# Investigation of correlation between characteristics of Raman spectra and parameters of data-scattering obtained from phase coherence theory

R.A. Moore**[a], S. Unnikrishnan*[a], T.S. Perova[a], N.D. McMillan[b], S. Riedel[b], M. O'Neill[b] and G. Doyle[b]

[a]Department of Electronic and Electrical Engineering, Trinity College, Dublin 2, Ireland; [b]Institute of Technology Carlow, Carlow, Ireland

## ABSTRACT

The importance of sensitive monitoring of changes in Raman spectra in particular for microelectronic applications is discussed here. We explore the practicality of using a data-scattering method to analyse Raman spectra, and to establish the dependence of changes observed in all the spectral function characteristics on the parameters of data-scatter such as scatter closeness and scatter radii using "Trace Miner" software. In addition to the analysis performed on model data, analysis on experimental Raman data is also discussed. The sensitivity of the approach is fully appreciated.

**Keywords**: phase coherence theory, data-scattering theory, Raman spectroscopy

## 1. INTRODUCTION

Micro-Raman spectroscopy is being increasingly used for composition, stress and defect analysis in semiconductor structures [1,2]. During these analyses, all the characteristics of Raman spectra, *i.e*. the peak position ($\omega_{max}$), the peak intensity ($I_{max}$), and the full width at half maximum (FWHM), can be used to gain information on different properties of materials, and in particular, at the stage of material fabrication. However, in many cases the changes in aforementioned properties is not so easy to detect. For example, the shift of the phonon lines in stressed semiconductors from the unstressed value can be as small as 0.05 $cm^{-1}$ (at maximum spectral resolution of ~1 $cm^{-1}$). We must note that this shift is directly related to the stress value, so it is important to estimate the shift with higher accuracy. Until now the fitting of the phonon bands with Lorentzian or Gaussian functions has been employed in order to determine the peak position with an accuracy of 0.1 $cm^{-1}$ [2]. Nevertheless, for certain types of semiconductors (in particular for nanostructured ones) the shape of the phonon line is also distorted due to the confinement effect and stress. This makes the proper fitting of these lines quite difficult. New methods of analysis are required in order to quantify small changes observed in the registered spectral function. One possible way to improve the accuracy of the results obtained is to use a data-scatter analysis based on Phase Coherence Theory (PCT). This method was suggested for the first time in Ref. [3]. The phase coherent algorithm offers a set of new measurands such as Scatter Coherence, Scatter Closeness and Scatter Radii. Based on this method a software package named "Trace Miner" was recently developed [4,5].

The technique of data-scatter was introduced initially by McMillan and Riedel in 1999 [3] principally for the fingerprinting of tensiograph measurements, although they pointed out that the technique had more general applications. In later studies the technique was used for UV-visible spectroscopic applications to fingerprint and quantify food colour spectral data. In the light of the obviously wide number of applications of the technique McMillan et al. [4,5] introduced a

**Alan.Moore@tcd.ie; phone 3 531 608-1738; fax 3 531 677-2442; *saritha@mee.tcd.ie; http://www.mee.tcd.ie.

Opto-Ireland 2005: Optical Sensing and Spectroscopy, Byrne, Lewis, MacCraith, McGlynn, McLaughlin, O'Sullivan, Ryder, Walsh, Eds., Proc. of SPIE Vol. 5826 (SPIE, Bellingham, WA, 2005) • 0277-786X/05/$15 • doi: 10.1117/12.606758

'flexible data mining toolkit' which was developed with a view to use in various areas of spectroscopy, chromatography, and electronic instrumentation applications. The present study is the first such studies with application in the area of Raman spectroscopy. Only the scatter closeness and scatter radii parameters of the phase coherence theory have been used in this study.

## 2. THEORETICAL CONSIDERATIONS

This work has been reposed on the theory of data-scatter named phase coherence theory [3]. The algorithm customised to meet the specific need to produce scatter of data points that gives visual representations of instrumental performance has been termed Phase Coherent Data-Scatter. The term signifies the phase relationships in data-scatter between the ordinate and abscissa of the two data sets that are being studied. The phase coherent algorithm offers a set of new measurands such as scatter coherence, scatter closeness and scatter radii. According to phase coherence theory, the most significant measurement of data scatter is 'Coherence' which gives a numerical measure of the individual error components for the two principal co-ordinates.

Based on this theory, a data mining software has been developed [4,5]. The software allows the user to obtain values of coherence scatter for two different spectra (Reference and Test traces) with respect to their centroid. The centroid, computed from the well-known area-averaging algorithm, is found to be a stable point with respect to spectral variations. But this technique has a drawback that there is a possibility of two different spectra being projected to the same centroid. Hence, it is impossible to distinguish between traces using only the centroid. Exchange operator is employed as a solution to this problem [4,5]. Phase coherence theory also explains Scatter Closeness, which gives the Euclidean distance between the two centroids of the unexchanged data in the Test and Reference traces. This parameter proves to be a significant one, which provides a single measure of scatter-points. The new coherence theory gives birth to a novel interesting phenomenon, which arises from combinations of digital and propagation errors arising from the algorithm used, called micro-scatter [4,5]. This phenomenon can be explained as the pattern observed with zooming magnification of what appeared to be a single centroid. Scatter radii, which gives a quantification of the scatter, is also given importance in the new coherence theory.

Raman scattering is being widely used as a fast and non-destructive tool for composition, stress and defect analysis in semiconductor structures. During these analyses, all the characteristics of Raman spectra, *i.e.* the peak position, the peak intensity and FWHM can be used to obtain information on different properties of materials under investigation. The goal of this work is to explore the practicality of phase coherence theory to analyse the changes in Raman spectra and to establish the dependence of changes obtained in all the spectral function characteristics on the parameters of data-scatter mentioned above, using "Trace Miner" software.

Phase Coherent Data-Scatter (PCD-S) is only one of a number of algorithmic possibilities. This approach has been used here only because of convenience. The original developers of this technique are developing a commercial package for use in UV-visible spectroscopy and this application needs above all measurements of photometric and wavelength reproducibilities. PCD-S has been developed specifically to deliver a visualisation of these instrumental performance factors in the first instance, but also for other spectroscopic measurements here. The real potential of data-scatter is that algorithms can be devised to answer specific problems. In the case of Raman spectral characteristics, PCD-S is not the optimum toolbox approach and in later work the authors will be hoping to present the optimised algorithm for the study of Raman spectra for microelectronic applications. However, we believe that even at this stage of the Programme developing it is important to show the correlation existed between different characteristics of Raman spectra and Phase Coherence Theory measurands.

## 3. EXPERIMENTAL

For the purpose of data-scatter analysis, we have chosen a model oscillator with spectral characteristics close to that of longitudinal optical (LO) phonon mode of bare Si at 520 $cm^{-1}$. Raman spectra can in general be fitted best by a Lorentzian function

$$I = I_0 / [(\omega-\omega_0/w)^2+1] \qquad (1)$$

where, $\omega_0$ is the peak frequency, $I_0$ is the intensity at $\omega_0$ and w is the half width at half maximum (HWHM). A family of Lorentzian spectral functions have been modelled using variation in peak position (Raman shift), peak intensity and FWHM. The analysis of these three sets of model spectra is done in "Trace Miner" in order to extract the parameters of data scattering analysis. In such a way, we have established the dependence of each of the data scatter parameters on changes observed in Raman spectral characteristics such as

$$\text{Raman shift, } \omega_{RS} = \omega_{Test} - \omega_{Ref} \quad (2)$$

$$\text{Raman Intensity difference, } I_{RID} = I_{Test} - I_{Ref} \quad (3)$$

$$\text{Raman FWHM difference, } FW_{RFD} = FW_{Test} - FW_{Ref}. \quad (4)$$

## 4. RESULTS AND DISCUSSIONS

### 4.1 Frequency monitoring

The analysis has been carried out by modelling a set of Lorentzian spectral functions by varying the peak frequency from 520 $cm^{-1}$ to 530 $cm^{-1}$ and keeping the peak intensity constant at 10000 A.U. and FWHM at 4 $cm^{-1}$. A step of 0.1 $cm^{-1}$ from 520 $cm^{-1}$ to 521 $cm^{-1}$ and a step of 1 $cm^{-1}$ from 521 $cm^{-1}$ to 530 $cm^{-1}$ were used (Fig. 1).

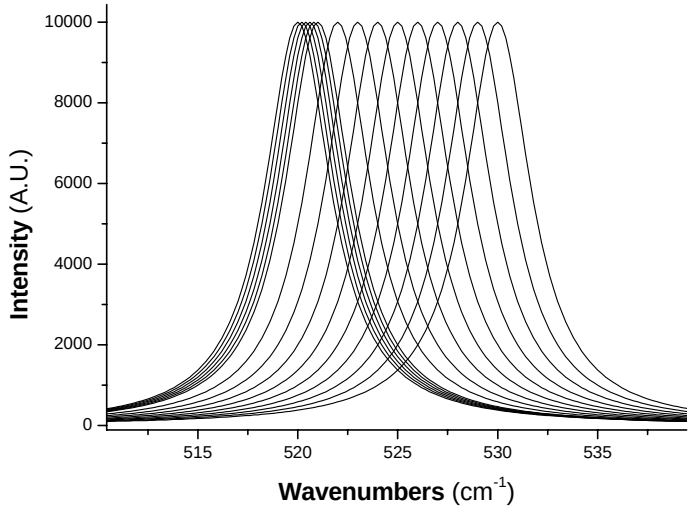


Figure 1: Set of the Lorentzian spectral functions with peak position ranging from 520 $cm^{-1}$ up to 530 $cm^{-1}$.

For the above set of model spectra, we have computed values of data-scatter parameters such as scatter coherence, scatter radii and scatter closeness using phase coherence theory by using the model curve with peak frequency at 520 $cm^{-1}$ as a reference. Fig. 2 shows the dependencies of data-scatter parameters on Raman shift ($\nu_{RS}$).

Scatter closeness gives perfect linear dependence (Fig. 2a) on Raman Shift in the whole frequency range from 520 $cm^{-1}$ to 530 $cm^{-1}$ for different step values, which is given by the equation.

$$Y = 0.989 * X \quad (5)$$

In the case of scatter radii (Fig. 2b), a linear dependence on Raman shift has been observed for frequency values ranging from 520 $cm^{-1}$ to 521 $cm^{-1}$ with a step of 0.1 $cm^{-1}$. When the shift is increased to 1 $cm^{-1}$, the data points of scatter radii for frequencies ranging from 521 to 530 $cm^{-1}$ can be best fitted by a polynomial function.

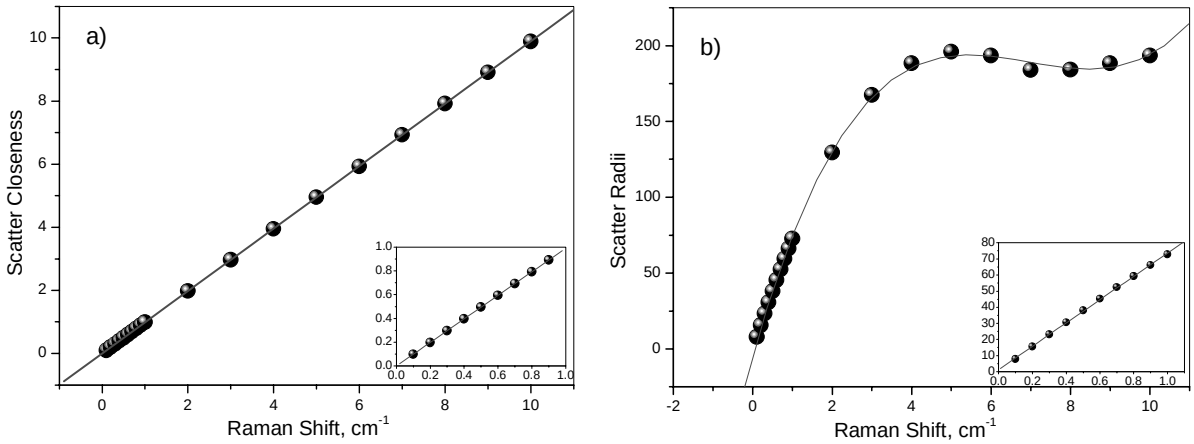


Figure 2: Dependence of Scatter Closeness (a) and Scatter Radii (b) on Raman Shift (peak frequency varied from 520 to 530 $cm^{-1}$). Inset shows the corresponding plots for frequencies ranging from 520 to 521 $cm^{-1}$ with a step of 0.1 $cm^{-1}$ respectively.

**4.2 Intensity monitoring**

A group of Lorentzian spectral functions has been modelled by varying the peak intensity from 1000 to 12000 (A.U.) initially with a step of 200 from 1000 to 2000 and then a with a step of 2000 from 2000 to 12000 A.U. (Fig. 3). Peak frequency is kept constant at 520 $cm^{-1}$ and FWHM at 4 $cm^{-1}$.

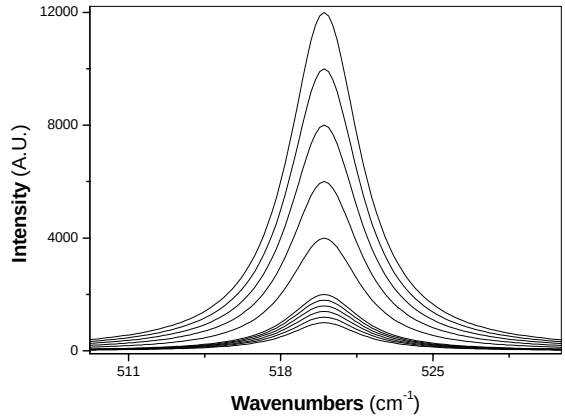


Figure 3: Set of Lorentzian spectral functions with peak intensity ranging from 1000 to 12000 A.U.

The curve with peak intensity 1000 A.U. is kept as the reference. Fig. 4 shows the correlation between the change in peak intensity and the data-scatter parameters obtained for the above set of spectral functions during analysis with "Trace Miner".

Both the data-scatter parameters *viz*. scatter closeness (Fig. 4a) and scatter radii (Fig. 4b) gives linear dependence on Raman intensity difference, which proves to be quite useful in further studies of nanostructures using Raman spectroscopy.

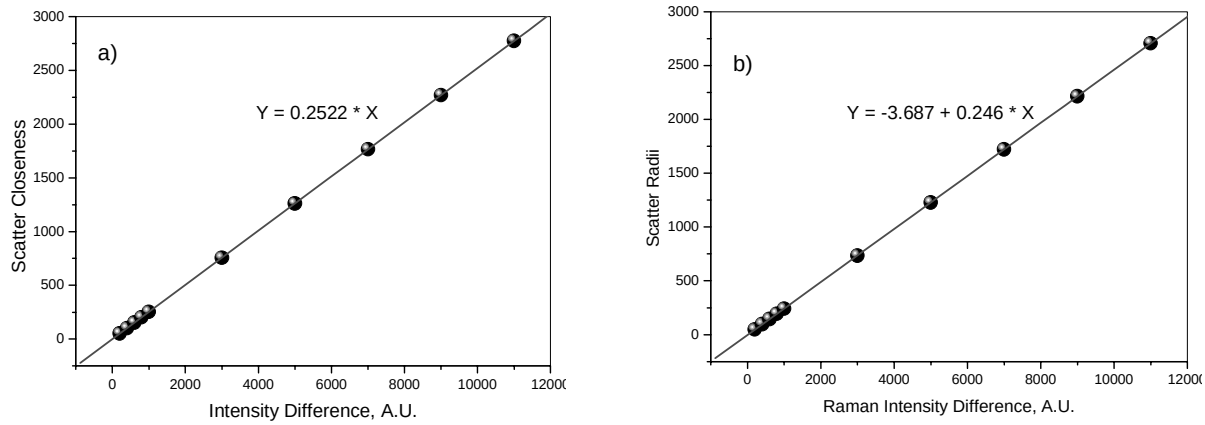


Figure 4: Dependence of Scatter Closeness (a) and Scatter Radii (b) on Raman Intensity Difference.

**4.3 FWHM monitoring**

By varying the FWHM values from 4 to 15 $cm^{-1}$ with different steps of 1 $cm^{-1}$, 2 $cm^{-1}$ and finally 3 $cm^{-1}$, a set of Lorentzian spectral functions has been created (Fig. 5). The peak position and peak intensity are kept constant at 520 $cm^{-1}$ at 10000 A.U. respectively.

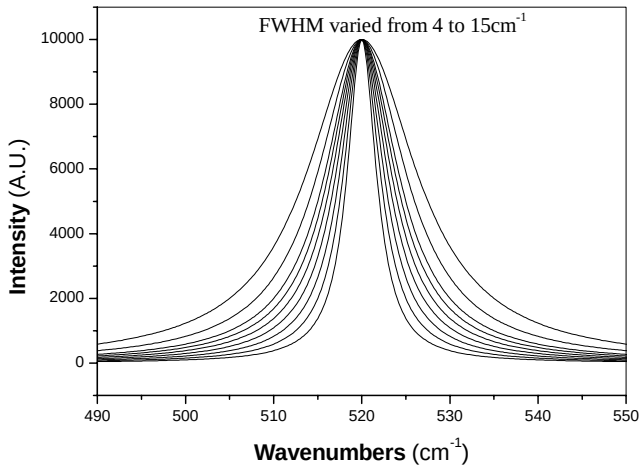


Figure 5: Lorentzian Spectral Function with FWHM varied from 4 $cm^{-1}$ to 15 $cm^{-1}$.

The dependence of data-scatter parameters on changes in FWHM of Raman spectra is shown in Fig. 6. Scatter closeness (Fig. 6a) shows perfect linear dependence, which is on line with linewidth variation, while scatter radii (Fig. 6b) values fit exactly into the sigmoid-Boltzmann function given by the following expression.

$$y = 884.898 + (-7182.081-884.898)/(1 + \exp((x+32.39)/15.403)) \quad (6)$$

However, as can be seen from Fig. 6b, the deviation of scatter radii from non-linearity is practically negligible up to FWHM difference of ~6 $cm^{-1}$.

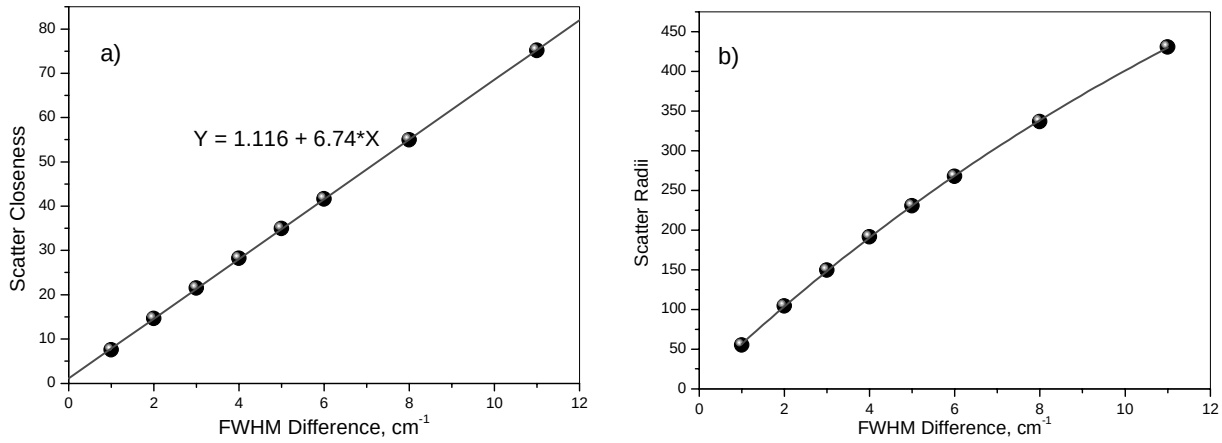


Figure 6: Dependence of Scatter Closeness (a) and Scatter Radii (b) on Raman FWHM Difference (varied from 4 to 15 $cm^{-1}$).

### 4.4 Real instrumental data

We have performed similar data scatter analysis on experimental Raman spectra of bare Si wafer. The experimental curve with peak frequency at 520 $cm^{-1}$ is taken as a reference. By shifting the reference curve from its initial position with a step of 0.2 $cm^{-1}$ and 1 $cm^{-1}$ in the regions of 520-521 $cm^{-1}$ and 521-530 $cm^{-1}$ respectively, a set of experimental Raman spectra similar to that shown in Fig. 7 have been created.

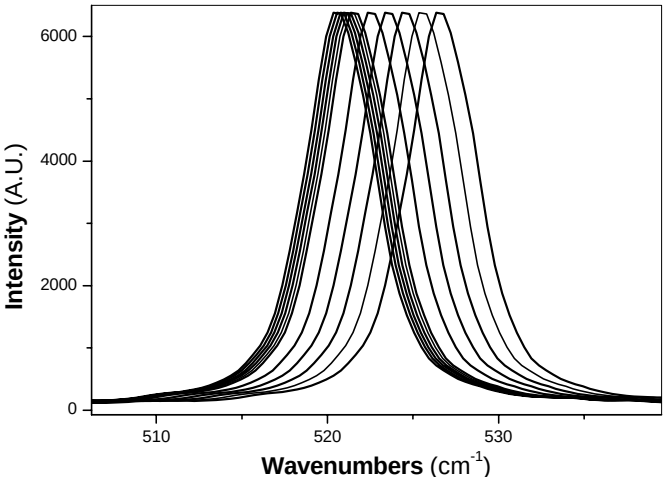


Figure 7. Experimental Raman spectra of bare Silicon at 520 $cm^{-1}$ and the test curves shifted from the experimental curve with a step of 0.2 $cm^{-1}$ and 1 $cm^{-1}$ respectively.

As a result of the analysis of these test traces with the reference trace using phase coherence theory, the dependence of scatter closeness on Raman shift has been obtained. As shown in Fig. 8, the Scatter closeness obtained from the above set of data, exhibits perfect linear dependence on Raman shift ($\omega_{RS}$ from 0.1 up to 6 $cm^{-1}$), which is very close to that obtained for model data shown in Fig. 2(a). On comparison with equation (5), the dependence of scatter closeness obtained for both model and experimental Raman data are approximately coinciding.

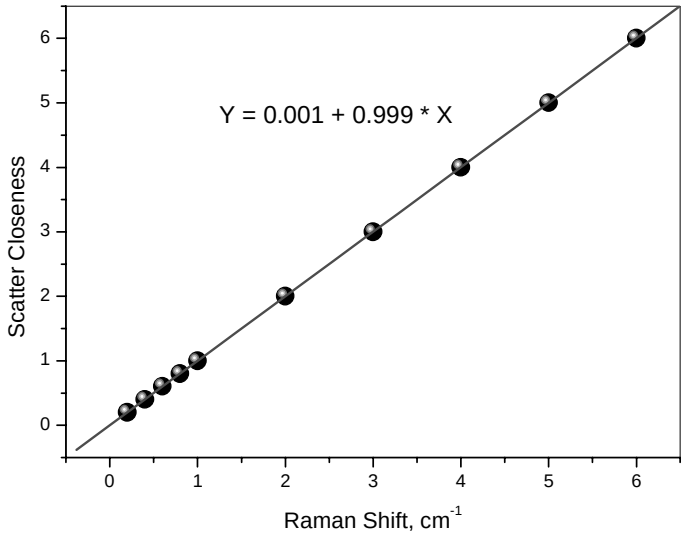


Figure 8: Dependence of Scatter Closeness on experimental Raman shift (in the range 520cm$^{-1}$ to 527cm$^{-1}$).

Based on previous analyses, scatter closeness is found to be the data-scatter parameter, which gives perfect linear dependence on all Raman spectral characteristics.

## 5. CONCLUSION

The correlation between variations in peak position, peak intensity and width of Raman scattering spectra with a number of measurands (scatter closeness and scatter radii) of "Trace Miner" toolkit based on data-scatter approach have been established. The advantage of this data-scatter approach is that it allows the user to experimentally investigate a data analysis problem and to find the most appropriate tool for the problem in hand. The paper illustrates this point very clearly as it shows that from a very large number of possible measurands that are available in the toolkit, only scatter closeness and scatter radii have been used. We found that scatter closeness gives a perfect linear dependence for all Raman spectral characteristics under investigation. This is a very promising result to use PCT for the analysis of Raman data. It is important to point out here that another novel aspect of the toolkit is that it is designed to produce measurements from statistics, rather than asking the usual statistical question as to whether or not two data sets are the same or different. The software package has been developed to provide statistical measurements based on what is in fact a comprehensive set of statistical approaches including all the parametric and non-parametric tests usually found in standard statistics books [6]. The relationship between data-scatter parameters and standard statistical measures will be investigated in the near future.

## ACKNOWLEDGMENTS

The authors thank HEA Ireland (PRTLI Grant) for partial financial support towards this work. Professor John M. Kelly is greatly acknowledged for useful discussions.

## REFERENCES


1. G. Turrell, and J Corset, *Raman microscopy*, Academic Press, (1996).

2. I. De Wolf, Micro-Raman spectroscopy to study local mechanical stress in silicon integrated circuits, *Semicond. Sci.Technol.* **11**, 139-154 (1996).

3. N.D. McMillan, et al. *Irish Machine Vision Conference Proceedings*. 330-346 (1999).

4. N.D. McMillan, , S.M. Riedel, , B. O'Rourke, J. Hammond, G. Doyle, F. Murtagh, , M. Kökűer, N. Whyte, A. O'Neill, D.G.E. McMillan, K. Beverley, A. Augousti, J. Mason, H.S Bertelsen,. and S. Asbjørnsen, A Flexible Data Mining Toolkit for UV-Visible Spectrophotometry, Tensiography and Signal Matching. *Chemometrics and Intelligent Laboratory Systems* (in press).

5. N.D. McMillan, O'Rourke, S.M. Riedel, D.O. Skelly, M. O'Neill, A.E. O'Neill, D. Boller, A.C. Bertho, G. Doyle, J. Hammond, and A.T. O'Neill, A new phase-coherent data-scatter fingerprinting and archiving technique for qualitative ultraviolet-visible spectroscopy and a new improved quantitative method for multi-component colorimetric determinations of food colours, *Analytica Chemica Acta* (In Press)

6. R.A. Johnson, and G.K. Bhattacharyya, *Statistics Principles and Methods*, John Wiley & Sons, (1992).