

LETTER

Superpixel Raman spectroscopy for rapid skin cancer margin assessment

Xu Feng¹  | Matthew C. Fox² | Jason S. Reichenberg² |
Fabiana C.P.S. Lopes² | Katherine R. Sebastian² | Andrew K. Dunn¹ |
Mia K. Markey^{1,3} | James W. Tunnell^{1*}

¹Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas

²Department of Internal Medicine, Dell Medical School, The University of Texas at Austin, Austin, Texas

³Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas

*Correspondence

James W. Tunnell, Department of Biomedical Engineering, The University of Texas at Austin, 107 W. Dean Keeton Street, Austin, TX 78712.
Email: jtunnell@mail.utexas.edu

Funding information

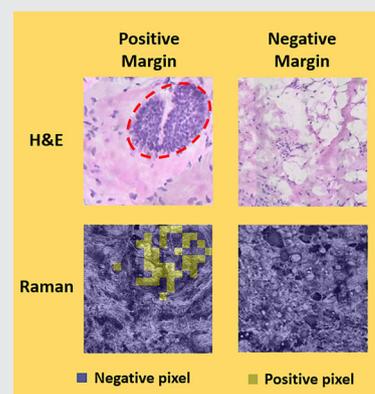
Cancer Prevention and Research Institute of Texas, Grant/Award Number: RP130702

Abstract

Spontaneous Raman micro-spectroscopy has been demonstrated great potential in delineating tumor margins; however, it is limited by slow acquisition speed. We describe a superpixel acquisition approach that can expedite acquisition between $\sim \times 100$ and $\times 10\,000$, as compared to point-by-point scanning by trading off spatial resolution. We present the first demonstration of superpixel acquisition on rapid discrimination of basal cell carcinoma tumor from eight patients undergoing Mohs micrographic surgery. Results have been demonstrated high discriminant power for tumor vs normal skin based on the biochemical differences between nucleus, collagen, keratin and ceramide. We further perform raster-scanned superpixel Raman imaging on positive and negative margin samples. Our results indicate superpixel acquisition can facilitate the use of Raman microspectroscopy as a rapid and specific tool for tumor margin assessment.

KEYWORDS

basal cell carcinoma, Raman imaging, Raman spectroscopy, rapid acquisition, skin cancer, tumor margin



1 | INTRODUCTION

Raman spectroscopy is a sensitive in detecting molecular differences between tumor and healthy tissue without requiring tissue sectioning or staining. Raman micro-spectroscopy offers the potential for accurate tumor margin detection for many cancer types, including skin [1], oral cavity [2], breast [3] and brain [4]. The detection is usually implemented in an imaging mode with a spatial resolution

of $\sim 1\ \mu\text{m}$. However, point-by-point scanning is extremely time-consuming: a single point typically requires ~ 1 second in tissue; thus, to scan the complete surface of a multi-centimeter tissue would take hours or even days.

A few approaches have been proposed to speed up Raman acquisition. Stimulated Raman techniques such as CARS [5] and SRS [6] produce images akin to hematoxylin and eosin (H&E) stained histology slides without staining, but under most implementations trade imaging speed for

spectral content. Several studies have accelerated spontaneous Raman measurements using sparse sampling techniques (sampling at 10-20 μm increments [1, 7–9]) collecting a full spectrum at each pixel at the expense of limited coverage of the tissue surface area (2% or less). Other methods that have been developed include line-scanning confocal [10], 2D multifocal arrays [11] and Wiener estimation [12], with speed up factors reported between $\times 10$ and $\times 100$). One attractive approach is superpixel acquisition whereby a spectrum is averaged over a larger pixel (aka a superpixel of $\sim 25 \times 25 \mu\text{m}^2$ - $100 \times 100 \mu\text{m}^2$) on the sample surface while integrating only once on the detector. Thus, the acceleration mainly comes from the reduction of total detector reading time. This approach was recently used for small superpixel sizes from $1 \times 1 \mu\text{m}^2$ to $30 \times 30 \mu\text{m}^2$ for human bone characterization [13].

In this letter, we aim at the proof of concept that superpixel acquisition performs statistically the same as our previous point-by-point scanning approach in classifying basal cell carcinoma (BCC) from normal skin structures while substantially speeding up acquisition. The speed up factor is proportional to the area of the superpixel. For a superpixel size between $25 \times 25 \mu\text{m}^2$ and $100 \times 100 \mu\text{m}^2$, the maximum speed up factor varies between $\times 625$ and $\times 10\,000$, as compared to point-by-point scanning with $1 \mu\text{m}^2$ point when sampling the complete surface area. We demonstrate the equivalence of superpixel acquisition in tissue simulating phantoms and human skin cancer specimens. Furthermore, we demonstrate raster-scanned superpixel Raman classification images of both positive and negative margin samples, and emphasize the need to determine the optimum superpixel size for the application of skin cancer diagnosis.

2 | EXPERIMENTAL

2.1 | System description

The system is based on a custom-built confocal Raman microspectroscopy integrated with a reflectance confocal microscope as described previously [14]. Superpixel acquisition still utilizes a confocal setup. The $1 \mu\text{m}$ laser spot is rapidly scanned across a user-defined region by using galvanometer scanners, while the CCD camera collects one average confocal Raman spectrum. As shown in Figure 1A, the speed up factor is proportional to the superpixel area and results in a trade-off in resolution. In this way, one can trade-off spatial resolution for speed and acquire an image that covers the complete surface. For this study, we mainly focus on a superpixel size of $100 \times 100 \mu\text{m}^2$, as that closely matches the resolution of a dermatopathologist analyzing frozen section histopathology.

2.2 | Comparison between point-by-point scanning and superpixel acquisition

The spectral difference and signal-to-noise ratio (SNR) were compared between point-by-point scanning and superpixel acquisition using phantom experiments [13]. All raw spectra underwent wavenumber calibration, dark noise subtraction, cosmic ray removal, spectral response calibration, smoothing and fluorescence background removal.

For the spectral difference test, mixed phantoms of synthetic collagen type I and elastin were measured using either point-by-point scanning (2 μm step size, 1 second per step) or superpixel acquisition (1 second per spectrum, repeated 10 times).

For the SNR test, SNR was calculated as follows [16]:

$$\text{SNR}(\nu) = \frac{I_{\text{mean}}(\nu)}{\text{SD}(\nu)} \quad (1)$$

where $I_{\text{mean}}(\nu)$ is the mean intensity and $\text{SD}(\nu)$ is the SD at a given wavenumber (ν). Because $\text{SD}(\nu)$ is not only influenced by the signal variations from the experimental setup but also the inhomogeneity of the sample [13], we used a pure collagen type I phantom. We measured the SNR of point-by-point scanning by taking 20 spectra only at the central position, and the SNR of superpixel acquisition by taking 20 spectra across the entire region. Integration time is 1 second per spectrum.

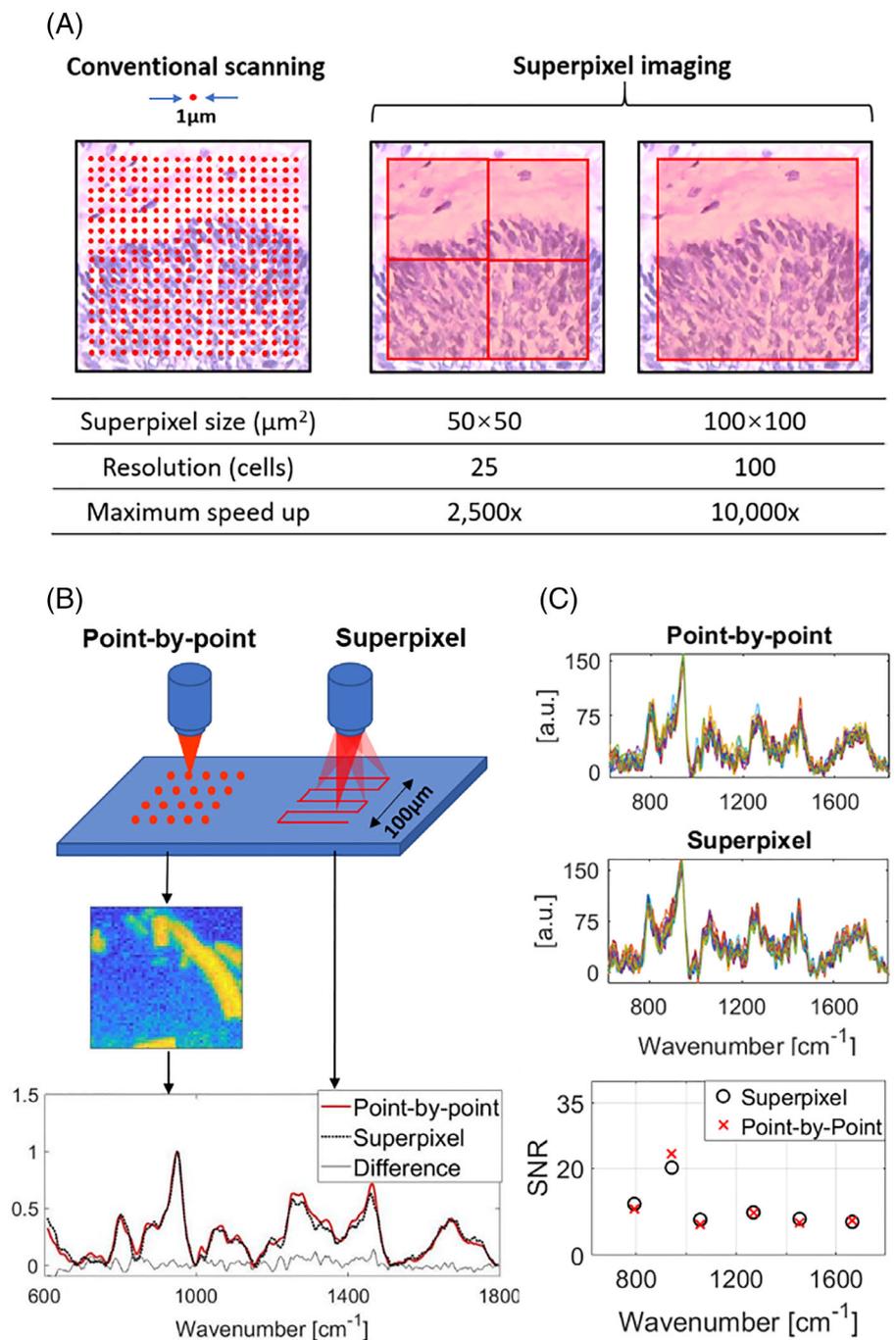
2.3 | Sample preparation

Ten skin tissue samples were obtained from eight patients undergoing Mohs micrographic surgery. Seven samples were found to have both BCC and normal tissue, and three contained only normal tissue. For each sample, a skin section of 20 μm thickness was cut at -22°C with a microtome, and then transferred onto an MgF_2 substrate for the experiment. The serial section went through H&E staining for histopathological diagnosis by a board-certified dermatologist. This study was approved by the Institutional Review Board at The University of Texas at Austin and the Seton Healthcare Family.

2.4 | Superpixel acquisition experiment and model establishment

Multiple locations were sampled on each skin section as shown in Figure 2A. By visual comparison of reflectance, bright-field and histopathology images, each superpixel was annotated as either BCC or one of the seven primary normal skin structures, including epidermis, dermis,

FIGURE 1 Comparison between superpixel acquisition and point-by-point scanning. A, Superpixel acquisition samples the complete area while point-by-point scanning samples only a fraction of the area (2%). A trade-off exists between superpixel size/resolution and maximum speed-up factor. B, Spectral difference test. Red line: mean spectrum from point-by-point scanning. Black line: mean spectrum from superpixel acquisition. Spectra were normalized to the maximum peak. Bottom: difference spectrum. The schematic of the point-by-point and superpixel setup is also shown. C, Signal-to-noise ratio (SNR) test. Top: 20 repeated spectra taken at the central position. Middle: 20 repeated spectra taken from superpixel acquisition. Bottom: SNR plot. The main bands of interest are assigned to collagen at 792, 942, 1056, 1269, 1454 and 1665 cm^{-1} [15]



inflamed dermis, hair follicle, hair shaft, sebaceous gland and fat. The corresponding average spectra were also saved in the database.

A previously established biophysical inverse model [14] was fitted to the average tissue spectra by nonnegative least squares fitting. The model has been recently applied to in vivo skin cancer diagnosis [17], and ex vivo skin tumor margin assessment [15]. The model describes the tissue spectra as a linear combination of the basis spectra of collagen, elastin, triolein, nucleus, keratin, ceramide and water. The fit coefficients were then used as the input variables of a logistic regression classifier to

discriminate BCC from normal skin structures. Receiver operator characteristic (ROC) curve was built on leave-one-out cross validation. The optimal combination of input model components was determined by maximizing the area under the ROC curve.

2.5 | Raster-scanned superpixel imaging

Two samples from two new patients were used to test the raster-scanned superpixel imaging. One sample contains both BCC and normal dermis (positive margin), while

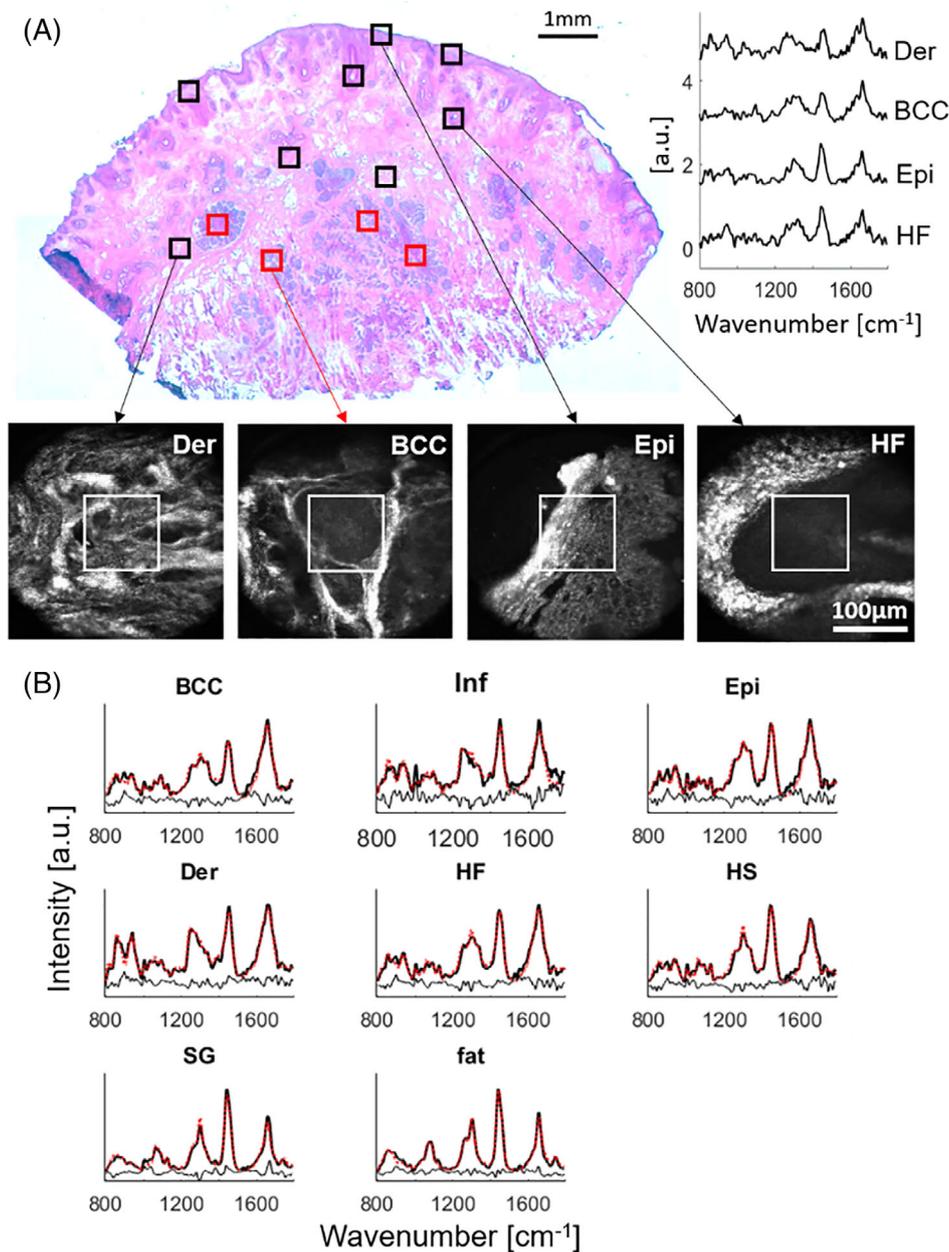


FIGURE 2 A, Superpixel acquisition experiment conducted on multiple locations on a basal cell carcinoma (BCC) skin section. Red squares: BCCs. Black squares: normal structures. Bottom: reflectance confocal microscope images of dermis, BCC, epidermis and hair follicle. The white square labels the location of superpixels. Top right: the corresponding average spectra. B, The biophysical model was fitted to the mean tissue spectra of BCC, inflamed dermis (Inf), epidermis (epi), dermis (Der), hair follicle (HF), hair shaft (HS), sebaceous gland (SG) and fat. Black line: mean tissue spectra. Red line: model fits. Residuals are plotted on the same scale

the other sample contains only normal tissue, mainly dermis and fat (negative margin). Superpixel imaging was performed by translating the sample in two dimensions using a linear motorized stage. Two superpixel sizes were compared: $100 \times 100 \mu\text{m}^2$ and $50 \times 50 \mu\text{m}^2$. The classification model established in Section 2.4 was applied to each superpixel, labeling it as positive or negative. A binary tumor heat map was generated by prioritizing high specificity.

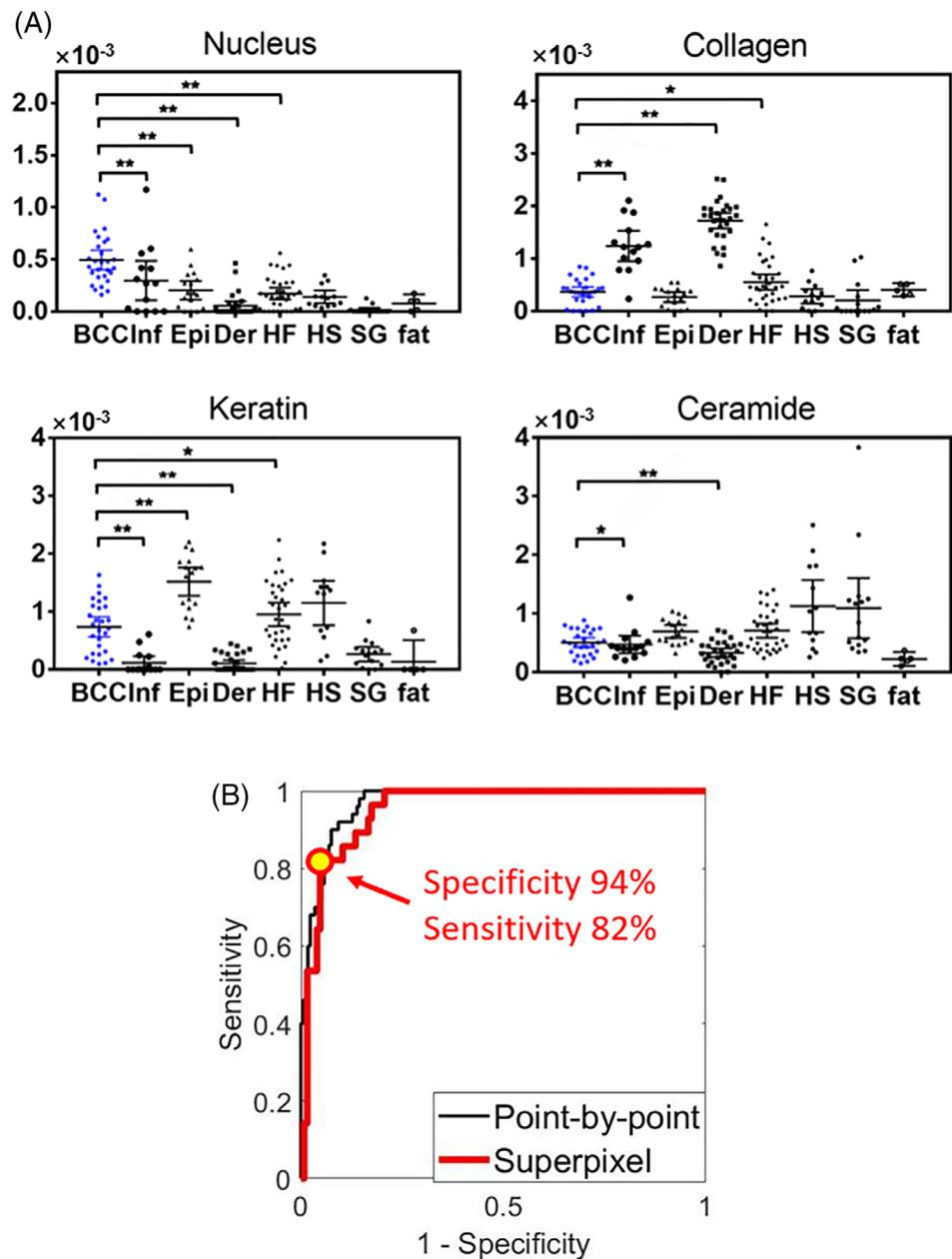
3 | RESULTS AND DISCUSSION

We observed that point-by-point scanning and superpixel acquisition exhibit visually similar spectra (Figure 1B).

Signal intensity is similar for both the point scanning setup and superpixel setup, and the SNRs are also similar (Figure 1C). The major reason that SNRs are similar is because when the shot noise is dominant, SNR should be proportional to the square root of the signal intensity [16]. Although the signal intensity of individual point fluctuates under the superpixel setup, the overall intensity retains the same after integrating over the entire region.

Figure 2A shows a typical example of a superpixel acquisition experiment. In total, we collected 154 annotated tissue spectra, including 28 spectra from BCC and 126 spectra from normal skin structures (epidermis [N = 17], dermis [N = 29], inflamed dermis [N = 14], hair follicle [N = 33], hair shaft [N = 13], sebaceous gland

FIGURE 3 A, Scatter plots of the fit coefficients of nucleus, collagen, keratin and ceramide for basal cell carcinoma (BCC), inflamed dermis (Inf), epidermis (epi), dermis (Der), hair follicle (HF), hair shaft (HS), sebaceous gland (SG) and fat. Each point represents a spectrum data. The fit coefficients of BCC and individual normal structures were statistically compared using *t* tests. Linear fixed-effects models were used to account for the nonindependencies of multiple measures per lesion [18]. $**P \leq .01$, $*P \leq .05$. B, Comparison of receiver operator characteristic curves between superpixel acquisition (red thick line) and point-by-point scanning (black thin line)



[$N = 15$] and fat [$N = 5$]). The model fitting results are shown in Figure 2B.

Figure 3A shows the scatter plots of the fit coefficients of primary model components. We found that nucleus is the most important component to discriminate BCC from normal tissues. BCC has a significantly larger amount of nucleus compared to normal structures. BCC also has a significantly higher amount of keratin compared to dermis/inflamed dermis, and a lower amount of keratin compared to epidermis and hair follicle. In addition, BCC has significantly lower collagen and higher ceramide than dermis/inflamed dermis.

The optimum classification result was achieved by combining the fit coefficients of nucleus, collagen, keratin and ceramide as the input parameters, leading to an

area under the ROC curve of 0.95, as shown in Figure 3B. By selecting a balanced tumor score threshold, the specificity and sensitivity reached 94% and 82%, respectively. This ROC was then statistically compared to our recent study in 30 patients [15] using R software [19] to determine if the superpixel acquisition provides equivalent potential for BCC classification compared with the point-by-point scanning.

Our results show that the areas under the two ROC curves are not statistically different (P value = .34). Because the P value does not convey the statistical power of the comparison, we also estimated the sample size required to demonstrate a difference in the area under the ROC curves at the level observed in our study. Since the required sample size is very large ($N > 1500$), we

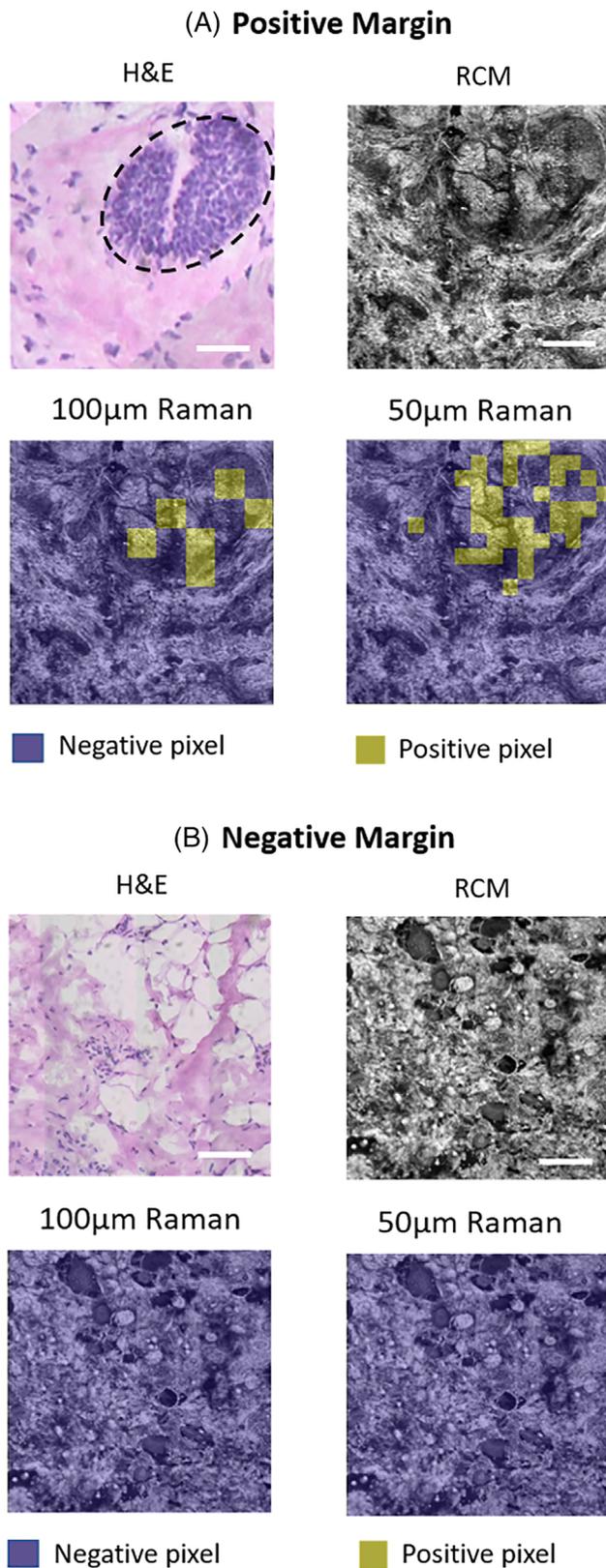


FIGURE 4 Raster-scanned superpixel Raman classification images of, A, positive and, B, negative margin samples. H&E, reflectance confocal microscope and Raman classification images of the same region are compared. The positive region is marked on the hematoxylin and eosin (H&E) image. Two superpixel sizes were compared: $100 \times 100 \mu\text{m}^2$ and $50 \times 50 \mu\text{m}^2$. Scale bar: $200 \mu\text{m}$

conclude that the effect size is small, that is, the observed difference in the area under the ROC curves is not meaningful. Therefore, the discriminant capabilities of superpixel acquisition and point-by-point scanning were not statistically distinguishable for BCC classification.

Figure 4 demonstrates an example of raster-scanned superpixel imaging of partial tissue samples. For these two samples, the negative and positive margins are correctly classified. In Figure 4A, BCCs are identified in both the $100 \times 100 \mu\text{m}^2$ and $50 \times 50 \mu\text{m}^2$ superpixel Raman classification images. Although prioritizing high specificity would lead to discontinuous positive pixels, it guarantees that only true positive pixels are classified as positive (high positive predictive value). In Figure 4B, the whole image is classified as normal tissue.

4 | CONCLUSION

In this study, we demonstrate a significant speed advantage of superpixel acquisition Raman spectroscopy for BCC tumor margin assessment. When compared to point-by-point scanning, the speed up factor is dictated by the ratio of the superpixel area to the laser spot size. Our superpixel approach provides an alternative speed up method to sparse sampling approach in applications where such high resolutions are not needed, enabling more complete sampling of the tissue surface area. It is worth mentioning that sparse sampling techniques [1, 7–9] have also demonstrated capability in detecting BCC margin. The speed up factor would be between $\times 25$ and $\times 100$ compared with sparse sampling using a step size between 10 and $20 \mu\text{m}$. Further experiments are needed to compare between the performance of superpixel acquisition and point-by-point scanning with a matching step size (eg, $100 \mu\text{m}$) for the same tissue surface area.

The total estimated scan time for tissue samples with areas of $1 \times 1 \text{ cm}^2$ would be around 2.7 hours ($100 \times 100 \mu\text{m}^2$ superpixel, 1 second per step). Future developments could combine this approach with other speed up approaches (eg, line-scanning confocal [10], multifocal [11] and Wiener estimation [12]) aimed to further reduce the acquisition time for intraoperative use (eg, < 1 hour). When compared to gold standard diagnostics such as histopathology, the sensitivity of 82% is still not acceptable for cancer detection. The goal of this study was demonstrating equivalence between the superpixel approach and traditional point scanning approach. It is our hope that combining this technique with other imaging modalities (such as autofluorescence imaging [8, 9] and reflectance confocal microscopy imaging [20]) the sensitivity can be improved upon without sacrificing speed.

Our next step involves fully characterizing the trade-offs of the superpixel approach by expanding to imaging of samples and more diversity of samples. We are currently collecting larger superpixel data sets containing diverse tissue structures (eg, hair follicle, inflamed dermis and epidermis), and including more tissue types per image to help refine our diagnostic algorithm and improve sensitivity further. We will also compare between different superpixel sizes to determine the optimum size for the application of skin cancer diagnosis that is larger than a single cell but smaller than the point when mixed microanatomies begin to obscure the identification of tumor cells.

ACKNOWLEDGMENTS

This work was funded by Cancer Prevention & Research Institute of Texas (CPRIT) (RP130702). We thank the Seton Healthcare Family and the Austin Dermatological Surgical Center for collaboration. We thank Sandra Esparza and Leandra Turner for their help in collecting fresh tissue specimens, Prof. Aaron Baker for providing the histology facility, and Greg Lyness and Esther Maier for H&E staining. We also thank UT Statistical consulting service and Erika Hale for providing statistical guidance.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

ORCID

Xu Feng  <https://orcid.org/0000-0002-9048-7144>

REFERENCES

- [1] A. Nijssen, T. C. B. Schut, F. Heule, P. J. Caspers, D. P. Hayes, M. H. A. Neumann, G. J. Puppels, *J. Invest. Dermatol.* **2002**, *119*, 64.
- [2] F. L. Cals, T. C. B. Schut, J. A. Hardillo, R. J. B. De Jong, S. Koljenović, G. J. Puppels, *Lab. Invest.* **2015**, *95*, 1186.
- [3] D. W. Shipp, E. A. Rakha, A. A. Koloydenko, R. D. Macmillan, I. O. Ellis, I. Notingher, *Breast Cancer Res.* **2018**, *20*, 69.
- [4] N. Bergner, A. Medyukhina, K. D. Geiger, M. Kirsch, G. Schackert, C. Krafft, J. Popp, *Anal. Bioanal. Chem.* **2013**, *405*, 8719.
- [5] J.-X. Cheng, X. S. Xie, *Coherent Raman Scattering Microscopy*, CRC Press, Boca Raton, FL, **2016**.
- [6] D. A. Orringer, B. Pandian, Y. S. Niknafs, T. C. Hollon, J. Boyle, S. Lewis, M. Garrard, S. L. Hervey-Jumper, H. J. Garton, C. O. Maher *Nat. Biomed. Eng.* **2017**, *1*, 0027.
- [7] K. Kong, C. J. Rowlands, S. Varma, W. Perkins, I. H. Leach, A. A. Koloydenko, H. C. Williams, I. Notingher, *Proc. Natl. Acad. Sci.* **2013**, *110*, 15189.
- [8] F. Sinjab, K. Kong, G. Gibson, S. Varma, H. Williams, M. Padgett, I. Notingher, *Biomed. Opt. Express* **2016**, *7*, 2993.
- [9] R. Boitor, K. Kong, D. Shipp, S. Varma, A. Koloydenko, K. Kulkarni, S. Elsheikh, T. B. Schut, P. Caspers, G. Puppels, *Biomed. Opt. Express* **2017**, *8*, 5749.
- [10] J. Qi, W.-C. Shih, *Appl. Opt.* **2014**, *53*, 2881.
- [11] L. Kong, M. Navas-Moreno, J. W. Chan, *Anal. Chem.* **2015**, *88*, 1281.
- [12] D. Wei, S. Chen, Y. H. Ong, C. Perlaki, Q. Liu, *Opt. Lett.* **2016**, *41*, 2783.
- [13] G. Falgayrac, B. Cortet, O. Devos, J. Barbillat, V. Pansini, A. Cotten, G. Pasquier, H. Migaud, G. Penel, *Anal. Chem.* **2012**, *84*, 9116.
- [14] X. Feng, A. J. Moy, H. T. M. Nguyen, J. Zhang, M. C. Fox, K. R. Sebastian, J. S. Reichenberg, M. K. Markey, J. W. Tunnell, *Biomed. Opt. Express* **2017**, *8*, 2835.
- [15] X. Feng, M. C. Fox, J. S. Reichenberg, F. C. Lopes, K. R. Sebastian, M. K. Markey, J. W. Tunnell, *Biomed. Opt. Express* **2019**, *10*, 104.
- [16] J. Desroches, M. Jermyn, K. Mok, C. Lemieux-Leduc, J. Mercier, K. St-Arnaud, K. Urmey, M.-C. Guiot, E. Marple, K. Petrecca, *Biomed. Opt. Express* **2015**, *6*, 2380.
- [17] X. Feng, A. J. Moy, H. T. Nguyen, Y. Zhang, J. Zhang, M. C. Fox, K. R. Sebastian, J. S. Reichenberg, M. K. Markey, J. W. Tunnell, *J. Biomed. Opt.* **2018**, *23*, 057002.
- [18] D. Bates, M. Mächler, B. Bolker, S. Walker arXiv preprint arXiv:1406.5823. **2014**.
- [19] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, *BMC Bioinform.* **2011**, *12*, 77.
- [20] E. S. Flores, M. Cordova, K. Kose, W. Phillips, A. Rossi, K. S. Nehal, M. Rajadhyaksha, *J. Biomed. Opt.* **2015**, *20*, 061103.

How to cite this article: Feng X, Fox MC, Reichenberg JS, et al. Superpixel Raman spectroscopy for rapid skin cancer margin assessment. *J. Biophotonics*. 2020;13:e201960109. <https://doi.org/10.1002/jbio.201960109>