5 The *en face* views of the tubular phantom (in section 2.1) obtained without (a) and with (b) the mechanical motion artifact suppression. (Rot: rotational)

### 3.3.2 Evaluation of surface detection and overall registration

For evaluation of the mucosal surface detection, the automatically detected boundaries in images from the 3-D sequence *in vivo* were compared to expert human segmentation. The measured mean and standard deviation of the segmentation error were -0.2 and 2.8 pixels, respectively (i.e. -1.8 and 25 μm in air or -1.3 and 19 μm in tissue).

For evaluation of overall motion artifact correction, MNSD was calculated in both the original and registered 2-D sequences *in vivo*. In the LBM process, images were divided into \( N \) (1 to 20) ROIs for comparison and the resulting MNSDs are plotted in Fig. 3.6. The original sequences (\( N=0 \) in Fig. 3.6) have much higher MNSDs than the registered sequences. When \( N=1 \), the registration
was similar to the global rigid transformation in Ref. [133], and resulted in a reduction of MNSD of 21% on average. As $N$ increased, the images in each sequence became more and more similar, so that for $N=20$, the MNSD was reduced by 26% on average. This improvement is attributed mainly to the correction for NURD. However, the improvement reached 25% at $N=10$, and only trivial additional improvement was measured with more ROIs. The MNSD was also measured in a 2-D sequence ex vivo in order to exclude motion artifact. The measured MNSD ex vivo was 0.28, meaning that the residual error in the corrected in vivo sequences is attributable mostly to noise in the images, not to uncorrected motion artifact. The registrations demonstrated in Fig. 3.7 were computed with $N=10$.

![Graph showing MNSD vs. number of ROIs](image)

Fig. 3.6 Evaluation of automatic registration using 2-D sequences recorded in 3 swine. The mean normalized standard deviation (MNSD) evaluated how the pixels varied between subsequent frames in the sequence. MNSD of the registered sequences was calculated as a function of the number of ROIs used for LBM. The MNSD scale begins at 0.28, which was the reference baseline MNSD attributable to image noise only, with no sample motion.
3.3.3 Demonstration of registration

The registration results are demonstrated in Fig. 3.7 using 2-D and 3-D sequences. Fig. 3.7(a) shows an en face view of the muscularis mucosa taken from a 2-D sequence from the excised esophagus, where only mechanical (not physiological) motion artifact is present. Note that the distortion along the time axis is subtle compared to Fig. 3.7(b), from an in vivo 2-D sequence where physiological motion artifact is prominent. The mechanical motion was not a significant cause of distortion. Fig. 3.7(c) displays the motion-corrected en face view of Fig. 3.7(b). The rotational distortion is significantly corrected, although the image pattern along the time axis shows variability compared to Fig. 3.7(a). This residual variability along the time axis was probably caused predominantly by longitudinal tissue movement. For illustration purposes, a straight line was selected in the corrected image (red line shown in Fig. 3.7(c)) and the position of the A-scans on that line is shown in the unprocessed image (red line shown in Fig. 3.7(b)). Figs. 3.7(d) and (e) represent typical unprocessed and corrected, respectively, radial-longitudinal cross-sections from a 3-D sequence in vivo. Their corresponding en-face views within the muscularis mucosa layer (after radial correction in both cases) are displayed in Figs. 3.7(f) and (g), where the yellow arrows indicate the rotational positions of Figs. 3.7(d) and (e), respectively. The radial position of Fig. 3.7(g) is indicated by the yellow arrow in Fig. 3.7(e). The saw-tooth patterns in both the radial and rotational directions were corrected and the tissue layers were aligned. The muscularis mucosa has a rich vascular network, which can be visualized more clearly in the corrected en face view.
3.7(h) and (i) are enlarged views of the regions indicated in the dashed yellow rectangles in Figs. 3.7(f) and (g), respectively. The registration algorithm was able to recover detailed features in Fig. 3.7(i) which were otherwise obscure in Fig. 3.7(g).
Fig. 3.7 (a) The en face view of the muscularis mucosa layer in the excised esophagus from a 2-D time sequence of images. (b) The en face view of the in
vivo 2-D sequence where physiological motion artifact is prominent. (c) The registered en face view of (b). (d) A radial-longitudinal view from the original 3-D sequence in vivo. (e) The registered view of (d). The arrow indicates the position of (g). (f) The en face view from the original 3-D sequence. Arrows indicate the position of (d). (g) The registered view of (f). Arrows indicate the position of (e). (h) The enlarged view of the square in (f). (i) The enlarged view of the square in (g). (Rad: radial; L: longitudinal; Rot: rotational. Bars: 1mm)

3.4 Discussion

The mechanical motion artifact was a significant source of image distortion if not suppressed. It was observed in the system that motion magnitude was on the order of hundreds of microns at a frequency higher than the heart rate. It probably originated from the eccentric manufacturing and/or assembly of the rotary joint, catheter probe and outer sheath. The suppression method in this work assumed that the motion artifact at both the distal and proximal ends of the probe was identical. By triggering the data acquisition at constant angular increments in the rotary joint, the mechanical motion artifact at the probe was suppressed.

The esophagus in the swine model was imaged with ventilation control and heart rate monitoring. This setting allowed for the identification of physiological origins of motion artifacts in chapter 4. A limitation of the swine model was that peristalsis had not been observed in either endoscopy or OCT. The inhibition of peristalsis was likely caused by the Isoflurane [138-139]. General anesthesia of the animal was necessary to avoid the otherwise intense body movement.
However, mild anesthetics, such as Midazolam, are used in human [140]. Peristalsis can be observed by endoscopy in patients undergoing BE surveillance. Patient’s hiccups, coughs or other random body movements due to mild sedation, as well as peristalsis, are likely to make the motion artifacts more complicated [141]. The performance of the reconstruction algorithm should be further examined on images from a clinical study.

The automated registration algorithm was efficient for suppressing in-plane motion artifacts for single balloon imaging scheme. Dynamic programming reliably detected the balloon-tissue interface because the contrast between the lumen and the tissue was consistently higher than other layered structure in the image. LBM assumed uniform rotational motion within each individual ROI \((N=10)\). The assumption was valid, because the duration to obtain such an ROI (10 micro-second) was much less than the period of the heart beating \((\sim0.6\) second) or respiration \((\sim5\) second). LBM in this work was a simplified deformable registration technique which only corrected the rotational deformation (or NURD). The tissue was compressed by the balloon and therefore considered no deformation with respect to the balloon radially. More sophisticated deformable registration technologies are applicable at the cost of intensive computation load [142-143]. However, because the noise inherent in the images contributes to approximately two third of the total MNSD of the unregistered sequence, new registration technologies may only improve a limited amount of MNSD.

The motion artifact correction algorithm was also applied to the esophageal images obtained from Barrett's esophagus in clinical trial. The detailed protocol
will be described in chapter 5. Fig. 3.8(a) and (b) are parts of the *en face* views of the Barrett's esophagus with high grade dysplasia before and after motion artifact correction. The entire data set was obtained in 1 cm long esophagus by expanding the esophagus with a single balloon. However, the single balloon was relatively small compared to human esophagus. Therefore the esophagus was not fully expanded, leaving gap between the balloon and the tissue as shown in Fig. 3.8(c). The *en face* views are approximately 300 mm from the tissue surface. The dark region in Fig. 3.8(a) on the lower right region was the gap between tissue and the balloon. The small dark holes numbered from 1 through 3 are Barrett's glands, whose appearance are similar to Ref. [144-145]. Barrett's gland is Barrett's epithelium buried in the mucosa, which may develop dysplasia or even carcinoma. These glands usually appear oval in the cross-sectional images [94]. Identifying these glands and counting the numbers were proposed as a way to estimate the eradication rate of Barrett's esophagus by using radio-frequency ablation [144-145]. However, manually counting these glands is time-consuming when there are hundreds of cross-sectional images. Especially when the glands are not well aligned between subsequent cross-sections. Registering the glands is necessary to avoid counting the same gland multiple times. The misalignment of glands along the longitudinal direction can be clearly visualized by comparing Fig. 3.8(a) and (b). The correction recovers the oval shape of glands in the *en face* view, which is the consequence when the glands are well aligned among subsequent cross-sections. The motion artifact algorithm relieves the burden of
manual registration, and also allows for developing automated algorithm to segment and count Barrett's glands.

Fig. 3.8 (a) The en face view of the Barrett's esophagus with high grade dysplasia before motion artifact correction. The regions numbered 1, 2 and 3 are regions where Barrett's glands significantly distorted by motion artifacts. (B) The en face view after motion artifact correction. The oval shape of gland are mostly recovered in regions of corresponding numbers. (c) An cross-sectional images of the Barrett's esophagus. The balloon (on top of the tissue) was not large enough to expand the esophagus, leaving a gap between the tissue and the balloon. The gap is the reason why there is a dark region in the en face view in (a). (Rad: radial; Rot: rotational; L: longitudinal. Bars:1mm)
Similar registering approach was tested for image sequences acquired with the double balloon imaging scheme. Fig. 3.9(a) is a 3-D reconstruction of the double-balloon sequence. The tissue surface was segmented. Each image was then transform into the circular anatomic view without aligning the surface to a circle. The reconstruction was sectioned along the longitudinal direction to expose the mucosal wall. Fig. 3.9(b) are the single balloon image whose reconstruction underwent the same processing. These two figures clearly show that the magnitude of motion artifacts in double-balloon sequence is much greater than that in the single-balloon sequence. Fig. 3.10(a) is the en face view of Fig. 3.9(a). Fig. 3.10(b) shows the registered en face view of Fig. 3.10(a) by choosing $N=20$ in the LBM step. The vasculature features were partially recovered. But there are a few observations showing the limitations of the registration approach. The first observation is the discontinuity of image intensity along the longitudinal direction. Such artifact was not corrected by considering only in-plane motions. Similar effect can be found at the blood vessel wall (yellow arrow in Fig. 3.10). The alignment of the left vessel wall is worse than the right wall. This observation clearly show that the limitation that assuming solely in-plane motion is inefficient for the violent motion in double-balloon imaging. The second observation is that LBM introduced uncontrolled deformation which significantly altered the true microstructure. The blood vessel indicated by the yellow arrow in Fig. 3.10(b) has a much larger diameter than that in the original view in Fig. 3.10(a). The difference became more obvious as $N$, the number of
ROIs, increased. The LBM step has the limitation that it does not consider the correlation of rotational motion between neighboring ROIs within the same image. It may result in artificial dilation or contraction in the rotational direction. Between the two limitations, the primary problem lies in the motion model. A more sophisticated motion model should be adopted by taking into account the longitudinal motion, so that an efficient registration approach may be developed for the new motion model.
Fig. 3.9 (a) A 3-D reconstruction of the double-balloon sequence. Each image was then transform into the circular anatomic view without aligning the surface to a circle. The reconstruction was sectioned along the longitudinal direction to expose the mucosal wall. (b) A 3-D reconstruction of the single-balloon sequence.
Fig. 3.10 (a) *En face* view of sequence acquired by the double-balloon imaging scheme. (b) *En face* view after the registration algorithm applied. The algorithm was not able to correct the motion artifact sufficiently. (L: longitudinal. Rot: rotational.)

### 3.5 Conclusion

An accurate and fully automated algorithm was proposed to correct the in-plane motion artifacts in OCT esophageal imaging, and evaluated in a porcine esophageal model. The algorithm successfully revealed the otherwise distorted microstructure in the esophageal mucosa, such as microvasculature network and the layered structure.
Chapter 4. Optimizing the balloon

4.1 Introduction

The balloon is a key component in 3-D EOCT. The reason is that placing the tissue within the Rayleigh range of the probing beam is critical for high quality OCT images. Fig. 4.1(a) shows an OCT imaging probe in a naturally collapsed esophagus. The tissue is not always placed within the vicinity of the beam focus. The balloon in Fig. 4.1(b) centers the probe and expands the mucosal wall to the circumference near the beam focus. 3-D images can therefore be acquired from the mucosal wall where BE occurs.

Fig. 4.1 (a) The catheter probe without a balloon fails to always focus the beam into the tissue. (b) The balloon-based catheter probe consistently maintains the focus on the mucosa. Dash circles indicate the working distance of the probe.

There are at least two criteria to select the appropriate balloon design and parameters for BE imaging in clinical trial. The first criterion is safety. Balloons have been widely used for endoscopic dilation for the treatment of esophageal
strictures [146-148], photodynamic therapy [149-152] and RFA for BE treatment and palliation [153-156]. For example, The system (HALO\textsuperscript{360} system, B\textregistered RRX Medical, Inc, Sunnyvale, CA) in the RFA used balloons with fixed diameter. The balloons were inflated to a pressure of 0.5 atm for ablative energy delivery [155-156]. Recently, Suter et al. demonstrated 3D OCT scanning of human esophagus with a single-balloon (2.5 cm in diameter) catheter with a rigid balloon inflated [40]. However, the rigid balloon does not allow the catheter inserting through the accessory channel of the endoscope. The flexible balloon (see section 2.2.3) was utilized in our EOCT system for the purpose of endoscopic guidance. The balloon pressure should be as low as possible because the balloon diameter increases as the pressure raises. The lowest pressure should be sufficient to support the esophagus and dampen motion artifact. The second criterion for choosing the balloon is diagnostic efficiency. In our previous clinical trials, OCT images were acquired without balloon-tissue contact [90, 92]. Diagnostic features, such as layered structure, penetration depth and texture, were proposed for diagnostic criteria. However, it has been demonstrated that catheter pressure significantly altered the image appearance in colonic tissue [120]. The single-balloon design was commonly used in 3-D EOCT, which raised the question whether the balloon pressure/contact would alter the tissue appearance for dysplasia diagnosis. The utility of balloon should enhance, or at least not impair the efficiency of diagnosis. In addition, biopsy concordance should be provided as a gold standard to determine the diagnostic efficiency. Obtaining biopsy with an inflated balloon was not feasible. Laser marking has
been proposed to mark the suspicious region while performing OCT imaging [128]. The fiducial then guided the biopsy after the balloon deflated. A simpler approach will be obtaining the biopsy first and then identifying the divots in the 3-D OCT images.

In this chapter, a quantitative measurement was first conducted to understand the dependence of motion artifact on the balloon pressure. The purpose was to determine the minimal pressure required for stabilizing the single balloon and centering the probe. Second, the image features utilized for dysplasia diagnosis were compared between the single-balloon and double-balloon imaging schemes. This was simply a comparison in tissue appearance with a certain pressure and without pressure. The study was conducted in swine in vivo, therefore the features were only calculated in normal epithelium. If difference in image features were observed, it at least established the rationale to utilize the double-balloon design in the clinical trial. The feasibility of identifying divots in 3-D OCT images for biopsy concordance was also examined. These results will guide the appropriate balloon design and parameters for the clinical trial.

4.2 Methods

4.2.1 Image acquisition

The imaging system described in chapters 2 was utilized. The imaging procedure was similar to that in chapter 3. The ventilation and heart beating caused esophageal movement, which were potential origins of the motion artifact. Therefore, the ventilation rate (13 breaths/min) and heart rate (80 to 103
beats/min) were specifically monitored. Using the single balloon imaging scheme, the same cross-section was imaged for 25 seconds at pressure levels 30, 50, 70, 90, 110 and 130 mmHg in each of the 3 swine (2-D sequence). For each sequence, the images were ideally identical if no motion artifacts existed. When the pressure was lower than 30 mmHg, the esophagus could not be fully expanded. When the pressure was higher than 90 mmHg, the diameter of the esophageal lumen varied little. 130 mmHg was the maximal pressure provided by the pump.

Another dataset was acquired for feature comparison. 1 cm long esophageal segment was imaged with both the single-balloon and double-balloon schemes (3-D sequence). The pressure for single-balloon scheme was kept at 100 mmHg. For double-balloon images, the balloons were manually over-inflated so that the tissue between the balloons could be imaged near the beam focus. The pressure was approximately 250 mmHg. Seven swine were imaged totally.

In order to test whether biopsy divots were visible in OCT, a biopsy was first obtained from the mucosal wall by using a jumbo forceps with endoscopic guidance. Either the single-balloon or double balloon probe was inserted into the esophagus and placed close to the divot. A 3-D volumetric scanning of 1cm long segment, which covered the divot, was obtained and reconstructed for visualization.
4.2.2 Quantifying motion artifacts

The quantification process made use of the registration algorithm described in chapter 3. The algorithm only detected the in-plane motion artifact, therefore 2-D sequences was utilized for the motion artifact quantification. The algorithm was first applied to generate registered sequences. The first image in each sequence was considered to have no distortion. The middle A-scans of the N ROIs (for LBM) were back-tracked to find their positions in original sequences. The trajectories of these A-scans were generated. The trajectories were then decomposed into the radial and rotational components and analyzed by Fourier transform. Fourier transform of the trajectories enabled one to link the motion to respiration and heart beating. The frequency components correlated with the respiration were further confirmed by the sequence when the ventilation paused. The components of the motion trajectories corresponding to heart beat and respiratory motion were extracted from the frequency spectrum and inverse Fourier transformed. The motion magnitudes were defined as the peak-to-peak amplitude of the reconstructed motion divided by 2. One-way analysis of variance (ANOVA) [157] was applied to determine whether the mean magnitudes of each component were significantly different among six pressure levels.

4.2.3 Comparing image features

Different image features were utilized for human diagnosis [90] and computer-aided diagnosis [92-93]. Ref. [90] defined the diagnostic criteria for HGD to be the lack of layered structure and reduced penetration depth. Refs. [92-93]
quantified the speckle pattern, the stripe, and the single scattering coefficient. Among these five criteria, the layered structure in single-balloon and double-balloon images were compared qualitatively. The strip pattern was not observed in any of the swine esophageal images, and therefore not calculated. The penetration depth and the single scattering coefficient were highly correlated. A higher single scattering coefficient caused a less penetration depth. In this work, only the penetration depth and texture features were quantified and compared.

**Penetration depth**

The penetration depth was quantified by a simplified attenuation model. Assuming the tissue was homogeneous, the light propagated through the tissue and its intensity followed the Beer’s law [158], given by

\[ I_z = I_0 e^{-\alpha z} \]

where \( I_0 \) and \( I_z \) are the intensity at zero depth and a depth of \( z \) in the tissue. \( \alpha \) is the attenuation coefficient. The backscattering intensity which OCT detected (\( I_b \)) was directly proportional to the incident intensity (\( I_z \)). Because the OCT images were displayed in logarithmic scale, the backscattering profile was linearly decreasing and became

\[ \log_{10} I_b \propto \log_{10} I_z = \log_{10} I_0 - az \cdot \log_{10} e \]

When the light propagated beyond the single backscattering regime, OCT would not generate interferometric signal. The image became flat noise background. The backscattering profile could therefore divided into the tissue region and the
background, as shown in Fig. 4.2. This profile was fitted with two line segments, given by

\[ y_z = u(z_i - z) \ast (y_0 - az) + u(z - z_i) \ast (y_0 - az) \]  

(4.1)

where \( y_z \) and \( z \) are the image intensity and the depth, \( y_0 \) the intensity of the tissue surface (which is consider \( z=0 \)), -\( a \) the slope of the tissue profile and \( z_i \) the intercept of the tissue profile and background profile. \( u(z) \) is the step function, given by

\[ u(z) = \begin{cases} 
1 & \text{if } z \geq 0 \\
0 & \text{if } z < 0 
\end{cases} \]

The penetration depth was defined to be \( z_i \). Fitting equation 4.1 to the actual backscattering profile was formulated as an optimization problem, which sought to optimize three variables (e.g. \( a, y_0, \) and \( z_i \)) so that the fitting had the minimal mean square error.

Fig. 4.2 Illustration of the backscattering profile. The tissue surface was considered the zero depth. The image intensity in the tissue region (\( 0 < z < z_i \)) decreased linearly, and became flat noise in the background region. The penetration depth was defined to be \( z_i \).
The penetration depth was calculated in the 3-D sequences. For each of the 1 cm long esophageal segment, 5 cross-sectional images were randomly selected. Small regions of 1 mm in width were selected from 4 quadrants of each image. The tissue surface was manually segmented. A mean backscattering profile was then calculated by averaging the A-scans, considering the tissue surface to be zero depth for each A-scan. The averaging step suppressed the noise in the A-scans and generated a smooth backscattering profile for the fitting. For each swine, 20 representative mean backscattering profiles were generated for both single-balloon and double-balloon images. Equation 4.1 was fitted to each profile to calculate the penetration depth. Two-way ANOVA [157] was applied to determine whether the mean penetration depths were significantly different between the two imaging scheme. The reason to apply two-way ANOVA was that there were two major factors which might affect the penetration depth, e.g., individual swine and the balloon design.

Texture analysis
Image texture was analyzed with parameters defined based on variance and linear center-symmetric auto-correlation (CSAC) [92-93], listed in table 4.1. These parameters were defined in a small region of 3 by 3 pixels (Fig. 4.3(a)), given by:

\[ VAR = \left[ \frac{1}{8} \sum_{l=1}^{4} (g_l^2 + g_l'^2) \right] - \mu^2 \]
where $g_i$ is the intensity, and $\mu$ the local mean intensity. VAR is a measure of local gray-level variation and is the sum of BVAR and WVAR. BVAR is a measure of between-pair intensity variance. WVAR is a measure of within-pair intensity variance. SVR, the symmetric variance ratio between the within-pair and between-pair variances, is a statistic equivalent to the auto-correlation measure SAC. These 4 parameters measure the homogeneity of the intensity within the ROI. It can be expected that a stronger speckle pattern results in a higher VAR, and probably a higher BVAR and WVAR. SAC is a normalized gray-level invariant version of the texture covariance SCOV. It is invariant under linear gray-level shifts such as correction by mean and standard deviation [159]. SCOV and SAC detect the bright or dark lines that pass through the 3 by 3 neighborhood. Fig's. 4.3(b), (c) and (d) illustrate how the line pattern is quantified by SAC. In Fig. 4.3(b), the four center-symmetric pairs have difference values. SAC for this neighborhood equals -1. The center-symmetric pairs in Fig's. 3 (c) and (d) have identical values. The pairs with value 1 are bright lines and those with value 0 are
dark lines. The SAC’s for these two figures equal 1. SCOV and SAC quantify the distinct speckle boundaries, and therefore indirect indicators of the speckle pattern.

Table 4.1 Abbreviations of texture analysis measures

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Texture features</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR</td>
<td>Local variance</td>
</tr>
<tr>
<td>BVAR</td>
<td>Between-pair variance</td>
</tr>
<tr>
<td>WVAR</td>
<td>Within-pair variance</td>
</tr>
<tr>
<td>SVR</td>
<td>Variance ratio</td>
</tr>
<tr>
<td>SCOV</td>
<td>Gray level texture covariance</td>
</tr>
<tr>
<td>SAC</td>
<td>Normalized SCOV</td>
</tr>
</tbody>
</table>

![Fig. 4.3](image)

Fig. 4.3 (a) The region of 3 by 3 pixels used to calculate the texture parameters. $g_i$ and $g_i'$ are the intensities of the center-symmetric pair. (b) An example when each center-symmetric pair has different values. SAC equals -1 in this case. (c) &
(d) Examples when each center-symmetric pair has identical values. SAC equals 1 in both cases.

These six parameters were calculated in the 3-D sequences. For each of the 1 cm long esophageal segment, 5 cross-sectional images were randomly selected. Small regions of 1 mm in width were selected from 4 quadrants of each image. The epithelium was manually segmented. The analysis of each parameters was identical. The analysis of VAR is described as an example. VAR was calculated for each pixel in the epithelium. The speckle pattern within the region was considered consistent. Therefore, a representative VAR was calculated by taking the mean of VAR's for all pixels. For each of the seven swine, there were 20 representative VAR's for both single-balloon and double-balloon images. Two-way ANOVA was applied to determine whether VAR was significantly different between the two imaging schemes.

All the analysis work presented here was carried out using standard functions available in the image processing, optimization and statistics toolboxes in MATLAB (The MathWorks, Inc., Natick, MA).

4.3 Results

4.3.1 Tissue motion analysis

Representative radial and rotational motions of swine 2 at a balloon pressure of 90 mmHg (only the first 100 images in a 2-D sequence) are plotted in Fig. 4.4(a). The amplitude spectra of the Fourier transform are plotted in Fig. 4.4(b). The peaks at 0.22 Hz and 1.32 Hz correlate well with the respiration rate (0.22 Hz)
and heart rate (1.36 Hz). The 0.22 Hz peaks and their harmonics (at 0.44 Hz and 0.68 Hz) disappeared when the ventilation paused, confirming that they are caused by respiration. It is evident that the physiological movement in the living animal is the main origin of motion artifacts in balloon-stabilized EOCT esophageal images. The mean and standard deviation of both physiological motions (respiration and heart beat) in both the radial and rotational directions, at six levels of balloon pressure are plotted in Fig. 4.4(c). One-way ANOVA showed no significant difference in mean magnitudes at six pressure levels (p>0.05) with the exception of the rotational component of respiration. In the latter case, the mean magnitudes remained above 0.2 mm. Tukey analysis of the rotational component of respiration showed that that the means at 30 and 50 mmHg were significantly lower than the means at 90, 110 and 130 mmHg. The magnitude plot in Fig. 4.4(c) shows neither respiratory nor cardiac related motion artifacts were suppressed by increasing the balloon diameter/pressure.
Fig. 4.4 (a) After the trajectory was generated, it was decomposed into radial (in black) and rotational (in red) components (first 100 images) (b) The amplitude

(c)
spectra of (a). The frequency components at 0.22, 0.44 and 0.68 Hz were related to respiration. Those at 1.32, 2.62, 3.94 Hz were related to hearting beating. (c) The motion magnitudes at six pressure levels. There were no significant difference in mean magnitudes at six pressure levels (p>0.05) except the rotational component of respiration.

4.3.2 Layered structure

The layered structure in single-balloon (Fig. 4.5(a)) and double-balloon (Fig. 4.5(b)) images were compared qualitatively. In the single-balloon image, the natural surface morphology was smoothed by the balloon. The balloon appeared as a thin line on top of the mucosa. The weak scattering muscularis mucosa was identified. However, the blood vessels were compressed and became no longer visible as pressure increased. The lamina propria and submucosa layers were hardly identified without knowing the muscularis mucosal layer. In the double-balloon image, the lamina propria and submucosa appeared to be distinct layers of high scattering. The blood vessels were readily identified by both the oval shape of weak scattering and the shadow below. It should be noted that the single balloon image was obtained at balloon pressure of 100 mmHg. The layered structure was easier to be identified at lower pressure than at higher pressure.
Fig. 4.5 (a) Esophageal mucosal wall imaged with the single balloon. (b) Mucosal wall imaged with the double balloon. The layered structure and blood vessels in the double-balloon images are readily identified.

4.3.3 Penetration depth

An averaged backscattering profile is plotted in blue in Fig. 4.6(a) along with the image. The surface was considered zero depth. The red dash line segments indicates the fitting result of equation 4.1. The penetration depth in the image was indicated in white curve. The profile in the tissue region appeared more linear when the light penetrated deeper in the tissue. The reason was that the mucosa had a rich layered structure, which caused sudden changes in backscattering. The muscularis propria was mainly muscle with a homogeneous refractive index. All the penetration depths calculated from seven swine imaged with two balloon designs were plotted in Fig. 4.6(b). Two-way ANOVA showed
that the penetration depth has no significant difference between the two balloon designs (p=0.03).

Fig. 4.6 (a) The mean backscattering profile is plotted in blue. The corresponding images from which the profile was generated is shown at the bottom. The fitting line segments is plotted in red. (b) All penetration depths calculated from seven swine imaged with both single-balloon and double-balloon designs. Two-way ANOVA showed that the penetration depth has no significant difference (p=0.03).
4.3.4 Texture analysis

The epithelial speckle pattern in the two imaging scheme exhibit different characteristic. The epithelium appears more homogeneous in the single-balloon image. The speckle size is small and consistent. The speckle in the double-balloon image varies in size and well-defined by the clear boundary. The six texture parameters measured in seven swine with the two balloon designs are plotted in Fig. 4.7. Two-way ANOVA shows that all parameters exhibit a significant difference between the two balloon designs (p<0.001) except WVAR (p=0.59). Fig. 4.7(a) shows that VAR in the double-balloon images is higher, which is consistent with the observation that the epithelium in double-balloon image appears less homogeneous in Fig. 4.6. The double-balloon images have a higher SAC, which indicates that there are more line segments present in the double-balloon images. These line segments are probably the boundaries defining the speckle.
Fig. 4.7 Six texture parameters calculated from seven swine imaged by two balloon schemes. Two-way ANOVA showed that all parameters exhibit significant difference between the two balloon designs (p<0.001) except WVAR (p=0.59).
4.3.5 Visualization of the biopsy divot

The biopsy divot was visualized by both balloon schemes. Fig. 4.8(a) shows the divot imaged with the single-balloon scheme. The tissue surface is smoothed by the balloon. The epithelium layer becomes thinner as it is closer to the divot, and eventually disappears due to the lost of tissue. The muscularis mucosa is exposed and high backscattering signal is observed. The strong backscattering is probably due to the blood in the divot. Fig. 4.8(b) shows the divot imaged with the double-balloon scheme. The missing epithelium is clearly visualized. Fig. 4.8(c) is a 3-D reconstruction of the divot imaged with the single-balloon scheme. The strong backscattering can be visualized from the en face view.
Fig. 4.8 (a) The divot imaged with the single-balloon scheme. The epithelium is not visible at the divot due to the lost of tissue. (b) The divot imaged with the double-balloon scheme. (c) 3-D reconstruction of the divot imaged with the single-balloon scheme. The strong backscattering in the en face view is probably due to the blood from the muscularis mucosa at the divot.

4.4 Discussion

In this chapter, the double-balloon and single-balloon designs with different specification were tested for imaging quality analysis. It has been demonstrated
in the chapter 3 that imaging with the double-balloon design resulted in significant motion artifacts. The single balloon contributed to the suppression of the excessive motion artifacts. In this chapter, it has been shown that a single balloon at lower pressure level can suppress the motion artifact as efficiently at higher pressure level. Therefore, lower pressure level is preferred to avoid perforation. It should be noted that the initial diameter of the low durometer balloon is 18 mm for swine imaging, which is the average diameter for swine esophagus [96, 98]. Balloon with an initial diameter of 25 mm should be used to match the size of human esophagus [40]. A pressure around 70 mmHg is a reasonable choice in the BE clinical trial in terms of motion artifact suppression. On the other hand, balloon pressure/ contact eliminates the layered appearance and alters the speckle pattern. It does not necessarily mean that the single balloon will deteriorate the diagnostic performance of 3-D EOCT for BE dysplasia. However, the diagnostic criteria for both human observer and computer-aided diagnosis should be re-investigated for single-balloon imaging. The double-balloon design provides a similar imaging circumstance to the previous TD-OCT study [90, 92-93]. Therefore, similar diagnostic criteria should be applicable to the double balloon images. Both imaging schemes should be included in the clinical trial so that the diagnostic performance for dysplasia can be compared.

A practical pressure range of 30 to 130 mmHg was studied to demonstrate that the single balloon design was not able to suppress motion artifact completely. It was noticed that the actual diameter of the balloon in vivo was smaller than that in the open air because of esophageal tension. Even with balloon technology, the
magnitude of the physiological motion artifacts (hundreds of microns) was comparable to the epithelial thickness (about 200 microns on average in our case) or the dimension of glandular structure [102, 160-161]. The 3-D reconstruction was therefore distorted. The information in the tissue that could be extracted was limited. Post-acquisition motion artifact correction method, similar to the one proposed in chapter 3, is necessary. Understanding the motion artifact may also facilitate the clinical application of laser marking technology [128], which uses a high-power laser beam to ablate abnormal tissue identified by OCT screenings for later biopsies. The radial esophageal motion is on the order of the Rayleigh range of the laser beam, and the rotational motion is much larger than the spot size at the focus. These motions should be taken into account when optimizing the exposure time in a clinical setting.

The reason why the layered architecture was less visible in single balloon images is not fully understood. It has been well studied that the contrast in OCT originates from the difference in refractive indices [49], the optical property of specific material. The pressure can not alter the composition of the tissue, but may possibly rearrange the alignment microscopically. The backscattering from a misaligned layer can be easily attenuated by the speckle nature of OCT imaging. The previous study of the pressure influence on colonic tissue demonstrated that increasing pressure allowed for the better visualization of the layered structure [120], which seemed controversial to this study. In fact, the penetration depth of the previous EOCT imaging probe was about 0.5mm, which could not even penetrate the colonic epithelium. The reason for the low penetration depth was
that the beam converged and diverged quickly in a short working distance probe.

By applying a large amount of pressure (about 1N on a few mm², on the order of thousands mmHg), the layers were compressed thinner so that the submucosa and muscularis propria were brought closer to the probe. It should be noted that the balloon pressure was much lower than the pressure applied to the colonic tissue. Also, the probe for esophageal imaging has a long Rayleigh range. The beam penetrates deep in the muscularis propria. Imposing higher pressure will not reveal any layers in the muscularis propria.

The reason why the speckle pattern altered when pressure applied is not fully understood either. The pressure might alter the arrangement of the squamous epithelial cell which affected the speckles. It should be noted that the speckle size and intensity also depends on the position of the beam focus with respect to the tissue. Tissue at the vicinity of focus could easily generate strong speckle pattern. The double-balloon images were obtained between two balloons where the lumen diameter was smaller than the balloon diameter (Fig. 2.1). In order to place the tissue around the beam focus, the double balloons had to be over-inflated manually so that the tissue could be imaged by the same probe used for the single balloon. However, the distance from the tissue surface to the probe had a significant variation at different rotational angle without balloon supporting (on the order of millimeters). This variation in the single-balloon scheme was on the order of hundreds of microns, which could be neglected. Therefore, the tissue was not images at the same position along the probing beam in the two balloon scheme, which brought uncertainty to the statistics of speckle pattern. Over-
inflation is not preferred in BE clinical trial. For human BE imaging, the diameters for the double balloons should still be 25 mm. But the working distance of the probe should be smaller. 7.5 mm and 9 mm probes are existing designs and planned to be utilized in the clinical trial.

Biopsy divot was visualized in both imaging schemes. The divot was identified by the loss of epithelial layer, which was based on the layered structure in the mucosal wall. The lack of layered structure has been a diagnostic criterion for dysplasia in BE [90-91]. Therefore, it is unlikely to identify the epithelium or muscularis mucosa in dysplastic mucosa. Identifying the divot in dysplastic mucosa may be challenging. The accumulation of blood or surface morphology may be potential signatures for the divot, but requires a clinical trial for further investigation.

4.5 Conclusion
The single-balloon and double-balloon designs were examined for motion artifact suppression and their influence on diagnostic features for dysplasia. In future clinical trial, balloons of 25 mm in diameter should be inflated to 70 mm Hg. Image probes with a working distance of 12.5 mm should be utilized for single-balloon imaging. Probes with working distances of 7.5 mm and 9 mm should be utilized for double-balloon imaging.
Chapter 5. Clinical trial

5.1 Introduction

It has been reported that OCT was able to detect SIM for the purpose of BE screening [88-89], diagnose dysplasia for BE surveillance [90-93] and identify subsquamous Barrett's epithelium [94] or glandular structure [38]. Dysplasia diagnosis remains the primary goal for the OCT clinical trial at Case Western Reserve University and University Hospitals Case Medical Center (UHCMC), because the current BE management depends on the degree of dysplasia. Isenberg et al. demonstrated that TD-OCT detected dysplasia with 68% sensitivity; 82% specificity; and 78% diagnostic accuracy when using histology as the standard [90]. Evans et al. reported the sensitivity and specificity for diagnosing intramucosal carcinoma/HGD were 83% and 75%, respectively [91]. Qi et al. developed a computer-aided diagnosis algorithm to differentiate dysplastic tissue from non-dysplastic tissue, and achieved a sensitivity of 82%, specificity of 74%, and accuracy of 83%. As mentioned in chapter 1, these studies utilized TD-OCT system which obtained images from a small region in BE. They did not fully explore the potential of using OCT in BE screening and surveillance.

The hypothesis of this clinical trial is that with newly-developed SD-OCT and balloon technology, OCT can provide 3-D comprehensive esophageal images, which may potentially improve the detection of dysplasia in BE and eliminate
sampling error. The first aim is to acquire 3D volumetric images of the esophageal wall with the balloon-based catheter probe in BE patients undergoing endoscopic surveillance, and to determine if volumetric OCT has the potential to detect dysplasia or early adenocarcinoma. The second aim is to correlate 3-D EOCT imaging with histopathology of the esophageal epithelium. This study is expected to test the feasibility of 3-D EOCT clinical trial, identify potential image features for dysplasia diagnosis and obtain the EOCT images with the single-balloon and double-balloon designs using biopsy divots for histopathologic marker.

5.2 Methods
The original protocol and amendment of the clinical trial was approved in February and November, 2010 by the Institutional Review Board (IRB) at UHCMC, Cleveland, OH (IRB protocol number 11-09-43). The amendment was made to use the appropriate size of the balloon and working distance of the imaging probe for human esophagus. The methods reported in this section are from the amendment.

5.2.1 Subject recruitment
Patients undergoing endoscopic surveillance for BE at the UHCMC will be eligible for inclusion in the study. The exclusion criteria are: 1) patient under 18 years of age; 2) patient unable to give informed consent themselves; 3) patient who is pregnant; and 4) patient who has a contraindication to endoscopy.
5.2.2 Imaging system

The system described in chapter 2 is the system described in the amendment. A marker attached to the proximal end of the balloon was added to indicated the projected length of the scanning probe. The position of the marker was 5 cm away from the probe. The endoscopist can estimate and adjust the position of the probe by looking at the marker via endoscope. The probe with 12.5mm working distance was utilized for single-balloon imaging. All the balloons were for single use only.

5.2.3 Image acquisition

All procedures are performed by or under the direct supervision of an attending gastroenterologist with experience in endoscopic surveillance of BE as well as OCT imaging. After obtaining informed consent, patients are sedated in standard procedure for endoscopy. An initial endoscopic evaluation is performed with a 2-channel endoscope (GIF-2T240, Olympus Optical, Inc., Tokyo, Japan). The extent of BE is noted including any focal mucosal lesions that raise a suspicion of dysplasia or cancer.

After the endoscopic image is taken, single-balloon or double-balloon probe is deflated and inserted through the accessory channel such that the balloon projects beyond the tip of the endoscope. The balloon is positioned endoscopically just below the distal end of the Barrett’s mucosa and expanded, and the tip of the endoscope is placed against the balloon. The imaging probe is
rotated at 10 revolutions per second and pulled at 1 cm per minute for 1 minute within the outer sheath. Thus, 3-D EOCT images without biopsy divots will be acquired. Image acquisition takes less than 2 minutes for a 1 cm segment of BE.

After EOCT imaging, standard care biopsies are obtained. It is standard practice to place the biopsies from 4 quadrants of the same circumference of Barrett's epithelium in a single bottle of fixative for subsequent histologic evaluation. Biopsies of focal lesions are usually placed in separate bottles. However, to correlate histology with endoscopic and EOCT findings, only one biopsy will be obtained from the circumference. This biopsy is kept in a separate bottle. After that, another 3-D EOCT scanning will be taken at the same 1 cm segment. This time the biopsy site should be imaged. It is assumed that the tissue near the biopsy divot is similar to that of biopsy divot due to field effect present with BE dysplasia. By observing the region near the biopsy divot in OCT images, the OCT images can be correlated with histology result. The OCT probe is then extracted and the rest of biopsies are taken according to the standard Seattle protocol. It was anticipated that about additional 15 minutes for obtaining the OCT images. However, if the OCT image acquisition procedure takes more than 15 minutes, the OCT image acquisition is terminated. The standard biopsy procedure is continued according to the Seattle protocol.

5.2.4 Pathology

Pathologists diagnoses the absence or presence of dysplasia for the specimen obtained, per standard protocol. Pathologists are blinded with respect to
endoscopic and 3-D volumetric EOCT findings. Specimens are evaluated according to standard criteria [162] and diagnosed as: 1) no dysplasia; 2) indefinite for dysplasia; 3) low grade dysplasia; 4) HGD; and 5) intramucosal cancer. A digital image is saved of representative pathologic sections of the specimens. The diagnosis of the pathologist is the clinical diagnosis used in the care of the patient, as appropriate for standard clinical care. Endoscopists blinded to the clinical and pathologic information then rate each OCT image using the same criteria. If marked discrepancies occur between endoscopic, EOCT and pathologic diagnoses, the coded digital pathologic images are reviewed by a separate pathologist. The final diagnosis for purposes of the study is determined by majority vote.

5.2.5 Statistics

The performance of EOCT with respect to the detection of the presence of dysplasia or malignancy is based on the pathologic diagnosis as the criterion standard. With the biopsy site as the experimental unit, it is possible to calculate values of sensitivity, specificity, positive predictive value, negative predictive value, accuracy, and their 95% confidence intervals by the modified Wald Method [163].
5.3 Results

5.3.1 Patient enrollment

All the patients enrolled are listed in table 5.1 with the result of previous surveillance, balloon type and OCT procedure. The entire procedure was finished in patients 3, 4 and 9. OCT imaging in the first patient took much longer than expected and exceeded 15 minutes. EAC, polyps and stricture were identified in patient 2, 7 and 10, respectively. Expanding the esophagus with an inflated balloon might cause bleeding. Therefore the procedure was halted. Biopsy was not obtained from patient 5 because EAC was diagnosed. In patient 6, HGD occurred at the gastro-esophageal junction (GEJ). The contraction of the lower esophageal sphincter made the positioning and inflation of the balloon very difficult. Malfunction occurred during the procedure in patient 8. The images had very low SNR but the reason is still not understood.
### Table 5.1 Patient enrollment

<table>
<thead>
<tr>
<th>Patient number</th>
<th>Previous surveillance</th>
<th>Balloon type</th>
<th>OCT procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HGD</td>
<td>Single</td>
<td>Not completed, not enough time for biopsy</td>
</tr>
<tr>
<td>2</td>
<td>EAC</td>
<td>Single</td>
<td>Halted, possibility of causing bleeding</td>
</tr>
<tr>
<td>3</td>
<td>HGD</td>
<td>Single</td>
<td>Completed, biopsy not visualized</td>
</tr>
<tr>
<td>4</td>
<td>HGD</td>
<td>Single</td>
<td>Completed, biopsy not visualized</td>
</tr>
<tr>
<td>5</td>
<td>BE</td>
<td>Single</td>
<td>Completed, biopsy not obtained</td>
</tr>
<tr>
<td>6</td>
<td>HGD</td>
<td>Double</td>
<td>Failed, balloons unable to expand the gastro-esophageal junction</td>
</tr>
<tr>
<td>7</td>
<td>BE</td>
<td>Double</td>
<td>Halted, polyps and bleeding</td>
</tr>
<tr>
<td>8</td>
<td>BE</td>
<td>Single</td>
<td>Failed to obtain images</td>
</tr>
<tr>
<td>9</td>
<td>BE</td>
<td>Single</td>
<td>Completed, biopsy not visualized</td>
</tr>
<tr>
<td>10</td>
<td>RFA follow-up</td>
<td>Single</td>
<td>Halted, esophageal stricture exists</td>
</tr>
</tbody>
</table>

#### 5.3.2 Images of normal and non-dysplastic tissue

Fig. 5.1 shows the images of normal and non-dysplastic tissue from patient 10. Fig. 5.1(a) was an rectangular view of the proximal esophagus. The typical normal esophagus has a thick and homogenous layer of epithelium with consistent thickness. The layered architecture can be clearly visualized, especially the boundary between epithelium and lamina propria. In non-
dysplastic Barrett's esophagus (Fig. 5.1(b)), the epithelium becomes thinner. The surface is no longer smooth but exhibits the finger-like topology highly assembling intestinal metaplasia. The layered architecture is less obvious but visible.

Fig. 5.1 (a) Normal esophageal image from patient 10. The typical normal esophagus has a thick and homogenous layer of epithelium with consistent thickness. The layered architecture can be clearly visualized, especially the boundary between epithelium and lamina propria. (b) Non-dysplastic Barrett's esophageal image. the epithelium becomes thinner. The surface is no longer smooth but exhibits the finger-like topology highly assembling intestinal metaplasia. The layered architecture is less obvious but visible. (SE: squamous epithelium; LP: lamina propria, MM: muscularis mucosa, SM: submucosa; MP: muscularis propria. Bars: 1mm in two directions. Rad: radial. Rot: rotational.)

5.3.3 Images of dysplastic tissue

Layered structure
Fig. 5.2 shows images obtained from a segment (patient 3) where both BE (possibly HGD) and normal esophagus existed. The layered structure in BE (Fig. 5.2(a)) is no longer visible within the entire image, which agreed with our previous observation in HGD. On the other hand, the lamina propria is clearly identified in Fig. 5.2(b) as expected. The difference in penetration depth in the two images is not obvious. Fig's. 5.2(c) and (d) are the enlarged views of the dash squares in Fig's. 5.2(a) and (b), respectively. The difference in layered structure can be visualized.
Fig. 5.2 (a) A cross-sectional image of human BE (possibly HGD), where the layered structure cannot be identified. (b) A cross-sectional image of normal esophagus, where the epithelium (EP), lamina propria (LP) and muscularis mucosa (MM) are clearly identified. (c) An enlarged view of the dash square in (a). (d) An enlarged view of the dash square in (d). (Bars: 1 mm in two directions.)

**Surface morphology**

Surface morphology is a characteristic feature observed that may potentially differentiate normal esophagus and HGD. Fig. 5.3(a) is an image of normal esophagus obtained from patient 5 when the balloon was not in full contact with
the tissue. Note that the surface is smooth compared to the finger-like morphology in the BE images (possibly HGD, patient 1) in Fig. 5.3(b). The detailed surface morphology was not visible when the balloon compressed the tissue. The finger-like pattern has a width of approximately a few hundred microns and highly resembles intestinal villi, which is the abnormal architecture in BE [164]. A 3-D reconstruction in the vicinity of Fig. 5.3(b) is shown in Fig. 5.3(c). The raw data was first registered by using the motion artifact correction algorithm in chapter 3, so that the microstructure in the mucosa was well aligned. The image of the balloon is removed to expose the tissue surface. The finger-like pattern in Fig. 5.3(b) seems more likely to be a cross-sectional view of the elongated folding of the epithelium (indicated by arrows in Fig. 5.3(c)). The longitudinal folding was commonly observed in 3-D reconstruction of BE. However, the reason to form such folding is not understood at this point.
Fig. 5.3 (a) A cross-sectional image of the normal esophagus with a smooth surface. The balloon was not in fully contact with the tissue. (b) A cross-sectional image of BE with finger-like pattern on the surface. (c) 3-D reconstruction of (b) suggests that the finger-like pattern may be a cross-sectional view of the longitudinal folding (indicated by arrows). (Bars: 1mm in two directions. L: longitudinal. Rot: rotational.)

**Glandular structure**

Small glandular structure close to the tissue surface is also observed. Fig. 5.4(a) is an *en face* view of the 3-D reconstruction in Fig. 5.3(c). The *en face* view was
at a depth of 300 μm from the tissue surface. The dark region at the top middle of the image is the air gap between the tissue surface and the balloon. Yellow arrows indicates representative BE glands. Compared to normal esophageal glands in the muscularis mucosa, these glands are too shallow (only a few hundred microns from the tissue surface) and too small (approximately 100 μm in diameter). Fig. 5.4(b) is a cross-sectional view of Fig. 5.4(a) indicated by the red arrows. The corresponding glands are also indicated by yellow arrows.

Fig. 5.4 (a) An en face view of Fig. 5.3(c) showing the glandular structure. (b) The corresponding cross-sectional view of the red arrows in (a). Yellow arrows indicate the corresponding Barret's glands found in the two views. (Bars: 1mm to two directions. L: longitudinal. Rot: rotational.)
**Biopsy concordance**

The biopsy divots in patients 3 and 4 were not identified. Therefore, no biopsy concordance is found at this point.

**5.4 Discussion**

Patients 1 through 5 were imaged with single-balloon design but a probe with 9 mm working distance. Human esophagus has a diameter of 25 mm on average. Therefore, the tissue was not imaged at the beam focus for those patients. The images did not have the highest resolution and largest penetration depth. Texture analysis was not reliable because the speckle pattern was not observed. The 12.5 mm working distance probe was proposed in the amendment after these patients were imaged. Images of better quality will be acquired and may provide more details for dysplasia diagnosis.

The biopsy divots were not identified in the images from patients 3 and 4. One of the possible reason was that the divot was flatten by the balloon. Without the layered structure, the loss of epithelium was difficult to be identified. Another possibility was that the segment imaged did not cover the divot. When the biopsy was obtained, the endoscope was pulled back 5 cm to make space for balloon insertion. The divot was no longer visible by the endoscope. Therefore, positioning the balloon close to the divot was difficult. The balloon marker which indicated the 5 cm insertion length was proposed only in the amendment. The purpose of the marker was to felicitate balloon positioning. Unfortunately, no
patients have been enrolled to test the new method since the amendment was approved.

Placing the double-balloon at the GEJ was challenging, which showed a limitation of the double-balloon design. When imaging the GEJ, only one balloon stayed in the esophagus. The contraction of the lower esophageal sphincter made the tissue collapse onto the outer sheath between the two balloons, which was too close for the probe to image. The single balloon that can expand the GEJ may be the only choice for such cases.

The expected layered appearance in both normal and Barrett's esophagus was observed, as well as new features including surface morphology and glandular structure. The lamina propria has a much higher scattering coefficient than normal epithelium from the previous swine study. The lost of layered structure in dysplastic tissue observed in Ref. [90] was probably due to the increase of scattering coefficient of the dysplastic epithelium. It has been reported that dysplastic cells in cervical epithelium has a much higher scattering coefficient than normal cells in squamous epithelium [165-166]. The boundary between dysplastic epithelium and lamina propria may be weaker due to the reduced contrast in backscattering. The surface morphology observed in the BE images raised another reason for using the double-balloon design for further investigation. When using the single balloon, the tissue was usually compressed by the balloon. Double-balloon design allows for the visualization of the surface to investigate whether it contains diagnostic information for dysplasia. The number of glands in BE has been proposed as an indicator of eradication when
using RFA for dysplasia treatment [144-145]. Visualization of the glands with the OCT system allows for investigating its potential as a tool for monitoring the eradication of BE in real time and detecting residual Barrett's epithelium. The motion artifact algorithm described in chapter 3 accurately registered the otherwise mis-aligned small glands in 3-D, which may potentially facilitate automated gland counting and analysis.

5.5 Conclusion

Clinical study has been carried out to test the feasibility of the 3-D EOCT system for human esophageal imaging. Differences in layered structure, surface morphology and glandular structure between normal and Barrett's esophagus were observed. Amendment was made to improve image quality, which required further examination in clinical trial.
Chapter 6. Summary and future plan

6.1 Summary

The dissertation presents the step-by-step work from initially building the 3-D EOCT system for esophageal imaging, to eventually utilizing the system to conduct the clinical trial. The work was motivated in chapter 1 by the unmet clinical need to eliminate sampling error in BE surveillance. The motivation was based on the success that OCT obtained interpretable images of GI mucosal microstructure, differentiate GI mucosal types and detect dysplasia in BE. The advent of FD-OCT technology enables rapid imaging speed which allows for comprehensive scanning of the long segments of BE in a clinical setting. It is expected that utilizing such a system in further clinical trial may eliminate the sampling error and determine the role of OCT in BE surveillance.

Chapters 2 and 3 describe the hardware and software specifically designed for 3-D EOCT. In chapter 2, the sample arm with the rotary-joint-pullback unit and balloon-based catheter were the main innovations for the capability of 3-D imaging. The rotary-joint-pullback unit generated the helical scanning pattern. The balloon-based catheter allowed for endoscopically accessing the expanded esophagus, Both single-balloon and double-balloon imaging schemes allowed for the investigation on the balloon influence on esophageal imaging. Miniature optical probe delivered a well-shaped Gaussian beam with a long working distance, and collected backscattering photons that carried signals. The EOCT
system was capable of obtaining volumetric images with high resolution and sufficient imaging range. Motion artifacts existed in volumetric images, which was the reason why the automated registration algorithm was proposed in chapter 3. The algorithm was based on the similarity between successive cross-sectional images, and utilized the previous frame as an reference to correct the current frame. Motion artifacts were decomposed into radial and rotational components, which were detected by surface detection and local block matching, respectively. The algorithm successfully corrected the motion artifacts in single-balloon images. The high quality 3-D EOCT images represented the accurate tissue structure after motion artifact correction, which eventually fulfilled the role of comprehensive imaging. Unfortunately, the motion artifacts in double-balloon images were so severe that the registration algorithm had limited success.

Animal study was necessary to test the feasibility, safety and efficiency of the imaging system. The feasibility of 3-D imaging was demonstrated in swine in vivo in chapter 2, and the registration algorithm was evaluated in images in vivo in chapter 3. The balloon designs were discussed in terms of safety and diagnostic efficiency for clinical trial in chapter 4. The relation between the balloon pressure and motion artifacts was investigated. The result showed that a low pressure level was sufficient to suppress motion artifact with a low risk of perforation. Images appearance such as layered architecture and texture feature was significantly influenced by balloon pressure/contact, which suggested that both single-balloon and double-balloon design should be tested in clinical trial for comparison of diagnostic efficiency.
Human clinical trial is challenging. The first cases in the clinical trials were reported in chapter 5. The feasibility of 3-D EOCT imaging system was further demonstrated in BE patients. 7 patients have enrolled and the procedure was completed in 2 patients. The lack of layered structure in BE tissue was observed as expected. The size and distribution of BE glands was identified, which might potentially be new features for BE treatment evaluation. Surface morphology was also new features which might potentially benefit dysplasia diagnosis. Imaging protocol was revised for utilizing a longer working distance probe and positioning the probe close to the biopsy divots for histological concordance for further clinical trial.

The presented work did not answer all questions about esophageal OCT imaging, but it provides a platform allowing for comprehensive imaging of BE in high quality. The technology has the potential to answer critical questions about what role OCT can play in dysplasia diagnosis during BE screening/surveillance as well as BE management.

6.2 Future plan

The work in the dissertation sets the path to a number of innovations in the future. Top goals discussed in the following sections can be divided into technical and clinical categories.

6.2.1 Technical plans

The dog-bone balloon
The dog-bone balloon is a single balloon with a dog bone shape. It allows for adjusting balloon contact and pressure with a smaller diameter at the middle section and larger diameters at both ends. The idea of such a balloon is an further extension of the double-balloon design. The two ends of the dog-bone balloon stabilize the catheter and center the balloon. The diameter and the length of the middle section will be carefully determined so that features of different tissue are preserved or enhanced. This design seeks to reduce the balloon influence on tissue features by imposing minimal pressure, and meanwhile suppress motion artifacts which severely distort the microstructure in the double balloon design.

*Ultra-high axial and transversal resolution*

In an OCT system, the axial and transversal resolutions are determined by the bandwidth of the light source and the numerical aperture of the objective lens, respectively. The ultra-broad band light source will be incorporated in an EOCT configuration for the first time to achieve 3 µm axial resolution, improve the image quality and visualize finer structural details [126]. EOCT scanning probes, utilizing multiple gradient index (GRIN) lenses instead of a single GRIN lens design, will be examined to increase the transversal resolution of the system [167].

*Spectral-domain EOCT with full imaging range*
As mentioned in chapter 1, only half of the axial range can be used for imaging in a typical OCT system, because the conjugate images of the negative frequency after Fourier transform may overlap with the images of the positive frequency. Full imaging range is desirable because it will be easier to adjust the axial position of the tissue. Also the tissue can be placed near the zero frequency where the image will not significantly affect by the sensitivity falloff of the spectrometer. Several methods have been proposed to solve this problem on a bench-top SD-OCT system and the feasibility will be tested in the spectral-domain EOCT system [168-170].

Computer-aided diagnosis

New computer-aided diagnostic algorithms based on 3D reconstruction will be developed to segment the esophageal mucosa and stage dysplasia. The image features with diagnostic power will be chosen and quantified from human interpretation experience and previous reports [90, 92-93]. Algorithms will also be investigated to identify BE glands and evaluate residual BE in post-RFA tissue. The result of diagnosis and evaluate will be mapped in the above 3D reconstruction for the reader’s reference.

6.2.2 Clinical plans

Dysplasia diagnosis using the 3-D EOCT

Further clinical trial will lead to a new understanding whether the 3-D EOCT improves dysplasia diagnosis in BE surveillance. The 3-D EOCT image data will
be acquired from patients undergoing BE surveillance following the amendment in chapter 5. Diagnostic criteria previously reported will be re-examined, and new 3D image features will be investigated as potential criteria. Diagnosis of different stages of dysplasia will be made by at least two clinicians based on solely EOCT images or correlated with endoscopic snapshots. The performance of diagnosis, especially the sensitivity and specificity of differentiating HGD from other stages, will be the evidences whether the comprehensive EOCT improves dysplasia diagnosis in BE surveillance.

Ablated tissue characterization and identification

3-D EOCT images of esophagus before and after RFA will be obtained from human patients. The image feature will be carefully analyzed by experts and computers to identify the potential indicators of ablated tissue. The performance of utilizing EOCT to differentiate whether or not the tissue is from the ablation zone will be evaluated. This evaluation serves as the first step to eventually understand whether 3-D EOCT can be utilized to identified untreated tissue within the post-RFA zone and evaluate the BE eradication.

Influence of pressure and contact of the balloon

The influence of balloon on dysplasia diagnosis will be understood by comparing the single-balloon and double-balloon images. These features are layered structure, surface topology, scattering coefficient, penetration depth, center-symmetric auto-correlation (CSAC) parameters, etc [90, 92-93]. The influence of
balloon contact/pressure on these features will be analyzed to determine their diagnostic powers in the case of balloon-based catheters. More importantly, the knowledge of the influence will provide the imaging basis for the new balloon design optimized for dysplasia detection.
Bibliography


<table>
<thead>
<tr>
<th>Reference</th>
<th>Authors</th>
<th>Title</th>
<th>Journal</th>
<th>Pages</th>
</tr>
</thead>
</table>


