Dynamic vocal fold imaging by integrating optical coherence tomography with laryngeal high-speed video endoscopy

Gopi Maguluri, Ernest Chang, and Nicusor Iftimia
Physical Sciences Inc., 20 New England Business Center, Andover, MA 01810

Daryush Mehta, James Kobler
Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, 55 Fruit Street, Thier Building - 620, Boston MA 02114

Abstract: We demonstrate three-dimensional vocal fold imaging during phonation by integrating optical coherence tomography with high-speed videoendoscopy. Results from ex vivo larynx experiments yield reconstructed vocal fold surface contours for ten phases of periodic motion.
OCIS codes: (120.3890) Medical optics instrumentation, (170.4940) Otolaryngology, (170.4500) Optical coherence tomography

1. Introduction

Voice disorders due to trauma (e.g., intubation injuries, vocal abuse) and disease (e.g., dysplasia, cancer, recurrent respiratory papillomatosis, nodules, polyps, and scar) are currently evaluated in the clinic using endoscopic imaging techniques such as videostroboscopy [1] or high-speed videoendoscopy (HSV) [2]. Clinicians couple these visual observations of vocal fold tissue motion with auditory-perceptual judgments of voice quality as part of a comprehensive assessment of the health and function of the larynx during phonation [3]. Endoscopic imaging, however, only provides two-dimensional (2D) spatial information to yield measures of vocal fold symmetry and lateral tissue motion, critically lacking vertical axis information. Various high-speed imaging techniques have attempted to capture vocal fold surface motion during phonation in three spatial dimensions [4–6]; however, they either lack adequate spatial resolution or have not been validated in vivo. Integrating HSV with optical coherence tomography (OCT) shows promise due to the complementary information captured by each modality. Whereas HSV can yield color images at rates of 4000 frames per second and higher of 2D vocal fold vibratory patterns, OCT adds the missing vertical dimension to image mucosal wave propagation across the vocal fold surface. OCT provides a non-contact, high-speed imaging technique to quantitatively measure vertical motion of the vocal fold surface during phonation.

In this paper, we describe the design and the preliminary testing of an integrated OCT-HSV system. Ten phases of vocal fold tissue motion are reconstructed using this system to assess the appearance, shape, and anatomy of the vocal folds during phonation.

2. System setup and design

The design of the OCT-HSV endoscopic instrument is shown in Figure 1. The OCT component of the system is a spectral domain–based interferometer that is illuminated by a superluminiscent diode with a full-width at half maximum bandwidth of 50 nm at a center wavelength of 1310 nm. A maximum A-scan rate of 76 kHz is achieved with a commercial line-scan camera (Sensors Unlimited Inc., Princeton, NJ) to provide the speed necessary to perform dynamic imaging during phonation. A frame grabber digitizes and transfers the data to a computer with a graphic processing unit that processes and displays the OCT data. The HSV component of the system constitutes a Phantom v7.3 color camera (Vision Research Inc., Wayne, NJ), typically set to record at 1000 frames per second at a spatial resolution of 256 x 256 pixels. An important step is the temporal synchronization of the OCT and HSV data. A photograph of the experimental setup for imaging airflow-driven phonation of an ex vivo calf larynx is shown in Figure 2.

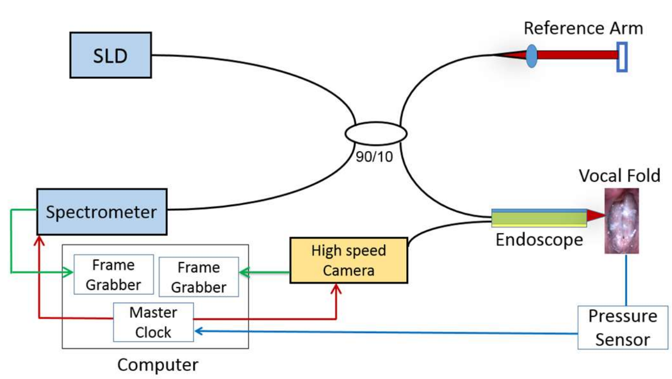

Fig. 1. Schematic of the integrated OCT-HSV instrument.

Figure 3 illustrates the use of a master clock generated from a pressure sensor near the vocal folds to generate phase-locked signals that trigger both OCT and HSV cameras. The reset/acquisition trigger occurs at the beginning of each phonatory cycle and starts a raster OCT beam scan spanning 10 mm x 10 mm. Each fast scan (B-scan) sends a trigger to the HSV camera to capture a high-speed image. For a typical phonating frequency of 100 Hz, 10 fast scans and 100 slow scans occur in a span of ~10 seconds to provide a series of 1000 HSV frames with 100 OCT B-scans per HSV frame. The data set provides 3D tissue surface information depicting temporal phases of laryngeal motion for a representative phonatory cycle.

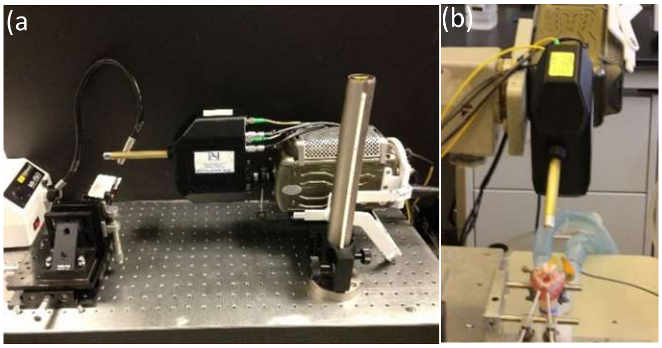

Fig. 2. OCT-HSV instrument. Setup showing (a) the OCT probe head mounted on the high-speed camera for co-linear imaging through a rigid endoscope and (b) the ex vivo calf larynx being imaged.

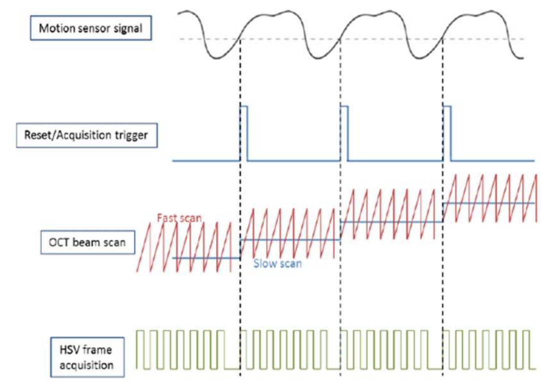

Fig. 3. OCT-HSV timing and synchronization diagram.

3. Initial experimental results

Measurements on vibrating calf vocal folds were performed combining OCT and HSV data. Both static anatomical images and dynamic images (10 phases during a phonation cycles) were recorded. Figure 4 shows the 3D reconstruction of the vocal fold surface at rest with an axial area of 10 mm x 10 mm and imaging depth of 3.5 mm, while Figure 5 shows the vertical position of the vocal fold surface during a phonatory cycle. The 3D coordinates will be used to provide a 3D surface contour for the HSV frames to enable 4D visualization of vocal fold kinematics.

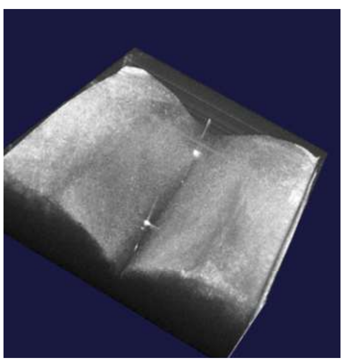

Fig. 4. 3D reconstruction of vocal fold integrating OCT and HSV images while the larynx is at rest.

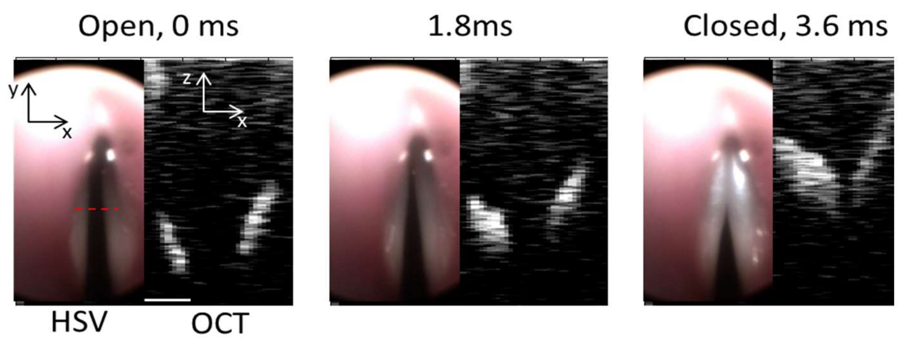

Fig. 5. Preliminary data displaying co-registered images acquired using swept-source OCT and HSV imaging modalities. The OCT beam scan path is indicated by a red dashed line. Scale bar: 1 mm.

4. Conclusion

In conclusion, we have demonstrated a novel imaging modality combining optical coherence tomography (OCT) and high-speed videoendoscopy (HSV) to image laryngeal motion at high speeds. This technology has the potential to enable systematic studies describing and developing acoustic correlates of irregularities in vocal fold vibration and to provide otolaryngologists with a clinically viable tool to assess the detailed motion of the musical wave affected by vocal pathology.

5. References

[1] D. D. Mehta and R.E. Hillman. “Current role of stroboscopy in laryngeal imaging,” Curr Opin Otolaryngol Head Neck Surg, **20**, 429–436 (2012).

[2] D. D. Deliyski, P. P. Petrushev, H. S. Bonilha, T. T. Gerlach, B. Martin-Harris, and R. E. Hillman, “Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and evolution,” Folia Phoniatr Logop, **60**, 33–44 (2008).

[3] S. M. Zeitels, A. Blitzer, R. E. Hillman, and R. R. Anderson, “Foresight in laryngology and laryngeal surgery: A 2020 vision,” Ann Otol Rhinol Laryngol, **116 Suppl 198**, 1–16 (2007).

[4] C.G. Tsai, J.H. Chen, Y.W. Shau, T.Y. Hsiao, “Dynamic B-mode ultrasound imaging of vocal fold vibration during phonation,” Ultrasound Med Biol, **35**, 1812–1818 (2009).

[5] G. Luegmair, S. Kniesburges, M. Zimmermann, A. Sutor, U. Eysholdt, M. Doellinger, “Optical reconstruction of high-speed surface dynamics in an uncontrollable environment,” IEEE T Med Imaging, **29**, 1979–1991 (2010).

[6] D. E. Sommer, I.T. Tokuda, S.D. Peterson, K. Sakakibara, H. Imagawa, A. Yamauchi, T. Nito, T. Yamasoba, and N. Tayama, “Estimation of inferior-superior vocal fold kinematics from high-speed stereo endoscopic data in vivo,” J Acoust Soc Am, **136**, 3290–3300 (2014).