PROCEEDINGS OF SPIE

Concept for high speed vocal cord imaging with swept-source optical coherence tomography

Zabic, Miroslav, Sharifpourboushehri, Sara, Müller-Wirts, Lennart, Benecke, Hannes, Heisterkamp, Alexander, et al.

Concept for high speed vocal cord imaging with swept-source optical coherence tomography

Miroslav Zabic*[a], Sara Sharifpourboushehri[b], Lennart Müller-Wirts[b], Hannes Benecke[b], Alexander Heisterkamp[a], Heiko Meyer[b], and Tammo Ripken[b]

[a]Institute of Quantum Optics, Leibniz University Hannover, Welfengarten 1, 30167 Hannover, Germany
[b]Industrial and Biomedical Optics Department, Laser Zentrum Hannover e.V., Hollerithallee 8, 30419 Hannover, Germany

ABSTRACT

Optical coherence tomography (OCT) enables non-invasive depth-resolved investigation of laryngeal tissue. However, with conventional systems, OCT cross-sectional images of vibrating vocal cords always suffer from motion artifacts. This is the case even at low phonation frequencies of about 100 Hz. Motion artifacts of predictable repetitive movements can be avoided with carefully timed acquisitions. Irregular, non-repetitive movements, e.g. disturbed vocal cord vibration caused by laryngeal disorders, require different strategies, such as the use of high frame rates. We present a novel concept for dynamic vocal cord imaging with a high speed 1.6 MHz swept-source OCT system. Due to the high image rate, a graphics processing unit (GPU) based signal processing software has been developed in order to obtain real time OCT images. To demonstrate the feasibility of our approach on vibrating samples, we present a laboratory setup which includes a MHz swept source for OCT. To enable the transfer of our setup to clinical applications a concept for a curved rigid laryngoscope design, integrating the optical components for high-speed OCT, is proposed.

Keywords: optical coherence tomography, laryngoscopy, motion artifacts, high-speed imaging, vocal cords

1. INTRODUCTION

For the treatment of many laryngeal pathologies, e.g. the Reinke edema, surgery is usually necessary.[1] To maintain the mechanical properties of the vocal cords and thus the voice of the patient, it is important in such procedures not to cut too deeply into the tissue. The superficial and deep layers of the lamina propria should be preserved to prevent scarring.[2] Depth-controlled laser cutting is possible with utilization of optical coherence tomography (OCT)[3] and could be used to avoid this issue. Due to the relatively slow acquisition speed of conventional OCT systems, OCT cross-sectional images, so called B-scans, of vibrating vocal cords suffer from motion artifacts.[4] Those motion artifacts are probably a hindrance to reliable depth control and can be avoided by triggered OCT capturing.[5] However triggered OCT capturing requires vocal cord oscillation with perfect periodicity.[5] This is usually not possible with pathological vocal cords.[6] Hence we propose a different approach, namely the use of MHz-OCT for high-speed examination of vibrating vocal cords.

1.1 Required OCT Acquisition Speed for Oscillating Sample

Consider a sample in periodic one-dimensional sinusoidal motion along the optical axis of an incident OCT beam. The beam is scanned laterally to acquire $N_{\text{A-scans}}$ adjacent one dimensional depth scans, so called "A-scans", and generate a cross-sectional image (so called "B-scan") of the scanned sample area. During maximum displacement of the sample the change of position is smallest, so this period of time τ is most suitable to acquire a B-scan. The displacement Δz during τ is equal to $A_0 - A_0 \cos(2\pi f \frac{\tau}{2})$, whereby A_0 is the amplitude of the oscillating sample and f the frequency. Figure 1 illustrates this situation.
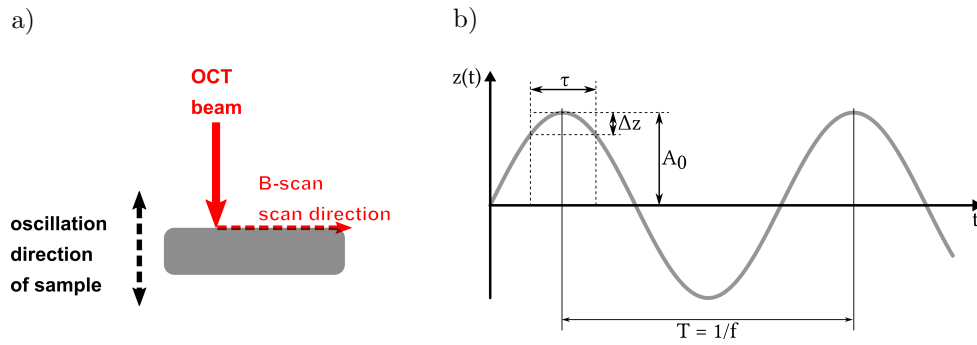
---

*zabic@iqo.uni-hannover.de

Figure 1. a) Sample in periodic one-dimensional sinusoidal motion along the optical axis of an incident transversely scanned OCT beam. b) Sample position as a function of time. $z(t)$: sample position at time t, A_0: oscillation amplitude of sample, T: oscillation period of sample, f: oscillation frequency of sample, τ: time interval in which sample displacement Δz occurs.

The required B-scan rate $f_{\text{B-scan}}$, which guarantees that at least one B-scan is acquired during the oscillation maximum, is the reciprocal of $\frac{\tau}{2}$. With the number of A-scans per B-scan $N_{\text{A-scans}}$ it is now easy to calculate the required A-scan rate $f_{\text{A-scan}}$ to capture a "snapshot" of the oscillating sample during maximum displacement:

$$f_{\text{A-scan}} = \frac{N_{\text{A-Scans}} \cdot 2\pi f}{\arccos(1 - \frac{\Delta z}{A_0})} \tag{1}$$

To determine realistic A-scan rates for vibrating vocal cords, plausible values for A_0 and f are needed. Exemplary values for a female person ($A_0 = 600\,\mu\text{m}$, $f = 286\,\text{Hz}$) and a male person ($A_0 = 1600\,\mu\text{m}$, $f = 143\,\text{Hz}$) can be found in 4. A common value for $N_{\text{A-scans}}$ is 512. Figure 2 shows $f_{\text{A-scan}}$ for different Δz for the just mentioned values and additionally for an optimistic case with $A_0 = 500\,\mu\text{m}$ and $f = 100\,\text{Hz}$.
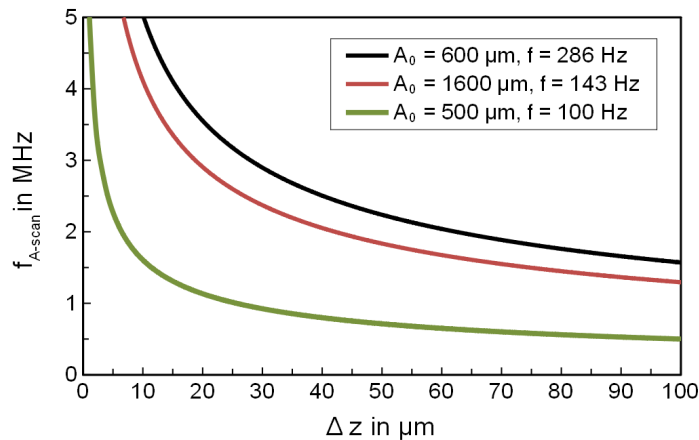

Figure 2. A-scan rates $f_{\text{A-scan}}$ for different sample displacements Δz that will occur during a single B-scan acquisition of an axially oscillating sample. Curves for three sample oscillations are depicted, black: $A_0 = 600\,\mu\text{m}$, $f = 286\,\text{Hz}$ (similar to a female test person from 4) , red: $A_0 = 1600\,\mu\text{m}$, $f = 143\,\text{Hz}$ (similar to a male test person from 4), green: optimistic oscillation values with $A_0 = 500\,\mu\text{m}$, $f = 100\,\text{Hz}$.

For a "snapshot" of an oscillating sample, Δz needs to be very small. A reasonable Δz is the typical OCT axial resolution of roughly 15 µm. With $\Delta z = 15\,\mu\text{m}$ an A-scan rate of about 4.1 MHz would be necessary to cover all three cases depicted in figure 2. Currently the fastest commercially available light sources for swept source OCT will generate A-scan rates of roughly 1.5 MHz.[7] This gives following displacements during a B-scan

acquisition for the three cases:

$A_0 = 600\,\mu\text{m}$, $f = 286\,\text{Hz}$ (female test person): $\quad \Delta z = 108\,\mu\text{m}$

$A_0 = 1600\,\mu\text{m}$, $f = 143\,\text{Hz}$ (male test person): $\quad \Delta z = 74\,\mu\text{m}$

$A_0 = 500\,\mu\text{m}$, $f = 100\,\text{Hz}$ (optimistic case): $\quad \Delta z = 11\,\mu\text{m}$

2. MATERIAL AND METHODS

2.1 MHz-OCT Setup

A schematic diagram of our laboratory setup for MHz-OCT is depicted in figure 3.
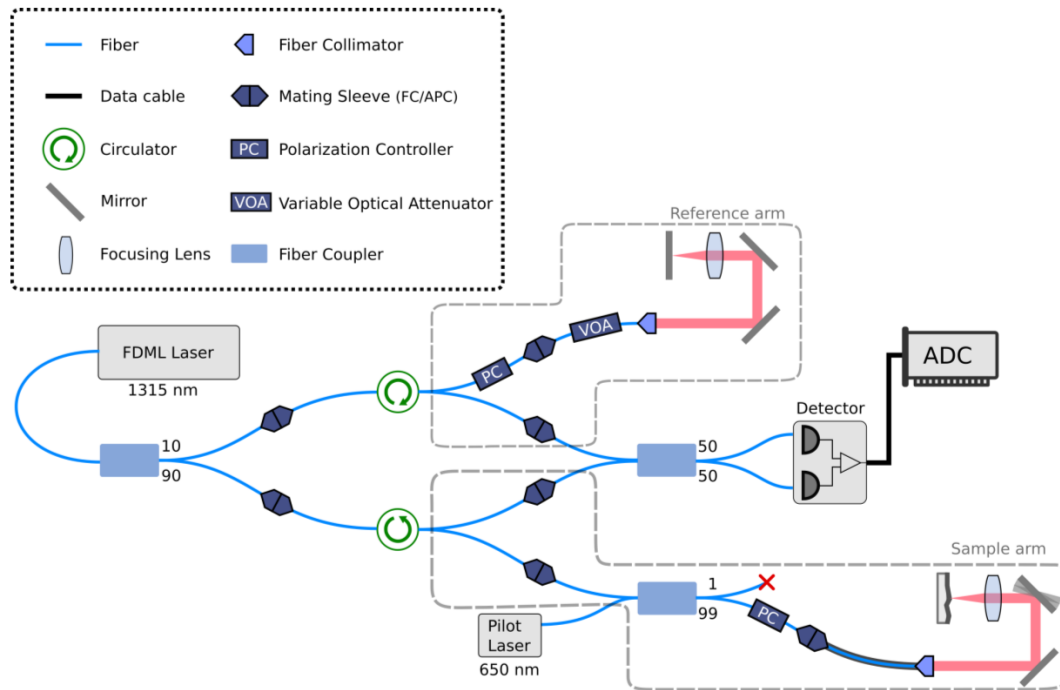

Figure 3. Fiber based MHz-OCT setup with a Fourier domain locked (FDML) Laser. PC: polarization controller, VOA: variable optical attenuator, ADC: analog-to-digital converter.

The MHz-OCT system is a fiber based Mach-Zehnder interferometer with a Fourier domain mode locked (FDML) laser ($\lambda_\text{c} = 1315.4\,\text{nm}$, $\Delta\lambda = 102\,\text{nm}$, $f_\text{sweep} = 1.678\,\text{MHz}$, NG-FDML, Optores GmbH, Germany) as OCT light source. The light is coupled into a 90:10 beam splitter. One tenth of the light is directed through a circulator to the reference arm that contains a fiber paddle polarization controller, a variable optical attenuator, a movable deflection unit to adjust path length difference and a reflecting end mirror. 90% of the light is directed through another circulator to the sample arm that contains a 1:99 fiber coupler to couple a visible pilot laser into the system, a fiber paddle polarization controller, galvanometer scanners (6210H Galvo with 677 servo amplifiers, 3 mm mirros, Cambridge Technology Inc., USA) and a scan lens (LSM03, Thorlabs, USA). The beam reflections from sample arm and reference arm pass the respective circulators and are superimposed by a 50:50 fiber coupler into a balanced photodetector (WL-BPD1GA, Wieserlabs UG, Germany). The detected fringe signal is acquired by a high speed analog-to-digital converter (ATS9373, Alazar Technologies Inc., Canada), which is installed in a high-end computer. Key components of the computer are: AMD Ryzen Threadripper 1900X CPU, ASUS PRIME X399-A mainboard, Corsair DIMM 32GB DDR4-2666 RAM and MSI GeForce GTX 1080 TI Armor 11G OC GPU.

2.1.1 MHz-OCT Processing Software

Live signal processing and visualization is performed by an in-house developed software written in C++ that utilizes CUDA for GPU parallel computing. The cross-platform application framework Qt is used for the graphical user interface (GUI). A screenshot of the GUI is depicted in figure 4b. The software consists of three parts (see figure 4a): the main application (OCTproZ), a development kit (DevKit) and dynamically loadable modules (OCT systems). Those dynamic libraries represent physical or virtual OCT systems and must be implemented by using the DevKit. OCTproZ accesses the raw data of the currently loaded system and processes it entirely on the GPU.

Raw data from the OCT system is first directly transferred to RAM until a user-defined number of B-scans is acquired. Via direct memory access the raw data set is then copied asynchronously to GPU memory where conventional OCT signal processing is executed, which includes background subtraction, resampling, numerical dispersion compensation, windowing, fast Fourier transform (FFT), logarithmic scaling and gray-scale conversion. The OCT processing pipeline is depicted in figure 5. Asynchronous data transfer makes it possible to process a raw signal set while simultaneously the next data set is transferred to GPU memory. This technique is also used in other high speed OCT signal processing programs.[8,9] For live visualization of B-scans, CUDA-OpenGL interoperability is used. This avoids unnecessary data transfer of the processed OCT data to host memory. Data transfer to host memory occurs only when the recording function is used to store the acquisition on the hard disk drive.
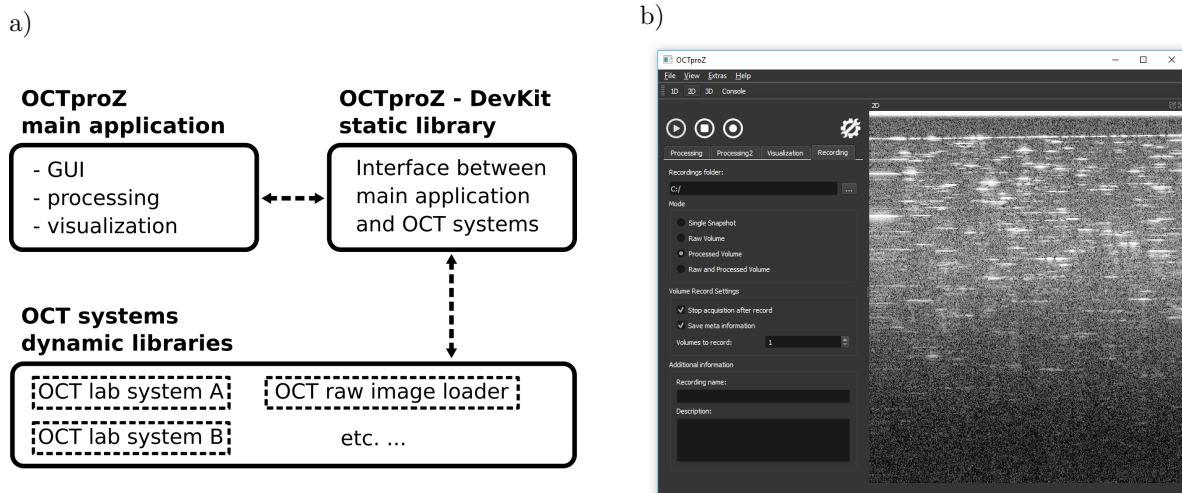

Figure 4. a) Basic overview of our OCT processing software that consists of three parts: main application, development kit and dynamically loadable implementations of OCT systems. The software is written in C++ and utilizes CUDA for GPU signal processing and Qt for the graphical user interface. b) Screenshot of the graphical user interface with B-scan live view of multiple layers of adhesive tape. Our software is a cross-platform application and may run on Microsoft Windows as well as Linux.
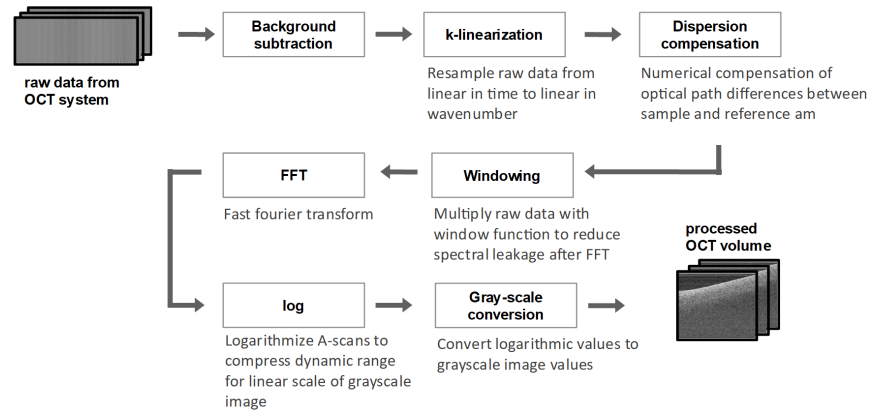
Figure 5. Conventional OCT processing pipeline. When using the FDML laser from Optores GmbH, k-linearization is essential. Without it, the axial resolution deteriorates considerably. OCTproZ supports numerical dispersion compensation, but it is not needed with our current optical setup.

2.2 Proposed Curved Rigid Laryngoscope

We propose a curved rigid laryngoscope design (fig. 6), integrating the optical components for a high-speed OCT sample arm, to enable the transfer of our setup to clinical applications.
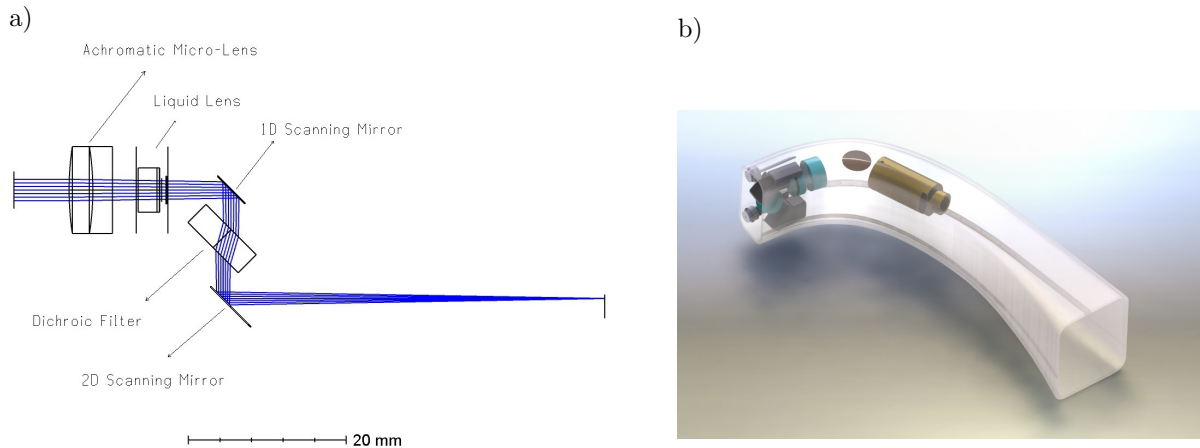

Figure 6. a) Proposed sample arm optics for MHz-OCT laryngoscopy. An achromatic micro-lens in combination with a liquid lens enables axial focus control within the desired working range. A fast 1D scanner is necessary for B-scan rates in the kHz range. The dichroic filter allows coupling of a surgical laser. A 2D scanner enables beam steering to desired position on sample. b) 3D rendering of optical components including two chip-on-tip cameras inside a curved rigid laryngoscope. The dimensions of the mechanical model roughly resemble the regular sized Airtraq SP video laryngoscope (Prodol Meditec S.A., Spain).

The proposed design uses a fiber collimator (60FC-4-M15-37, Schäfter + Kirchhoff GmbH, Germany) to direct the OCT beam into the tip of the curved laryngoscope. The tip contains an achromatic micro-lens (5M-M60-26, Schäfter + Kirchhoff GmbH, Germany) and a liquid lens (EL-3-10 , Optotune, Switzerland) for working distance adjustment and tremor compensation of the operator. A fast 1D microelectromechanical system (MEMS) scanner is placed after the liquid lens for high-speed OCT image formation. Directly after the scanner a dichroic filter is placed that enables coaxial coupling of a surgical laser. Surgical laser and OCT beam are then directed towards a 2D MEMS scanner, which can direct the beams to the desired location at the vocal cords. Both MEMS scanners are restricted in space and need to be custom designed. This setup, including two chip-on-tip cameras, fits inside

a mechanical model of a curved rigid laryngoscope (see figure 6b). The dimensions of the mechanical model are roughly based on the regular sized Airtraq SP video laryngoscope (Prodol Meditec S.A., Spain).

3. PRELIMINARY RESULTS

3.1 B-scans of Vibrating Sample

As first test measurements we acquired B-scans (with 512 A-scans per B-scan) of a vibrating sample with our MHz-OCT setup and with a commercial OCT setup (OCS1300SS, Thorlabs, 16 kHz A-scan rate). To generate sample motion, we attached the sample to a loudspeaker which was connected to a function generator via an amplifier. The function generator was used to generate a sinusoidal signal with 50 Hz. OCT images of the MHz-OCT setup did not show motion artifacts. However, the OCT images of the conventional system did. Figure 7 shows the acquired B-scans.

Despite the not existing motion artifacts, image quality is not optimal. Further investigations are necessary to explain the modest image quality of the MHz-OCT images. Possible causes could be that the sample was not positioned in focus, the reference arm power was too low, or incorrect parameters were used for k-linearization.
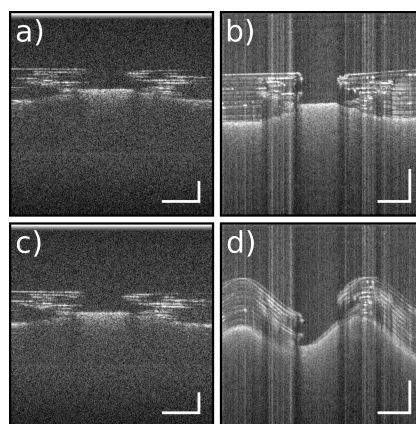

Figure 7. Comparison of motion artifacts of an oscillating sample in B-scans from our MHz-OCT system (1.678 MHz A-scan rate) and in B-scans from a conventional OCT system (OCS1300SS, Thorlabs, 16 kHz A-scan rate). The sample consists of two sets of several layers of adhesive tape separated by a small gap. a) Stationary sample, MHz-OCT system. b) Stationary sample, conventional OCT. c) 50 Hz axial sample oscillation, MHz-OCT. d) 50 Hz axial sample oscillation, conventional OCT. The scale bars have a length of 0.5 mm in horizontal direction as well in vertical direction.

3.2 Optics Simulation of Curved Rigid Laryngoscope

We used Zemax OpticStudio 16.5 (Zemax LLC, USA) for optics simulation of the proposed laryngoscope design. Figure 8a shows the Airy disk radius and the RMS spot radius as function of working distance. The working distance is the distance between the tip of the laryngoscope and the focus of the OCT beam. It can be varied by changing the curvature of the liquid lens. Ideally, the focus point is on the sample to be examined (e.g. vocal cord). The optics simulation shows that the RMS spot radius is always smaller than the Airy disk radius. This indicates that the optical setup is diffraction limited for every working distance.

Figure 8b shows the spot diagram for a working distance of 40.0 mm. The circle represents the Airy disk and the colored areas in the center are spots calculated by geometric ray tracing. Each spot color represents a wavelength of the simulated beam (blue: 1260 nm, green: 1310 nm and red: 1360 nm).
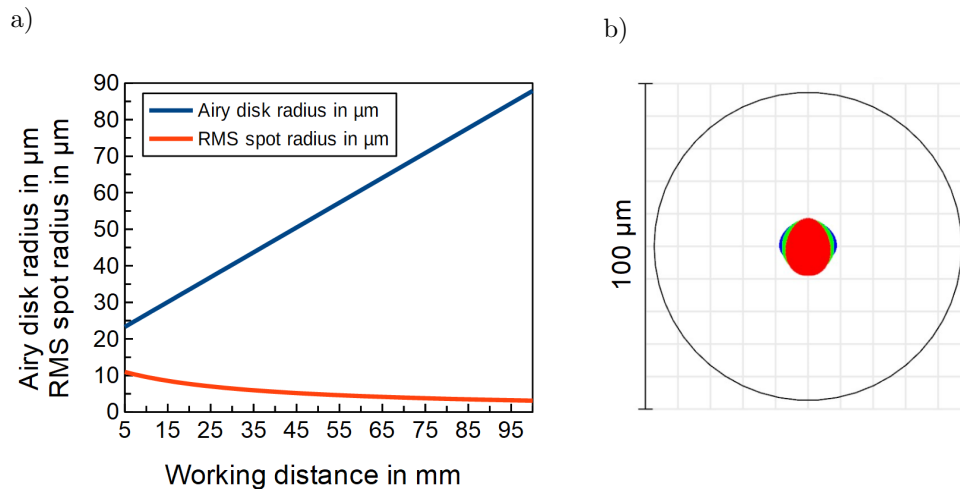
Figure 8. Results of optics simulation with Zemax. a) Airy disk radius and the RMS spot for different working distances. b) Spot diagram for working distance of 40.0 mm. The circle represents the Airy disk. Each spot color represents a wavelength of the simulated beam (blue: 1260 nm, green: 1310 nm and red: 1360 nm).

4. CONCLUSION AND OUTLOOK

We experimentally demonstrated that motion artifacts due to 50 Hz sample oscillation in axial direction do not occure in cross-sectional images acquired with MHz-OCT. However, with conventional OCT, motion artifacts are clearly visible. To use MHz-OCT for the examination of vocal cords, we proposed a curved rigid laryngoscope design, integrating the optical components for the sample arm of a MHz-OCT system. The proposed design includes a liquid lens for axial focus control.

Before the proposed laryngoscope will be realized, a laboratory setup will be first developed to verify the basic functionality of the combined components. We plan to combine stereo camera, surgical laser and MHz-OCT in the near future which will eventually be transferred into a curved rigid laryngoscope.

ACKNOWLEDGMENTS

REFERENCES

[1] Goswami, S. and Patra, T. K., "A clinico-pathological study of reinke's oedema," *Indian Journal of otolaryngology and Head and Neck surgery* **55**(3), 160–165 (2003).

[2] Benninger, M. S., Alessi, D., Archer, S., Bastian, R., Ford, C., Koufman, J., Sataloff, R. T., Spiegel, J. R., and Woo, P., "Vocal fold scarring: current concepts and management," *Otolaryngology—head and neck surgery* **115**(5), 474–482 (1996).

[3] Fuchs, A., Pengel, S., Bergmeier, J., Kahrs, L. A., and Ortmaier, T., "Fast and automatic depth control of iterative bone ablation based on optical coherence tomography data," in [*European Conference on Biomedical Optics*], 95420P, Optical Society of America (2015).

[4] Donner, S., Bleeker, S., Ripken, T., Ptok, M., Jungheim, M., and Krueger, A., "Automated working distance adjustment enables optical coherence tomography of the human larynx in awake patients," *Journal of Medical Imaging* **2**(2), 026003 (2015).

[5] Chang, E. W., Kobler, J. B., and Yun, S. H., "Triggered optical coherence tomography for capturing rapid periodic motion," *Scientific reports* **1**, 48 (2011).

[6] Dollinger, M., Hoppe, U., Hettlich, F., Lohscheller, J., Schuberth, S., and Eysholdt, U., "Vibration parameter extraction from endoscopic image series of the vocal folds," *IEEE Transactions on Biomedical Engineering* **49**(8), 773–781 (2002).

[7] Klein, T. and Huber, R., "High-speed oct light sources and systems," *Biomedical optics express* **8**(2), 828–859 (2017).

[8] Wieser, W., Draxinger, W., Klein, T., Karpf, S., Pfeiffer, T., and Huber, R., "High definition live 3d-oct in vivo: design and evaluation of a 4d oct engine with 1 gvoxel/s," *Biomedical optics express* **5**(9), 2963–2977 (2014).

[9] Jian, Y., Wong, K., and Sarunic, M. V., "Graphics processing unit accelerated optical coherence tomography processing at megahertz axial scan rate and high resolution video rate volumetric rendering," *Journal of biomedical optics* **18**(2), 026002 (2013).