Spatiotemporal correlation of optical coherence tomography *in-vivo* images of rabbit airway for the diagnosis of edema

DongYel Kang
Alex Wang
Veronika Volgger
Zhongping Chen
Brian J. F. Wong
Spatiotemporal correlation of optical coherence tomography in-vivo images of rabbit airway for the diagnosis of edema

DongYel Kang, a, * Alex Wang, b Veronika Volgger, c Zhongping Chen, b,d,e and Brian J. F. Wong b,d,e

a Hanbat National University, College of Engineering, School of Basic Sciences, 125 DogSeoDaeho, YuSeong-Gu, Daejeon 305-719, Republic of Korea
b University of California Irvine, Beckman Laser Institute and Medical Clinic, Irvine, California 92617, United States
c Ludwig-Maximilians-University Munich, Department of Otolaryngology-Head and Neck Surgery, Marchioninistr. 15, Munich 81377, Germany
d University of California Irvine, Department of Biomedical Engineering, Irvine, California 92617, United States
e University of California Irvine, Department of Otolaryngology-Head and Neck Surgery, Irvine, California 92617, United States

Abstract. Detection of an early stage of subglottic edema is vital for airway management and prevention of stenosis, a life-threatening condition in critically ill neonates. As an observer for the task of diagnosing edema in vivo, we investigated spatiotemporal correlation (STC) of full-range optical coherence tomography (OCT) images acquired in the rabbit airway with experimentally simulated edema. Operating the STC observer on OCT images generates STC coefficients as test statistics for the statistical decision task. Resulting from this, the receiver operating characteristic (ROC) curves for the diagnosis of airway edema with full-range OCT in-vivo images were extracted and areas under ROC curves were calculated. These statistically quantified results demonstrated the potential clinical feasibility of the STC method as a means to identify early airway edema.© The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JBO.20.7.076015]

Keywords: optical coherence tomography; airway stenosis; edema; spatial and temporal correlation; tissue characterization; image analysis.

Paper 150134PR received Mar. 6, 2015; accepted for publication Jun. 30, 2015; published online Jul. 29, 2015.

1 Introduction

Critically ill intubated neonates or pediatric patients are at risk for airway injury that progresses through phases of wound healing starting with edema and ending with scar formation. This process in the subglottic region of airway may lead to stenosis, abnormal narrowing of airway.1,2 As edema is the first physiologic event that occurs with injury, the early detection of edema in the subglottis may lead to improvement of airway management, early detection of tissue injury, and hence the prevention of stenosis. The gold standard for diagnosing subglottic edema in vivo is surgical endoscopy, but this is limited to the visual inspection of the airway surface or gentle palpation with blunt surgical instruments cantilevered off the surgeon’s fingers. Mucosal changes in subepithelial tissue layers cannot be identified using this approach. Further, surgical endoscopy requires general anesthesia, which is a high-risk operation in the neonate, and may lead to additional mechanical injury to the airway.

Imaging modalities, such as computed tomography, magnetic resonance imaging, and ultrasonography, also have multiple limitations due to low resolution, radiation exposure, high cost, and/or complexity.3,4 Optical coherence tomography (OCT) has been proposed as an alternative imaging modality to noninvasively image the delicate mucosal layers of the upper airway.5–9 The key component of airway OCT systems is a rotating fiber-based probe that enables noninvasive acquisition of subepithelial tissue anatomy. Preclinical and clinical studies using the full-range OCT system have been reported with three-dimensional (3-D) airway structures reconstructed from hundreds of cross-sectional images acquired by helical scanning.8,9 Initial attempts at differentiating edematous from normal airway tissue included analysis of signal intensity variations, signal-to-noise ratios, and gray scale co-occurrence matrices10,11 in full-range OCT images, but none of these identified any reliable measures for edema.12 Alternatively, temporal correlation between consecutively measured full-range OCT images on the same airway region was evaluated for the task. Although characteristics of temporal correlation showed meaningful differences between edematous and normal states, there was the lack of an objective decision mechanism with observers that process the data to derive statistically quantified task performance.13–16

As noise is inherent in all medical images, OCT airway images are also compromised by noise, such as speckle and physiological noise, which suggests that the task of differentiating the edematous from normal is a statistical decision problem. In order to resolve this binary statistical problem, we develop an “observer” that is applied to OCT data to generate a scalar “test statistic.”13 This test statistic output can then be “objectively” compared with a user-defined threshold number to classify the data with respect to one of physiologic states (i.e., normal or edema). The observer requires to fulfill some conditions.13,14 Every observation for all tested data results in a distinct test statistic value that leads to a unique decision based upon threshold value selection. Also, repeated observations of the exact same data by a given observer
must result in the same test statistic. But in reality, the same object does not produce the same data. For example, images acquired from exactly the same airway region may generate different test statistic values because they are not identical due to noise (e.g., caused by speckle, physiological, and/or systematical variations). An observer applied to binary object data sets generates two probability distributions in a test statistic region, which typically overlap due to the statistical nature of the measured data. Classification of an image as either normal or edematous is established by careful selection of the threshold for those probability distributions. The process where clinicians evaluate medical images could be similarly interpreted as a human observer “operating” on medical images.13 The degree of separation between probability distributions of two states reflects the overall performance of the observer and image quality for a given task, which can be quantitatively evaluated by a “figure of merit” (FOM).13,15,16 A FOM is a scalar number that allows the performance of different imaging systems and observers to be objectively evaluated and compared to one another. By systematically adjusting parameters to maximize a FOM, an imaging system can be optimized for a specific task.

One of the widely used FOMs is the area under a receiver operating characteristic (ROC) curve (AUC), applicability and reliability of which has been verified in various statistical decision problems.13–15 The ROC curve is extremely useful as it summarizes the difficulty of a task, the performance of a decision strategy including the imaging system, and the quality of the data for enabling the observer to perform the specified task.13 Additionally, the ROC curve and AUC are independent of prevalence, the main complaint raised in the usage of other FOMs, such as accuracy. There have been numerous studies that have quantitatively evaluated OCT images and systems for various biomedical estimation and classification tasks by calculating ROC curves and AUCs.18–20 Also, the approach of optimizing OCT hardware has investigated based on the use of ROC curves and AUCs for specified tasks, such as detecting the refractive index changes in multiple layers.21,22

Here, we establish an observer for the task of diagnosing edema, which is called spatiotemporal correlation (STC) that is applied to OCT images using a rabbit airway surgical model. To the authors’ knowledge, this is the first attempt to statistically quantify the differentiable characteristics of airway edema from in-vivo OCT images. In order to demonstrate the feasibility of the STC observer, we performed several in-vivo experiments with live rabbits and acquired a number of full-range OCT images for normal and surgically induced edematous airway areas. The newly developed observer is applied to the OCT data to produce the test statistics that construct ROC curves, and thus calculate AUCs. The overall performance of the observer (i.e., detectability of edematous tissue), quantified by ROC curves and AUCs, reflects how well the applied observer to those in-vivo OCT images performs for the early diagnosis of edema.

2 Spatiotemporal Correlation

A full-range swept source OCT system (central wavelength of 1310 nm) was used for in-vivo experiments in rabbits under general anesthesia. The experiments were performed under the aegis of the University of California, Irvine Institutional Animal Care and Use Committee. A Hopkins rod endoscope was inserted through a pediatric surgical laryngoscope, and edema was simulated by injecting saline beneath the epithelial layer of the subglottic mucosa. Since this study’s main goal is to initially demonstrate the feasibility of the STC method for the diagnosis of edema in vivo, the edema was created in specific locations of the airway known to the investigators. The generated edema was confirmed by visual observation with a rigid surgical endoscope. The fiber-based probe of the full-range OCT system was inserted into the rabbit airway through the fluorinated ethylene propylene sheath that is transparent to the wavelength of the system. The OCT probe can be mechanically rotated and translated within the sheath, so multiple rotations of the probe could generate a series of images for the same airway section or a spiral data set for some longitudinal segment of the airway. For in-vivo measurement of data required for this study, the system was set to 25 rotations/s and 2000 A-lines per sectional image that was acquired by 360 deg rotation. Figure 1(a) illustrates rectangular-coordinate in-vivo OCT airway images acquired from the same rabbit airway section. Figure 1(b) shows the same data set in polar-coordinates for an anatomic representation. Since these images are acquired for the same airway section in vivo, the temporal variation caused by actual physiological movements of the live rabbit can be observed in Fig. 1(b) (Video 1). The more detailed information about the in-vivo experiment using the built-in-laboratory full-range OCT system to obtain rabbit airway images as in Fig. 1 can be found elsewhere.23–7,12

As the first step to investigate the STC for the task of characterizing edema in the airway, temporal correlation coefficients (TCCs) between consecutively measured OCT images on the exact same airway section are analyzed. Under the assumption that the condition of the OCT system and imaged object does not change abruptly, it is expected that these series of images are highly correlated with one another due to nearly identical tissue structures and speckle patterns. However, because of physiological movements of the airway (e.g., movement from respiration or circulation) as observed in Fig. 1(b), the TCCs temporally vary. This variation is shown in Fig. 2(a) where TCC functions between the first and subsequently acquired OCT images are calculated for selected local regions, as indicated as the rectangular box in Fig. 1(a). Here, the periodic temporal variation in the TCC functions is likely related to respiration of the rabbit.12 However, for OCT images measured in edematous regions, where accumulation of interstitial fluid was simulated surgically, TCC functions change randomly even with the respiratory airway motion, as indicated in Fig. 2(b). This could be conjecturally explained: photons reflected from the simulated edematous airway region experience larger random phase changes than photons from the normal regions.

Although TCC functions reflect different characteristics between the normal and edematous airway regions, as shown in Fig. 2, they are functions, not scalar values, so making a decision using TCCs remains a subjective process. For example, if temporal behaviors of TCC functions are ambiguous, such as a trace that is intermediate between those illustrated in Figs. 2(a) and 2(b), the decision could be classified as either edematous or normal or unidentified depending upon the interpreter even for the same OCT data set. Furthermore, it might be necessary to know a priori the dominant frequency components of airway movements. Even with this priori information, unexpected airway movements easily compromise the task of diagnosing edema with TCC functions. Most importantly, the task performance cannot be objectively and quantitatively measured with those randomly variable functions, so it is a formidable challenge to characterize and optimize the OCT system performance.
to diagnose edema. Hence, it is imperative to establish an appropriate observer for the proposed task.

Since two different TCC functions are obtained from two different spatial locations on the airway, the effect of airway movement (e.g., normal respiration, etc.) on the pair of TCC functions is highly synchronous. This implies that TCC functions of two adjacent (or closely spaced) normal airway locations ought to strongly correlate with one another almost regardless of physiological motions or other movements. This synchronization is observed with red dotted and blue solid lines in Fig. 2(a). For an edematous airway region, however, the synchronization is likely to be very weak due to large, random phase variations of reflected photons even though the measured TCC locations are adjacent to one another. Based on this observation, the scalar test statistic for diagnosing edema can be proposed as a cross-correlation coefficient between two spatially separated TCC functions, which will be referred to as the STC coefficient. The STC coefficients calculated from edematous regions would be statistically smaller than those derived from normal tissue, so setting a threshold to statistically distributed STC coefficients could objectively differentiate edema from normal airway tissue areas.

This STC approach is conceptually different from speckle decorrelation that has been studied for recognizing rapidly variable portions in biological tissues. For example, speckle decorrelation between temporally measured OCT images abruptly decreases in blood vessel areas due to rapid movements of hemoglobin molecules. Unlike the STC concept, any additional noise, such as physiological movements, adversely affects the task performance with speckle decorrelation because they quickly deteriorate the ability to differentiate between speckle-boiling and speckle-stable regions. Of course, speckle decorrelation in subepithelium airway regions would partially affect the decision mechanism with STC coefficient. However, it can be


**Fig. 1** In-vivo optical coherence tomography (OCT) images of rabbit subglottic airway: (a) experimentally measured stack of serial OCT images (Cartesian coordinates) and (b) polar coordinate images (anatomic representation). The valley in (a) is the near airway area of the vocal cord. The arrow indicates the airway portion of simulated edema, which is located to the right lower quadrant of (b). The central circles in (b) are the images of the fluorinated ethylene propylene sheath. Temporal correlation coefficients (TCCs) are calculated for user-defined local airway regions, one example of which is indicated as the rectangular box in (a). (Video 1, MPEG, 27.5 MB) [URL: http://dx.doi.org/10.1117/1.JBO.20.7.076015.1]


**Fig. 2** Temporal variations of cross-correlation coefficients between OCT images are shown for (a) normal and (b) edema states, respectively. Red dotted and blue solid lines are for different local airway regions for each state.
stated that the STC method measures the relative response between adjacent local tissue areas to naturally introduced physiological motions and/or movements that generally occur in in-vivo experiments.

3 Procedure and Result

For investigating the feasibility of the STC method, a total of 14 data sets were acquired from five in-vivo experiments with three rabbits, where four of those data sets contained simulated edematous areas, as described in Fig. 1. Each data set is composed of 100 to 150 OCT images consecutively measured on the same airway sections. First, TCC functions are calculated for all those data sets considering two differently sized local regions to investigate the possible effect of a local tissue size on the statistical decision process. One is three A-lines of 200-pixel length (~2-mm depth in the actual airway) and the other is five A-lines of 200 pixels. The starting points for those 200-pixel A-lines are set to the 10th pixels below airway surfaces on the first OCT image of each data set, which were preselected before TCC calculations. Since a full-range OCT image consists of 2000 A-lines, 666 and 400 TCC functions are generated for three and five A-line bundle cases, respectively. Second, archetype edematous and normal airway areas are selected to determine the ground truth data for the statistical decision process. Although the location of the surgically created edematous tissue is always confirmed by visual observation of the tissue surface with a rigid surgical endoscope and analysis of relevant OCT images, it is almost impossible to identify a distinct boundary between edematous areas and normal tissue in in-vivo situations. Therefore, the central portions rather than the periphery of the surgically created edematous areas were selected to serve as archetype regions. The remaining pristine regions of the rabbit airway were considered to be normal. Third, normalized cross-correlation coefficients between TCC function pairs (i.e., STC coefficients) are calculated for those selected edematous and normal airway areas. The spatial distance between TCC function pairs is set by the selected number of TCCs between them, which is defined as $D_{TCC}$, and this is also varied to investigate its impact on the statistical decision process. For example, $D_{TCC} = 3$ indicates STC coefficients are extracted between first and fifth TCC functions, second and sixth TCC functions, etc. We assume that the ground truth state for the STC coefficient for a given TCC pair is determined by the airway region from where the central TCC function between those paired TCC functions is selected. For example, the ground truth state for the STC coefficient from second and sixth TCC functions is considered normal if the fourth TCC function is located in the selected normal region. For a single airway image set as shown in Fig. 1, several hundreds of independent test statistic outputs (i.e., STC coefficients) may be extracted because TCC functions are calculated for small local airway areas consisting of only three or five A-lines. Applying the process to all 14 in-vivo OCT data sets, the overall numbers of STC coefficients from selected edematous and normal areas are 417 and 7861 for three A-line bundled TCCs and 250 and 4720 for five A-line bundled TCCs.

From the previously described process, thousands of STC coefficients extracted as test statistics for both three and five A-line bundled TCC cases to evaluate the feasibility of the STC observer to derive ROCs and AUCs. By applying a threshold to those test statistics, the statistical decision process between the binary states (e.g., normal versus edematous) can be performed. The hypothesis is that airway regions across two spatially separated TCC functions that generate STC coefficients less than the threshold, are considered to be edematous, otherwise they are normal. This decision strategy applied to the selected archetype edematous and normal airway areas statistically quantifies the feasibility of the STC observer by comparison with ground truth data. (If some local airway region is composed of the TCC functions, say fifth, sixth, seventh, eighth, ninth, and only the fifth or ninth one is in an edematous region, the STC coefficient would be very small so that it is likely to be determined as edema although the local airway region is considered as ground truth normal. However, this discrepancy between the decision strategy and ground truth selection would rarely happen in this study because the statistical decision process was performed only for separated archetype edematous and normal airway areas.) If decisions for some STC coefficients are incorrect (i.e., designating a tissue region as edematous for a known normal airway area and vice versa), these are false positive and negative by the threshold, respectively. This partitioning process calculates the true positive fraction (TPF) and true negative fraction (TNF) from the test statistics, which are defined as:

$$TPF = \left( \frac{N_{TP}}{N_{TP} + N_{FN}} \right)$$

and

$$TNF = \left( \frac{N_{TN}}{N_{TN} + N_{FP}} \right),$$

where $N_{TP}$, $N_{TN}$, $N_{FP}$, and $N_{FN}$ indicate the random numbers of STC coefficients generating true positive, true negative, false positive, and false negative decisions with a given threshold. The angle brackets in Eq. (1) indicate statistical expectation of those random variables. Notice that $N_{TP}$ and $N_{FN}$ and $N_{TN}$ and $N_{FP}$ remain constant during the entire statistical testing process. Calculating Eq. (1) is equivalent to finding false negative fraction (FNF) and false positive fraction (FPF) from 1-TPF and 1-TNF by definition, respectively.

The quality of OCT images and the performance of the STC observer for the task of diagnosing edema can be objectively evaluated by investigating the thresholding separation between two states on the test statistics. This is achieved by plotting ROC curves with different combinations of TPF and FPF that are generated from Eq. (1) by changing the threshold from 0 (infinitely strict) to 1 (infinitely lenient). Figures 3(a) and 3(b) show ROC curves calculated for test statistics of three and five A-line bundles, respectively, where $D_{TCC}$ is set to 1, 3, 5, 7, and 9 for comparison. For the worst performance (i.e., edematous and normal test statistics completely overlap), the ROC curve becomes a straight diagonal line. As the classification between edematous and normal test statistics improves, the ROC curve will approach the left-upper quadrant of the FPF–TPF rectangular box. Therefore, the ROC curves shown in Fig. 3 are reasonably meaningful in terms of the task of classifying the edematous from normal airway region.

The overall performance of those ROC curves can be quantitatively represented by the AUCs that are also shown in Fig. 3. Considering an AUC ranges from 0.5 to 1 in a classification task, the calculated AUC values in Fig. 3 indicate that the STC method shows quite meaningful performance in differentiating edema from normal airway tissue. Also, as shown in Fig. 3, AUC values increase as the $D_{TCC}$ increases, but the increasing rates saturate. This trend is understandable as TCC functions calculated from adjacent regions correlate at
least to some degree regardless of the tissue characteristics of edema or normal. As the $D_{TCC}$ increases, the effect of morphological correlation on STC coefficients decreases and the effect of random phase variation caused by the temporal variation of tissue on STC coefficients becomes dominant, which contributes to the AUC increasing. When the distance continues to increase, only tissue characteristics such as what in normal or edematous tissues, are crucial to STC coefficients, so the AUC values would nearly saturate. Figure 3 also shows that calculating TCC functions for the bundle of three A-lines is better than five A-lines for the present *in-vivo* data sets processed by the STC method.

Although it is generally true that acquiring more data enhances the task performance more for any given observer, one of the benefits of the STC method is that the number of consecutively measured OCT images on a targeted airway region need not be too large to effectively perform the decision process. The reason is that the diagnosis performance of the STC method is mainly related to the correlation of spatially separated TCC functions regardless of a single TCC function’s temporal variation. Therefore, even if there are unexpected variations on a single TCC function caused by patient movement, physiologic motion, or other factors, the feasibility of the STC method would be still valid as long as the selected TCC pairs experience the same temporal variations simultaneously. Since it is almost impossible to formulate the general effect of unexpected nonstationary noise on STC results, the variation of AUCs to the number of OCT images is investigated. Figure 4 shows calculated AUC values with fewer OCT images and different $D_{TCC}$ values. The result suggests that the STC observer applied to $\sim 50$ OCT images
could effectively perform the diagnosis task with the present in-vivo OCT data. Considering 25 frames/s of the OCT probe rotation, the measurement time for 50 images is about 2 s.

4 Conclusion and Future Work

A novel STC observer was introduced to diagnose early airway edema from full-range OCT images. The initial feasibility of the observer was investigated as a means to detect and classify surgically simulated edema in the subglottic airway of rabbits incorporating statistical decision theories. The ROC curve is an extremely useful tool that: (1) can measure image quality, (2) evaluate the performance of decision strategies, and (3) summarize the difficulty of such classification tasks. Therefore, it is essential to first extract ROC curves from full-range OCT airway images in vivo. In this paper, ROC curves were derived from test statistics that were generated by applying the STC observer to in-vivo OCT images. AUCs calculated from the ROC curves were around 0.8, which implies meaningful differentiation of surgically created edema from normal airway tissue. As mentioned previously, ROC curves are independent of prevalence so the estimated performance of this approach would not be affected by the occurrence rate of edema in a given airway. Furthermore, since the detectability is quantified by AUCs, it is possible to use this information as a performance metric to optimize the parameters for OCT systems. Likewise, the comparison between different full-range OCT systems (e.g., differently designed OCT probes) can be studied by this task-based assessment method.

While the STC observer was developed specifically to diagnose subglottic airway edema using full-range OCT in this paper, this metric could be applied to other OCT applications. For example, there are numerous studies aimed at diagnosing macular edema, where OCT is used to measure retinal thickness as the main indicator for this disease. Combining the STC concept with retinal thickness estimation tasks might improve the diagnostic accuracy and contribute to development of enhanced computer-aided detection algorithms. Also, OCT has recently emerged as the modality to estimate cerebral edema as well as to image cerebral cortex and subcortical structures, where the proposed STC concept may also have significant utility.

With respect to the clinical applicability of a full-range OCT system using the STC observer, we highlight a couple of points that are related to possible future research steps. Since the purpose of this study is to confirm and evaluate the feasibility of the STC method, the simulated edematous tissues were verified with surgical endoscopy and their locations were known a priori. Such would not be the case in actual clinical implementation, especially for the purpose of diagnosing an early stage of airway edema. In order to supplement these shortcomings on the clinical applicability of an STC observer, STC observer sensitivity to the extent of the edema in vivo must be studied; this is controlled by the amount of injected saline. Also, it is necessary to reconstruct 3-D airway with the functionality for detecting arbitrarily located edema, so detectability can be performed with location estimation. One potential approach to achieve these tasks would be to acquire OCT image sets for multiple segments of the airway through the entire subglottic area (e.g., 100 images for every 0.5 cm spatially separated airway segment), and then perform the STC diagnostic tests for each image set as well as reconstructing the entire airway in 3-D. Further developing the STC observer as a clinical tool would require broad clinical studies, and our group has already performed full-range OCT imaging in nearly 100 neonates with comprehensive data sets. We hope to pursue such studies in the future and develop a means to better detect, diagnose, and monitor the formation of edema in the neonatal airway.

Acknowledgments

The authors express gratitude to Ashley Hamamoto, BS and Erica Su, BS for their assistance with in-vivo experiments. This work was supported by the National Institutes of Health under Grants R01HL-105215, R01HL-103764, and P41EB015890, and the newly appointed professor research fund of Hanbat National University in 2013.

References


**DongYel Kang** is an assistant professor at Hanbat National University, Daejeon, Korea. He received his BS and MS degrees in physics from Pusan National University and KAIST, and his PhD degree in optical sciences from College of Optical Sciences, University of Arizona. His research interests include characterization and optimization of the performance of imaging systems, such as frequency-domain diffusive imaging, OCT, and photoacoustics, considering image science approaches. He is a member of SPIE and OSA.

**Zhongping Chen** is a professor of biomedical engineering and the director of the F-OCT Laboratory at the University of California, Irvine. He is a co-founder and the board chairman of OCT Medical Imaging Inc. His research group has pioneered the development of Doppler optical coherence tomography. He has published more than 200 peer-reviewed papers. He is a fellow of the American Institute of Medical and Biological Engineering, a fellow of SPIE, and OSA.

Biographies for the other authors are not available.