Objective analysis of collagen organization in thyroid nodule capsules using second harmonic generation microscopy images and the Hough transform

JUAN M. BUENO,1,* FRANCISCO J. ÁVILA,2 RADU HRISTU,3 STEFAN G. STANCIU,3 LUCIAN EFTIMIE,3 AND GEORGE A. STANCIU3

1Laboratorio de Óptica, Instituto Universitario de Investigación en Óptica y Nanofísica, Universidad de Murcia, Campus de Espinardo (Ed. 34), 30100 Murcia, Spain
2Departamento de Física Aplicada, Universidad de Zaragoza, 50009 Zaragoza, Spain
3Center for Microscopy-Microanalysis and Information Processing, University Politehnica of Bucharest, 060042 Bucharest, Romania

*Corresponding author: bueno@um.es

Received 26 March 2020; revised 23 June 2020; accepted 29 June 2020; posted 29 June 2020 (Doc. ID 393721); published 4 August 2020

Papillary carcinoma is the most prevalent type of thyroid cancer. Its diagnosis requires accurate and subjective analyses from expert pathologists. Here we propose a method based on the Hough transform (HT) to detect and objectively quantify local structural differences in collagen thyroid nodule capsules. Second harmonic generation (SHG) microscopy images were acquired on non-stained histological sections of capsule fragments surrounding the healthy thyroid gland and benign and tumoral/malignant nodules. The HT was applied to each SHG image to extract numerical information on the organization of the collagen architecture in the tissues under analysis. Results show that control thyroid capsule samples present a non-organized structure composed of wavy collagen distribution with local orientations. On the opposite, in capsules surrounding malignant nodules, a remodeling of the collagen network takes place and local undulations disappear, resulting in an aligned pattern with a global preferential orientation. The HT procedure was able to quantitatively differentiate thyroid capsules from capsules surrounding papillary thyroid carcinoma (PTC) nodules. Moreover, the algorithm also reveals that the collagen arrangement of the capsules surrounding benign nodules significantly differs from both the thyroid control and PTC nodule capsules. Combining SHG imaging with the HT results thus in an automatic and objective tool to discriminate between the pathological modifications that affect the capsules of thyroid nodules across the progressions of PTC, with potential to be used in clinical settings to complement current state-of-the-art diagnostic methods.

© 2020 Optical Society of America

https://doi.org/10.1364/AO.393721

1. INTRODUCTION
Thyroid cancer represents the most common endocrine malignancy. If treatment is well timed, patients affected by the most frequent versions of this pathology can be efficiently cured, thus timely diagnosis is very important [1]. Tumors (or nodules) originated from thyroid follicular cells are mainly divided into papillary (PTC), follicular (FTC), and anaplastic (ATC) thyroid carcinoma [2]. PTC accounts for 85%–90% of all thyroid cancer cases. Unlike ATC (an inoperable tumor with a mean survival time of ∼6 months after diagnosis), PTC and FTC are less aggressive, and the survival rate is noticeably higher [3].

The analysis of morphological features of hematoxylin and eosin (H&E) stained samples under bright-field microscopy is currently regarded as the gold standard to distinguish normal and tumor thyroid tissue [4]. However, it is well known that this is a complex and highly time-consuming process, and the final diagnosis is based on a subjective opinion that strongly depends on the pathologist’s experience [5].

In recent years, different optical techniques dealing with the characterization and diagnosis of thyroid tissues have been introduced. These include infrared spectroscopy [6–8], Raman spectroscopy [9,10], and multiphoton microscopy, including both two-photon excitation fluorescence microscopy (TPEF) and second harmonic generation (SHG) microscopy [11]. Whereas TPEF signals allowed visualizing thyroid gland follicles, the collagen content and associated structures were only observable in SHG images. Changes in these microstructures were clearly identified when comparing normal, nodular goiter,
and PTC tissue. Texture analyses were used to differentiate capsules surrounding the nodules in thyroid follicular adenoma (FA) and PTC in SHG images [12]. In addition, in an experiment dealing with polarization-resolved SHG, the degree of linear polarization was found to decrease in cancerous thyroid tissue (compared to non-tumoral tissue), suggesting a collagen disorder increase [13]. More recently, this technique has provided additional information about the capsule surrounding thyroid nodules [14].

The combination of Fourier transform infrared spectroscopy and a canonical discriminant analysis method was reported to accurately discriminate benign from malignant thyroid nodules [6]. This tool was later used not only to compare normal and FTC tissues [15] but also to distinguish thyroid carcinoma subtypes (in particular FTC from follicular variant papillary carcinoma) [16] and to differentiate FA (i.e., benign) from widely invasive FTC (WI-FTC) tissues [8]. Raman spectroscopy was also demonstrated to provide an accurate tissue classification to discriminate between normal thyroid parenchyma and follicular patterned nodules and between adenoma and carcinoma [10,17].

Collagen structure and content are essential in the analysis of tumor progression and development [18]. Because collagen is known to yield strong optical signals resulting from the generation of second-order harmonics [19], SHG microscopy represents a powerful tool for the diagnosis of different types of cancer, such as thyroid, breast, ovarian, and skin, among others (see [18,20] as general references).

The fibrous capsule surrounding the thyroid gland is made of collagen fibers. This capsule has been reported to suffer invasion in the presence of PTC [21]. In addition, a FA is a benign encapsulated tumor of the thyroid gland [22]. Because both PTC and FA might modify the capsule’s collagen arrangement in a non-controlled manner, the purpose of this study has been to establish objective methods to differentiate malignant from benign thyroid nodules through the analysis of the capsular collagen distribution.

2. METHODS AND MATERIAL

A. SHG Microscope

The experimental setup (Fig. 1) used in this work combines a mode-locked Ti:sapphire laser (Tsunami, Spectra-Physics, Santa Clara, California, USA) and a modified laser scanning microscope (Leica TCS-SP, Wetzlar, Germany) equipped with an infrared port.

The excitation wavelength was set to 780 nm, and the laser beam was focused on the sample through a 40× objective (numerical aperture [NA] = 0.75). The average power at the sample’s plane was 15 mW, and the polarization of the light beam was circular. The nonlinear signal from the sample was collected in the forward direction (i.e., in transmission mode) using a 0.9 NA condenser lens. Before reaching the photomultiplier tube (PMT) used as detection unit, this signal passed two spectral filters, a short-pass filter (FF01-750/SP-25, Semrock, Rochester, New York, USA) to reject the fundamental beam and a narrowband filter (FB390-10, Thorlabs, Newton, New Jersey, USA) to isolate the SHG signal emitted by the tissues under study. For all samples, the size of the SHG images acquired and analyzed in this work was 250 × 250 µm². Three individual frames were averaged to reduce noise.

B. Samples

This study was conducted at the University Politehnica of Bucharest. Signed informed consent was obtained from each patient. The samples were gently provided by the Pathology Department of the Central University Emergency Military Hospital of Bucharest, Romania. All samples were treated following the tenets of the Declaration of Helsinki.

A set of thyroid fragments collected from different patients were involved in the present experiment. Tissues were obtained by means of surgical excision. These were fixed with formalin and embedded in paraffin. For each sample, a pair of thin tissue sections (4 µm thick) were cut with a microtome and mounted on glass slides. One tissue section was stained with H&E and another was left unstained. An expert pathologist visualized the H&E stained sections with a bright-field microscope to identify the capsules surrounding both the thyroid gland (that we regard as control) and the nodules. Some of these nodules were diagnosed as benign FA, while others corresponded to PTC. Whereas these H&E samples served as reference, the unstained tissue pairs were used for SHG imaging. For the purpose of this work, a total of 24 capsule areas were imaged: 10 from the normal thyroid, 4 from FA, and 10 from PTC.

C. Image Analysis: SHG Signal and Density of Fibers

Once the SHG images were acquired, image processing was carried out to compute two parameters: the SHG intensity across the images and the density of the collagen fibers. Three sub-areas with a size of ~25 × 25 µm² were randomly chosen across each sample (the operator made sure all these sub-areas included collagen). The total SHG signal within these sub-areas was calculated.

The collagen fiber density was computed as follows. For each sub-area, the grayscale image was thresholded and converted into a binary image, where white pixels represent the collagen fibers. The open-source ImageJ software was used to perform this image analysis. A cross section across the fiber bundles

Fig. 1. Schematic diagram of the SHG microscope in transmission mode. PMT, photomultiplier tube. See text for further details.
D. Quantitative Analysis of the Collagen Distribution: The Hough Transform

The spatial organization of the collagen fibers in the studied samples was analyzed by using the Hough transform (HT) [23]. This is a well-established image-processing method for quantitative analysis able to detect aligned segments (i.e., straight lines) within an image, although it can also be used to localize circles and ellipses [24,25].

Because of its robustness to noise, gap tolerance, and the overall superb performance, it has rapidly been implemented into microscopy imaging applications [26,27]. In particular, this algorithm has been used to automatically detect nanoparticles in both atomic force microscopy and transmission electron images [28–30]. The HT was also employed in various applications dealing with biomedicine topics. For example it has been proposed for quantitative assessment of collagen type I gels [31], segmentation of blood cells [32], and, more recently, to detect breast cancer from mammogram images through the detection of circular features [33]. To the best of our knowledge, it has never been used before to detect collagen remodeling in SHG images of thyroid capsules.

In brief, the method is based on the pixel-by-pixel identification of straight lines described in polar coordinates. When the evidence of a straight line is found, the corresponding polar coordinates are filed in a two-dimensional array called an accumulator matrix. Each time a new straight line is detected, the accumulator’s bin increments the corresponding value. The local maximum values (peaks) in this accumulator space provide the preferential orientations ($\theta_i$) found across the image. The structural dispersion (SD) of the collagen fibers is defined as the standard deviation of the $\theta_i$.

Figure 2 presents an illustrative example of the use of the HT in an artificial image. Colored arrows within the accumulator matrix indicate the peaks corresponding to the bands marked in the original image. When the bands (collagen fibers in the case of a real image) are quasi-parallel, the peaks in the accumulator are almost aligned along the vertical of a particular $\theta_i$. Because the image presents features quasi-aligned along a similar direction, the SD takes a low value ($7^\circ$). The more apart from a particular vertical location (i.e., $\theta_i$ value), the lower the degree of organization. Preliminary tests (not included here) on calibrated samples allowed establishing the intervals assigned to the different organization groups. When $\text{SD} < 20^\circ$, a quasi-organized structure is present. If $\text{SD} > 40^\circ$ the sample is non-organized. Partially organized distributions correspond to values in between. A custom-built MATLAB software was used for image processing and HT calculations.

3. RESULTS

A. SHG Images of the Thyroid Capsule

Figure 3 shows representative SHG images of collagen-based capsules of the thyroid gland and FA (benign) and PTC (malignant) nodules from different patients. Bundles of collagen fibers can be clearly observed in the three types of investigated capsules.

A direct observation reveals qualitative differences in the collagen organization between PTC samples and both FA and control specimens. In particular, differences in the collagen distribution of the capsule surrounding the healthy nodule [Fig. 3(a)] and that surrounding a PTC nodule [Fig. 3(c)] are readily visible. The former presents a non-organized pattern with local undulations. For the latter, adjacent collagen fibers are organized along a preferential direction and show mainly similar orientations. From a qualitative point of view, the morphological structure of the benign nodule capsule is similar to that found in the thyroid capsule. However, for a better understanding and description of the three experimental conditions involved, a quantitative analysis must be carried out.

B. SHG Intensity

Because the SHG signal efficiency depends on both the density and the arrangement of the collagen fibers, it is interesting to compare the SHG intensity values for the different specimens involved in the present experiment. Results for all samples are shown in Fig. 4. Each symbol represents the mean SHG intensity for the three sub-regions of each sample (see Section 2.C for details).

The observed differences in SHG intensity between the control capsules and those surrounding malignant nodules were statistically significant ($t$-student, $p < 0.0001$). In addition, differences were not statistically significant when comparing control and benign samples ($p = 0.37$). This fact can be better observed in Fig. 5, where the averaged SHG intensity values for each group are depicted.

![Fig. 2. Representative example of the use of the HT in an artificial image presenting a structural organization along a preferential direction. As a general idea, the higher the SD, the lower the degree of organization of the collagen distribution.](image_url)
**C. Collagen Fiber Density**

Collagen fiber density was computed as indicated previously. For the three sets of tissues, Fig. 6 presents the mean values together with the standard deviations. It can be observed that the averaged density for control tissues is around 40% higher than that of PTC ones. Moreover, the statistical analysis shows that this difference is significant ($p = 0.005$). Significant differences were also revealed when comparing PTC and FA specimens ($p = 0.001$). No statistical differences were found between control and FA samples ($p = 0.052$).

**D. SD through the HT**

The HT was used to determine the SD of the samples under analysis. As stated in Section 2.D, the goal of the developed algorithm was to find straight lines (i.e., collagen fibers) across the SHG image. Once each line is found, its position and orientation are filled in the so-called accumulator matrix.

From a mathematical point of view, those lines can be treated as vectors. Consequently, two parameters are required to be calibrated before the analysis of the samples: (1) the vector length and (2) the number of orientations (i.e., the maximum number of different orientations $\theta_i$ that we want to file and take into account to compute SD). Figure 7 depicts the SD values for two representative samples (control and PTC) as a function of both parameters. It can be observed that, the sample’s SD does not depend on the length of the vector used [Fig. 7(a)]. However, Fig. 7(b) shows that the SD presents some variations when the number of orientations is below 60, but when the number increases, the values of SD remain fairly constant.

Taking into account this calibration, the HT procedure was applied to all specimens using a vector length of 10 pixels and a number of orientations of 60. Figure 8 presents the corresponding individual SD values. The averaged data for the three sets of tissues are presented in Fig. 9. The statistical analysis shows significant differences among the three groups (see labels in the figure).
Stained capsules of thyroid neoplasms were analyzed early with polarized light by Koren and co-authors [34]. Although no significant differences between minimally and widely invasive carcinomas were found, these were consistently different compared to the capsules surrounding adenomas.

In our work, we have investigated the ability of SHG imaging to objectively distinguish between the capsules surrounding the healthy thyroid gland and nodules (FA and PTC). Given that our approach deals with label-free (unstained) samples, its potential utility is threefold: it can address (i) fixed, (ii) freshly excised, and (iii) in vivo tissues. The latter might be implemented into devices for endomicroscopy. In cancer-related clinical environments, an objective classification and characterization of these changes is crucial for certain diagnoses. The performance of the mathematical algorithm known as HT (as well as other parameters) has been successfully tested here. This allowed detecting changes produced in the collagen structure of the non-stained capsules surrounding both PTC and FA nodules, which were compared with the collagen structure in capsules surrounding normal thyroid glands.

It is well known that SHG microscopy is a powerful tool to assess and visualize the morphology of collagen-based tissues. Pathologies and physical damages modify the collagen architecture and distribution in tissues. Thus, the presence of structural abnormalities or alterations of the intrinsic organization might compromise biological functions. In particular, those related to the early stages of certain diseases, or pathological conditions with not readily visible modifications, are very important in the context of diagnostics. The combination of useful imaging techniques (as the one used herein) with quantitative indices/parameters might simplify observations, reduce the time of analysis, and avoid erroneous (or controversial) assessment.

Although some mathematical algorithms have been reported to be very efficient in discriminating healthy from diseased tissues [6,12–16,35,36], the results (or alternatively, the diagnosis) are usually based on global information from the features of the entire image. However, local sub-areas might provide relevant information, and the final score could be underestimated and an erroneous diagnosis might be reached. Moreover, although widely applied, some of those algorithms are (partially) operator dependent, often combine analytical and manual schemes (image filtering, contrast and smoothing adjusts), and might fail when complex structural patterns (e.g., undulations, interweaving, etc.) are present.

The main quantitative method here proposed is the HT. This is based on the local analysis of SHG images to identify collagen fibers within the thyroid capsule. No previous literature on this topic has been found by these authors. This mathematical tool eliminates the drawbacks of other methods because it is a totally automated procedure and operates in a local-based manner, storing/accumulating the corresponding information. In addition, apart from the numerical values that can be obtained, the visualization of the accumulator matrix provides direct and intuitive information on the spatial arrangement of the tissue under study.

In the present study, the capsules surrounding the healthy thyroid were found to exhibit a non-regular distribution of their collagen fibers. In terms of SD, the values were always above 20°, indicating the absence of a preferential orientation. On the contrary, for all PTC nodule capsule samples, SD values were below 11°, which indicates that a fairly well-organized structure is present. However, in our opinion the most interesting finding of this study is the ability of this tool to differentiate between the capsules surrounding benign nodules from those surrounding cancerous nodules and the thyroid gland (see Fig. 9).

The collagen patterns found in the present work for the considered thyroid tissues are consistent with those found by Hristu et al. [12]. The present study consolidates thus the list of SHG intensity-based metrics identified in this previous study as being feasible for quantitatively differentiating between the capsules surrounding benign and malignant thyroid nodules.

The metrics used herein, namely the SHG intensity and the density of fibers, have also been used in studies addressing other types of cancers (see [37,38] as general references). For the samples involved in this study, these quantitative parameters were able to significantly differentiate the capsules surrounding the thyroid glands and FA thyroid nodules from those surrounding PTC ones. However, success in differentiating the thyroid capsules from those surrounding FA nodules was limited. The decrease in both SHG signal and collagen density of cancerous tissues is a direct result of the fiber reorganization [20]. This agrees well with previous measurements, and in the case of mammary carcinoma tumors, it can be used to stage progression levels [38,39].

Although the HT presents a series of advantages over alternative image analysis methods reported as solutions for the quantitative assessment of the collagen architecture in tissues, this tool should not be regarded as a competitor to those, and hence a thorough comparison of the proposed method over others lies outside the scope of this paper. Different complementary strategies could jointly play an important role in the problem of collagen organization assessment for diagnostics purposes. In the present case, the HT could potentially be exploited for this purpose in tandem with alternative classifiers based on other image parameters [12].

We believe that it is feasible to consider the HT as a valuable tool to complement other image features (and vice versa). In this regard, it is interesting to recall that the machine-learning community is placing significant efforts for developing methods that can jointly exploit the outputs of distinct classifiers [40].

**4. DISCUSSION AND CONCLUSIONS**

The performance of the mathematical algorithm known as HT is based on the local analysis of SHG images to identify collagen fibers within the thyroid capsule. Although some mathematical algorithms have been reported to be very efficient in discriminating healthy from diseased tissues [6,12–16,35,36], the results (or alternatively, the diagnosis) are usually based on global information from the features of the entire image. However, local sub-areas might provide relevant information, and the final score could be underestimated and an erroneous diagnosis might be reached. Moreover, although widely applied, some of those algorithms are (partially) operator dependent, often combine analytical and manual schemes (image filtering, contrast and smoothing adjusts), and
Such fusion methods are successful in generating higher classification accuracy compared to any of the constituent classifiers. Furthermore, we find important to highlight that exploiting various ways of representing images is known to enable powerful data augmentation strategies [41]. In this context, image-processing operations that simulate specific acquisition conditions or sample features bring an additional level of efficiency in the case of classifying multiphoton microscopy images for disease diagnostics [42]. Considering this, the image representations yielded by the HT method are also likely to play an important role for efficiently augmenting SHG image datasets, a subject that we plan to investigate in the near future.

In conclusion, to visualize and understand the changes suffered by the thyroid collagen-based nodule capsules under pathological conditions (neoplastic), SHG images have been acquired and analyzed. Whereas the collagen matrix in the normal thyroid capsule revealed a wavy structure, the pattern of capsules surrounding PTC nodules displayed a fine and relative linear distribution with an absence of local orientations. On the contrary, the collagen of capsules surrounding benign nodules revealed an undulated architecture with local orientations. Furthermore, the collagen distribution was assessed by employing a series of metrics to extract quantitative information from the collected SHG images. Among these metrics, the HT was found to be able to characterize collagen realignment in PTC specimens and establish quantitative and significant differences among the three types of tissues explored. Unlike the HT, the SHG total intensity and the density of fibers were only able to objectively differentiate the capsule of PTC nodules from that surrounding both FA nodules and the thyroid gland.

The usefulness of the HT method might have clinical applications given its potential for enabling fast and accurate objective diagnostic procedures. These could either be used for automated screening or could facilitate the work of pathologists in computer-aided assays by forecasting an overall tissue state for a subject that we plan to investigate in the near future.

Funding. Fundación Séneca (19897/GERM/15); UEFISCDI (PN-III-P1-1.1.-TE-2016-1756 (NANOSHG)).

Acknowledgment. The authors thank D. Párraga, MSc, for her useful help during the early steps of this work.

Disclosures. The authors report no conflicts of interest and have no proprietary interest in any of the materials mentioned in this paper.

REFERENCES


