



Novel Curve-Fitting Algorithm Based on Particle Swarm Optimization for Optical Spectrum Analysis

Chuttima Wongpa, Atimon Teerasoradech, Suvicha Sasivimolkul, Chayanisa Sukkasem, Phitsini Suvarnaphaet and Suejit Pechprasarn*

College of Biomedical Engineering, Rangsit University, Pathum Thani, Thailand

* Corresponding author E-mail: suejit.p@rsu.ac.th

Abstract

A spectrophotometer is a scientific instrument used to separate and measure spectral components of biochemical compounds and molecules, which are related to the amount of light intensity absorbed by the substance. The Spectrometric system consists of a light source, optical lenses, monochromators, and detectors. The monochromators are required to separate optical wavelengths using refraction through a prism or diffraction through a diffraction grating. In this paper, the authors demonstrate a novel curve-fitting algorithm based on particle swarm optimization (PSO) for fitting a spectral response with 5 Gaussian distributions. There are 15 variables in total describing the Gaussian equations. To demonstrate the capability of the proposed algorithm, the spectra of samples including the DI water, ethanol, FeCl₂, and the plasmonic gold sensor were collected by UV/Visible spectrophotometer operated with SP-8001 of Metertech for testing the PSO algorithm. As a result, the Gaussian functions based on the PSO model showed the most predictive and well-fitting with the plasmonic gold sensor where the spectrum matched to the single peak at 527.44 nm with 0.9987 R². The peak performed specific modeling equations of peaks and valleys related to its spectral nature mostly using in plasmonic applications.

Keywords: *Spectrophotometer, Monochromator, Particle swarm, Curve fitting, Gaussian distributions, Biochemical sensing*

1. Introduction

A spectrophotometer is an instrument measuring light absorption of a chemical substance or solution as a function of varying incident optical wavelengths. Organic substances or complex compounds have different optical spectra in terms of reflectance, transmittance, and absorbance (Richards-Kortum & Sevcik-Muraca, 1996). These optical quantities are related to the amount of light intensity absorbed by the substance. Optical spectroscopy is a gold standard for quantitative biochemical measurement allowing mixture compounds and their concentrations to be determined. The spectrum response can be used as a fingerprint for identifying each type of chemical substance (R. A. Lewis et al., 1980).

There are multiple wavelength ranges used in spectroscopy ranging from X-ray (0.01 nm -10 nm), ultraviolet (4 nm - 400 nm), visible light (400 nm - 700 nm), and near-infrared light (700 nm - 2500 nm). A selection of wavelength range depends on the nature of molecular vibration, translation, and rotation of the chemical molecule (McDonald, 1979). A typical spectrophotometer system consists of a light source, optical lenses, monochromator to separate each wavelength of the light source using refraction through a prism or diffraction through a diffraction grating and photodetectors (Fastie, 1952).

Suppose that there is a database for an optical spectrum of different chemical compounds and concentrations available. This database can be employed to trace back the compounds and separate their concentrations by curve-fitting the experimental spectral response from a spectrophotometer to the database. The process involves a curve-fitting the spectrum to identify the local peaks, local minima and maxima of the range, and denoising the spectrum (Islam, Moniruzzaman, Misran, & Shakib, 2009). There are two types of curve-fitting methods, which are direct search and random search. The direct search is the most common method of analysis that demonstrates nonlinear optimization. This method solves the problem by substituting every variable, and it is possible to answer certainly for most cases. However, this method has a few limitations. Firstly, the user's requirements are minimal, and the algorithms themselves require the setting of a few parameters. Secondly, it is not suitable for complex optimization problems with many parameters that cannot finish it manually. Thirdly, it is time-consuming when solving distant problems (R. M. Lewis, Torczon, & Trosset, 2000). The random search is another quick method to carry out the problems. This



method is simple and can random variables to represent values suitable for nonlinear optimization and complex optimization. However, the answer may not be precise (Karnopp, 1963). For this reason, particle swarm optimization (PSO) was chosen for this research. This paper proposes a novel curve-fitting method based on the PSO algorithm (Kennedy & Eberhart, 1997). The PSO algorithm has been put forward for solving complex optimization with the multi-model such as complex equation and the multidimensional problem with discrete or integer variables (Ahmad Rezaee Jordehi, Jasni, Wahab, & Abd Kadir, 2013). The curve-fitting ability of PSO was robust compared with the direct search method, and it is less time-consuming (A Rezaee Jordehi, 2015). An optical spectrophotometer with the optical wavelength range of 190 nm to 1100 nm was employed to measure the spectral responses is 46 nm gold-coated glass thin film. The proposed curve-fitting algorithm was validated using the experimental spectrum from the Metertech SP-8001 UV/Visible spectrophotometer.

2. Objectives

- 1) To collect the spectral data from the spectrophotometer.
- 2) To create a particle swarm optimization model to analyze spectrum from a plasmonic gold sensor and curve-fit the spectral data to 5 Gaussian equations.

3. Materials and Methods

3.1 The design of the control system through the program

Once the transmittance is obtained, the transmittance is then analyzed using curve-fitting with 15 variables. These 15 variables were expressed in a summation of 5 Gaussian distributions as shown in Equation (1) to Equation (6).

$$G_0(\lambda) = A_0 \exp(-(\lambda - \lambda_0)/2\sigma_0^2) \quad (1)$$

$$G_1(\lambda) = A_1 \exp(-(\lambda - \lambda_1)/2\sigma_1^2) \quad (2)$$

$$G_2(\lambda) = A_2 \exp(-(\lambda - \lambda_2)/2\sigma_2^2) \quad (3)$$

$$G_3(\lambda) = A_3 \exp(-(\lambda - \lambda_3)/2\sigma_3^2) \quad (4)$$

$$G_4(\lambda) = A_4 \exp(-(\lambda - \lambda_4)/2\sigma_4^2) \quad (5)$$

$$\text{Model}(\lambda) = G_0(\lambda) + G_1(\lambda) + G_2(\lambda) + G_3(\lambda) + G_4(\lambda) \quad (6)$$

where G_0 to G_4 is the Gaussian distribution functions
 A_0 to A_4 is the peak amplitude of the Gaussian distribution functions
 λ_0 to λ_4 is the central wavelengths of the Gaussian distribution functions
 σ_0 to σ_4 is the standard deviations of the Gaussian distribution functions
 Model is the summation of the 5 Gaussian distribution functions

The model was then evaluated with the 15 variables to predict a spectral response. The curve-fitted spectral response accuracy for each set of the 15 variables was assessed using residue R^2 , comparing the experimental spectrum with the modeled spectrum.

Of course, the Equation (6) optimum solution occurs when R^2 is zero or least square problem. In other words, it perfectly matched the experimental data and the modeled data in the absence of noise. However, there are two main issues here. Firstly, solving a nonlinear equation of 15 variables to determine the lowest value of R^2 is tedious and time-consuming for direct search (R. M. Lewis et al., 2000) and random search algorithms (Ali, Törn, & Viitanen, 1997). Secondly, in the presence of noise like in the experimental condition, the R^2 value cannot be zero, and it leads to an optimization problem to find the lowest R^2 value.

Here in this paper, the particle swarm algorithm, an iterative method, has been employed to optimize for the lowest R^2 values and determine the corresponding 15 variables. The reason to use 5 Gaussian distribution functions is that the spectrum's nature usually has peaks and dips on a continuous curve. The limited number of Gaussian distribution functions also allows us to filter out the high-frequency components, hence noise filtering (Shi & Eberhart, 1998).

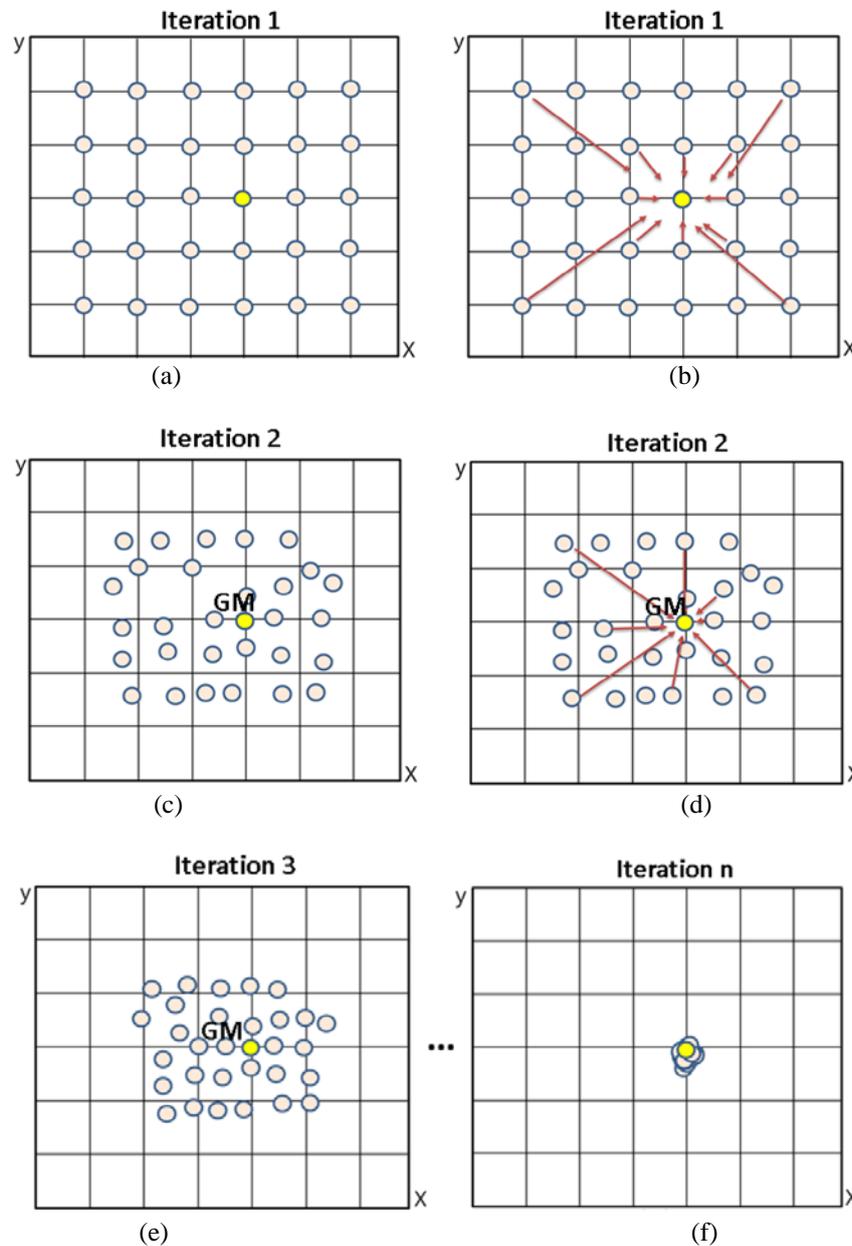


Figure 1 shows (a) particles' positions for the 1st iteration (b) calculated particles' velocities using Equation (8) for the 1st iteration (c) particles' positions for the 2nd iteration (d) computed particles' velocities using Equation (8) for the 2nd iteration (e) particles' positions for the 3rd iteration, and (f) particles' positions for the nth iteration

3.2 Find the minimum by PSO

$$\text{Objective function} = R^2 = \sum(\text{Measured signal} - \text{Model})^2 \quad (7)$$

where the objective function or R^2 is the square of the difference between the value of the measured signal and the model,



The measured signal is the spectral data from the spectrophotometer, and

The model is the modeled value of the 5 Gaussian distribution functions.

Here, the swarm size for optimization was set to 10,000 particles, and the iteration was assigned to the maximum iteration of 1,500.

To explain how the PSO works, let's consider a simplified scenario with an objective function comprising two variables: x and y , assuming there were 30 particles, six by five on the XY grid as depicted in Figure 1(a). The PSO operates by:

(1) On the first iteration, velocities of all the particles were set to V_0 of 0; in other words, there was no moving particle in the first round. Then, the objective function values are evaluated corresponding to all the independent variables on the grid. In this example, 25 (x,y) coordinates are corresponding to the grid position. For demonstration propose, assuming the minimum objective function occurred at the particle highlighted by the yellow color in Figure 1(a). The value corresponding to the minimum is set as the global minima (g) and the local minimum (p).

(2) The program will then work out the velocities of each other particles apart from the particle stationed at the current global minimum value, which will not be moved in the next iteration. The velocities are calculated using Equation (8), as shown in Figure 1(b). It is essential to point out that based on Equation (8), the particles whose position is further away from the global minimum or the local minimum will move at a greater velocity in the iteration, hence longer distance. On the other hand, the particles with a shorter length to the local minimum and the global minimum will move at a slower velocity in the next iteration, hence a shorter distance (Adewumi & Arasomwan, 2015). The iteration is depicted in Figure 1(b).

$$V_{t+1} = V_t + r_1 \text{Dist}(p, (x, y)) + r_2 \text{Dist}(g, (x, y)) \quad (8)$$

where V_{t+1} is the particles' velocity for the next iteration,
 V_t is the particles' velocity for the current iteration,
 r_1 and r_2 are the learning rates,
 $\text{Dist}((x_1, y_1), (x_2, y_2))$ is the displacement between two points (x_1, y_1) and (x_2, y_2) ,
 g is the current global minima, and
 p is the current local minima.

(3) Upon entering the next iteration, the positions of the particles are then updated using Equation (9) with its velocity as described in Equation (8) (Barrera, Álvarez-Bajo, Flores, & Coello Coello, 2016) as shown in Figure 1(c).

$$(x_{t+1}, y_{t+1}) = (x_t + V_t, y_t + V_t) \quad (9)$$

where x_{t+1} is the particles' x position for the next iteration,
 x_t is the particles' x position for the current iteration,
 y_{t+1} is the particles' y position for the next iteration,
 y_t is the particles' y position for the current iteration, and
 V_t is the particles' velocity for the current iteration.

(4) The algorithm then evaluates the objective function values corresponding to the position of all the particles. Here, if an objective function value is less than the global minimum (g), the global minimum is updated with the lower objective function value, and the previous global minima are now kept as the local minima (p) instead. Otherwise, the local minima stay the same, and the local minima are updated with the second value to the lowest objective function value.

(5) The steps (2) to (4) are repeated until all particles reach the same position as depicted in Figure 1(c) to Figure 1(f).



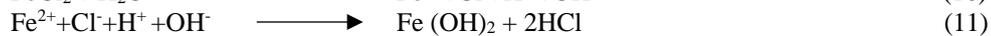
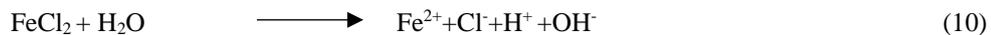
Here, the PSO algorithm operates by determining the global minimum of 15 variables starting from a 15-dimensional grid, which, of course, cannot be visualized in 3D axes.

3.3 Gold coated glass slide

The authors fabricated a gold-coated glass slip using ultra-high vacuum electron beam evaporation operated with Syskey Co., Ltd at Nanophotonics Research Center of Shenzhen University. Firstly, glass slips of Fisher Scientific were prepared by washing in acetone and isopropyl alcohol by sonication each for 5 min, rinsing with deionized water, and then blow-drying with nitrogen gas. Gold with 99.9999% purity was deposited on the glass slips with 0.17 mm thick by electron beam evaporator at a controlled pressure of 10^{-6} mbar to obtain the gold thin film with 46 nm thick on the glass slip. The samples were mounted on a rotating holder to ensure that the thin film was deposited uniformly. Herein, the gold-coated glass slip used as the plasmonic gold sensor was measured Ultraviolet-visible (UV/Visible) transmittance (%) by the Metertech, SP-8001 model spectrophotometer in a range of 300-1100 nm of wavelength.

3.4 Synthetic Ferrous Chloride ($FeCl_2$)

Synthesis $FeCl_2$ for comparing the efficiency of PSO when curve-fitting with another example is prepared by adding 25 mg of $FeCl_2$ (198.81 g/mol) to 10 ml of deionized water (DI water) (10). After that, the solution is mixed by high-frequency sonication for about 15 min (11). The solution was then measured its UV/Visible transmittance by the Metertech, SP-8001 model spectrophotometer in a range of 300-1100 nm of wavelength.



4. Results and Discussion

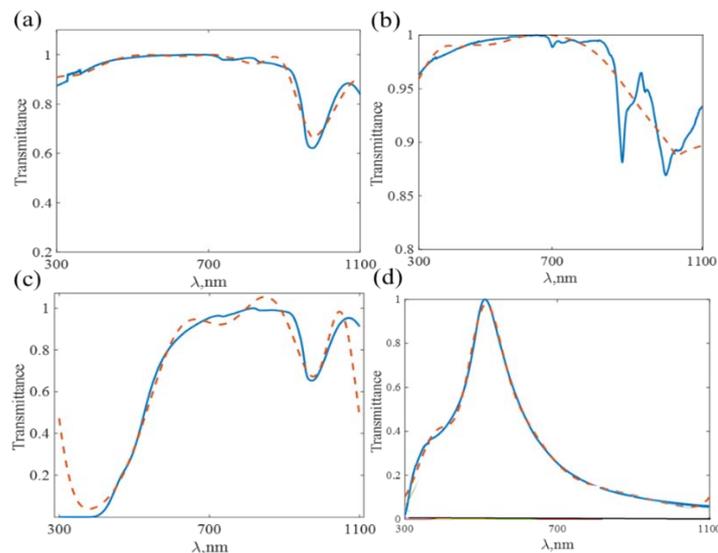


Figure 2 shows a transmittance (blue line) and curve-fitting by PSO (red dash line) of a) DI water, b) ethanol, c) $FeCl_2$, and d) gold-coated glass thin film.

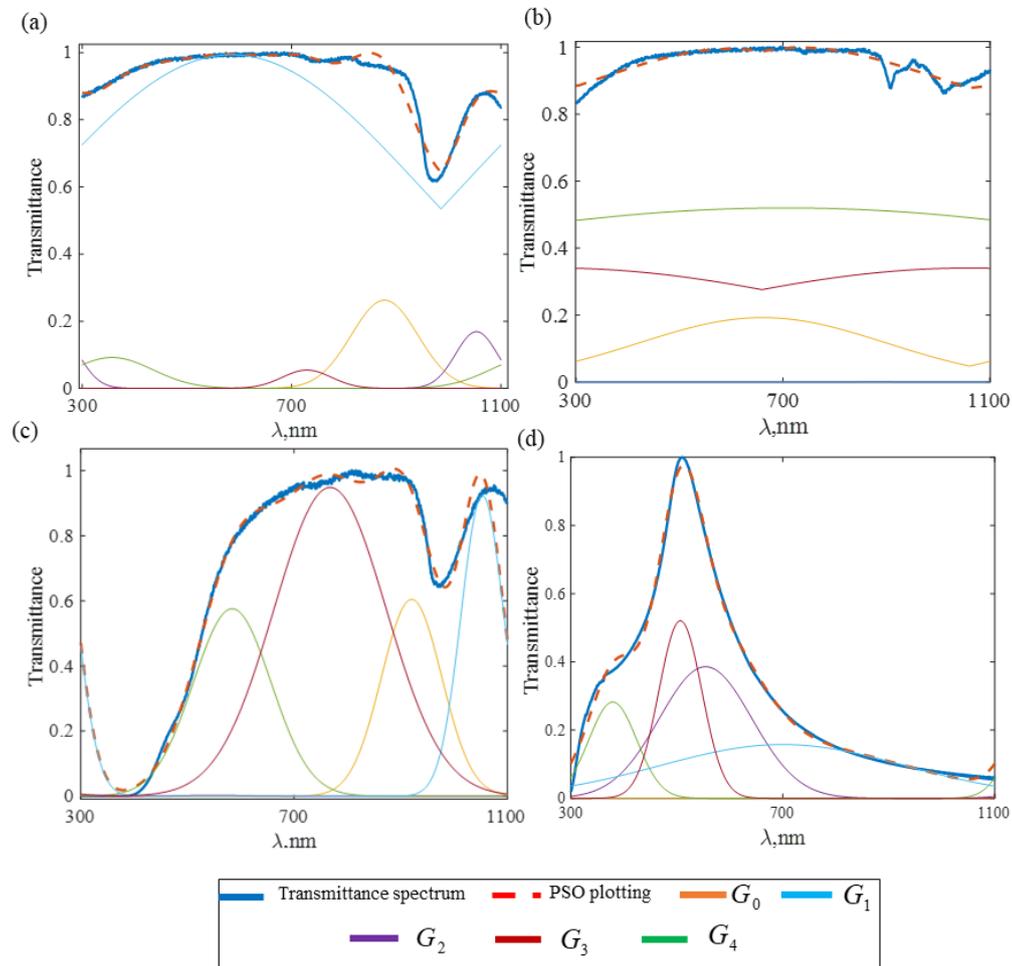


Figure 3 shows curve-fitted results obtained from the proposed particle swarm-based curve-fitting algorithm of a) DI water, b) ethanol, c) FeCl₂, and d) gold-coated glass thin film, respectively.

The results shown in Figure 2 were the transmittance of DI water, ethanol, FeCl₂, and 46-nm-thick gold-coated glass thin-film measured by the spectrophotometer. Firstly, the minimum characteristic dip of DI water was at 902 nm with 61% transmittance. Secondly, the characteristic dips of ethanol were at two positions, 908 and 1012 nm, about 95% transmittance. Thirdly, the minimum characteristic dip of FeCl₂ was 980 nm with 60% transmittance, and the maximum characteristic peak of 46-nm-thick gold-coated glass thin film was at 527.44 nm with 25.28% transmittance. The proposed particle swarm-based curve-fitting algorithm analyzed this sample spectrum. The corresponding fitted curves are shown in Figure 2. It can be observed in Figure 2 that the proposed particle swarm-based curve-fitting algorithm can be employed to model and denoise the experimental data (blue line) compared with the fitting curve (red dash line). The sum of the 5 Gaussian distributions is shown in Figure 3. The results show that the value of the square of residue R^2 in the gold sample gives a better result than the other samples at 0.9987.

5. Conclusion

This paper has demonstrated that the particle swarm-based curve-fitting algorithm can be employed to optimize a large-scale optimization problem. The UV-VIS spectrophotometer was used to measure DI



water's transmittance, ethanol, FeCl_2 , and the gold-coated glass thin film. The authors have described in detail the procedures and explained how the proposed algorithm works. The proposed particle swarm-based curve-fitting algorithm analyzed the gold-coated glass thin film's spectral responses. It was shown that the proposed algorithm can preserve all the main spectral features and successfully denoise the curve.

6. Acknowledgements

The authors would like to acknowledge the Biophysics and Medical Optic laboratory, College of Biomedical Engineering, for supporting the research. We also would like to thank the Research Institute of Rangsit University.

7. References

- Adewumi, A. O., & Arasomwan, A. M. (2015). Improved particle swarm optimizer with dynamically adjusted search space and velocity limits for global optimization. *International Journal on Artificial Intelligence Tools*, 24(05), 1550017.
- Ali, M. M., Törn, A., & Viitanen, S. (1997). A numerical comparison of some modified controlled random search algorithms. *Journal of Global Optimization*, 11(4), 377-385.
- Barrera, J., Álvarez-Bajo, O., Flores, J. J., & Coello Coello, C. A. (2016). Limiting the velocity in the particle swarm optimization algorithm. *Computación y Sistemas*, 20(4), 635-645.
- Fastie, W. G. (1952). A small plane grating monochromator. *JOSA*, 42(9), 641-647.
- Islam, M. T., Moniruzzaman, M., Misran, N., & Shakib, M. (2009). Curve fitting based particle swarm optimization for UWB patch antenna. *Journal of Electromagnetic Waves and Applications*, 23(17-18), 2421-2432.
- Jordehi, A. R. (2015). Enhanced leader PSO (ELPSO): a new PSO variant for solving global optimisation problems. *Applied Soft Computing*, 26, 401-417.
- Jordehi, A. R., Jasni, J., Wahab, N. I. A., & Abd Kadir, M. Z. A. (2013). Particle swarm optimisation applications in FACTS optimisation problem. Paper presented at the 2013 IEEE 7th international power engineering and optimization conference (PEOCO).
- Karnopp, D. C. (1963). Random search techniques for optimization problems. *Automatica*, 1(2-3), 111-121.
- Kennedy, J., & Eberhart, R. C. (1997). A discrete binary version of the particle swarm algorithm. Paper presented at the 1997 IEEE International conference on systems, man, and cybernetics. Computational cybernetics and simulation.
- Lewis, R. A., Austen, K. F., Drazen, J. M., Clark, D. A., Marfat, A., & Corey, E. (1980). Slow reacting substances of anaphylaxis: identification of leukotrienes C-1 and D from human and rat sources. *Proceedings of the National Academy of Sciences*, 77(6), 3710-3714.
- Lewis, R. M., Torczon, V., & Trosset, M. W. (2000). Direct search methods: then and now. *Journal of computational and Applied Mathematics*, 124(1-2), 191-207.
- McDonald, J. (1979). Creation and Disposal of Vibration Energy in Polyatomic Molecules. *Annual Review of Physical Chemistry*, 30(1), 29-50.
- Richards-Kortum, R., & Sevick-Muraca, E. (1996). Quantitative optical spectroscopy for tissue diagnosis. *Annual Review of Physical Chemistry*, 47(1), 555-606.
- Shi, Y., & Eberhart, R. (1998). A modified particle swarm optimizer. Paper presented at the 1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360).