Improvement of tissue analysis and classification using optical coherence tomography combined with Raman spectroscopy

Chih-Hao Liu*, Ji Qi†, Jing Lu†, Shang Wang*, Chen Wu*, Wei-Chuan Shih‖§¶ and Kirill V. Larin*‡¶
*Department of Biomedical Engineering
University of Houston
3605 Cullen Blvd., Houston, Texas 77204-5060, USA
†Department of Electrical and Computer Engineering
University of Houston
4800 Calhoun Rd., Houston, Texas 77204, USA
‡Department of Molecular Physiology and Biophysics
Baylor College of Medicine
One Baylor Plaza, Houston, Texas 77030, USA
§wshih@uh.edu
¶klarin@uh.edu

Received 15 March 2014
Accepted 10 April 2014
Published 12 June 2014

Optical coherence tomography (OCT) provides significant advantages of high resolution (approaching the histopathology level) real-time imaging of tissues without use of contrast agents. Based on these advantages, the microstructural features of tumors can be visualized and detected intra-operatively. However, it is still not clinically accepted for tumor margin delineation due to poor specificity and accuracy. In contrast, Raman spectroscopy (RS) can obtain tissue information at the molecular level, but does not provide real-time imaging capability. Therefore, combining OCT and RS could provide synergy. To this end, we present a tissue analysis and classification method using both the slope of OCT intensity signal vs depth and the principle components from the RS spectrum as the indicators for tissue characterization. The goal of this study was to understand prediction accuracy of OCT and combined OCT/RS method for classification of optically similar tissues and organs. Our pilot experiments were performed on mouse kidneys, livers, and small intestines (SIs). The prediction accuracy with five-fold cross validation of the method has been evaluated by the support vector machine (SVM) method. The results demonstrate that tissue characterization based on the OCT/RS method was superior compared to using OCT structural information alone. This combined OCT/RS method

†Corresponding authors.

This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution 3.0 (CC-BY) License. Further distribution of this work is permitted, provided the original work is properly cited.
is potentially useful as a noninvasive optical biopsy technique for rapid and automatic tissue characterization during surgery.

Keywords: OCT signal slope; principle component analysis; multi-support vector machine; Raman spectra.

1. Introduction

Surgical resection is currently one of the most effective solutions for cancer treatment. Minimizing the removal of normal tissues can maintain the functions of the organs, cosmetic appearance, and further reduce the possibility of chronic complication such as lymphedema. However, during the resection surgery, incomplete resection is the main factor associated with the risk of the local recurrence. Thus, accurate intraoperative margin assessment is required for reducing the prevalence of repeated surgeries. Currently, intraoperative gross examination (macroscopic) for tumor margin is performed by the naked eye, and, if a suspicious region is identified, more accurate diagnosis can be obtained via microscopic analysis of frozen section. However, this is time-consuming and can take up to an hour and can be biased by the small area of tissue examination. Thus, the outcome of the resection surgery can be significantly enhanced by availability of a real-time, noninvasive and cross-sectional imaging tool capable of the intraoperative detection of tumor margins.

Optical coherence tomography (OCT) is a non-invasive imaging technique that provides high spatial resolution and cross-sectional images of tissues and cells. The imaging depth of OCT can reach up to 3 mm in biological tissues. Compared with ultrasound imaging, OCT provides better spatial resolution (\( \approx 2-10 \mu m \)) at least an order of magnitude. Images obtained using OCT have also been shown to be well correlated with histological sectioning. Because of the superior resolution and real time advantages, the integrated system of laparoscopic OCT has been successfully developed to detect ovarian cancer. In addition to classifying tissues based on their structural image, it has been demonstrated that OCT has the abilities to differentiate diseased region from normal tissues based on their optical attenuation and elastic properties. However, sensitivity and specificity of OCT imaging for detection of tumor boundaries need to be improved before they can be clinically accepted.

Raman spectroscopy (RS) provides a way to measure molecular composition of tissues through inelastic light scattering. The frequency shift of the vibrational mode produces the specific characteristic peak corresponding to the given molecular vibration state. Thus, the distinct fingerprint for many molecules can be formed. RS has been widely used on the examination of cancer pathology, the results show that RS is capable of the tumor detection with high classification accuracy. However, because of the weak nature of inelastic scattering, RS is typically implemented as point measurements instead of a rapid spectral imaging over a large spatial area. Consequently, the combination of structural analysis from OCT and the molecular information from RS can be a good way to compensate for the limitation of both OCT and RS. In previous works, the preliminary results show that integrated OCT and RS system is a promising tool for classifying skin cancer. However, the multimodal analysis of combining the morphological and molecular information can be as important as the integration of the hardware. Recently, Ashok et al. has reported a multimodal analysis applied on colonic adenocarcinoma, where they combined the morphological indicators from OCT with RS to have improved accuracy compared with simply the Raman results. The morphological properties were quantified from the gray level co-occurrence matrix (GLCM) and the results of combination method were evaluated by a binary support vector machine (SVM) classification.

In this study, we introduce OCT attenuation parameter for a combined analysis of OCT/RS with the proof-of-principle experiments performed on mouse abdominal organs, such as the intestine, the kidney and the liver. With RS, the distinct fingerprints from these type of organs are well studied. Due to the similar features shown in the OCT images of these organs, we utilize the molecular information from RS to improve the performance of the OCT method. For the tissue characterization, the indicators from RS were obtained through
principal component analysis (PCA), and the indicators from OCT were selected as the OCT signal slope (OCTSS) and the standard deviation (SD) of intensity fluctuation. Compared with the OCT detection only, our results demonstrate improvements in the classification accuracy of OCT by combining with the RS spectrum analysis. Our study suggests that this combined OCT/RS method is potentially helpful to further assist in the intraoperative tumor margin delineation efforts.

2. Materials and Methods

2.1. OCT and RI systems

The OCT imaging is performed using a time-domain OCT system schematically shown in Fig. 1(a). The system applies a low-coherence light source with a central wavelength of 1310 nm, and a bandwidth of ~30 nm. The axial resolution of the system is ~25 μm in air. Light in the sample arm of the interferometer is directed into the tissues through a miniature scanning endoscopic probe with the diameter of ~3 mm (Imalux Corporation, Cleveland, Ohio). The in-depth (z-axis) information of the sample is achieved by the piezoelectric modulation of the optical path in the reference arm. Two-dimensional images of samples were obtained with the size of 450 pixels × 450 pixels and the dimension of 2.2 mm (axial) × 2.4 mm (transversal).

The B-mode imaging rate was ~0.25 Hz and the operation of the OCT system was fully controlled by PC.

The Raman imaging (RI) system used in this study is described in Fig. 1(b). We employ a CW Titanium: Sapphire laser (Spectra-Physics 3900S) pumped by a diode-pumped solid-state green laser (Spectra-Physics Millennia X). The laser output filtered by a 785 nm laser line filter (Semrock LL01-785-12.5) is transformed by a line-generating optical system consisting of a Powell lens and two cylindrical lenses to form a uniform line, which is 133 μm long. This line is relayed to the side-port focal plane of an inverted microscope (Olympus IX71). A dichroic beamsplitter (Semrock LPD01-785RU-25) is placed on the beam path to reflect the laser light. Epi-Raman is collected by a microscope objective (Olympus UPLSAPO 60XW) and sent through the dichroic beamsplitter and a long-wave pass filter (Semrock LP02-785RS-25) for laser intensity reduction, and finally, imaged at the entrance slit of a spectrograph with a CCD detector (Princeton LS785). The spatial resolution of the system is 0.6 μm × 0.8 μm, and the spectral resolution is 8 cm⁻¹.

2.2. Animal manipulation

All the animal manipulation procedure described below has been approved by the Institutional
Animal Care Committee (IACUC) of University of Houston. Six mice were housed under specific pathogen-free conditions in the animal facility. In OCT experiments, three mice were utilized and these mice were deeply anesthetized through the use of isoflurane gas mixed with oxygen for the in vivo experiments. Another three mice were used for the RS experiments with ex vivo tissue conditions. For RS experiments, the animals were euthanized and organs such as kidney, small intestine (SI), and liver were harvested. The RS study followed immediately after extracting the organs from the mice. The total process of organ extraction was within 20 min.

2.3. Data acquisition and processing

The OCT experiments were performed on three mice in vivo with three different types of organs: SI, kidney and liver. We quantify the slope and the SD of the OCT signal as the indicators for the tissue characterization. The procedure workflow is shown in Fig. 2. First, the surface of the tissue image was flattened before further processing. Because the large mismatch of the refractive index at the tissue-air interface causes a strong reflection on the surface of the tissue sample, we chose the region \(0.14 \text{mm} \) below the sample surface and \(0.42 \text{mm} \) in width for analysis as shown in Fig. 2(b). Also, the analysis was performed on the image parts that do not involve the motion artifacts associated with animal breathing (see examples in Fig. 3). In order to suppress speckle noise of the A-line signals, the intensity of the A-lines were averaged in the lateral direction approximately every \(53 \mu\text{m} \) (10 A-lines). The typical A-line intensity resulted as a function of depth (rescaled from dB unit) is shown in Fig. 2(c). Linear fit was applied to the selected axial region to obtain the slope of the intensity A-line as the first parameter for the tissue characterization. For SD calculation, linear fit is applied to the A-line signal directly without the averaging. The slope is then removed by performing a subtraction of the linear fit, resulting in an intensity signal with the mean of approximately zero, as shown in Fig. 2(d). The SD of this slope-removed signal was then quantified as the second parameter. The OCT signals from 5 positions (5 frames) were measured randomly from the surface layer of each organ. In each mouse organ, we acquired 5 frames and each frame included 45 slope values and 450 SD values. These values from one position were averaged to one value for the statistical process. Finally, the two-dimensional computational plot of the slopes and SD values was achieved and the performance of the method was computed by multi-SVM method.

In the Raman experiment, the measurement was performed ex vivo in the same organ as with the
OCT experiment. Tissue samples were placed on fused silica coverslips and fixed at the sample holder of microscope. The laser line projected at the sample was 133 μm long. 133 spectra were collected simultaneously with a total acquisition time of 60 s at a power density of 1 mW/μm² to obtain excellent SNR. The spectra were based line corrected using adaptive iterative reweighted penalized least squares algorithm (airPLS) and the 133 corrected spectra were averaged as one spectrum representing one position. Totally, 5 positions were randomly measured from the surface of each organ, and the mean spectra of three organ types were plotted (see example in Fig. 5(a)). Finally, we applied PCA to quantify the variation of Raman spectra, and used multi-SVM method to assess the performance of the Raman data.

2.4. Characterization of the method

Multi-SVM is a common machine-learning tool used to classify new unknown data (testing data) based on experiences from a dataset of known data (training data) among multiple classes. We selected one-versus-all (OVA) method with a radial basis function (RBF) kernel. The OVA method is one of the methods to decompose a multiclass problem into two-class problems. The aim of OVA is to classify the data into the class that has the highest decision value compared with the remaining classes. The 5-fold cross validation is applied to verify this model. An example of this method is shown in Fig. 4(c). First, all the data points in Fig. 4(c) are split randomly into five groups and the classification model (RBF kernel) is trained with four groups. The fifth group is considered as the testing sample for predicting the accuracy. This is repeated for 5 times. During each time, classifiers are internally constructed with M times if there are M classes of the training data. In the prediction stage, the classifiers follow the “winner takes all” principle that the testing sample is classified into the class with highest decision value. In addition, the classifiers with different cost and gamma parameters indeed influence the prediction accuracy, thus, we search the best cost and gamma parameters by a grid search method. Also, 5-fold cross validation is utilized to optimize best parameters. Finally, the confusion table is achieved, and then, the accuracy can be computed according to the confusion table.

3. Result and Discussions

Typical OCT images of the SI, kidney and liver are shown in Fig. 3. It is fairly difficult to find the structure difference based on visual observation. The slope values of SI, kidney and liver were quantified as 0.67 ± 0.102 μm⁻¹, 0.84 ± 0.038 μm⁻¹ and 0.95 ± 0.046 μm⁻¹, respectively, as shown in Fig. 4(a). The SD values were computed as 112.06 ± 2.89, 112.97 ± 1.49, and 106.20 ± 0.70 in an arbitrary unit. From the statistical results, the mean slope value of liver is the largest because of much denser distribution of erythrocytes within connective tissues and capillary vessels than the others. On the other hand, based on the aspect of image speckles, the result of the SD values in Fig. 4(b) suggests that the variations of refractive index in the structure of SI and kidney are greater than liver. To further verify the significant difference, two-dimensional Kolmogorov–Smirnov test (K–S test) has been performed on the OCT data in Figs. 4(a) and 4(b), and it turns out that the P values of the slope among the three types of samples are all smaller than 0.001, but the P values of the SD between kidney and SI are more than 0.05, which suggests that the SD values of intensity A-lines cannot correctly differentiate SI from kidney.

Three mice from the OCT group were used for the combined OCT/RS analysis. The two-dimensional scatter plot with 95% confidence region reveals that SI and kidney cannot be well differentiated, as shown in Fig. 4(c). In Table 1, the prediction accuracy of simply OCT texture analysis is 84.4%. Also, Table 1 demonstrates that OCT analysis has relatively better performance to classify liver because of its distinct optical and structural properties while SI and kidney show worse prediction than liver. This result demonstrates that OCT analysis alone is not sufficient to identify SI and kidney although it provides high-resolution images and optical properties. Consequently, the performance of the organ classification can possibly be further improved by means of the assistance of RS due to its superior molecule specificity.

In Fig. 5(a), the mean RS spectrum of normal liver indicates the prominent peaks at 1132, 1169, 1207, 1307, 1395, 1450, 1549, 1601 and 1659 cm⁻¹. Similar observance can be found in the literature, e.g., in Ref. 26. The characteristic Raman peaks of kidney are 1132, 1276, 1306, 1456, 1549, 1617, 1662 cm⁻¹. This fingerprint is similar to the previous
Fig. 3. Typical OCT images from different organs, (a) SI, (b) kidney and (c) liver, the sharp pulses in (a)–(c) were caused from the body motion. Scalar bars are equal to 0.25 mm.

Fig. 4. Box plots of (a) the slope and (b) the SD for the organs of small intestine, kidney, liver; \( N = 3, n = 15 \) (\( N \) and \( n \) are the sample size of mice and measured position respectively). The error bars and squares in the boxplots are one statistical SDs and the mean values of each type of the organs. (c) Scatter plot of the two-dimensional distribution of the values of slope and SD.
studies as well.\textsuperscript{27,28} Also, the predominant peaks of SI are 1166, 1213, 1306, 1457, 1660 cm\textsuperscript{-1}, which are in accord with the literature.\textsuperscript{18} Among the comparison of the three in Fig. 5(a), the Raman bands of liver show the identifiable peaks based on molecular information at 1169 cm\textsuperscript{-1} (C-H in-plane bending mode of tyrosine), 1207 cm\textsuperscript{-1} (Amid III), and 1601 cm\textsuperscript{-1} (phenylalanine). Particularly, in high frequency band, the strong Raman intensity at 1601 cm\textsuperscript{-1} is observed as the most obvious feature for identifying the liver from the others. Besides, the characteristic peaks of kidney at 1549 cm\textsuperscript{-1} (C-N stretching mode/Amid II), 1617 cm\textsuperscript{-1} (Tyrosine), and 1662 cm\textsuperscript{-1} (Amid I) have the major difference compared with the liver and SI while the Raman band of SI has the strongest peak at 1660 cm\textsuperscript{-1} but no peaks from the range of 1500 cm\textsuperscript{-1} to 1620 cm\textsuperscript{-1}. All the discussed Raman peaks are organized and shown in Table 2.

The data of Raman spectra were analyzed using PCA to extract the principal components and their associated scores. In Fig. 5(b), the spectral variances of PC1 and PC2 were computed as 66.8\% and 15.3\%, respectively. The principal component

![Image](image-url)

Fig. 5. (a) Shows the mean spectrum of the different types of organs and (b) shows the loading plot of PC1 and PC2.

### Table 1. Confusion table for the SVM classification model based on OCT detection only. The optimized cost value and gamma are 5 and 0.25, respectively for the training model.

<table>
<thead>
<tr>
<th>Predicted labels</th>
<th>SI (#sample)</th>
<th>Kidney</th>
<th>Liver</th>
</tr>
</thead>
<tbody>
<tr>
<td>True label SI</td>
<td>11</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Kidney</td>
<td>2</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Liver</td>
<td>1</td>
<td>0</td>
<td>14</td>
</tr>
</tbody>
</table>

### Table 2. Observed Raman frequencies with corresponding band assignments for liver, kidney and intestine.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Raman shift (cm\textsuperscript{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver</td>
<td>1132 1132 1134</td>
</tr>
<tr>
<td>Kidney</td>
<td>1169 1169 1166</td>
</tr>
<tr>
<td>Intestine</td>
<td>1207 1213 1601</td>
</tr>
<tr>
<td></td>
<td>1281 1276 1617</td>
</tr>
<tr>
<td></td>
<td>1307 1306 1659</td>
</tr>
<tr>
<td></td>
<td>1395 1450 1614</td>
</tr>
<tr>
<td></td>
<td>1450 1455 1659</td>
</tr>
<tr>
<td></td>
<td>1549 1549 1662</td>
</tr>
<tr>
<td></td>
<td>1601 1617 1660</td>
</tr>
<tr>
<td></td>
<td>1659 1662 1660</td>
</tr>
</tbody>
</table>

### Table 3. Confusion table for the SVM classification model based on the combination of OCT and RS. The cost value and gamma are 1 and 0.1 respectively for the RBF training model.

<table>
<thead>
<tr>
<th>Predicted labels</th>
<th>SI (#sample)</th>
<th>Kidney</th>
<th>Liver</th>
</tr>
</thead>
<tbody>
<tr>
<td>True label SI</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Kidney</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Liver</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
</tbody>
</table>
loadings suggest that liver can be clearly differentiated by PC1 at the peaks 1169, 1207, 1460, 1601, 1665 cm\(^{-1}\) as well as kidney at the peaks 1142, 1169, 1222, 1281, 1395, 1460, 1549, 1616, 1653 by PC2. The score plot of PC1 vs PC2 is shown in Fig. 6(a). Compared with the scores of PC1, the scores of PC2 are more confident to identify intestine and kidney. By calculating the two-sample unequal-variance \(t\)-test, all the P values of PC2 among intestine, kidney and liver have significant difference (<0.001), demonstrating that the PC2 values can successfully detect intestine, kidney and liver. The computed result of multi-SVM shows that PC2 combined with PC1 has 100% prediction accuracy for classifying these three types of organs. From Fig. 6, the scores of PC2 are suitable to improve the inability of OCT images for differentiating SI, kidney and liver. Besides, the scores of PC1 can be utilized to further assist the performance of OCTSS for liver classification. In comparison with the SD result, the values of OCTSS have less overlap area between SI and kidney. Thus, we combined the scores of PC2 and PC1 with OCTSS based on the homogeneous assumption on the sample surface and achieved Fig. 6(b). The classification results among these three types of organs are listed in Table 3.

According to the results listed in Tables 1 and 3, the comparison between the OCT and combined RS/OCT analysis of tissues indicates that the prediction accuracy is increased from 84.4% to 100%. This clearly demonstrates that the proposed combination method can be used to detect SI, kidney and liver successfully. Although RS alone provide sufficient accuracy for classifying these three types of organs, RS is unable to provide rapid spectral mapping over large area. Instead, OCT can provide morphological imaging in real-time, while the classification accuracy of OCT is not as sufficient as RS. Our results suggest that using combined OCT/RS successfully enhances the classification accuracy for different tissues due to the better separation capability of RS. In future, the integrated system of OCT and RS can also benefit the image-guided optical biopsy in clinical use.

The results of this feasibility study demonstrate that the proposed combined OCT/RS method is promising for the classification of SI, kidney and liver. In addition, our study suggests that this combined OCT/RS method is potentially helpful to further assist in the intraoperative tumor margin delineation efforts. Currently, we are proceeding to increase the sample size and applying this method on a greater variety of organs, such as heart, spleen, stomach, and colon to fully assess this combined method for tissue characterization.

4. Conclusion

OCT technique is widely considered as a real-time intraoperative tumor margin assessment because of its high-resolution images, rapid scanning, and optical properties. However, although OCT provides
images with high-resolution, the combination of OCTSS and SD is still insufficient to effectively classify different types of internal organs through the images. The main reason is that OCT images are simply composed of the reflectivity of light (elastic scattering property), which can only reflect the texture information instead of molecular information. The similarity in our OCT images among SI, kidney and liver reveals that the accuracy of OCT needs to be improved for accurate margin delineation.

Based on our results that RS performs the high accuracy to differentiate the organs, we proposed a method to combine the analysis of OCT images with Raman spectrum for the differentiation of different types of internal organs. The proposed method combined three indicators, the slope of intensity A-line from OCT and the PCs (PC1, PC2) of PCA from RS for the classification among kidney, SI and liver. Compared with the simply OCT analysis, the proposed method improved the accuracy from 84.4% to 100%, demonstrating proposed tissue analysis achieved by combined OCT/RS is potentially useful as a quantitative and qualitative technique for rapid and automatic tissue characterization during, e.g., surgical tumor resection. After this pilot study of tissue analysis using combined OCT/RS, in future we are planning to integrate the system of OCT and RS for detecting the diseased margin of tumor surrounded by normal tissues (e.g., muscle). Based on the integrated system, OCT and RS can acquire the measurement with more similar experimental condition. Furthermore, the performance of tumor margin detection will be evaluated with different types of tumor organs.

Acknowledgments
The OCT part of this work was supported in part by the grants to Kirill Larin from NIH 1R01EY022362, 1R01HL120140, U54HG006348, and DOD PRJ71T; the RS part of this work was supported by grants to Wei-Chuan Shih from NSF CAREER Award (CBET-1151154), NASA Early Career Faculty Grant (NNX12AQ44G) and Gulf of Mexico Research Initiative (GoMRI-030). Chih-Hao Liu and Ji Qi are equally contributed.

References


