The analysis of hyperspectral broadband coherent anti-Stokes Raman scattering (CARS) microscopic images

Alexander Khmaladze

Department of Chemistry, University of Michigan, 930 N. University, Ann Arbor, MI, USA 48109

Coherent anti-Stokes Raman scattering (CARS) microscopy is a powerful nonlinear optical technique, which allows chemically selective imaging of microscopic samples. In broadband CARS, the application of a spectrally narrow pump and probe beams and a spectrally wide Stokes beam results in multiple Raman transitions occurring simultaneously, generating a broad CARS spectrum with several orders of magnitude higher intensity relative to the spontaneous Raman scattering. This allows generation of three dimensional data sets (one dimensional spectral data for each pixel of a two dimensional image), known as hyperspectral images. Here, the applications of several techniques to hyperspectral CARS image analysis are illustrated. In particular, singular value decomposition (SVD) is shown to be a highly efficient method of distinguishing between various components of a polymer bead mixture. SVD performance is demonstrated to be superior to the better-known k-means clustering analysis.

Keywords Coherent anti-Stokes Raman scattering (CARS); Singular value decomposition (SVD); K-means clustering analysis; Raman microscopy; Scanning microscopy; Spectroscopy

1. Introduction

Vibrational spectroscopy, coupled with scanning microscopic imaging provides a way to visualize unstained, live cells with chemical selectivity. Linear vibrational spectroscopic techniques based on infrared (IR) absorption and spontaneous Raman scattering have been applied to image such samples. However, the spatial resolution of IR microscopy is limited by longer excitation wavelengths (on the order of several micrometres), low sensitivity due to non-background-free detection, and strong water absorption of the mid IR light which limits its application to biological studies [1].

Linear Raman microscopy has been extensively explored and successfully applied to biomedical applications such as DNA and glucose detection [2,3], tumor diagnostics [4,5], as well as microendoscopy [6], and bone imaging [7]. Spontaneous Raman spectroscopy provides excellent spectral resolution. However, Raman imaging requires long acquisition times because of the non-directional light scattering pattern. As a result, it commonly takes several hours to acquire a Raman image of cells and tissues. The long exposure time limits the application of Raman microscopy for the dynamic living systems. This limitation is overcome by multi-photon, vibrational coherent anti-Stokes Raman scattering (CARS) microscopy.

CARS as a method of spectroscopic chemical analysis was first reported in 1965 by Maker and Terhune [8]. By using a pulsed ruby laser, they were able to observe the nonlinear optical effects of an induced optical polarization in electric field density. In 1982, at the Naval Research laboratory, this method was used for construction of a spatially scanning microscopy apparatus to collect CARS images [9]. However, significant advances were not made until 1999, when the Xie group [10] built a CARS microscope using a new near-infrared laser with ultrashort pulses. With these advances, biological living samples could be imaged by CARS without the requirement of exogenous labels that perturb the observed system [11].

CARS signal is produced as a result of a nonlinear optical four-wave mixing process. It involves three photons interacting with a sample, which are generated by pump (I_{pump}), Stokes (I_{Stokes}), and probe (I_{probe}) beams, producing the fourth, CARS photon (I_{CARS}). CARS signal intensity is greatly increased when the frequency difference between Stokes and pump beams match the frequency of a Raman resonance [12,13]. Using the same beam for pump and probe, the intensity of the CARS signal can be written as

$$I_{CARS} \sim |\chi^{(3)}_{CARS}|^2 I_{pump/probe}{}^2 I_{Stokes} \quad (1)$$

where $\chi^{(3)}_{CARS}$ is the third order nonlinear optical susceptibility. By recording the CARS intensity as a function of the frequency difference between pump/probe and Stokes wavelengths, it is possible to obtain spectroscopic information about the vibrational transitions of molecules within a sample.

Recent advances in the development of ultrafast lasers, as well as photonic crystal fibers (PCF), enable broadband CARS. The mixing of a spectrally narrow pump and probe beam and a spectrally wide Stokes beam results in multiple Raman transitions occurring simultaneously, generating a broad CARS spectrum. This technique, coupled with a scanning laser microscope and spectral detection by a highly sensitive back-illuminated cooled CCD camera, allows generation of three dimensional data sets (one dimensional spectral data for each pixel of a two dimensional image),

known as hyperspectral images. These hyperspectral images need to be analyzed in order to extract the chemical information for each pixel. Here, the application of singular value decomposition (SVD) technique to CARS hyperspectral CARS image analysis is illustrated and compared to the better-known k-means clustering analysis.

2. Broadband CARS System

A schematic of the CARS setup [14] is shown in Fig. 1(a). The light source is a Ti:sapphire oscillator (Spectra Physics Mai Tai), which produces femtosecond pulses, centred at 800 nm (repetition rate of 80 MHz). The laser output power is 2.95 Watts. After passing through a Faraday isolator, the beam is split by a Glan prism to the pump/probe and Stokes arms.

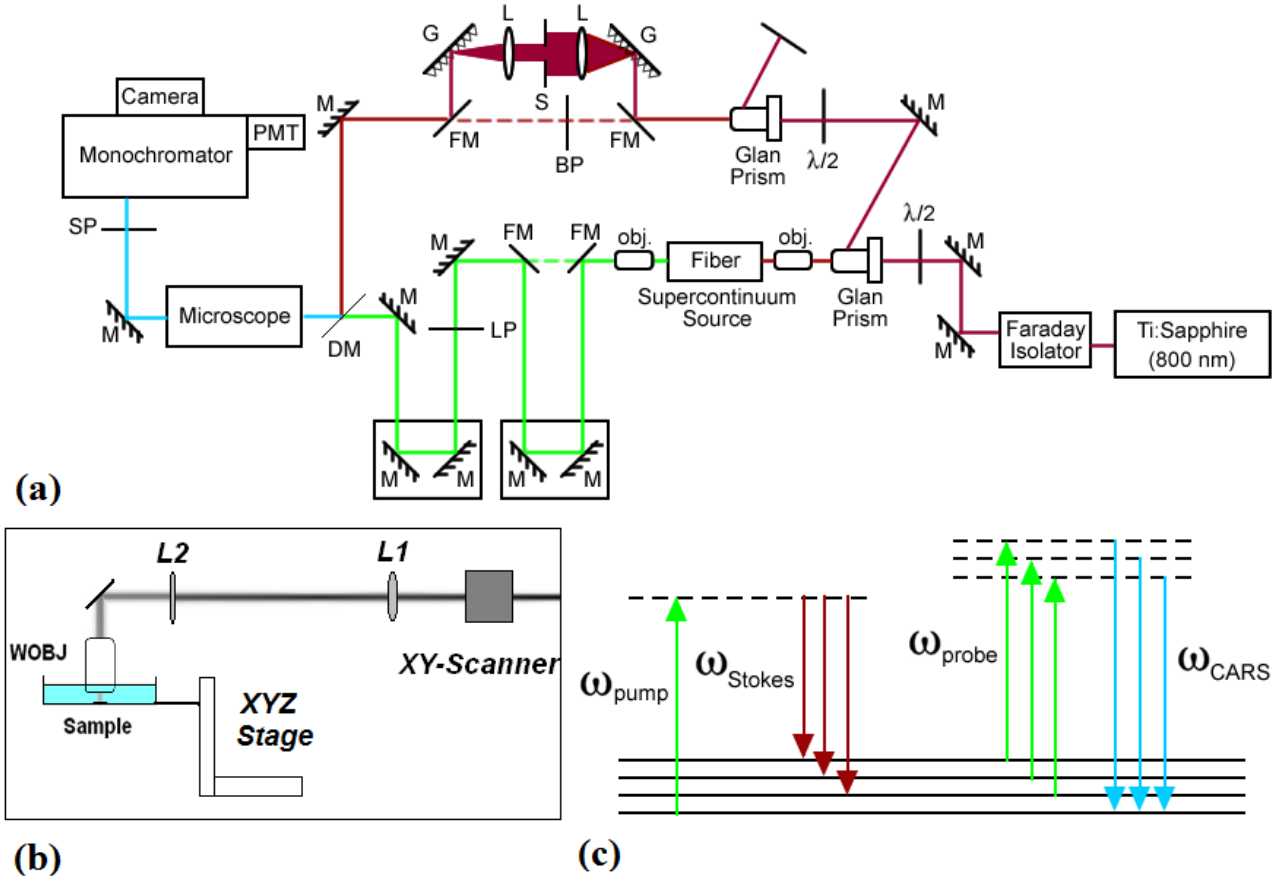

Fig. 1 Broadband CARS system: (a) CARS spectroscope: M – mirror, λ/2 – half-wave plate, OBJ. - 40x microscope objective, LP/SP/BP – long-pass/short-pass/band-pass filters, G - grating, FM – flip mirror, S – slit, L – lenses (b) CARS microscope diagram (c) Broadband CARS energy levels.

The spectral resolution of a CARS system is determined by the spectral width of the pump/probe beam. In this setup, pump/probe line was flexible to either allow for lower spectral resolution (but maximum power for fast imaging, as shown in the dotted line segment in the pump/probe line) obtained by using a 3 nm bandpass (BP) filter (Semrock LL01-808-25) or to higher spectral resolution obtained by using a 4*f*- system with two gratings (600 lines/mm).

Stokes beam is generated by focusing the remaining portion of the oscillator beam into a PCF (Newport SCG800). Stokes beam is passed through a delay stage to match the path lengths of that beam to the pump/probe beam. The second stage is used if the beam is passed through the 4*f*- system. The use of a motorized delay stage allows for broad tuning of the Stokes beam to minimize the amount of non-resonant background (arising from a purely electronic response of the material). Stocks beam is then spatially overlapped with the pump/probe arm by a dichroic mirror. Both beams are then sent into a laser scanning microscope.

The microscope (Fig. 1(b)) was assembled vertically using Thorlabs cage systems hardware. For CARS imaging, a XY-galvanometer scanner (Thorlabs) is used, as it allows for fast scanning. In this case, a CARS scanning microscope system was designed and built to produce a diffraction limited (~ 500 nm) scanning beam. The focal lengths of lenses L1 and L2 were 7.5 cm and 10 cm respectively, and the distance between the lenses L2 and a 40x microscope objective was equal to the sum of their focal lengths. Samples were positioned on a micrometer driven XYZ stage, which could be both computer and manually controlled, facilitating easy sample changing, zooming, focusing and moving. The

beams were focused onto the sample using either a regular or a water-immersion objective. The power at the sample was measured to be around 8 mW.

Here, a broadband CARS technique is employed. Its energy diagram is shown in Fig. 1(c). In broadband CARS, multiple Raman transitions are excited within the bandwidth of the Stokes pulse and probed with the narrow probe beam, which makes it possible to generate an entire CARS spectrum in one acquisition. The measured continuum beam profiles through the PCF and the pump/probe beam profiles are shown in Fig. 2(a). Of the supercontinuum light, the wavelengths higher than 800 nm were used to generate the Stokes beam (the rest is removed with a low-pass (LP) filter). When this spectroscopic signal generation is used together with a scanning microscope, each pixel within the image contains an entire vibrational spectrum at that point.

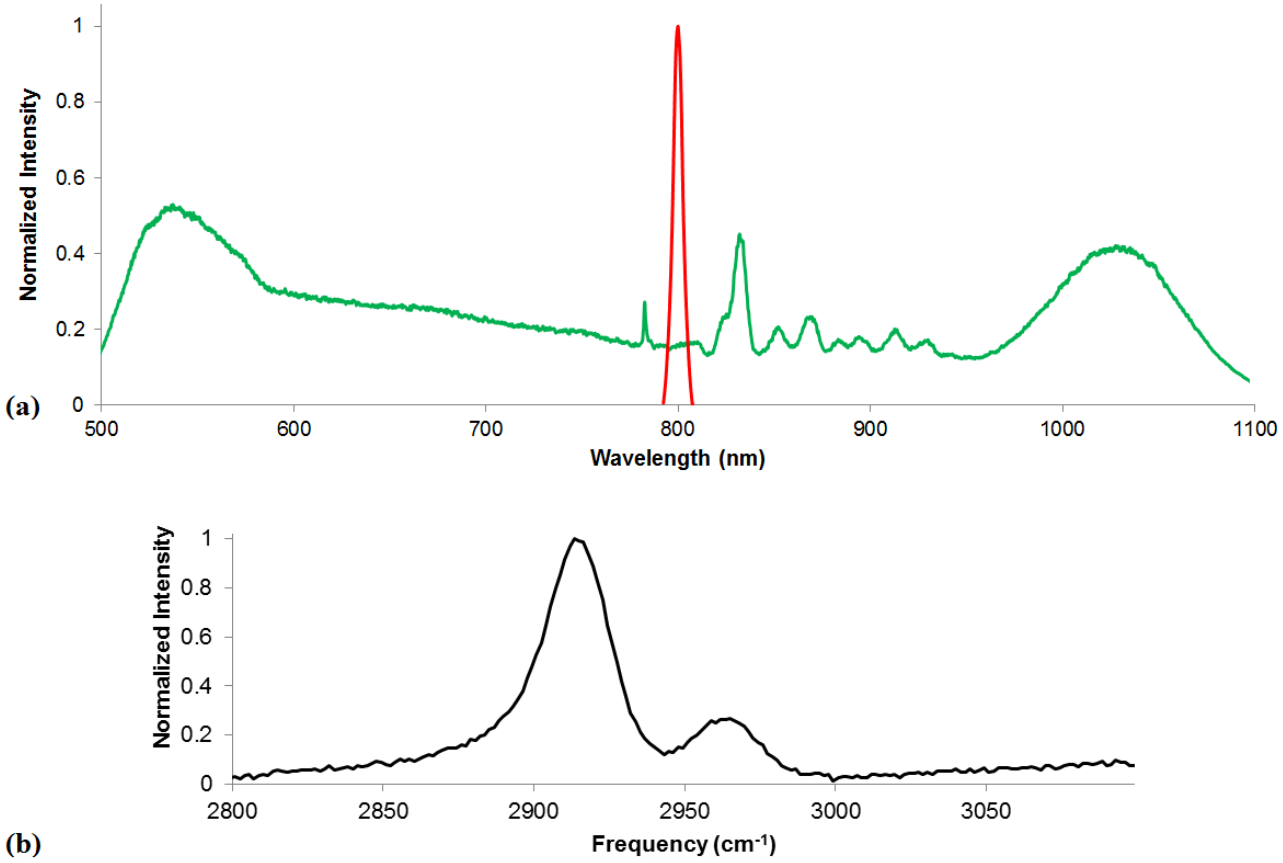

Fig. 2 Spectra (a) combined spectra of pump/probe (red) and Stokes (green) beams (b) CARS spectrum of a thin PDMS film. Peaks are centred at 2915 cm^{-1} and 2965 cm^{-1}.

The forward CARS signal is collected by a condenser lens, and after the remaining pump and Stokes light is filtered out by a shortpass filter (SP), is sent to a spectrometer (Oriel MS257 with adjustable entrance and exit slits), which houses both a PMT detector used for single-wavelength imaging and a CCD camera for multi-wavelength spectral acquisition. A set of Labview programs were developed to control scanning, collect the CARS signal from both PMT and CCD detectors, and analyze the hyperspectral data.

In order to use the system as a CARS spectrometer, scanning mirrors in the microscope are set to zero deflection and the spectrum can be collected by the monochromator and PMT or CCD. A CARS spectrum obtained from a thin film of polydimethylsiloxane (PDMS) is shown in shown in Fig. 2(b). The two peaks (C-H symmetric and asymmetric stretching of Si-CH_3 groups) at 2915 cm^{-1} and 2965 cm^{-1} are well separated and easily observed directly or via spectral fitting. This spectrum is consistent with the published spectrum [15] and shows that this CARS spectrometer is capable of resolving peaks separated by less than 50 cm^{-1}.

PMT-based image and spectra acquisition are performed similarly to any traditional scanning microscope system, i.e. the intensity on the PMT is read at a particular wavelength by adjusting the position of the diffraction grating inside the monochromator. The CCD camera (Oriel InstaSpec X model 78237 attached to the second monochromator port) permits simultaneous detection of a broad (80 nm) spectral range. The location of this spectral range is broadly adjustable by tuning the monochromator grating.

Although this CARS setup lacks the confocal pinhole typically used in commercial scanning microscope systems, the third order dependence of CARS signal on the input intensity results in the CARS signal intensity rapidly decreasing

away from the plane, where the laser spot is focused. Therefore, little out of focus signal is reaching the CARS detectors.

In addition to that, there is a sufficiently high amount of Rayleigh scattering generated by the pump beam to detect the scattered signal by a photodiode placed slightly off-axis from the CARS beam line past the microscope. This signal is then used to create a scanning optical image [16] in the same manner and at the same time as CARS.

Thus, such broadband CARS system is capable of acquiring hyperspectral CARS images in any of the following two ways. The first is a PMT-based detection method, which can acquire a sequence of CARS images at different wavelengths. If one is interested in a CARS spectrum at a particular point, the spectrum can be subsequently reconstructed from the stack of images taken at the different wavelength by recording the intensity at this particular point for each image as a function of wavelength/wavenumber [10].

The second method is based on the simultaneous detection of the portion of the spectra by a CCD camera, coupled with point-by-point XY scanning. Since for this setup all the CARS wavelengths are generated simultaneously, by scanning the beam over the sample and recording spectra at each point, a three-dimensional data set can be acquired – a set of spectra recorded for each pixel of a two-dimensional image. This hyperspectral image of the sample can be easily rendered as either a set of two-dimensional images corresponding to each wavelength/wavenumber or a set of spectral lines (1D spatial data of spectra for each point). One must note that if the sample is a dynamic system, where changes in morphology and/or chemical composition occur within a short period of time, only with the CCD based scanning is it possible to obtain accurate spectral information for each image point.

3. Hyperspectral Data Analysis and Classification

Once the entire three-dimensional set of hyperspectral data is acquitted, several approaches to data analysis can be undertaken. For example, the three dimensional data can be viewed as a series of images, by taking two-dimensional "slices" for a set wavenumber/wavelength. By selecting a set of several wavenumbers, which correspond to the spectral peaks of the various compounds of the sample, these images can be examined. While this method is sufficient to accurately classify the chemical composition of samples in many cases, it has some limitations, such as when multiple distinct peaks are present, or when the images are noisy or suffer from inconsistent brightness. This is especially evident when the sample contains two chemicals with the CARS spectral peaks close to each another. The two compounds will then both appear bright at the wavelengths corresponding to those peaks, making it impossible to distinguish between the two. However, by comparing the spectral lineshapes, the two compounds can easily be distinguished.

Figure 3(a) shows the CARS image of a mixture of polystyrene (PS) and larger poly(methyl methacrylate) (PMMA) beads. As expected, PS produced signal at the C-H aliphatic stretching region at 2850 cm^{-1} (Fig. 3(b) and the aromatic C-H stretching region at 3020 cm^{-1}(Fig. 3(c)), but PMMA produced a peak at around 2850 cm^{-1} as well (Fig. 3(d)), limiting the applications of single wavelength intensity measurements. Moreover, images acquired at only a single wavelength may show bright or dark spots due to factors other than changes in chemical composition. For example, the intensity of the pump/probe or Stokes beam may vary slightly within the image frame, either due to scattering or temporal variation of the laser intensity, producing an uneven distribution of CARS signal. For the image in Fig. 3, some of the increased brightness in the CARS image of the beads is seen in the first order scattered light as well, and may be attributable to variation of the transmission coefficient within the sample. Also, the edges of the beads appear dimmer than the center due to the sample's geometry (a change in the orientation of the bead surface with respect to the scanning beams). For all of the above reasons, a single-wavelength CARS image always contains both chemical and optical information.

Therefore, a different approach to data analysis is needed. In order to utilize all the available spectral information and classify the entire set of hyperspectral data, several algorithms exist. K-means clustering is one of the better known methods. It has been previously applied to the classification of spontaneous Raman data [17-20]. K-means clustering is a method of partitioning all the spectra from a hyperspectral image into several clusters, in which each spectra belongs to the cluster with the nearest mean. Once the spectra have been classified, they are mapped back on the two-dimensional image frame, showing the chemical map of the sample.

Figure 4 shows the result of this approach applied to the hyperspectral set shown in Fig 3. The program for k-means analysis was written in Labview and used a clustering function from Matlab. Prior to clustering, each spectum was normalised and background was subtracted. Since in this case it is known that the sample contains two kinds of beads and the aqueous background, three clusters are used. The average spectra for each cluster is shown in Fig. 4(a). The spectral shapes of the background, PMMA and PS are easily recognisable. The algorithm accurately identifies the locations of all PMMA and PS beads (Fig. 4(b)). The only shortcoming of the k-means clustering generated map is the fact that not all pixels belonging to both kinds of polymer beads were identified. The problem is especially obvious for the larger PMMA beads, where the ring pattern seen in a single wavelength CARS images (Fig. 3 (b-c)) is also visible in Fig. 4(b).

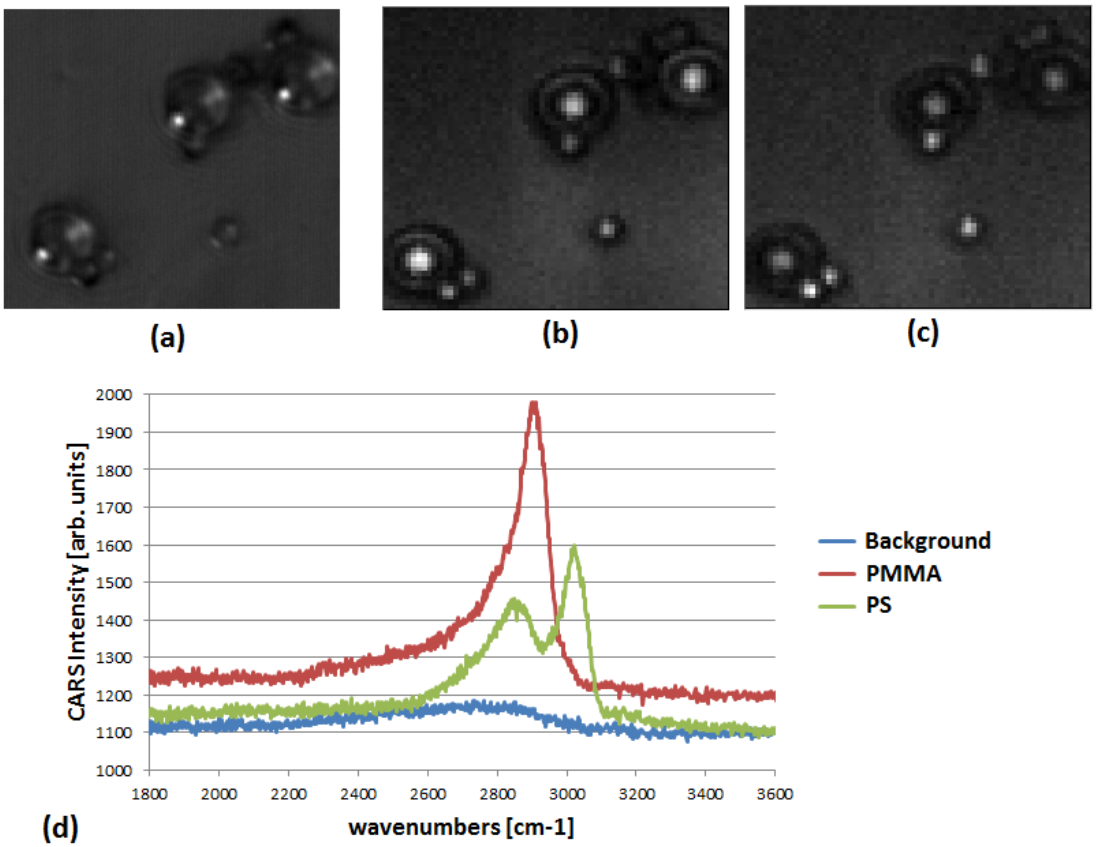

Fig. 3 CARS hyperspectral imaging by CCD camera: (a) optical scanning image, (b-c) single-wavelength images at (b) 2850 cm-1 and (d) 3021 cm-1 (d) spectra of background, PS and PMMA. Note that since the PMMA beads are bigger, they are not entirely in focus. This results in a "ring pattern" in CARS images.

A much better result can be obtained by using a singular value decomposition analysis. SVD is a computationally fast and noise-tolerant algorithm that decomposes a hyperspectral dataset into major contributing lineshapes. When each spectrum (corresponding to a single image pixel) is plotted as a point on a scatter plot (e.g. SVD component 1 vs. SVD component 2), the pixels corresponding to each different chemical compounds are grouped together.

When SVD algorithm (implemented in Labview) was applied to this hyperspectral data set, a scatter plot shown in Fig. 5(a) was obtained, for the two SVD components shown in Fig, 5(b). Once again, one can easily notice the similarity between the lineshapes of SVD components 1 and 2 and the spectra of PMMA and PS respectively. Within the SVD scatter plot, three areas can be identified. One is dominated by SVD component 1, which corresponds to the PMMA pixels. Another is dominated by SVD component 2, which corresponds to the PS pixels. Finally, the area centered around the point (0,0), contains background pixels, as it includes no significant contribution from either component 1 or component 2. By selecting these three groups and mapping them back to the image frame and pseudo coloring them accordingly, the chemical map of the sample was obtained (see Fig. 5(c)). As expected, the larger beads were identified as PMMA and the smaller beads as PS. Notice that even in the cases where the bead edges are blurry, the algorithm correctly identifies them as PS/PMMA and separates them from the background.

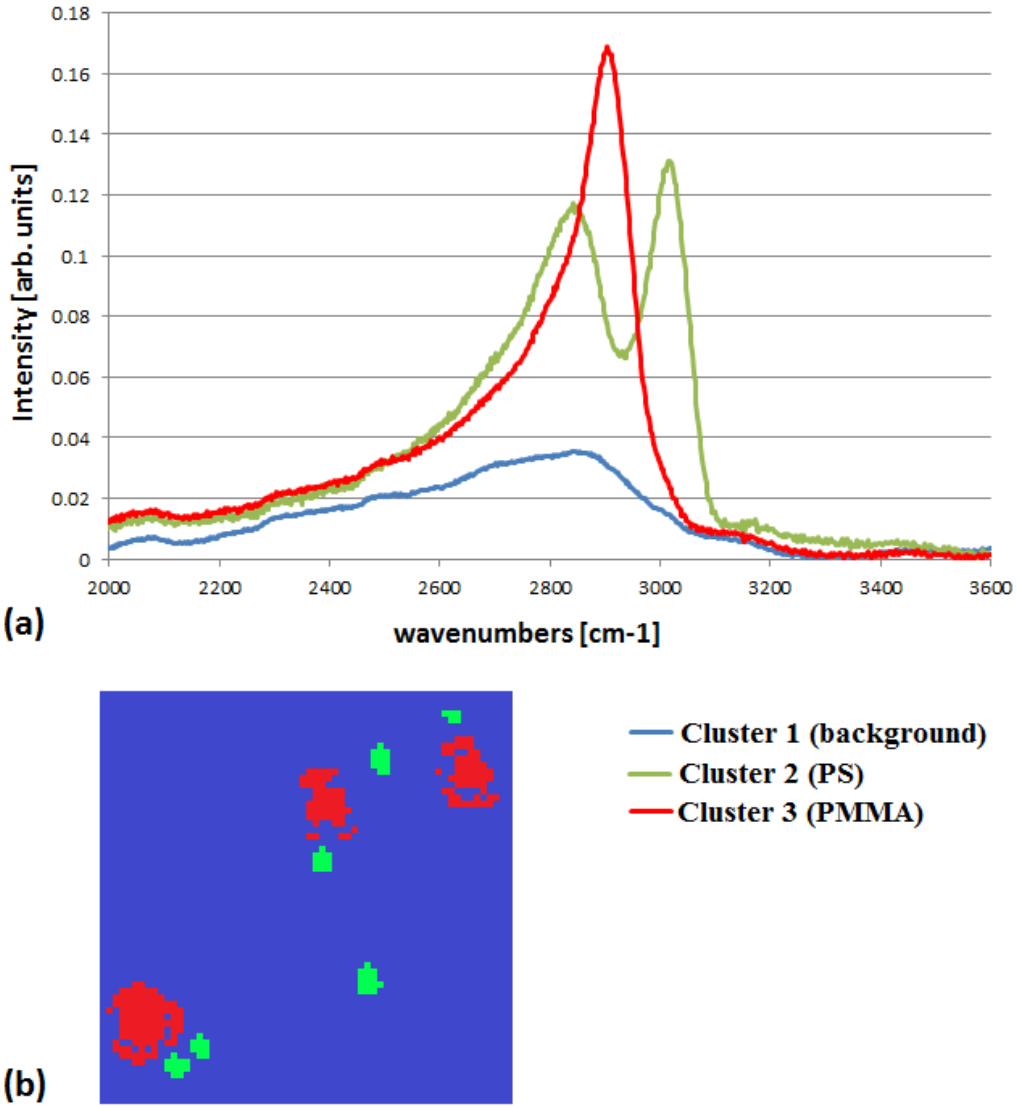

Fig. 4 K-means clustering analysis: (a) average spectra of each of the three clusters corresponding to PS, PMMA and aqueous background, (b) Pseudo-coloured image, based on k-means clustering analysis, clearly distinguishing between PS, PMMA and background.

In conclusion, this implementation of CARS spectroscopic/microscopic system has a number of important advantages. In comparison to spontaneous Raman imaging, the speed at which the hyperspectral data is generated is significantly increased. The entire hyperspectral image (64x64 pixels) presented here takes less than 5 minutes to generate, which potentially enables live cell imaging and allows determining the chemical composition of live cells without dye staining. The detection of Rayleigh scattered light by a photodiode allows for the acquisition of a scanning optical image simultaneously with CARS. Broadband signal generation and detection enables hyperspectral imaging, where each image pixel contains a significant portion of Raman vibrational spectrum. These spectra can be analyzed by various well-established hyperspectral imaging techniques, leading to better understanding of chemical composition of samples. In this case, in addition to producing the most accurate chemical map, SVD proved to be computationally faster and more noise-tolerant than the better known k-means clustering analysis.

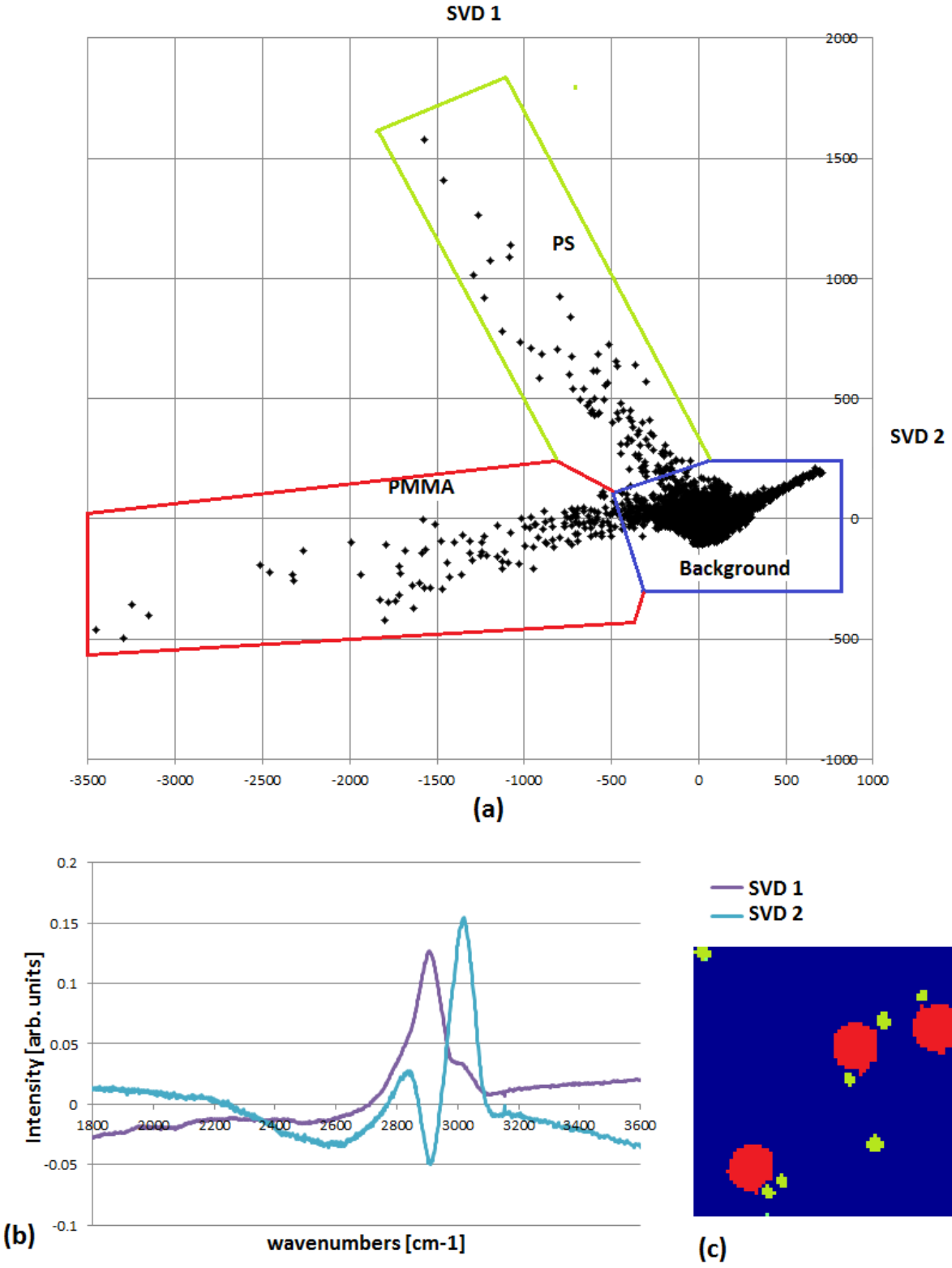

Fig. 5 Singular value decomposition analysis: (a) SVD scatter pot, showing the area corresponding to PS, PMMA and aqueous background (b) SVD components used to generate scatter plot, (c) Pseudo-coloured image, based on SVD analysis, clearly distinguishing between PS, PMMA and background.

Acknowledgements: I am grateful to Mr. Joshua Jasensky for his help in generating the PS/PMMA beads hyperspectral image used here and to Dr. Andrew Boughton for useful discussions about k-means clustering and SVD analysis.

References

[1] Evans CL, Xie XS. Coherent anti-Stokes Raman scattering microscopy: chemical imaging for biology and medicine. *Annu. Rev. Anal. Chem*. 2009; 1:883-909.

[2] Berger AJ, Itzkan I, Field MS. Feasibility of measuring blood glucose concentration by near-infrared regimes. *Spechtrochim. Acta A Mol. Biomol. Spectrosc*. 1997; 53:287-292.

[3] Cao Y, Jin R, Mirkin C. Nanoparticles with Raman spectroscopic fingerprints for DNA and RNA detection. *Science*. 2002; 297:1536-1540.

[4] Huang Z, McWilliams A, Lui H, McLean DI, Lam S, Zeng H. Near-infrared Raman spectroscopy for optical diagnosis of lung cancer. *Int. J. Cancer*. 2003; 107:1046-1052.

[5] Nijssen A, Bakker Schutt TC, Heule F, Caspers PJ, Hayes DP, Neumann MHA, Puppels GJ. Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy. *J. Invest. Dermatol*. 2002; 119:64-69.

[6] Shim MG, Song LMWK, Marcon NE, Wilson BC. In vivo near-infrared Raman spectroscopy: demonstration of feasibility during clinical gastrointestinal endoscopy. *Photochemistry Photobiol*. 2000; 72:146-150.

[7] Kohn DH, Sahar ND, Wallace JM, Golcuk K, Morris MD. Exercise Alters Mineral and Matrix Composition in the Absence of Adding New Bone. *Cells Tissues Organs*. 2009; 189:33-37.

[8] Maker PD, Terhune RW. Study of optical effects due to an induced polarization third order in the electric field strength. *Phys. Rev.* 1965; 137:A801-A818.

[9] Duncan MD, Reintjes J, Manuccia TJ. Scanning coherent anti-Stokes Raman microscopy. *Opt. Lett*. 1982; 7:350-352.

[10] Zumbusch A, Holtom GR, Xie XS. Three-dimensional vibrational imaging by coherent anti-Stokes Raman scattering. *Phys. Rev. Lett*. 1999; 82:4142-4145.

[11] Zipfel WR, Williams RM, Christie R, Nikitin AY, Hyman BT, Webb WW. Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generations. *Proc. Natl. Acad. Sci*. 2003; 100:7075-7080.

[12]Zheltikov AM. Coherent anti-Stokes Raman scattering: from proof-of-the-principle experiments to femtosecond CARS and higher order wave-mixing generalizations. *J. Raman Spectrosc*. 2000; 31:653-667.

[13] Tolles WM, Nibler JW, McDonald JR, Harvey AB. A review of the theory and application of coherent anti-Stokes Raman spectroscopy (CARS). *Appl. Spectrosc*. 1977; 31:253-271.

[14] Khmaladze A, Jasensky J, Zhang C, Han X, Ding J, Seeley E, Liu X, Smith GD, Chen Z. Hyperspectral microscopic imaging by multiplex coherent anti-Stokes Raman scattering (CARS). *SPIE Proc.* 2011; 8158:815805-1-815805-8.

[15] Majumdar P, He J, Lee E, Kallam A, Gubbins N, Stafslien SJ, Daniek J, and Chisholm BJ. Antimicrobial activity of polysiloxane coatings containing quaternary ammonium-functionalized polyhedral oligomeric silsesquioxane. *J. Coat. Technol. Res*. 2010; 7: 455-467.

[16] Khmaladze A, Kim M. Design and imaging properties of a laser scanning microscope with a position-sensitive detector. *J. Mod. Opt*. 2008; 55:2785-2796.

[17] Davidson B, Spears N, Murray A, Elfick, A., The changing biochemical composition and organisation of the murine oocyte and early embryo as revealed by Raman spectroscopic mapping. *J. Raman Spectrosc*., 2011; 43: 24–31.

[18] Krafft C, Knetschke T, Funk RHW, Salzer R. Identification of organelles and vesicles in single cells by Raman microspectroscopic mapping. *Vib. Spectrosc.* 2005; 38:85–93.

[19]Krafft C, Knetschke T, Funk RHW, Salzer R. Studies on Stress-Induced Changes at the Subcellular Level by Raman Microspectroscopic Mapping. *Anal. Chem*. 2006; 78:4424–4429.

[20] Wood BR, Chernenko T, Mattha C, Diem M, Chong C, Bernhard U, Jene C, Brandli AA, McNaughton D, Tobin MJ, Trounson A, Lacham-Kaplan O. Shedding New Light on the Molecular Architecture of Oocytes Using a Combination of Synchrotron Fourier Transform-Infrared and Raman Spectroscopic Mapping. *Anal. Chem*. 2008; 80:9065–9072.

