Characterization and three-dimensional localization of cancerous prostate tissue using backscattering scanning polarization imaging and independent component analysis

Yang Pu
Wubao Wang
Min Xu
James A. Eastham
Guicheng Tang
Robert R. Alfano
Characterization and three-dimensional localization of cancerous prostate tissue using backscattering scanning polarization imaging and independent component analysis

Yang Pu,
Wubao Wang,
Min Xu,
James A. Eastham,
Guicheng Tang,
and Robert R. Alfano

1 Introduction

Optical imaging technique using near-infrared (NIR) light provides an attractive approach for detecting human diseases. The key advantages of optical imaging are the inherent low-end sensitivity, the ability to monitor multiple independent optical reporters simultaneously, the absence of radioactive intermediates, and the relative simplicity of the imaging hardware as compared to magnetic resonance imaging (MRI) and positron emission tomography (PET) methods. The major disadvantage of optical imaging is that high scattering of biological tissue causes most photons to be diffused, and makes optical images blurry. Due to the high scattering of tissue, scientists have to explore optical tomography methods and/or inverse image reconstruction approaches to determine differences of spatial distributions of optical parameters between the normal and diseased tissues, and obtain the three dimension (3-D) locations of the diseased tissue. Since the light scattering depends on optical properties of the tissue, such as scattering coefficient ($\mu_s$), anisotropy factor ($g$), and absorption coefficient ($\mu_a$), the measured light intensity distribution on the boundary of the tissue can be used to generate a map of the $\mu_s$ and $\mu_a$ for the tissue areas using an inversion algorithm.

Many algorithms were developed and tested to reconstruct images of objects hidden inside a scattering medium. The use of independent component analysis (ICA) was introduced from information theory into the field of biomedical optics to obtain 3-D locations of objects embedded in thick turbid media. This information-theory-inspired approach is referred to as optical tomography using independent component analysis (OPTICA). The theory of the OPTICA algorithm was described elsewhere. The efficacy of the OPTICA approach has been demonstrated for the conditions of intralipid scattering medium and human breast tissues in transmission geometry. The OPTICA combining with the multiple source and multiple detector (S-D) measurements was used to extract ICs in the forward scattering model and obtain the 3-D locations of the targets in the host media. In this research, backscattering scanning imaging measurements were performed and the backscattering model of the OPTICA algorithm was used for image analysis. The light backscattered from a tissue sample in response to a scanning laser beam (working as multiple sources) was recorded with a CCD camera (working as multiple detectors), and was used to obtain multiple angular views of the objects embedded in the tissue. The OPTICA algorithm was improved specifically...
for the application of backscattering configuration. The difficulty arises in backscattering geometry because the incident beam profile and the surface property of the sample appreciably affect the spatial distribution of the backscattered light. Such issues are not as important in transmission geometry. We addressed this challenge by improving the OPTICA for the backscattering model with two procedures: 1. synthesizing a “clean” background image of the host normal prostate tissue from the recorded scanning images of the tissue sample with embedded cancerous prostate tissue; and 2. numerically marching the Green’s function of the light intensity distribution from the target to the surface of the medium until matching the retrieved ICs, incorporating both the beam profile and the surface property of the tissue sample. The existence and 3-D locations of a small piece of cancerous prostate tissue embedded in the host normal prostate tissue were determined using the recorded set of 2-D backscattering images and the OPTICA algorithm, and the results were compared with the known 3-D locations of the cancerous tissue.

The remaining parts of this paper will first briefly introduce the OPTICA algorithm, and then describe our scanning backscattering imaging methods and results with discussion.

2 Formalism and Algorithm of OPTICA for Backscattering Geometry

The theory and major theoretical formalism of the OPTICA have been described in the previous works. The ICA, core of the OPTICA, is a solution to the blind source separation problem and has emerged as a new paradigm in signal processing and data analysis. Blind source separation is a class of problem of general interest that no precise knowledge is available for neither the mixing channel nor the sources. In the OPTICA approach, the diffusion approximation (DA) was used to solve the radiative transfer equation (RTE) since the diffusion approximation is valid when the guest objects are within a highly scattering medium. The “guest objects” are defined as the objects whose absorption and scattering coefficients and/or other optical properties are different from the host medium. The solution of DA for the propagation of light from a point-source in the scattering media is a Green function. The measured array of images is given by

\[ I(\rho_d, \rho_s) = I_0 G_0(r_d, r_t) \]

\[-\psi I_0 \Delta z \int G_0(r_d, \rho', z') \delta \mu_0(\rho', z') G_0(\rho', z'; r_t) d^2 \rho', \]

where \( \rho_d \) stands for detection positions covering the whole 2-D array, \( \rho_s \) stands for scanning positions of incident light with the intensity of \( I_0 \) at \( z = 0 \), and \( G_0 \) denotes the Green’s function for light propagation in the host medium. It is assumed that one tiny object (with absorptive perturbation \( \delta \mu_0 \) from the host medium in volume \( \Delta V \)) is located at \( r' = (\rho', z') \) with the extension \( \Delta z \ll 1 \) along one direction, and \( \delta \mu_0 \) is a constant within \( \Delta V \) and 0 outside \( \Delta V \). The speed of light is set to be unity for clarity. Scattering targets can be treated in a similar fashion.

The background image was approximately generated by averaging all of the acquired images by shifting the scanning position for each image with a specific \( \rho_s \) to the origin 0 in the previous study. The average of the shifted images is given by Ref. 14:

\[ \bar{I}(\rho_d, 0) = I_0 G_0(r_d, 0) - h(\rho_d, 0), \]

where the term \( h \) stands for the error term. This approximation is based on the assumption of a large number of acquired images, and a small number and tiny size of objects. This “averaged” background image is called “dirty” background image since the object(s) with \( \delta \mu_0 \) and/or \( \delta \mu_0 \) is included in the averaging calculation. To discard the “dirty” background image that is influenced more than noise level by the embedded object(s), a simple and fast estimation of threshold value of \( h \) is needed to generate the “clean” background image. This can be achieved by selecting and averaging all shifted images that are minimally perturbed by the embedded objects. The average image can be expressed as

\[ I_c = \bar{I}(\rho_d, 0) + h(\rho_d, 0), \]

where the error term can be chosen by

\[ h(\rho_d - \rho_s, 0) = \frac{1}{N_B} \sum_{p \in B} \Delta I(\rho_d - \rho_s, 0), \]

where \( B \) denotes the perimeter of the scanning grid that contains \( N_B \) scanning positions. After \( h \) and the “clean” image of the host medium are obtained, the difference images namely perturbation images, can be calculated more accurately by subtracting the “clean” background image from the recorded set of 2-D images.

After the perturbation images have been obtained, ICA can be used to unmix the signals arising from individual targets, and obtain the independent intensity distribution contributed from each of the target(s). Suppose the certain ICs correspond to the targets, which contribute to the projection of the Green’s functions of \( G_0(r_t, r_t) \) and \( G_0(r_d, r_t) \) on the source-and-detector plane. In the backscattering geometry, the coherent backscattering (CBS) of light in random media will generate a peak of scattered light intensity in the backward direction. There is no analytical solution for the Green’s function in this case, and numerical simulations such as Monte Carlo methods have been used instead of generating the Green’s function. To address this challenge, light propagation from \( r_t \) to \( r_d \) was approximated by the fact that: 1. forward propagation light first bounces back from the bottom of the semi-infinite medium \( z > z_0 \); and 2. light undergoes diffusing from the layer \( z = z_t \) to the top surface \( z = z_d \). By performing Fourier transform over the lateral coordinates, the projection of the Green’s functions for the target, \( G_0(r_d, r_t) \), on the detector plane can be written as a simple relationship:

\[ G_0(r_d, r_t) = \frac{1}{2\pi} \int G_0(q_0, 0) \frac{\partial g(q, z_d, z_t)}{\partial z_t} \]

where \( G_0(q) \) is the Fourier transform of the intensity of the IC for the object on the detector plane, \( G_0(q, 0) \) is the Fourier transform of the clean image \( \bar{I}_c \), and \( g(q, z_d, z_t) \) is the Fourier transform of \( g(\rho, z_d, z_t) \) representing the propagation of a plane wave of spatial modulation frequency \( q \) from the \( z = z_t \) plane to the surface \( z = z_d \), and \( g(\rho, z_d, z_t) \) is the Green’s function inside an infinite homogeneous medium for a diffuse photon density wave using the first Rayleigh-Sommerfeld integral while the integration is performed over the \( z = z_t \) plane. Considering the detection condition that only photons
escaping from the medium in the normal direction can be detected, and using the center-moved diffusion model,\(^1\) the 
\[
g(\rho, z_d, z_t) = \exp(-kr)/r, \tag{6}
\]
where \(r = \sqrt{\rho^2 + (z_j - z_d)^2}\) and \(k = (\mu_\text{abs} - i\omega)/D_0\) indicates the attenuation coefficient for the incident beam with intensity modulation frequency of \(\omega\), in which \(D_0 = l_\text{t}/3\) is the diffusion coefficient, \(l_\text{t}\) is the transport mean free path, \(\mu_\text{abs}\) is the absorption coefficient for the host medium. If taking \(z^d = z_d + l_\text{t}\), which is the location of the effective source by displacing the incident point of the collimated beam by one \(l_\text{t}\) along the incident direction, Eq. (5) is then simplified as\(^1\)
\[
G_0(q) = G_0(q, 0) \exp\left(-Q(|z_t - z_d|)\right), \tag{7}
\]
where \(Q \equiv \sqrt{q^2 + k^2}\). The position of the target \((z_t)\) can be obtained by fitting \(G_0(r_d, r_s)\) from the inverse Fourier transform of Eq. (6) to the retrieved independent intensity distribution on the detector plane contributed by the target.

The graphical illustration of the OPTICA algorithm is shown in Fig. 1. Figure 1(a) shows the pathway of backscattered light in one direction and the formation of a backscattering image of the host scattering medium without guest object(s) acquired from photons surviving back from the medium, which carry the information of optical coefficients of the medium. Figure 1(b) exhibits the voyage of the backscattered light in one direction and their intensity distribution on the source-detector plane for the host medium with one embedded guest object. The light intensity distribution is affected by the photons having interaction with the guest object and the optical coefficients of the tissue. If the difference of the intensity distribution of images recorded at those two different conditions is detectable, ICA can be applied to retrieve the information of the object (IC).

A more real-world situation (multiple guest objects) is shown by Fig. 1(c), in which the 2-D intensity distribution of a backscattering image is contributed by photons having interaction with the multiple guest objects. The OPTCA can be used to recover each IC (target) from the mixture perturbation. This can be achieved by using the approach of multiple detectors (or CCD camera) and multiple sources (or scanning the illuminated laser source at different positions on samples in our case) as shown in Fig. 1(d). Each recovered IC is used to obtain the 3-D localization of the corresponding target by marching the propagation of the scattered light from the target to the surface until matching the retrieved IC. The OPTICA algorithm has high sensitivity to detect a small-sized target with detectable contrasts to the host medium because of its capability to separate the signal of each IC from huge background noise of the host medium.

ICA, the core of the OPTICA algorithm, can be used to extract independent intensity distributions (IID) contributed by each IC from the total perturbations of the spatial intensity distributions on the detector plane caused by guest objects.\(^3,4\) The procedures of the scanning imaging and OPTICA approach for obtaining 3-D locations of the guest objects in the host media can be described as six steps shown in Table 1.

3 Experimental Methods and Samples

The experimental arrangement for the NIR scanning polarization imaging system is schematically shown in Fig. 2. The sample is illuminated by a collimated laser beam with \(\lambda = 635\,\text{nm}\) in the direction close to the normal to the surface of the medium. A narrow band filter (NBF) was used to block the autofluorescence light contributed from the tissue spectral wing emission\(^1\) so that only the light backscattered from the samples was detected. Polarizer \(P_1\) was used to ensure the linear polarization of the input laser pulses and \(P_2\) was set to 90 degree with respect to that of \(P_1\) to record the image in perpendicular
polarization configuration. The perpendicular image was acquired to cancel the surface effect of the tissues. Two galvanometric mirrors and their controllers (General Scanning Inc. Lumonics-GSIL, Model DSC W/HCl) were used to scan the illuminating beam along the $x$- and $y$-directions on the front surface of the sample. The 2-D images formed by the light backscattered from the sample were recorded with a CCD camera (Photometrix CH350L, 1024×1024 pixels, 16 bit) after passing through the camera focus lens. A custom-made software developed using GSIL, WinMCL Plus was used to control the scanning speed, the 2-D scanning grid, and to trigger the recording of the backscattering image at each scanning position. The control board for the CCD camera (Photometrix PCI-X 01-490-400 with PVCAM driver) and the corresponding imaging software (Digital Optics V++) were used to record the 2-D backscattering images. To define the coordinate system, a front view of the sample position is shown in the insert of Fig. 2, indicating the case of the cancerous prostate tissue covered by a large piece of normal prostate tissue. The spot size of the scanning beam is $\sim 500 \mu m$. The scanning grid is $8 \times 8$ with a step size of 2.13 mm.

Two kinds of tissue samples were tested. The first sample used for testing the technique of NIR scanning polarization imaging and the algorithm of OPTICA was a black rubber absorber ($\sim 2.7 \times \sim 3.0 \times \sim 1.5 \ mm$) embedded in chicken breast tissue with different depths. The thickness of the back tissue is $\sim 8 \ mm$ with a lateral dimension of $38 \times 29 \ mm$. A transport length $l_i = 1.1 \ mm$ and an absorption coefficient $\mu_a = 0.007 \ mm^{-1}$ were used for chicken breast tissue in the OPTICA analysis. The second study was designed to distinguish cancerous prostate tissue from surrounding normal prostate tissue. The sample consists of a small piece of cancerous prostate tissue ($4 \times 4 \times 1.5 \ mm$) embedded inside large pieces of normal prostate tissue at the depth of $z = 3.0 \ mm$. The thickness of the whole sample is 10 mm. In the OPTICA analysis, the absorption and reduced scattering coefficients at 635 nm were taken as $\mu_a = 0.026 \ mm^{-1}$ and $\mu_s' = 0.53 \ mm^{-1}$ for the host normal prostate tissue whereas $\mu_a = 0.0025 \ mm^{-1}$ and $\mu_s' = 0.44 \ mm^{-1}$ for the cancerous prostate tissue estimated from our earlier spectroscopic investigations of ex vivo prostate tissues. The less absorption and reduced scattering coefficients of cancerous prostate tissue were also reported in our and other groups’ previous studies.

![Fig. 2](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics) A schematic diagram of the experimental arrangement used for scanning backscattering polarization imaging for a small piece of cancerous prostate tissue embedded in large pieces of host normal prostate tissue. $P_1$ and $P_2$: polarizer, NBF: 635 nm narrow band filter.

---

**Table 1** The procedures of the scanning imaging and OPTICA approach.

1. Record a set of 2-D ($x,y$) images by scanning an illumination laser beam on sample surface;
2. Obtain a clean image of the host medium by selecting and averaging all shifted images that are minimally perturbed by embedded guest objects;
3. Calculate the perturbation of 2-D intensity distributions of scanning images by subtracting the clean image from the recorded images;
4. Perform OPTICA analysis using the perturbation images to recognize leading independent components (ICs);
5. Obtain the contribution of each target to the perturbed 2-D intensity distribution by treating targets as independent components, and obtain their ($x$, $y$)-positions using the corresponding IC-perturbed 2-D intensity distributions; and
6. Estimate z-position of the target relative to boundary surface by numerically marching the propagation of the light from the target to the surface until matching the retrieved IC, incorporating both the beam profile and the surface property of the sample.
4 Experimental Results and Discussion

A set of scanning backscattering polarization images of a cancerous prostate tissue embedded in normal tissue were acquired and processed. Figure 3 shows some typical scanning images chosen from the total 64 (8 × 8) 2-D images recorded by scanning the incident beam on the x-y plane of the sample. The sample consists of a small piece of cancerous prostate tissue hidden in the host normal tissue. The markers of (1, 1), (1, 8), (8, 1), (8, 8), (3, 5), (3, 6), (4, 5), and (4, 6) indicate the (x, y) positions of the scanning illuminating beam for which corresponding 2-D images were recorded. There is no big difference among the incident light intensity distribution patterns shown in these recorded images. This is because the dominant peak intensity of light backscattered from the surface of the

Fig. 3 A part of the 64 (8 × 8) 2-D images recorded by scanning the incident beam on the x-y plane of the sample. The sample consists of a small piece of cancerous prostate tissue hidden in the host normal tissue. The markers of (1, 1), (1, 8), (8, 1), (8, 8), (3, 5), (3, 6), (4, 5), and (4, 6) indicate the (x, y) positions of the scanning illuminating beam for which corresponding 2-D images were recorded.

Fig. 4 (a) The clean host image shown in a 10-base logarithm scale obtained by selecting and averaging the minimally perturbed 2-D images; and (b) spatial intensity distribution of the clean image.
prostate tissue suppresses the intensity perturbation caused by $\delta \mu_a$ and/or $\delta \mu_s$ between cancerous and normal prostate tissues. To overcome this difficulty, the “clean” background image was synthesized, which stands for the image of the host normal prostate tissue without cancerous tissue embedded. In the synthesis of the “clean” background image, the incident light position of each of the recorded $8 \times 8$ images was set to origin. All 2-D images are then shifted to the origin and the size of each image was cropped at the boundary while incident light distribution reaches the noise level. The images minimally perturbed by the embedded targets were selected. The “clean” background image was generated by averaging all selected minimally perturbed images, and is shown in Fig. 4 in a 10-base logarithm scale.

After the “clean” background image was obtained, the perturbation 2-D ($x$-$y$) images were generated by extracting the “clean” background image from the recorded set of 2-D images. The generated $8 \times 8$ perturbation images were used to recognize ICs (targets) and locate their 3-D positions using the OPTICA algorithm. Figure 5 shows some typical perturbation images, located in the corner areas [(1, 1), (1, 8), (8, 1), and (8, 8)] and in the inner areas [(3, 5), (3, 6), (4, 5), and (4, 6)], chosen from the total 64 perturbation images. The perturbation produced by the cancerous tissue can be clearly observed. It can be seen from Fig. 5 that the perturbation of the intensity distribution shown in images located in the boundary areas is very small, indicating that the cancerous tissue was embedded far away from the boundary areas. At positions between rows 3 and 4, near column 6, the perturbation of the image intensity distribution can be clearly observed, indicating that cancerous tissue is located in these areas.

The ICA is then performed upon the $8 \times 8$ perturbation images to recognize leading ICs. The contribution of each target to the perturbed 2-D intensity distributions can be extracted by treating each target as an IC. The ($x$, $y$)-positions of the target(s) can be obtained using the corresponding IC-perturbed 2-D intensity distributions. Figure 6 shows the OPTICA-generated intensity distributions on the detector plane contributed from the ICs for the prostate tissue sample. Figure 6(a) and 6(c) are for the leading IC (the cancerous prostate tissue), and Fig. 6(b) and 6(d) are for the residual (noise) component. It is clear that the existence of the small piece of cancerous prostate tissue can be realized from Fig. 6(a). The $x$- and $y$-locations of the cancer tissue can be obtained from Fig. 6(c). The random intensity distribution of the residual (noise) shown in Fig. 6(b) and 6(d) indicates the validation of our scanning imaging and OPTICA approach for the detection of cancerous prostate tissue embedded inside the host normal prostate tissue. Furthermore, the signal strength shown in Fig. 6(c) is much stronger than that in Fig. 6(d), indicating that outstanding intensity distribution peak shown in Fig. 6(c) generated by the OPTICA is caused by the leading IC.

The OPTICA-generated intensity distributions of the leading IC on the detector plane can be also used to locate the $z$-position of the cancerous tissue by numerical matching. Figure 7 shows Green’s functions for (a) the light spatial distribution without targets obtained from the clean image, and (b) the light intensity distribution caused by the leading IC obtained from the IC-perturbed images. The $z$-location of the cancerous prostate tissue (the leading IC) can be obtained by numerically marching the Green’s function of the light intensity spatial distribution from the target to the surface of the medium using diffusion approximation as shown by Eq. (6). The fitting yields $z = 3.1$ mm.

Table 2 summarizes all of the experimental conditions, and lists the OPTICA-determined positions of the guest objects.
Table 2 shows that the retrieved 3-D locations of the small piece of cancerous prostate tissue embedded in the host normal prostate tissue are in good agreement with its known 3-D position. In the case of the black rubber hidden in the chicken breast tissue, the backscattering imaging experiments were carried out with an optical round-trip path up to \( \sim 12 \) times larger than the transport mean-free path of the tissue medium.

**Fig. 6** OPTICA-generated intensity distributions of the independent components on the detector plane: (a) and (c) are for the leading IC; and (b) and (d) are for the residual (noise) component.

**Fig. 7** Green’s functions of (a) the light spatial distribution without targets obtained from the clean image, and (b) the light intensity distribution caused by the leading independent component obtained from the IC-perturbed images.


5 Conclusion

Scanning polarization imaging measurements and the OPTICA-based image analysis were carried out to characterize and obtain the 3-D locations of cancerous prostate tissue embedded in normal prostate tissue in backscattering geometry for the first time. Two key steps of the OPTICA approach were developed and improved to address problems raised in the backscattering geometry: 1. synthesizing a “clean” background images from a set of measured images; and 2. localizing target positions by numerical marching the target to the surface of the host turbid medium. The retrieved 3-D locations are in good agreement with the known positions of the embedded objects. The approach is applicable to different medium geometries, and amenable to near-real-time imaging application for in vivo detection and localization of cancerous sites in the host prostate.

6 Future Work

The study presented here was focused on in vitro prostate tissue specimens only. In the future, we plan to perform in vivo tests to evaluate the potential of the scanning backscattering polarization imaging and the OPTICA-based image analysis for the clinical applications, and investigate the inhomogeneous properties of normal and cancerous prostate tissues.

Acknowledgments

This research is supported by U.S. Army Medical Research and Material Command grants of W81XWH-08-1-0717 (CUNY RF #47170-00-01), W81XWH-10-1-0526, and W81XWH-11-1-0335 (CUNY RF #47204-00-01). M. Xu acknowledges additional support from Research Corporation and the National Institutes of Health (1R15EB009224). The authors acknowledge the help of Cooperative Human Tissue Network (CHTN) and the National Disease Research Interchange (NDRI) for providing normal and cancerous prostate tissue samples for the measurements.

Table 2

<table>
<thead>
<tr>
<th>Objects</th>
<th>Covered tissue</th>
<th>Known (mm)</th>
<th>OPTICA (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancerous prostate tissue</td>
<td>Normal prostate tissue</td>
<td>(13.5, 10.1, 3.0)</td>
<td>(13.8, 10.3, 3.1)</td>
</tr>
<tr>
<td>Black rubber</td>
<td>Chicken breast tissue</td>
<td>(15.6, 10.4, 5.1)</td>
<td>(15.2, 10.2, 4.7)</td>
</tr>
<tr>
<td>Black rubber</td>
<td>Chicken breast tissue</td>
<td>(15.6, 10.4, 6.6)</td>
<td>(15.3, 10.8, 6.2)</td>
</tr>
</tbody>
</table>

References