Circuits and Systems:An International Journal (CSIJ), Vol.2, No.1 , January 2015

TREND REMOVAL FROM RAMAN SPECTRA WITH LOCAL VARIANCE ESTIMATION AND CUBIC SPLINE INTERPOLATION

Bidaa Mortada[1], El-Sayed M. El-Rabaie, Mohamad F. El-Kordy, Osama Zahran, and Fathi E. Abd El-Samie

[1]Department of Electronics and Electrical and Communications Engineering, Faculty of Electronic Engineering, Menofia University, Menouf, EGYPT.

Abstract

Trend removal is an important problem in most communication systems. Here, we show a proposed algorithm for trend (background) removal from Raman Spectra by merging local variance estimation and cubic spline interpolation methods. We found that Raman spectrum does not need a smoothing process to remove trend from noisy signals. Employing this technique results in more speedy and noiseless systems than other techniques that use wavelet transformation to suppress noise.

Keywords

Raman spectroscopy, Background correction method, Local variance, Cubic spline interpolation.

1. Introduction

Spectroscopy is the study of the interaction between matter (a particle that has rest mass) and radiated energy. Spectroscopic data is often represented by a spectrum, a (frequency & intensity) which is the response of intensity to the frequency [1].

Raman spectroscopy is an application of spectroscopy, and it has a lot of advantages. It can be used with solids and liquids. There is no need for sample preparation. It is non-destructive, and it is acquired quickly within seconds. Also, Raman spectroscopy has some disadvantages. It cannot be used for metals or alloys. The detection process needs a sensitive instrumentation. The Sample can be destroyed by heating through the intense laser radiation, and the fluorescence of impurities in the sample itself can hide the Raman spectrum [2].

Raman spectrum is defined as a plot of the intensity of Raman scattered radiation as a function of its frequency difference from the incident radiation. Due to the existence of the background affecting the main spectrum, the detection becomes very difficult. So, it is necessary for applying a background correction method (BCM) to the spectrum before performing analysis of the spectra obtained from Raman spectroscopy [3].

In signal processing software, it is necessary to be able to distinguish noise and background from the original signal. To express this mathematically, a sampled signal can be considered as an array (S) that can be given as:

$$S = (p_s + B) + N \tag{1}$$

Where, p_s , B and N refer to the noiseless signal without background, background, and noise, respectively.

It is important to remove noise and background signals from the experimental spectrum. In most cases, it is necessary to apply a proper background correction algorithm in order to increase the effective resolution for quantitative analyses [4].

The BCM algorithms can be categorized into two major groups, based on the type of information which needs to be extracted from original signals. The first group of BCMs includes methods requiring knowledge about background, blurring effect and noise that often deal with signals by using knowledge about the signal components such as background shape, position and SNR. This category includes the noise median method [5], signal removal method (SRM) [6] and threshold-based classification (TBC) [7].

The second group of BCMs includes those requiring knowledge about frequency of signal components, as it is well known that the noise and background would have completely different characteristics, because noise is generally a high-frequency phenomenon, while background has a low-frequency component of the signal. This type of signal processing includes Fourier transform (FT) [8] and wavelet transforms (WT) method [9].

In signal removal methodology (SRM), peaks are removed from the spectrum using the derivative of the spectrum to understand the position, starting, and finishing points of these peaks. We can use continuous WT (CWT) and discrete WT (DWT) as alternative approaches to get derivatives of noisy signals [10–13].

There are several algorithms used to solve the background problem. One of these algorithms is based on combining SRM and CWT methodologies. This algorithm is developed by *Kandjanih et al. [3].* This algorithm starts with employing CWT to calculate the second derivative of the noisy signal, which identifies signal peak positions in the experimental spectrum. To remove the signal peak component of the spectrum and fit the reminiscent spectrum the SRM method to be used to find the background, which is further subtracted from the original spectrum to obtain a background corrected signal.

In this paper, we present a proposed algorithm based on using the *local variance estimation and cubic spline interpolation* to make background correction, wherein the error was found considerably as the best value reported so far for similar studies. The major advantage of the current approach is that it does not involve any smoothing step which is a major challenge in obtaining background-corrected spectra.

We organize the paper as follows. Section two includes the problem formulation. Section 3 introduces the method of solution of the trend problem. Section 4 includes the discussion of the local variance and cubic spline interpolation concepts, and then an explanation of how to generate a simulated Raman spectrum with the types of trends (linear, sigmoidal, and sinusoidal) is given in section 5. Section 6 shows the results. Finally, section 7 gives the concluding remarks.

2. Problem Formulation

We aim to remove a type of noise called trend or background from noisy signal and extract the original noiseless signal with minimum error, high performance and less data processing time. This will be applied on Raman spectrum using local variance estimation and cubic spline interpolation to detect trend, and hence make some calculations to estimate variance and apply interpolation to remove the trend.

3. Method of Solution

We present a smoothing algorithm for trend removal by carrying out some calculations using the Matlab package starting from generating simulated Raman spectrum by Matlab, adding three types of trend (linear, sigmoidal, and sinusoidal) to the original simulated spectrum, estimating local variance to determine peak values and their vicinity to remove them, and applying cubic spline interpolation in the removed regions from the spectra determine the trend and subtract it. We can summarize the proposed method in the following steps:

- Simulate Raman spectrum.
- Add one of three types of trends (linear, sigmoidal and sinusoidal trend).
- Estimate local variance to estimate the peak regions for removal.
- Apply cubic spline interpolation to interpolate in removed peak regions.
- Subtract the estimated trend without peaks from the noise spectrum.
- Estimate RMSE between an original spectral without trend and the spectrum after trend removal.

4.Local Variance and Cubic Spline Interpolation Concepts

As mentioned in theory of probability and statistics, **variance** measures how far a set of numbers is spread out. When variance is equal to zero, it indicates that all the values are identical. Also, when variance has a small value, it indicates that the data is very close to the mean (expected value) and hence to each other, and a high value indicates that the data is very spreading out around the mean and from each other [14].

We can estimate the local variance of a signal $x(n)$ as follows [15]:

$$\hat{\sigma}_x^2(n) = \frac{1}{(2k+1)} \sum_{k=n-k}^{n+k} \left(x(k) - \hat{x}(n) \right)^2 \tag{2}$$

where (2K+1) is the number of samples in the short segment used in the estimation and   is the local mean defined as:-

$$\hat{x}(n) = \frac{1}{(2K+1)} \sum_{k=n-k}^{n+k} x(k) \tag{3}$$

Interpolation is defined simply as it is an informed estimate of the unknown. It is defined also as a model based recovery of continuous data from discrete data within a known range of abscissa. In digital signal processing, interpolation is considered as a digital convolution operation. This convolution operation can be implemented using the digital filtering approach, row by row and then column by column, separately.

Spline interpolation is a type of interpolation where the interpolant is a special type of piecewise polynomial called a spline. We preferred Spline interpolation over polynomial interpolation as it has a small interpolation error even when using low degree polynomials for the spline. Also, spline interpolation avoids the problem of Runge's phenomenon, which occurs only in high degree polynomials [15]. Figure 1 shows the shape of spline.

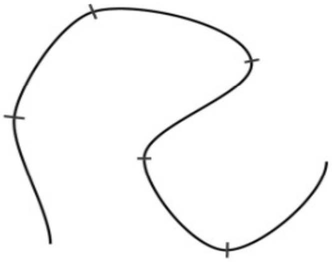

Fig.1 Spline shape

5.Generation of Simulated Spectra

We utilize in this paper Matlab for MS Windows, version 7.10 (R2010b) for the experimental steps. First, Raman peaks were simulated using a Gaussian function and it is expressed as:

$$f(x) = a.e^{\left(\frac{-0.5(x-c)^2}{\sigma^2}\right)} \tag{4}$$

where a is the intensity controller, c and σ are mean and variance of the Gaussian peak, respectively as shown in Fig 2.

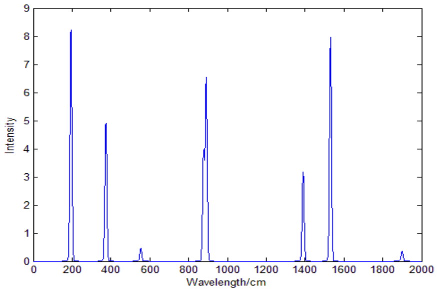

Fig. 2: Simulated Raman spectrum

For accuracy, we simulated Gaussian peaks of variable quantity with random positions, intensity and width with three trends as linear, sigmoid and sinusoid forms and variable background constants. Then we added a trend (background).

A.Linear Trend:

$$\text{Trend} = a.x+b \tag{5}$$

where, a is the slope of the linear trend and b is a constant. This is shown in fig 3.

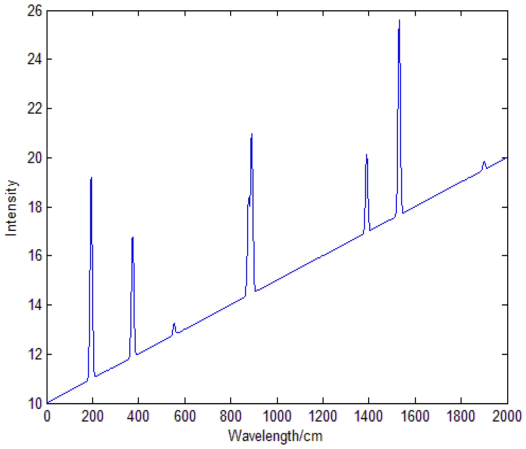

Fig. 3: Spectrum with linear trend

B.Sigmoidal Trend:

$$\text{Background} = \frac{1}{1+\exp(-a(x-c))} I + O \tag{6}$$

where a defines the gradient at the inflection point, c defines the location of the inflection point, I defines the intensity controller and O defines the offset, and this is shown in next figure.

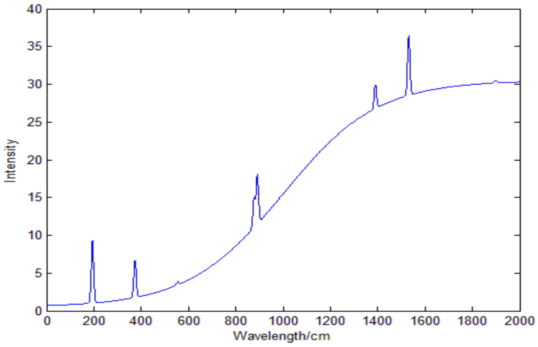

Fig. 4: Spectrum with sigmoidal trend

C.Sinusoidal trend:

$$\text{Background} = x^{1.5}\sin(\frac{x}{a}).I + O \tag{7}$$

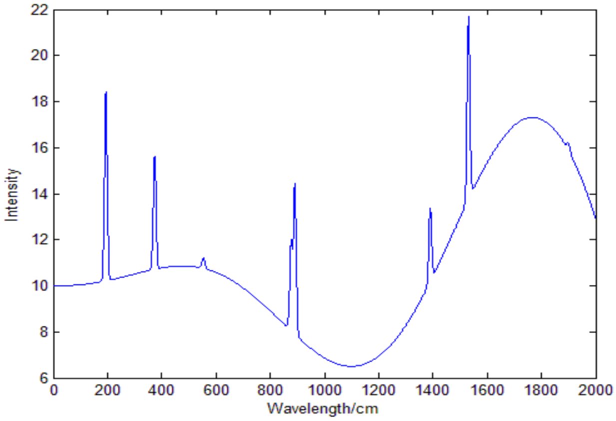

Fig. 5: Spectrum with sinusoidal trend.

6.Trend Removal Results

After these calculations, we can extract the trend from the noisy spectra.

A.Linear Trend

The estimated linear trend is shown in Fig. (6).

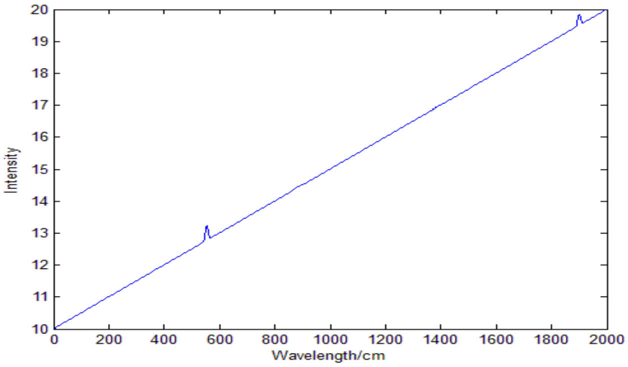

Fig 6: Estimated linear trend.

Then remove the estimated linear trend from the spectrum with linear trend fig (3), we use subtraction and get the spectrum with trend removal in Fig 7.

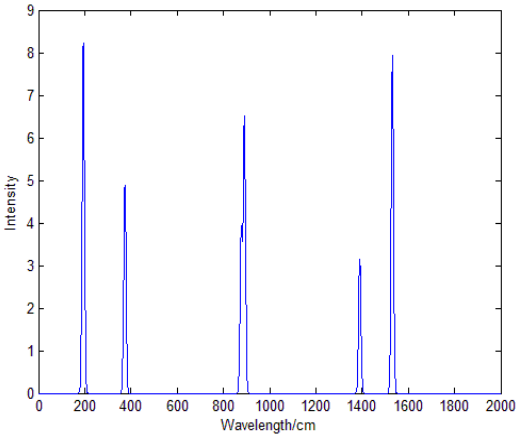

Fig7: Spectrum after trend removal for the linear case, MSE= 0.0015.

B.Sigmoidal Trend

The estimated sinusoidal trend is shown in Fig. (8).

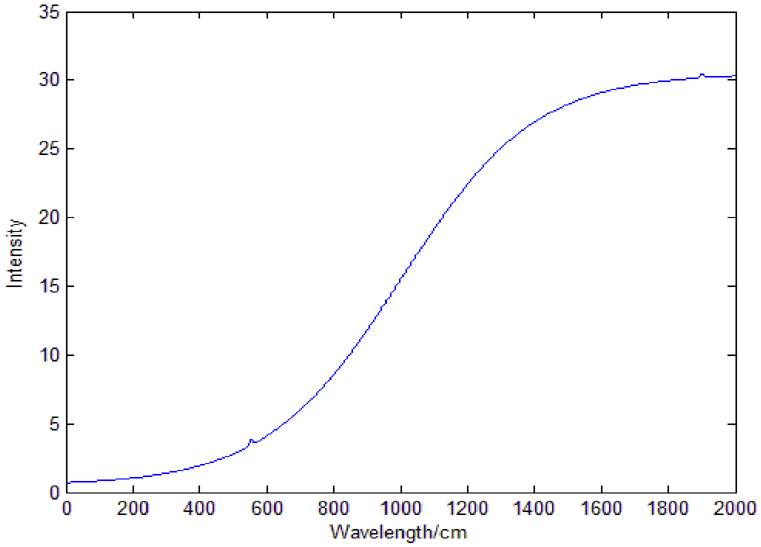

Fig8. Estimated sigmoidal trend

After trend removal, we get the spectrum in Fig 9.

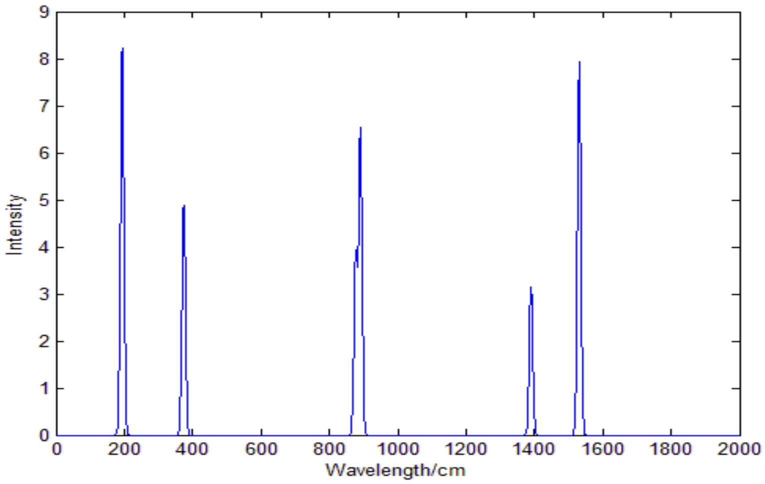

Fig9: Spectrum after trend removal for the sigmoidal case, MSE=0.001

C.Sinusoidal Trend:

The estimated sinusoidal trend is shown in Fig. (10).

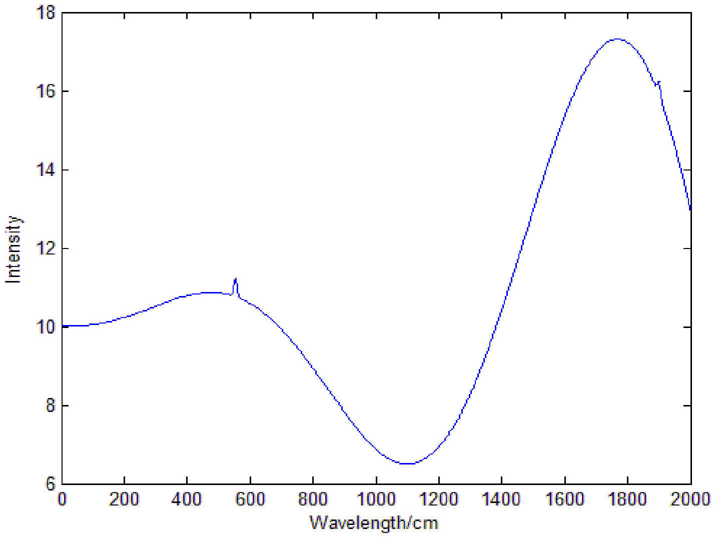

Fig10. Estimated sinusoidal trend

After trend removal, we get the spectrum in Fig 11.

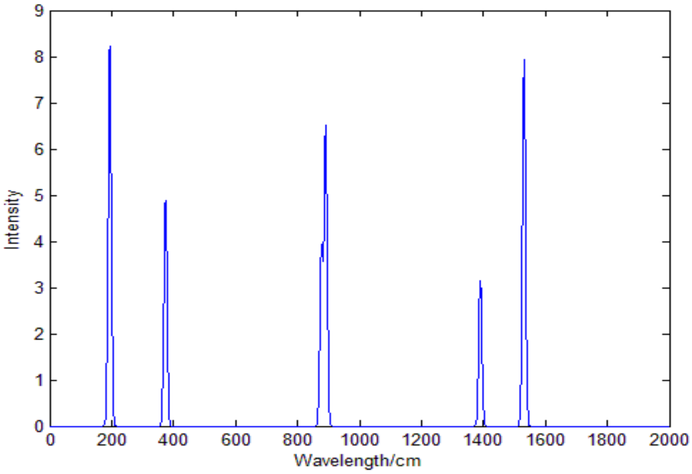

Fig11 .Spectrum after trend removal for the sinusoidal trend, MSE=0.0019

8. Conclusions

This paper presented an efficient trend removal algorithm from Raman spectra. This algorithm is based simply on local variance estimation, and cubic spline interpolation. Simulation results have revealed the success of the proposed trend removal algorithm with various types of trends and various levels of peaks in spectra. As compared to the *Kandjani's* algorithm which achieves MSE values in the range of 0.1 to 0.2, the proposed algorithm achieves MSE values in the range of 0.001 to 0.002.

References

[1] John Wiley & Sons Ltd, " MODERN SPECTROSCOPY Fourth Edition", The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England,2004.

[2] Wei-Chuan Shih1, Kate L. Bechtel2, and Michael S. Feld, "Intrinsic Raman spectroscopy for quantitative biological spectroscopy Part I: Theory and simulations" OPTICS EXPRESS 12726, (2008).

[3] Ahmad EsmaielzadehKandjani, MatthewJ. Griffin, Rajesh Ramanathan, Samuel J. Ippolito, Suresh K. Bhargavab and VipulBansala. A new paradigm for signal processing of Raman spectra using a smoothing free algorithm: Coupling continuous wavelet transform with signal removal method "Wiley Online Library: 25 January 2013DOI 10.1002/jrs.4232.

[4] G. Schulze, A. Jirasek, L. Yu, A. Lim, B. Turner, M. Blades, "investigating of selected baseline removal techniques as candidates for automated implementation "Application Spectroscopy 2005 , 59, 545.

[5] M. S. Friedrichs, J. Biomol. (A model-free algorithm for the removal of baseline) artifacts NMR 1995, 5, 147..

[6] M. A. Kneen, H. J. Annegarn, "Algorithm for fitting XRF, SEM and PIXE X-ray spectra backgrounds " Nucl. Instrum. Methods Physics. Res., Sect. B 1996, 109–110, 209.

[7] Dietrich W, RuK CH, Neumann (Fast and precise automatic baseline correction of one- and two-dimensional)NMR spectra. J MagnReson 1991;90:113.

[8] Marion D, Bax A. (Baseline correction of 2D FT NMR spectra using a simple linear prediction extrapolation of thetime-domain) data. J MagnReson 1989;83:205}11.

[9] B.-F. Liu, Y. Sera, N. Matsubara, K. Otsuka, S. Terabe,(Signal denoising and baseline correction by discrete wavelet transform for microchip capillary electrophoresis). Electrophoresis 2003, 22:28, 3260.

[10] J. R. Beattie, Optimizing reproducibility in low quality signals without smoothing; an alternative paradigm for signal processing J. Raman Spectroscopy. 2011, 40:45, 1419.

[11] L. Nie, S.W, X. Li, L. Zheng, L. Rui, J. Chem "Approximate derivative calculated by using continuous wavelet transform ".Inf. Comput. Sci. 2002, 43, 274

[12] X. Shao, C. Ma, Chemom. "A general approach to derivative calculation using wavelet transform"Intell. Lab. Syst. 2003, 69, 157.

[13] A. K. Leung, F. T. Chau, J. B. Gao "wavelet transform a method for derivative calculation in analytical chemistry" Anal. Chem. 1998, 70, 5223.

[14] Giovani Gomez "Estimating local variance for Gaussian filtering" Dept. of Computing ITESM - Campus Morelos, gegomez@campus.mor.itesm.mx ,2003.

[15] De Boor, C., "A Practical Guide to Splines, " Springer-Verlag, 1987.

AUTHORS

Bidaa mortada received the B.Sc. degree in communication engineering from Faculty of Electronic Engineering, Menoufia University, Egypt, in May 2009, and she is currently working toward the MSc degree in Electrical communication engineering. Her current research interests are in applications of spectroscopy in different media.

Prof. S. El-Rabaie (Senior Member, IEEE'1992-MIEE-Chartered Electrical Engineer) was born in Sires Elian (Menoufia), EGYPT in 1953. He received the B. Sc. degree with honors in radio communications from Tanta University, Egypt, 1976, the M.Sc. degree in communication systems from Menoufia University, Egypt, 1981, and the Ph.D. degree in microwave engineering from the Queen's University of Belfast, 1986. He was a postdoctoral fellow at Queen's (Dept. of Electronic Eng.) Up to Feb. 89. In his doctoral research, he constructed a CAD package used in nonlinear circuit simulations based on the harmonic balance techniques. He has been involved in different research areas including CAD of nonlinear microwave circuits, nanotechnology, communication systems, and digital image processing. He was invited in 1992 as a research fellow in the North Arizona University (College of Engineering and Technology) and in 1994 as a visiting Prof. in Ecole Polytechnique de Montreal (Quebec), Canada. Prof. El-Rabaie has authored and co-authored more than 120 papers and technical reports, fifteen books under the titles (Computer Aided Simulation and Optimization of Nonlinear Active Microwave Circuits, The Whole Dictionary for The Computer and the Internet Terminologies, Basics and Technologies of Data Communications in Computer Networks, Technologies and Internet Programming, The Distance Learning and its Technologies on the Third Millennium, Computer Principles and Their Applications in Education, Software Engineering (1), Management of Computer Networks(1,2), Advanced Internet Programming, Data-base Principles, Building of Compilers, Software Engineering (2), Ethics of Profession). In 1993, he was awarded the Egyptian Academic Scientific Research Award (Salah Amer Award of Electronics) and in 1995, he received the Award of the Best Researcher on (CAD) from Menoufia University. He has participated in translating the first part of the Arabic Encyclopedia. Now, he is a professor of Electronics and Communications Eng., Faculty of Electronic Engineering, Menoufia University.

Mohammad Elkordy received the B.Sc. (Honors), M.Sc., and PhD. from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1979, 1985, and 1991, respectively. He joined the teaching staff of the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1991. His current research areas of interest include SAW applications, radiation applications, image processing, and signal processing.

Osama zahran received the B.Sc. (Honors), M.Sc. from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1997, 1999 respectively, and the Ph.D. from Liverpool University, UK. He joined the teaching staff of the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt. He is a co-author of about 29 papers in national and international conference proceedings and journals. His current research areas of interest include Nano-scale devices, expert systems, artificial intelligence and hybrid intelligent systems.

F. E. Abd El-Samie was born in Tanta, Egypt, on May 12, 1975. He received the B.Sc. degree in communication engineering from Faculty of Electronic Engineering, Menoufia University, Egypt, in May 1998, the MSc. in Electrical Communications, Faculty of Electronic Engineering, Menoufia University, 2001. PhD. in Electrical Communications, Faculty of Electronic Engineering, Menoufia University, 2005. He has received the most cited paper award from Digital Signal Processing Journal in 2008 for the paper entitled: "Efficient Implementation of Image Interpolation As An Inverse Problem", authored and co-authored more than 120 papers and 2 Books, interested in Image Processing: (Enhancement of old images and images acquired under bad illumination conditions, restoration of degraded images, restoration of degraded and noisy images, multi-channel image processing, image interpolation and resizing, super resolution reconstruction of images, color image processing, image watermarking, encryption, and data hiding) , Signal Processing: (Spectral Estimation, Wavelet Processing, Signal Separation, and Speech Processing) and Digital Communications (CDMA, OFDM, Dynamic Spectrum Management, Channel Equalization and Channel Estimation).