

Office-based dynamic imaging of vocal cords in awake patients with swept-source optical coherence tomography

Lingfeng Yu

Gangjun Liu

University of California, Irvine
Department of Biomedical Engineering
Irvine, California 92697

and

University of California, Irvine
Beckman Laser Institute
Irvine, California 92617

Marc Rubinstein

University of California, Irvine
Beckman Laser Institute
Irvine, California 92617

and

University of California, Irvine Medical Center
Department of Otolaryngology–Head and Neck Surgery
Orange, California 92868

Arya Saidi

University of California, Irvine
Department of Biomedical Engineering
Irvine, California 92697

Brian J. F. Wong

University of California, Irvine
Department of Biomedical Engineering
Irvine, California 92697

and

University of California, Irvine
Beckman Laser Institute
Irvine, California 92617

and

University of California, Irvine Medical Center
Department of Otolaryngology–Head and Neck Surgery
Orange, California 92868

Zhongping Chen

University of California, Irvine
Department of Biomedical Engineering
Irvine, California 92697

and

University of California, Irvine
Beckman Laser Institute
Irvine, California 92617

1 Introduction

Laryngeal carcinoma is one of the most common primary head and neck tumors. The cardinal symptom of early laryngeal cancer is hoarseness, but because this complaint is relatively innocuous; laryngeal cancer often goes undiagnosed for many months, and referral to an otolaryngologist may be delayed up to nine months following the onset of symptoms.

Abstract. Optical coherence tomography (OCT) is an evolving non-invasive imaging modality that has been used to image the human larynx during surgical endoscopy. The design of a long gradient index (GRIN) lens-based probe capable of capturing images of the human larynx by use of swept-source OCT during a typical office-based laryngoscopy examination is presented. *In vivo* OCT imaging of the human larynx is demonstrated with a rate of 40 frames per second. Dynamic vibration of the vocal folds is recorded to provide not only high-resolution cross-sectional tissue structures but also vibration parameters, such as the vibration frequency and magnitude of the vocal cords, which provides important information for clinical diagnosis and treatment, as well as fundamental research of the voice itself. Office-based OCT is a promising imaging modality to study the larynx for physicians in otolaryngology. © 2009 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3268442]

Keywords: larynx; biomedical imaging; early diagnosis; laryngeal cancer; optical coherence tomography.

Paper 09191PR received May 12, 2009; revised manuscript received Aug. 31, 2009; accepted for publication Sep. 24, 2009; published online Dec. 2, 2009. This paper is a revision of a paper presented at the SPIE conference on Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XIII, January 2009, San Jose, California. The paper presented there appears (unrefereed) in SPIE Proceedings Vol. 7168.

Accurate clinical diagnosis and treatment of early-stage laryngeal cancer, based on an initial consult, is extremely difficult for a physician to determine even with proper imaging techniques, and a biopsy is required to differentiate between benign, premalignant, and malignant pathologies. Conventional laryngeal examination can be performed using a laryngeal mirror and flexible fiber-optic or rigid laryngoscopes (with or without videostroboscopy) to achieve a two-dimensional (2-D) view of the laryngeal structures.

Address all correspondence to: Lingfeng Yu, University of California, Irvine, Department of Biomedical Engineering, 305 Rockwell Engineering Center, Irvine, California 92697. Tel: 949-824-8975; Fax: 949-824-8413; E-mail: yulingfeng@gmail.com.

Although it is difficult to differentiate among the wide spectrum of diseases ranging from chronic laryngitis to pre-malignant and malignant lesions, the lack of basement membrane integrity is a key feature of early invasive cancer of the vocal cords. The basement membrane is a thin layer of collagen and other proteins that the surface epithelial cells rest on, and it is the dividing line between the epithelium and the larger layers of the tissue such as the *lamina propria*. Therefore, early laryngeal carcinoma can be diagnosed quickly if the basement membrane can be visualized. Currently, there is no reliable noninvasive or nonoperative method available for surgeons to make the diagnosis of laryngeal cancer without a biopsy. Using current flexible fiber-optic techniques, the endoscopic yield in terms of diagnosing sensitivity and specificity for visible lesions in patients is very low. Biopsies of the vocal cord aimed at diagnosing cancer require a full-thickness excision of superficial epithelium, basement membrane, and connective tissues. These biopsies may have a detrimental effect on the patient's vocal cord vibration and ultimately lead to a permanent change in voice. Repeated biopsies are common in order to ascertain a definite diagnosis and will bring even higher risks. The preceding difficulties demonstrate the need for improved noninvasive diagnostic technology, as well as improved abilities to determine margins and to perform definitively safe biopsies on patients with clinically suspected larynx malignancies.

Optical coherence tomography (OCT) is a noninvasive medical imaging method based on the principle of low-coherence interferometry.¹ Low-coherence interferometry is a system where a low coherent light source is split into two identical beams, one for reference purposes and the other for the actual imaging information (sample arm). When the reference and sample arm are the same length, any difference in the backscattered light from the sample should provide information about subepidermal tissue. OCT was developed for the exact purpose of performing *in vivo* cross-sectional tomography imaging of tissue structure and composition with high imaging speed and resolution. It has become a powerful tool for medical diagnostics.²⁻⁴ Recently Luerssen et al. reported *in vivo* OCT imaging of vocal folds in a contact mode with local anesthesia.⁵ Our group also reported an office-based laryngeal time-domain OCT imaging device.⁶ A rigid laryngoscope served as a platform to which a second device could be attached to perform simultaneous OCT imaging. However, the scanning mechanism was too slow at a rate of ~ 1 frame per second. A gradient index (GRIN) lens-based probe was also developed for office-based laryngeal imaging at a rate of ~ 8 frames per second in conjunction with a spectral domain OCT system.⁷ However, one of the biggest challenges with an office-based OCT laryngeal imaging device is the movement of the patient's head and the physician's hand in the examination process. The latter is further levered by the cantilevered design of the probe. The relative movements between the vocal cords and the probe tip can easily exceed several millimeters, thereby shifting the images outside the OCT imaging window (or A-scan imaging depth). Also, the working distance (in this case, the distance from the probe tip to the vocal cords) is different from one patient's anatomy to another; physicians have to practice their ability to adjust the working distance while holding the probe steady to capture an OCT image. This is no simple task and requires the full atten-

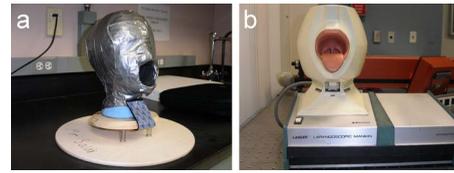


Fig. 1 Head models: (a) An *ex vivo* tissue head model made with wooden base plates and springs for simulating dynamic movement; the head is made of Styrofoam, clay, and duct tape, which surrounds the whole surface. This can be used for imaging any sample tissue. (b) A commercially available laryngoscopic manikin, for laryngeal examination practice.

tion and skill of the physician. This practice takes the form of the physician using the probe on an illuminating objective slide model or an *ex vivo* head model fashioned by our group with the use of swine vocal cords, as shown in Fig. 1. *In vivo* and noncontact imaging of the vocal folds in awake patients without anesthesia has not been reported.

In this paper, we demonstrate an office-based laryngeal swept-source OCT imaging system. Fast laryngeal imaging of 40 frames per second is realized to greatly eliminate motion artifacts caused by tremors (< 1 Hz) between the patients and the probe. *In vivo* noninvasive and noncontact OCT imaging of the vocal folds in awake patients without the use of anesthesia is reported here. Furthermore, dynamic vibration of the vocal folds is recorded to provide not only the high-resolution cross-sectional tissue structures but also important vibration parameters, such as the vibration frequency and amplitude of the vocal cord oscillations, which may provide additional helpful information for diagnosis.

2 Design

The schematic diagram of the fiber-based swept source OCT system is shown in Fig. 2. The output light from a swept light source at 1310 nm with a FWHM bandwidth of 100 nm and output power of 5 mW was split into the reference and sample arms by a 1×2 coupler (of 20/80 split ratio). The GRIN lens-based OCT probe was connected to the sample arm with 80% power from the source. The light source was operated at a sweeping rate of 20,000 Hz. The reference power was attenuated by an adjustable neutral density attenuator for maximum sensitivity. The measured sensitivity of the OCT system with an ideal partial reflector as the sample was 108 dB. Two circulators were used in both the reference and sample arms to redirect the backreflected light to a 2×2 fiber coupler (50/50 split ratio) for balanced detection. Dispersion compensation is important to achieve high resolution. The dispersion can be measured with a mirror as a sample by constructing the complex representation of the spectral fringe pattern and correcting the phase as a function of the wave number.⁸ The measured axial resolution of $8 \mu\text{m}$ was close to the theoretical axial resolution of $7.5 \mu\text{m}$ since the spectrum of the swept light source is nearly Gaussian shaped. The lateral resolution, which is determined by the OCT probe's focus spot, was measured to be $25 \mu\text{m}$.

In laryngeal endoscopy, the depth of the larynx varies remarkably from patient to patient. Changing the optical path length of the reference arm to match a variable working distance is difficult. The most convenient solution is to maintain

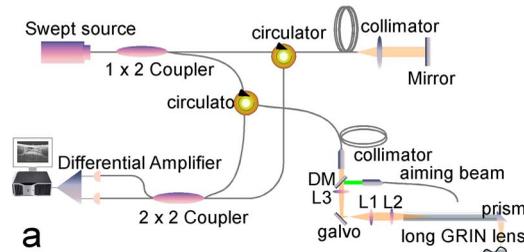


Fig. 2 (a) Schematic diagram of a GRIN lens rod-based dynamic focusing swept-source OCT system. (b) Design of the probe. DM—dichroic mirror; Ls are lenses.

a constant optical delay in the sample arm while tuning the working distance to ensure that the sample beam is always focused into the vocal cord. The device must quickly adjust to image the vocal cords as it changes position within the larynx. We use an enhanced version of a previously reported GRIN lens-based probe⁷ to fulfill constant optical delay dynamic focusing. The long GRIN lens used in this design can be considered as one pitch and an optical relay for visible wavelength. However, for a 1310-nm wavelength, which is the center wavelength of the OCT light source, the GRIN lens is closed to one pitch but cannot be considered as an ideal optical relay any more, especially when the average working distance of the probe (or the beam coming out of the probe tip) reaches about 65 mm for laryngeal imaging. In order to achieve an ideal optical relay, the GRIN lens is used with a group of lenses L_1 , L_2 to form a so-called optical-ballast within a $4f$ optical system. The composite $4f$ optical system has a magnification of one and can be considered as an optical relay; the optical delay of the focal point remains constant during adjustment of the working distance.

A carriage holds the OCT device and the rigid video endoscope (Carl Zeiss Ltd) together in a “double-barreled” configuration. In order to identify the scanning point (area) during an OCT examination, an aiming beam should be coupled into the system, because the OCT light source is in the nonvisible infrared spectrum. Previously, a 2×1 coupler was used in the sample arm to couple a green aiming beam from a 532-nm solid-state laser. The $1.3\text{-}\mu\text{m}$ OCT beam will need to pass through the 2×1 coupler back and forth (twice) before reaching the detector. In order to achieve the best image quality for OCT imaging without too much sacrifice of the OCT power,

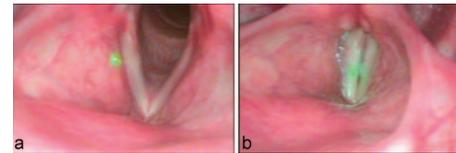


Fig. 3 (a) Image of the vocal cords while patient is breathing and green aiming beam. (b) Image of the vocal cords while the patient is phonating and green aiming beam.

the suggested coupling efficiency for the green light would have to be small. Normally, a very small portion ($< 10\%$) of the green light was allowed for aiming purpose, which was oftentimes not bright enough with the background incandescent lamp on for the video endoscope imaging (Fig. 3).

In our enhanced probe, a dichroic mirror-based design was used to solve the “compromise” between high OCT imaging power and bright illumination of the aiming beam. The sample beam from the OCT system is collimated, passes through a dichroic mirror and a focusing lens L_3 , and then reflects 90 deg to the fixed lens group by a scanning galvo [Fig. 2(b)]. The sample beam is now coupled to the probe with about 95% efficiency. Over 85% of the green beam is also coupled into the system through another channel of the dichroic mirror for aiming purposes. The fiber, the two collimators, and focusing lens L_3 are assembled as one component and can be moved back and forth along the propagation direction for distance adjustment by the physician during the examination. The typical range of scanning (working) distance is about 40 mm. Two customized prisms are attached at the proximal and distal tips of the GRIN lens for beam deflection. During the examination, both of the dual-channel rigid endoscope and OCT signals are digitized and displayed on a single monitor (Fig. 4).

3 Method

During the examination of the vocal cords, the patient is asked to sit up straight and hold his or her tongue out and downward with gauze; this is for the precise purpose of clearing the way for the probe. Before the probe is inserted, it is mildly heated with a small blow dryer to prevent the optical components from fogging up when exposed to internal body temperatures. The probe is then inserted through the oral cavity and centered several centimeters above the larynx. Once we obtain a clear vision of the vocal cords as well as a good position of the OCT aiming beam, the patient is asked, as with the conventional stroboscopic examination, to phonate in order to produce different movements and positions of the vocal

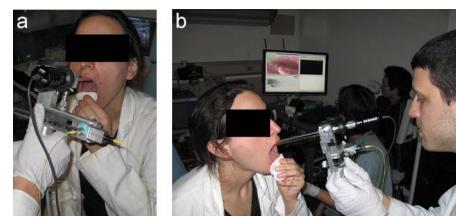


Fig. 4 (a) OCT probe attached to the laryngoscope for office-based laryngoscopic examination. (b) Dual-channel endoscope and OCT signals shown in a monitor.

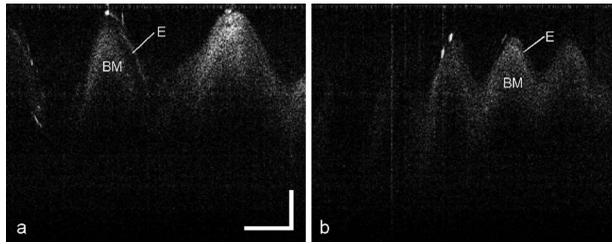


Fig. 5 Vibrating vocal cord with different frequencies: (a) ~ 120 Hz and (b) ~ 200 Hz. Es are epitheliums, and BMs are basement membranes. The scale bar represents $500 \mu\text{m}$.

cords. This is illustrated by the endoscope image in Fig. 3(b); it is evident that the vocal cords come together, which makes imaging easier for the otolaryngologist. During the whole procedure, we are recording both the OCT images as well as the laryngoscopic images.

Figure 5 shows the cross-sectional images of vibrating vocal cords of male and female volunteers during examination. The epithelium and basement membrane can be clearly identified. The images are comparable with images obtained in anesthetized patients during surgical endoscopy. The vocal fundamental frequency in females is approximately 200 Hz, whereas in males, it is approximately 120 Hz.⁹ This is a perfect fit for the measured OCT images shown in Fig. 5. Since the OCT imaging speed is 40 frames per second and 3 cycles are observed per frame, Fig. 5(a) corresponds to a frequency of approximately 120 Hz ($40 \times 3 = 120$). On the other hand, Fig. 5(b) contains up to 5 vibrations per cycle and corresponds to a vibration frequency of 200 Hz ($40 \times 5 = 200$). The precise dynamic vibration amplitudes can also be measured based on the OCT images. Since the total imaging depth in Fig. 5 is 2.6 mm, the estimated maximum vibration amplitudes in Figs. 5(a) and 5(b) are 1.2 mm and 0.59 mm, respectively.

4 Conclusion

We have demonstrated video-rate *in vivo* laryngeal imaging at 40 frames per second during a typical office-based laryngoscopy examination with a swept-source OCT system. Dynamic vibration of the vocal folds is recorded to provide not only the high-resolution cross-sectional tissue structures but also important vibration parameters, such as the frequency and amplitude of the vocal cord, which provide important information for clinical diagnosis and treatment as well as additional research in speech. Office-based OCT is a promising new imaging modality to image the larynx. Having the advantage of being performed without the need for general anesthesia or tissue removal, it adds a level of practicality and ease of use.

Office-based OCT has the potential to guide surgical biopsies, direct therapy, and monitor disease. The future success of this device is deeply rooted in the amount of clinical volunteers that we can use the device on. Therefore, it is suggested that the device be more aesthetically pleasing, because too often people are reluctant to participate in this study because of the rudimentary nature of the device, with fiber-optic cables protruding and DC power supply running.

Acknowledgments

This work was supported by the National Institutes of Health (DC 006026, CA 91717, EB 00293, RR 01192, RR00827), National Science Foundation (BES-86924), Flight Attendant Medical Research Institute (32456), State of California Tobacco Related Disease Research Program (12RT-0113), Air Force Office of Scientific Research (F49620-00-1-0371), Undergraduate Research Opportunities Program, and the Henry Samueli School of Engineering Undergraduate Research Fellowship. Support from the Beckman Laser Institute Inc. Foundation is also gratefully acknowledged.

References

1. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," *Science* **254**, 1178–1183 (1991).
2. B. J. F. Wong, R. P. Jackson, S. Guo, J. M. Ridgway, U. Mahmood, J. Su, T. Y. Shibuya, R. L. Crumley, M. Gu, W. B. Armstrong, and Z. Chen, "In vivo optical coherence tomography of the human larynx: normative and benign pathology in 82 patients," *Laryngoscope* **115**, 1904–1911 (2005).
3. A. M. Sergeev, V. M. Gelikonov, G. V. Gelikonov, F. Feldchtein, R. Kuranov, N. Gladkova, N. Shakhova, L. Snopova, A. Shakhov, I. Kuznetsova, A. Denisenko, V. Pochinko, Yu. Chumakov, and O. Streltsova, "In vivo endoscopic OCT imaging of precancer and cancer states of human mucosa," *Opt. Express* **1**, 432–440 (1997).
4. A. V. Shakhov, A. B. Terentjeva, V. A. Kamensky, L. B. Snopova, V. M. Gelikonov, F. I. Feldchtein, and A. M. Sergeev, "Optical coherence tomography monitoring for laser surgery of laryngeal carcinoma," *J. Surg. Oncol.* **77**, 253–258 (2001).
5. K. Luerssen, H. Lubatschowski, H. Gasse, R. Koch, and M. Ptok, "Optical characterization of vocal folds with optical coherence tomography," *Proc. SPIE* **5686**, 328–332 (2005).
6. S. Guo, R. Hutchison, R. P. Jackson, A. Kohli, T. Sharp, E. Orwin, R. Haskell, Z. Chen, and B. J. F. Wong, "Office-based optical coherence tomographic imaging of human vocal cords," *J. Biomed. Opt.* **11**, 030501 (2006).
7. S. Guo, L. Yu, A. Sepehr, J. Perez, J. Su, J. M. Ridgway, D. Vokes, B. J. Wong, and Z. Chen, "Gradient-index lens rod based probe for office-based optical coherence tomography of the human larynx," *J. Biomed. Opt.* **14**(1), 014017 (2009).
8. M. Wojtkowski, V. Srinivasan, T. Ko, J. Fujimoto, A. Kowalczyk, and J. Duker, "Ultra-high-resolution, high-speed, Fourier domain optical coherence tomography and methods for dispersion compensation," *Opt. Express* **12**, 2404–2422 (2004).
9. J. P. Noordzij and R. H. Ossoff, "Anatomy and physiology of the larynx," *Otolaryngol. Clin. North Am.* **39**(1), 1–10 (2006).