

## Multiplex coherent anti-Stokes Raman scattering microspectroscopy of brain tissue with higher ranking data classification for biomedical imaging

Christoph Pohling  
Thomas Bocklitz  
Alex S. Duarte  
Cinzia Emmanuello  
Mariana S. Ishikawa  
Benjamin Dietzeck  
Tiago Buckup  
Ortrud Uckermann  
Gabriele Schackert  
Matthias Kirsch  
Michael Schmitt  
Jürgen Popp  
Marcus Motzkus

Christoph Pohling, Thomas Bocklitz, Alex S. Duarte, Cinzia Emmanuello, Mariana S. Ishikawa, Benjamin Dietzeck, Tiago Buckup, Ortrud Uckermann, Gabriele Schackert, Matthias Kirsch, Michael Schmitt, Jürgen Popp, Marcus Motzkus, "Multiplex coherent anti-Stokes Raman scattering microspectroscopy of brain tissue with higher ranking data classification for biomedical imaging," *J. Biomed. Opt.* **22**(6), 066005 (2017), doi: 10.1117/1.JBO.22.6.066005.

# Multiplex coherent anti-Stokes Raman scattering microspectroscopy of brain tissue with higher ranking data classification for biomedical imaging

Christoph Pohling,<sup>a</sup> Thomas Bocklitz,<sup>b,c</sup> Alex S. Duarte,<sup>a</sup> Cinzia Emmanuello,<sup>a</sup> Mariana S. Ishikawa,<sup>a</sup> Benjamin Dietzeck,<sup>b,c</sup> Tiago Buckup,<sup>a</sup> Ortrud Uckermann,<sup>d</sup> Gabriele Schackert,<sup>d</sup> Matthias Kirsch,<sup>d</sup> Michael Schmitt,<sup>b</sup> Jürgen Popp,<sup>b,c,e</sup> and Marcus Motzkus<sup>a,\*</sup>

<sup>a</sup>Physikalisch Chemisches Institut, Heidelberg, Germany

<sup>b</sup>Institut für Physikalische Chemie and Abbe School of Photonics, Jena, Germany

<sup>c</sup>Leibniz Institute of Photonic Technologies, Jena, Germany

<sup>d</sup>Carl Gustav Carus Universitätsklinikum an der Technischen Universität Dresden, Klinik und Poliklinik für Neurochirurgie, Dresden, Germany

<sup>e</sup>InfectoGnostics Forschungscampus, Jena, Germany

**Abstract.** Multiplex coherent anti-Stokes Raman scattering (MCARS) microscopy was carried out to map a solid tumor in mouse brain tissue. The border between normal and tumor tissue was visualized using support vector machines (SVM) as a higher ranking type of data classification. Training data were collected separately in both tissue types, and the image contrast is based on class affiliation of the single spectra. Color coding in the image generated by SVM is then related to pathological information instead of single spectral intensities or spectral differences within the data set. The results show good agreement with the H&E stained reference and spontaneous Raman microscopy, proving the validity of the MCARS approach in combination with SVM. © The Authors.

Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.22.6.066005](https://doi.org/10.1117/1.JBO.22.6.066005)]

Keywords: multiplex coherent anti-Stokes Raman scattering; support vector machines; Raman; brain; cancer; classification.

Paper 160867RR received Dec. 21, 2016; accepted for publication May 24, 2017; published online Jun. 14, 2017.

## 1 Introduction

The application of coherent anti-Stokes Raman scattering (CARS) to microscopy provides access to structural and molecular information on a vast range of chemical and biological systems.<sup>1–6</sup> In general, structural information from CARS microscopy is easily obtained associating a single spectral intensity with a certain color scheme. Although this method allows analysis of spatial information with submicron resolution, it cannot specifically address overlapping Raman resonances and, therefore, lacks molecular specificity.

One way of circumventing this issue can be efficiently implemented if the single chemical constituents are known in advance: in this case, the measured spectra are fitted as the sum of the Raman spectra of the isolated constituents.<sup>7</sup> However, in a situation where the Raman spectrum of each component is not known *a priori*, a different approach must be applied. In this regard, a multistep analysis<sup>8</sup> of CARS data can be implemented by combining Raman reconstruction algorithms,<sup>9</sup> principal component analysis (PCA), and decomposition of the measured data using pure spectra. This approach can reveal quantitative information without any previous information.<sup>8</sup> This method is superior to an univariate approach since additional molecular features beyond single spectral intensity are analyzed, visualized, and matched to standard histological assessment of the brain tissues. Nevertheless, a PCA-based type of data analysis is limited to spectral differences within the individual data sets. In other words, the structural and molecular

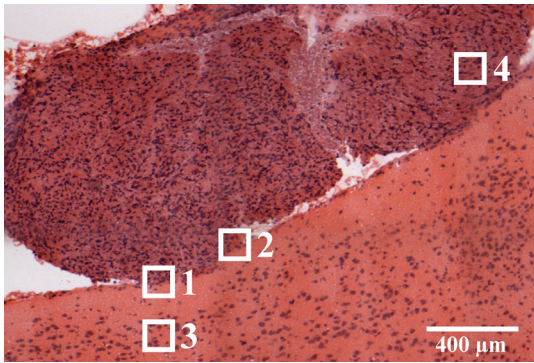
information obtained with a PCA-based analysis cannot be easily generalized and extended to other samples.

To transfer knowledge of characterized data sets to unknown spectra, higher ranking classification methods such as soft-independent modeling of class analogies or support vector machines (SVM) are required. In this case, new spectra are classified in terms of distances to groups of classified data within a multidimensional vector space.<sup>10,11</sup> As demonstrated for Raman and infrared microscopy, the visualization of class membership via color coding of spatially resolved spectral information then provides image contrast and pathological information as well.<sup>12,13</sup> The aim of this work is to illustrate the necessity of higher ranking data analysis also for multiplex CARS (MCARS) imaging by extracting pathological diagnosis on experimental tumors of mouse brain tissue. In this proof-of-principle-study, we worked on the differentiation of healthy brain tissue and one type of tumor (Fig. 1). In other words, biochemical differences between different types of brain tissue, e.g., originating from different lipid to protein ratios,<sup>14</sup> resulted in differences of the corresponding vibrational spectra that were used for image processing. As only two classes are present, we used a binary classifier; out of four method combinations (SVM alone, PCA–SVM, PCA–linear discriminant analysis, and feature forward selection–SVM), PCA–SVM is considered the best choice.

## 2 Experimental

A standard cryosection (10  $\mu\text{m}$ ) of fixed mouse brain tissue, containing a solid tumor induced by intracerebral stereotactic implantation of murine mK1735 melanoma cells as described previously,<sup>15</sup> was prepared. Tissue was prepared either as (1) an H&E-stained reference sample or (2) an air-dried sample used for the linear and nonlinear Raman experiments.

\*Address all correspondence to: Marcus Motzkus, E-mail: [marcus.motzkus@pci.uni-heidelberg.de](mailto:marcus.motzkus@pci.uni-heidelberg.de)



**Fig. 1** H&E-stained bright-field image of mouse brain tissue. A solid tumor is shown in the upper part. The white bordered fields were scanned with MCARS and spontaneous Raman microspectroscopy.

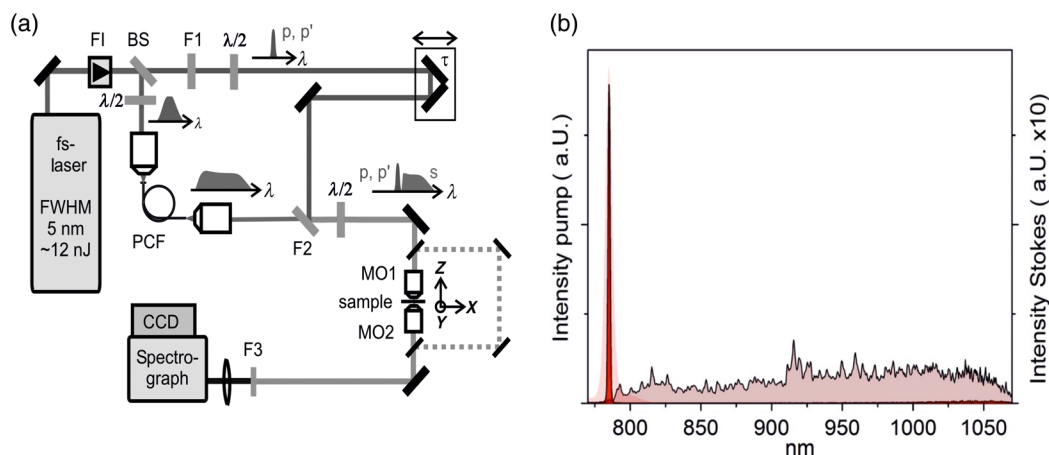
The MCARS experimental setup has been recently described,<sup>16</sup> and its scheme is shown in Fig. 2. The laser beams driving the MCARS process were obtained from a Ti:Sapphire oscillator (Coherent Mira) by narrowband filtration [for pump and probe beams, centered at 785 nm with a full width at half maximum (FWHM) of 1 nm] and using a photonic crystal fiber (PCF) for Stokes-beam generation (FWHM of pump and probe  $\sim 16 \text{ cm}^{-1}$ , Stokes-continuum ranging from 790 to 1100 nm). The beams were focused in the sample by a 50 $\times$  microscope objective (Olympus LMPlan50xIR). The pixel dwell time was set to 60 ms using a real-time environment (ADWin) controlling a piezodriven sample holder. Areas of  $97 \times 97 \mu\text{m}^2$  were raster scanned with a step size of  $1 \mu\text{m}$ . The signal passed a spectrograph (Princeton Instruments) and was spectrally resolved (Andor Newton CCD). We estimated the average power in focus volume to be  $\sim 30 \text{ mW}$ .

Spontaneous Raman microspectroscopy was performed using a commercial system (Witec, Alpha 300) with an excitation wavelength of 532 nm. A sample area of  $97 \times 97 \mu\text{m}$  was measured accordingly. With an irradiation power of 5 to 10 mW, the acquisition time was set to 500 ms per spectrum to balance between signal quality and prevention of photodamage at the selected laser power.

The data processing consisted of separate steps. After normalization of the MCARS raw data by a background signal to correct for experimental influences, the Raman information was extracted using the maximum entropy method (MEM) algorithm,<sup>9</sup> and an asymmetric least square (ALS)-baseline correction was carried out.<sup>17</sup> Before performing SVM, it is reasonable to include a dimension reduction step. Instead of using the original 800 spectral variables, 15 PCA scores were used after PCA. The processing of MCARS data with PCA has been described in a previous work.<sup>16</sup> Thereafter, the SVM algorithm with radial-base kernel was used to build a classification model for normal and cancerous tissue. The calculation of MEM and PCA were carried out in LabView Software while the R-program<sup>18</sup> with the package e1072 (Ref. 19) and baseline<sup>17</sup> was used for SVM and ALS-baseline correction, respectively. All results have been qualitatively correlated with spontaneous Raman microspectroscopy and with bright-field microscopy of the H&E-stained reference slide.

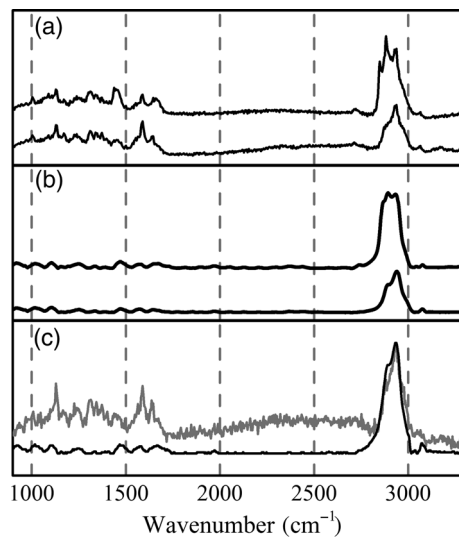
### 3 Results and Discussion

The mouse brain tissue consists of a solid brain tumor of the melanoma cell line mK1735 that is bordering the brain parenchyma of the cortex. The H&E stained reference is shown in bright-field contrast at 100 $\times$  magnification in Fig. 1. The melanoma appears in the upper part (high nuclear density, dark red), whereas the gray matter is shown in the lower part (light red). The border between tumor and normal tissue has been measured



**Fig. 2** (a) Optical setup for MCARS laser pulse generation is driven by an fs-laser source (Coherent Mira), protected by a Faraday isolator (F1). A beam splitter separates the beam into two pathways. The reflected part is used for supercontinuum generation in a PCF while the transmitted part passes a narrowband filter (F1) resulting in narrowband laser pulses with only 1-nm FWHM that are used as pump and probe beams (p and p'). A delay stage (t) allows temporal synchronization with the supercontinuum, which acts as Stokes pulse. Spatial overlap is realized by a long-pass filter (F2). The laser pulse package is coupled into a microscope objective (MO1) and collimated by a second objective (MO2) after transmitting the sample (mounted on a piezodriven three-axis stage). A long-pass filter separates the blue-shifted MCARS signal, which is detected and spectrally resolved by a CCD detector. (b) Pump/probe and Stokes spectra. Pale red: the initial fs beam (3-nm FWHM) at 785 nm. Red: the narrowband filtered part (785 nm), which is used as pump and probe. The ultrabroadband fiber continuum (790 to 1100 nm) is used as Stokes pulse.



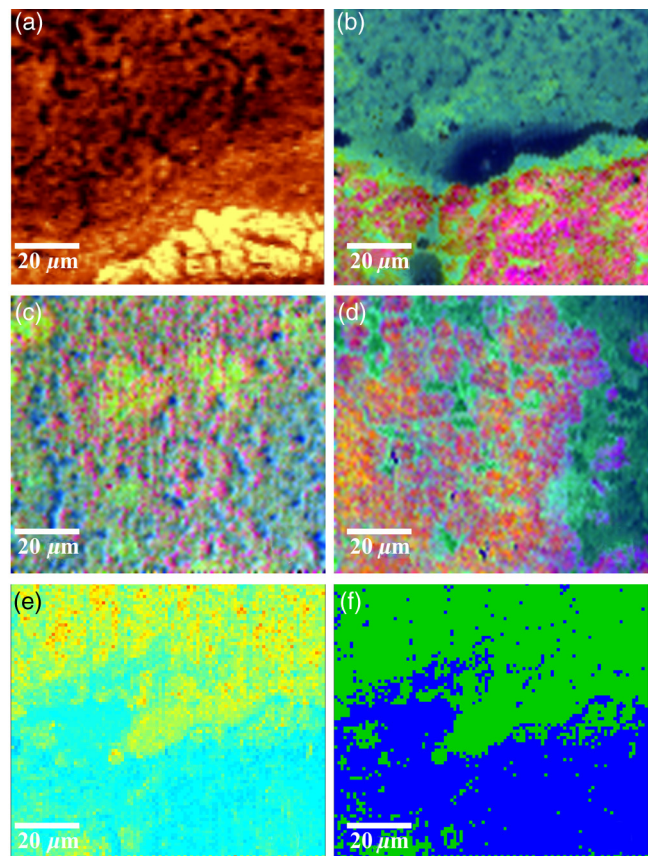


**Fig. 3** (a) Averaged Raman spectra of normal (upper line) and tumor tissue (lower line). Due to the weak fluorescence, the baseline was corrected by a standard polynomial fit and respective subtraction. (b) Averaged MCARS spectra of normal (upper line) and tumor tissue (lower line) after normalization and MEM extraction. (c) Raman spectrum of tumor tissue with 500-ms acquisition time (gray line) and MEM-extracted MCARS spectrum with an acquisition time of 60 ms (black line).

by MCARS and spontaneous Raman microscopy (regions 1 and 2 in Fig. 1). Additional MCARS-scans have been performed in normal tissue of gray matter (region 3) and tumor tissue (region 4).

The average of 390 spontaneous Raman spectra measured in tumor and normal tissue in region 2 of Fig. 1 is shown in Fig. 3(a). Both spectra show clear signal intensities between 2800 and 3000  $\text{cm}^{-1}$ , where the  $\text{CH}_2$  symmetric- and antisymmetric-vibrational modes are the most prominent features around 2890 and 2930  $\text{cm}^{-1}$ , respectively. The typical Raman signatures of proteins in brain tissue show up at 1440 and 1660  $\text{cm}^{-1}$ , originating from the  $\text{CH}_2$ -deformation and the amide vibrational modes.<sup>20</sup> Significant differences occur at 1442 and 2887  $\text{cm}^{-1}$ , where the normal tissue provides stronger relative signal intensities. Since tumors may consist of several subpopulations of cells with highly different metastatic behaviors,<sup>21,22</sup> we cannot exclude that tumors with different grades of proliferation might also show different CARS signatures. Therefore, biological aspects of spectral alterations are not a subject of this study and will be addressed in future experiments. The average MCARS-spectra of normal tissue and tumor tissue in Fig. 1 are shown in Fig. 3(b), after normalization and MEM processing of the raw data. Taking the lower spectral resolution of the CARS setup into account, the CH-stretching region shows good agreement with the data obtained from spontaneous Raman spectroscopy and is considered to have a higher influence on the later classification result compared with the fingerprint region that is not as prominent as in the Raman control in this study. The signal-to-noise ratio of both methods is compared in Fig. 3(c). While the spontaneous Raman spectrum was recorded with an acquisition time of 500 ms, the Raman-extracted MCARS spectrum of the same specimen is measured within just 60 ms, but it still shows a three times better signal-to-noise ratio.

Figure 4(a) shows the result of spontaneous Raman imaging performed at field 2 of Fig. 1, where tumor tissue borders the gray matter. The integrated signal between 2800 and 3100  $\text{cm}^{-1}$



**Fig. 4** (a) Intensity map of the CH stretching vibrational region (integrated between 2831 and 3032  $\text{cm}^{-1}$ ) using spontaneous Raman microscopy at the border between tumor (upper left) and normal tissue according to area 2 in Fig. 1. (b) MCARS imaging of a similar area using Raman extraction and PCA (area 1 in Fig. 1). (c) MCARS imaging of tumor tissue of area 4 in Fig. 1, using Raman extraction and PCA. (d) MCARS imaging of normal tissue (area 3 in Fig. 1) using Raman extraction and PCA. (e) Decision values of the SVM classification with tumor tissue (yellow to red) and bordering gray matter (cyan to blue) [same area as in (b)]. (f) Group value of the SVM classification with positive values (tumor tissue, green color) and negative values (normal tissue, blue color).

referred to a color map ranging from black to yellow. The normal tissue displays a higher absolute signal intensity and is consecutively encoded in yellow. The result reveals a near perfect overlap with the H&E stained reference image (Fig. 1). Figure 4(b) shows a similar area measured with MCARS. The image contrast was obtained using the two step approach consisting of Raman extraction and PCA.<sup>8</sup> The color coding refers to spectral differences within the scanned area. The scores arising from PCA were associated with different colors, so tumor tissue appears in green and normal tissue in red. This color coding was based on previous knowledge such as the H&E reference and manual recognition of spectra. The same method of data postprocessing was performed in the next two images, where the tissue of the fourth area is shown in Fig. 1(c) and the one from the third area in Fig. 1(d). The scores resulting from the three highest loadings were set to red, green, and blue. The combination of the color maps results in an image contrast showing the morphology of the dried sample, which is also different for tumor and normal tissue. On the other hand, the results of Figs. 4(c) and 4(d) clearly show the limitation of the method. Although image contrast is obtained, it is not possible to use

information of a previous measurement to apply the same color coding, for example, the color coding of Fig. 4(b). We address this issue using the spectral data sets of Figs. 4(c) and 4(d) to train the SVM model: these two datasets are used to establish a simple binary SVM model based on tumor and normal tissue. The algorithm constructs a separating hyperplane based on the largest margin between both groups. For every spectrum, the distance to that hyperplane is calculated. The distance is subsequently converted to a classification decision for a new spectrum. To build such a model, the spectra (9409 each) of the tumor (area 4 in Fig. 1) and the healthy region (area 3 in Fig. 1) represented the training set. The SVM model based on the spectral datasets of Figs. 4(c) and 4(d) was then applied to the tissue region containing the border of the tumor [area 1 in Fig. 1, Fig. 4(b)] as an independent test set. The accuracy of the PCA-SVM model on the training set was 99.985% estimated by a 10-fold cross validation. The results of the SVM analysis of the border of the tumor are shown in Figs. 4(e) and 4(f): In Fig. 4(e), the decision value of the region with the tumor border (area 1 of Fig. 1) is shown, e.g., the distance to the hyperplane. In that representation, the region of brain parenchyma exhibits negative values, while the tumor region shows positive decision values. The decision values can be seen as a spectral marker, which is afterwards converted to a classification for the tissue. This is shown in Fig. 4(f), where only two colors exist, indicating the two groups (green represents tumor tissue and blue indicates normal tissue). This procedure now relies on trained regions and automatically highlights the tumor border. Furthermore, it simplifies the information necessary for potential *in vivo* diagnostics. The image result of Figs. 4(e) and 4(f) basically agrees with Fig. 4(b) and the H&E stained reference. Minor differences occur at the border between the tissues due to non-resonant contributions such as glass not included in our training data sets. We also have to consider that data originating from highly complex biological tissue alterations might not be fully characterized by a single hyperplane in an SVM routine. Nevertheless, this model is a first step toward a CARS-based diagnostic procedure for brain metastases of malignant melanoma. Moreover, we would like to emphasize that the H&E stained information is only required to distinguish the tumor from normal tissue in the original training datasets. Later, when the SMV model is applied to the test dataset, no H&E stained information is used or required. H&E staining was shown here only to further confirm our analysis since it has been the gold standard in microscopy of biological systems.

## 4 Conclusion

The border between normal mouse brain tissue and a solid tumor has been visualized by MCARS microscopy, and an SVM model based on a PCA of training data was established to predict the border automatically. Therefore, the image contrast is based on class affinities and not limited to purely relative effects such as spectral differences within a single data set. This is an important advantage over PCA-alone-based analysis approaches (or similar methods like singular value decomposition): image contrast can be obtained, but the obtained information cannot be directly applied for the detection of tumor regions in an unknown sample. This is not the case with the SVM model based on a PCA illustrated in this work. We illustrate that MCARS microscopy can provide suitable information for higher ranking data analysis as known for linear techniques such as spontaneous Raman and infrared microscopy. In this study, MCARS also

turned out to be already one order of magnitude faster than spontaneous Raman microscopy, which was used to verify the results. Future studies will focus on improving spectral resolution and accommodating the higher ranking classification method introduced here.<sup>23</sup> For example, if more classes would be present as training data sets, a combination scheme for the binary classifier could be applied. Often used schemes are one-against-one or one-against-rest, while using the SVM-multiclass formulation would be an interesting option to proceed. Nevertheless, the question of which of these methods is best for MCARS is the focus of ongoing research. Moreover, the type of tumor used in this work shows very sharp tumor/normal margins and is the best-case scenario for the SVM-PCA modeling. It is still an open question whether more structured margins, such as an infiltrating glioma, would be reasonably predicted by our model. Finally, future tasks will also aim to explore the potential of MCARS for providing high-speed and Raman analogue spectral information suitable for supporting databases of various kinds of biological tissues for biomedical imaging and *in vivo* application.

## Ethical Standard

Animal experiments were conducted in accordance with the institutional and national guidelines in full agreement with the European Union directives. They were approved by the animal welfare committee of Saxony, Germany (Regierungspräsidium Dresden, AZ: 24-9168.11-1/2011-39). All surgery was performed under ketamine-xylazine anesthesia, and all efforts were made to minimize suffering.

## Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

## Acknowledgments

This work was supported by the BMBF Project (MediCARS, FKZ: 13N10778) and the seventh European Framework Programme (CROSS TRAP 244068).

## References

1. A. Volkmer, J. X. Cheng, and X. S. Xie, "Vibrational imaging with high sensitivity via epidetected coherent anti-Stokes Raman scattering microscopy," *Phys. Rev. Lett.* **87**(2), 023901 (2001).
2. C. L. Evans et al., "Chemical imaging of tissue *in vivo* with video-rate coherent anti-Stokes Raman scattering microscopy," *Proc. Natl. Acad. Sci. U. S. A.* **102**(46), 16807-16812 (2005).
3. M. Müller and A. Zumbusch, "Coherent anti-Stokes Raman scattering microscopy," *Chemphyschem* **8**(15), 2157-2170 (2007).
4. C. L. Evans and X. S. Xie, "Coherent anti-Stokes Raman scattering microscopy: chemical imaging for biology and medicine," *Annu. Rev. Anal. Chem.* **1**, 883-909 (2008).
5. J. L. Suhaim et al., "The need for speed," *J. Biophotonics* **5**, 387-395 (2012).
6. E. Belanger et al., "Live animal myelin histomorphometry of the spinal cord with video-rate multimodal nonlinear microendoscopy," *J. Biomed. Opt.* **17**(2), 021107 (2012).
7. B. von Vacano, L. Meyer, and M. Motzkus, "Rapid polymer blend imaging with quantitative broadband multiplex CARS microscopy," *J. Raman Spectrosc.* **38**(7), 916-926 (2007).
8. C. Pohling et al., "Chemoselective imaging of mouse brain tissue via multiplex CARS microscopy," *Biomed. Opt. Express* **2**(8), 2110-2116 (2011).
9. E. M. Vartiainen et al., "Direct extraction of Raman line-shapes from congested CARS spectra," *Opt. Express* **14**(8), 3622-3630 (2006).

10. K. V. Branden and M. Hubert, "Robust classification in high dimensions based on the SIMCA method," *Chemom. Intell. Lab. Syst.* **79**(1–2), 10–21 (2005).
11. C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discovery* **2**(2), 121–167 (1998).
12. B. Bird et al., "Spectral detection of micro-metastases and individual metastatic cells in lymph node histology," *Technol. Cancer Res. Treat.* **10**(2), 135–144 (2011).
13. C. Krafft et al., "Disease recognition by infrared and Raman spectroscopy," *J. Biophotonics* **2**(1–2), 13–28 (2009).
14. M. B. Ji et al., "Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy," *Sci. Trans. Med.* **5**(201), 201ra119 (2013).
15. M. Kirsch et al., "Angiostatin suppresses malignant glioma growth in vivo," *Cancer Res.* **58**(20), 4654–4659 (1998).
16. C. Pohling, T. Buckup, and M. Motzkus, "Hyperspectral data processing for chemoselective multiplex CARS microscopy of unknown samples," *J. Biomed. Opt.* **16**(2), 021105 (2011).
17. K. H. Liland, T. Almoy, and B. H. Mevik, "Optimal baseline correction for multivariate calibration using open-source software," *Am. Lab.* **43**(4), 13–16 (2011).
18. R. Ihaka and R. Gentleman, "R: a language for data analysis and graphics," *J. Comput. Graphical Stat.* **5**(3), 299–314 (1996).
19. E. Dimitriadou et al., "Misc functions of the department of statistics (e1071), TU Wien," 2008, [cran.r-project.org/web/packages/e1071/index.html](http://cran.r-project.org/web/packages/e1071/index.html), R Foundation for Statistical Computing, Vienna, Austria (6 April 2009).
20. N. Amharref et al., "Discriminating healthy from tumor and necrosis tissue in rat brain tissue samples by Raman spectral imaging," *Biochim Biophys Acta.* **1768**(10), 2605–2615 (2007).
21. I. J. Fidler and M. L. Kripke, "Metastasis results from preexisting variant cells within a malignant tumor," *Science* **197**, 893–895 (1977).
22. I. J. Fidler et al., "Demonstration of multiple phenotypic in a diversity murine melanoma of recent origin," *J. Natl. Cancer Inst.* **67**(4), 947–956 (1981).
23. M. B. Fenn et al., "Raman spectroscopy utilizing Fisher-based feature selection combined with support vector machines for the characterization of breast cell lines," *J. Raman Spectrosc.* **44**, 939–948 (2013).

Biographies for the authors are not available.