Cascaded Binomial Filter Algorithms for Noisy FTIR Spectra

Yessi Jusman[1] and Nor Ashidi Mat Isa[1], *Member*, IEEE

[1,1] Imaging and Intelligent System Research Team (ISRT)
School of Electrical &Electronic Engineering
Universiti Sains Malaysia (USM), Engineering Campus
14300 Nibong Tebal, Penang, Malaysia
anisah_rabbani02@yahoo.co.id; ashidi@eng.usm.my

Abstract

In this paper, a new algorithm has been developed for denoising a spectrum using cascade form filter. The algorithm is based on cascading binomial filter with other direct form filters to increase signal to noise ratio (S/N) of the spectrum. The proposed algorithm consists of three types namely, Cascaded Binomial-Savizky Golay (CBSG), Cascaded Binomial-Triangular (CBT), and Cascaded Binomial-Rectangular (CBR) filters. Different levels of additive white gaussian noises (AWGN) are added to cervical cell FTIR spectra (i.e. as case study) and processed it using the proposed algorithms. The algorithms are proven to minimize the effects of the AWGN and increased S/N as compared to other compared smoothing filters.

1. Introduction

Fourier transform infrared (FTIR) spectroscopy has been applied in numerous fields. Recently, this technique has been used to study the structural changes of cells at the molecular level in various human cancers [1], [2]. The advancement in science and technology has led to the development of the automated computer-assisted human cancers [3]. The systems are specifically designed to assist pathologists during the crucial screening procedures.

However, a measured spectrum of the FTIR often consists of spectral line, which has a certain line shape, and white noise. According to Quintero *et al* [4], during acquisition process using the FTIR spectrometer, noise may affect the spectrum more than once. The noise could disturb spectral analysis for classification of the cancers. Thus, smoothing process is an important part to suppress the noise to improve the poor appearance of spectrum before analyzing and interpreting the spectrum.

Savitzky-Golay (SG) filter is generally used for smoothing and differentiation of the spectroscopy spectra [5], [6], [7]. Almost every spectroscopic software package contains a standard smoothing named 'Savitzky-Golay smoothing'. The filter used least-squares polynomial (LSP) method to calculate filter coefficients [8]. The calculated coefficients of the SG filter could be used directly by applying a convolution between spectrum with the calculated coefficients [5]. However, Čupek *et al* [6] found that the SG smoothing algorithm caused additionally false negative signals at shoulders of each vibrational band. In addition, the SG smoothing algorithm can lead to the loss of weak signals and the reduction of spectral resolution [9].

The binomial smoothing filter [9], [10] is another type of filters for smoothing spectra, so called because the ($2Np + 1$) coefficient is defined by binomial coefficient formulae. The coefficient filter fulfills the conditions of a smoothing technique [10]. In addition, this filter is easy to implement in hardware and is as fast as the LSP method [9]. Meanwhile, according to Battistoni *et al* [9], this filter is proven of having better performance than the SG filter when the filter width does not exceed 17 points. Therefore, denoising based cascaded binomial with other direct form filters are proposed to increase S/N of the SG and the binomial filters.

2. Theory

Smoothing of a spectrum $H(f)$ is an operation where the spectrum is convoluted with a smoothing function $W(f)$, known as window function, in order to reduce the rapidly oscillating random noise of the data. A spectrum with noise can be written as $H(f) = H_0(f) + N(f)$, where $H_0(f)$ is the original spectrum and $N(f)$ is the white noise spectrum [11]. The smoothing operation in the frequency domain can be expressed as:

$$H'(f)=W(f)*H(f) \tag{1}$$

where $H'(f)$ is the smoothed spectrum.

According to Kauppinen *et al* [11], the aim of smoothing process is to enhance the S/N by reducing the noise as much as possible, but distorting the true spectral line shape as small as possible.

The smoothing process is assumed to be convolution of the experimental data sequence by filter coefficients. The smoothed output data Y_j is given by:

$$Y_j = \sum_{k=-N_p}^{N_p} b_k X_{j+k} \tag{2}$$

where X_j is a given equally spaced data sequence to be smoothed and b_k is the filter coefficients. The total number of filter coefficients $2N_p+1$ is known as the filter width.

One of the simplest ways to smooth fluctuating data is by a moving average, known as rectangular filter, where the coefficient is equal to $b_k = 1/(2N_p+1)$ [9]. However, according to [8], for many types of data, the set of all 1's, which yields the average, is not particularly useful. For example, during implementation on a sharp peak, the average would tend to degrade the end of the peak [8].

According to [9], the triangular window is the simplest method to be implemented as illustrated by Equation 3:

$$b_k = 1 - \left| \frac{i}{N} \right| \tag{3}$$

where i ranges from $-n$ to n and $N = 2n+1$.

The SG filter is one of the best smoothing filters for spectral data. The ($2Np + 1$) coefficient of SG filter is obtained in classical paper in 1964. Detail of SG filter could be found in [8].

Another group of filters comprises the binomial filters, so called because the ($2Np+1$) coefficient is defined by the binomial coefficient formulae

$$b_k = \frac{\begin{pmatrix} 2N_p \\ N_p + k \end{pmatrix}}{4^{N_p}} = \frac{\frac{2N_p!}{(N_p+k)!(2N_p - N_p - k)!}}{4^{N_p}}$$

$$= \frac{2N_p!}{(4^{N_p})(N_p+k)!(N_p-k)!} \quad (k = 0, 1,.., N_p) \tag{4}$$

For $Np = 1$ the sequence is 1/4, 1/2, 1/4, and for $Np = 2$ the sequence will be 1/16, 1/4, 3/8, 1/4, 1/16.

The smoothing performance of the filter function depends on its width. In fact, an increase in the width corresponds to an increase in the smoothing action. Using a filter with very large width results in broadening of the peaks, i.e. a reduction of the energy resolution of the spectrum [9]. In this paper, 5 points of filter coefficient are used as a limit of the filter width. Based on the aforementioned explanation, Table 1 shows the filter coefficients for different smoothing filters using five data points.

Table 1. Filter coefficients for different smoothing filters using five data points

Filters	Coefficients
LSP degree 2 (SG filter)	[-3, 12, 17, 12, -3]/35
LSP degree 1 (Rectangular filter)	[1, 1, 1, 1, 1]/5
Binomial filter	[1, 4, 6, 4, 1]/16
Triangular filter	[1, 2, 3, 2, 1]/9

3. Methods

3.1. Fourier Transform Infrared (FTIR) spectrum

In this study, FTIR spectra of cervical cells are used to test the proposed smoothing algorithms. These spectra are obtained by scanning the cervical cells from ThinPrep® solution with FTIR spectrometer. These ThinPrep® solutions are obtained from Gribbles Pathology Laboratory, Petaling Jaya, Selangor, Malaysia as in our previous work [12].

Noise are unwanted component of measured signal. They may originate from electronic circuits, environment and physicochemical process. In this study, the capability of the proposed smoothing technique is tested using different noise levels added to the cervical cell FTIR spectra, which include 20, 25, 30, 35, 40, and 45 dB (S/N).

3.2. Smoothing Algorithms

Cascade form filter is proven to produce better results than the direct form filter [13]. In this paper, three smoothing algorithms based on cascaded concept between the binomial filter and other direct form filters (i.e. SG, triangular, and rectangular filter) are proposed namely Cascaded Binomial-Savizky Golay (CBSG), Cascaded Binomial-Triangular (CBT), and Cascaded Binomial-Rectangular (CBR) filters. These algorithms are proposed to enhance the SG and the binomial filter performances.

The implemented of the proposed smoothing algorithms are as follow:

(i) Apply the 5 point coefficients of binomial filter as tabulated in Table 1 to be the first filter function.

(ii) Convolute the filter function to noisy spectrum and denote the output spectrum as $F_1(s)$.

(iii) As the second filter function, apply the 5 point coefficients of one of the other filters as tabulated in Table 1. For the combination of the binomial filter with the SG filter, binomial filter with triangular filter, and binomial filter with rectangular filter, the proposed algorithms are called CBSG, CBT, and CBR respectively.

(iv) Convolute each of the filter function (i.e. refer to the proposed smoothing algorithms in step (iii)) to $F_1(s)$ and denote the output spectrum as $F_2(s)$.

(v) Calculate the S/N of the $F_2(s)$.

4. Results and Discussions

In this Section, analyses of smoothing methods are presented on several cervical cell FTIR spectra both qualitatively and quantitatively. For qualitative analysis, the reference, noisy and smoothing spectra are plotted to show the smoothing effect on the spectra while the S/N values are compared for all methods for quantitative analysis [11].

4.1. Qualitative Analysis

In this evaluation, we use noisy spectra with different level of white noise to test the effectiveness and efficiency of the algorithms. The reference spectra are shown in Fig. 1(a) and 2(a) which the spectra are called as Spectrum 1 and Spectrum 2, respectively. Fig. 1(b) and 2(b) are the noisy spectra corrupted with 20 and 45 dB of S/N, respectively. Fig. 1(c) to 1(g) and 2(c) to 2(g) present the smoothed spectra with different smoothing algorithms.

The results depict that the smoothed spectra produced by the SG and binomial filters are still affected by noise, which indicate that the algorithms are less robust to the noise effect, as shown in Fig. 1(c), 1(d) and 2(c), 2(d). The proposed algorithms can significantly remove the effect of noise as compared to the SG and binomial filters. The smoothed spectra are less jagged. The curve of the output spectra are smoother than those produced by the SG and binomial spectra. Visually, the CBR algorithm produces the smoothest spectra, thus prove its ability on smoothing of the cervical cell FTIR spectra.

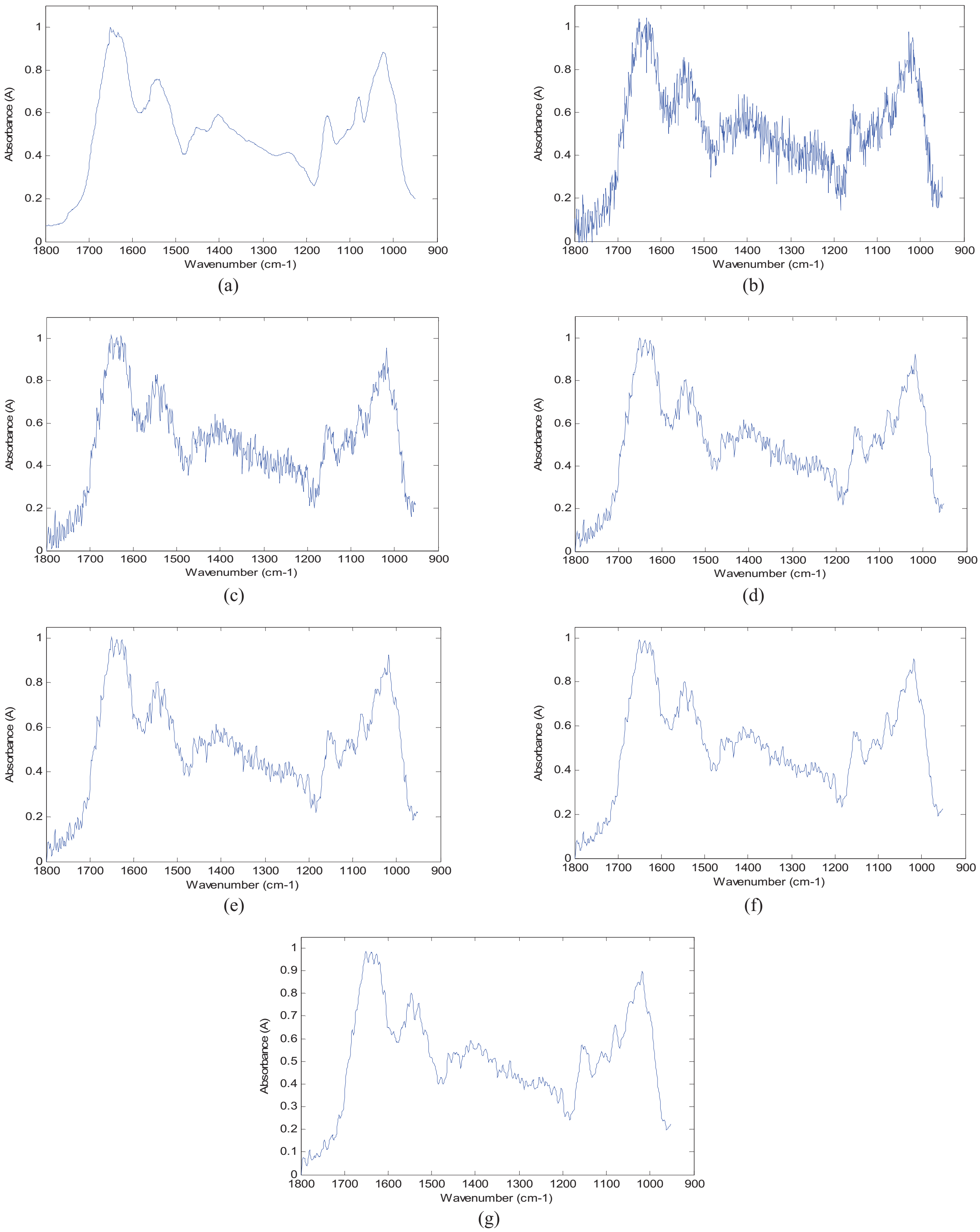

Fig. 1. Resultant FTIR Spectrum 1 after applying: (a) original spectrum 1, (b) noisy spectrum, (c) SG filter, (d) binomial filter, (e) CBSG, (f) CBT, (g) CBR

4.2. Quantitative Analysis

In this Section, we tabulate a quantitative evaluation of the smoothing results for Fig. 1(c) to 1(g), and 2(c) to 2(g). In this paper, we use S/N value as quantitative analysis. Higher value of S/N indicates better smoothing result. The S/N values obtained are tabulated in Table 2. The best results are bold.

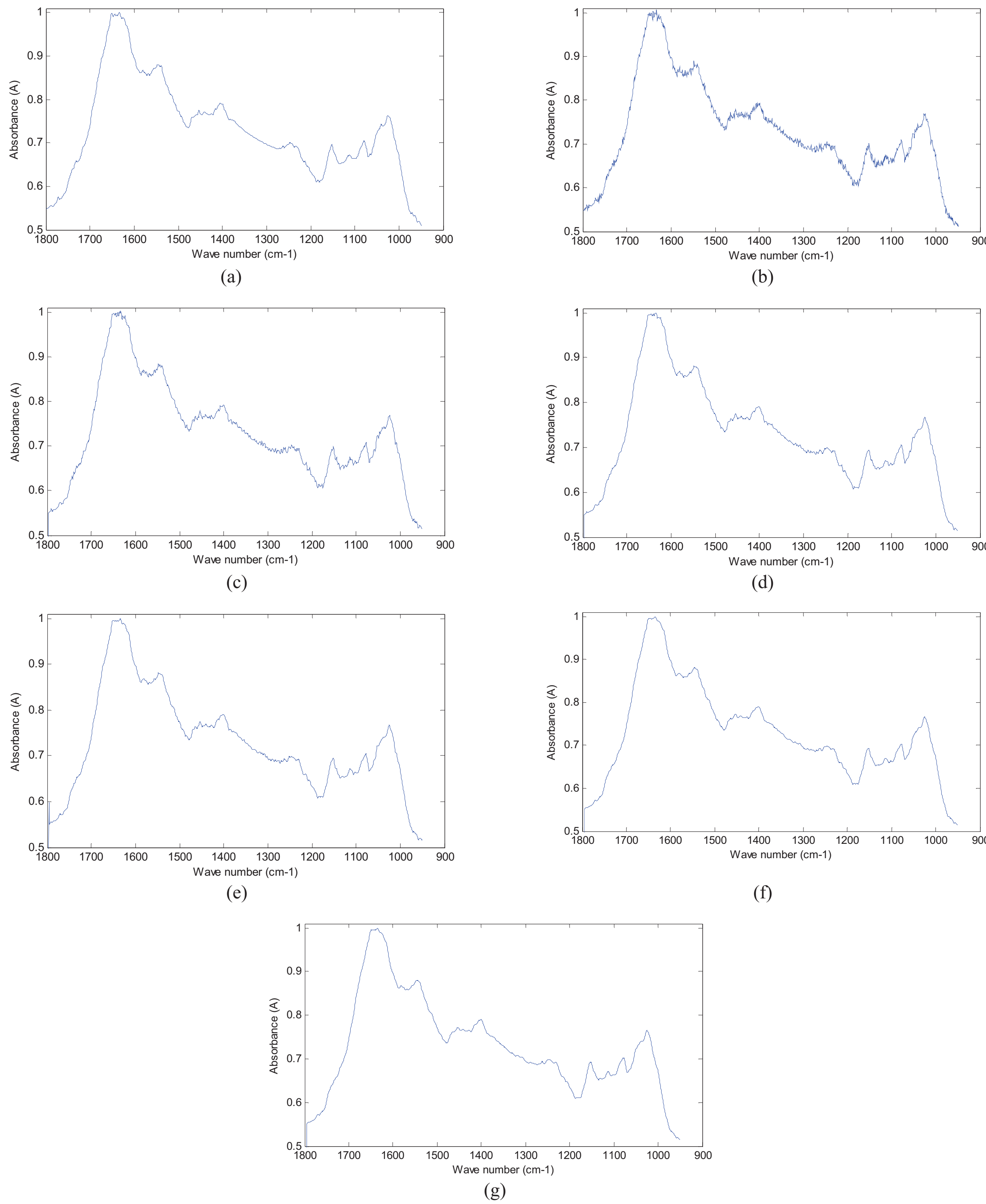

Fig. 2. Resultant FTIR Spectrum 2 after applying: (a) original spectrum 2, (b) noisy spectrum, (c) SG filter, (d) binomial filter, (e) CBSG, (f) CBT, (g) CBR

From Table 2, the proposed CBR algorithm consistently outperforms the other algorithms by producing the highest results for both tested spectra. These results strongly support the qualitative findings in Section 4.1. In addition, the processing time for each algorithm is also recorded. As tabulated in Table 2, it can be observed that the processing time of the cascade form filters is almost twice as compared to that of the direct form filters. This is due to the spectra need to be processed by two filters (as in the cascade form filters) instead of one filter (as in the direct form filters).

Table 2. Quantitative evaluations on smoothing selected spectra

Spectrum Types	Smoothing Algorithms	S/N Values	Processing Time (s)
Spectrum 1	SG	20.93	4.765
	Binomial	24.95	**4.687**
	CBSG	26.10	9.026
	CBT	29.45	8.846
	CBR	**32.03**	8.900
Spectrum 2	SG	51.02	4.428
	Binomial	53.53	**4.352**
	CBSG	53.53	8.926
	CBT	58.74	8.708
	CBR	**63.34**	8.939

This study further analyzes the capability of the proposed algorithms using 857 cervical cell FTIR spectra. These spectra are corrupted with white noise ranging from 20 to 45 dB of S/N. Using the aforementioned quantitative analysis, the average performance of all algorithms on the tested spectra is shown in Fig. 3. From the figure, it can be seen that the proposed algorithms provide a relatively higher value of S/N value which demonstrate their capability in reducing the noise effect of the cervical cell FTIR spectra. Overall, the CBR produces the best result as compared to others.

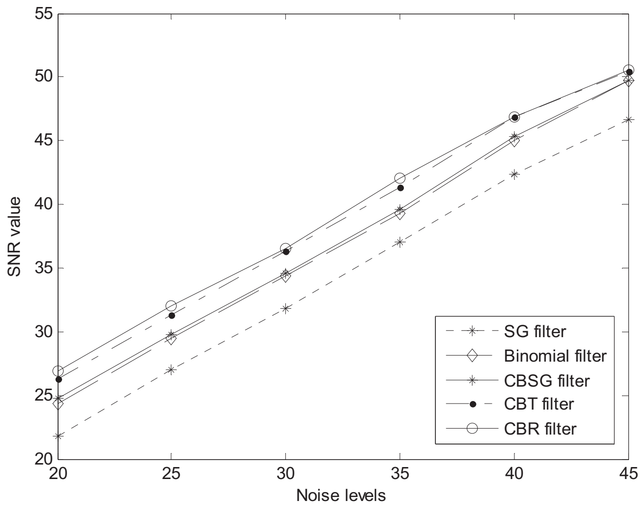

Fig. 3. Average the proposed smoothing evaluation on 857 cervical spectra

6. Conclusions

This paper presents new denoising-based cascaded binomial with other direct form filter algorithms for smoothing noise-corrupted cervical cell FTIR spectra. The qualitative and quantitative analyses favor the proposed algorithms as good smoothing algorithms. The proposed algorithms produce better results as compared to the conventional SG and binomial smoothing filter. This finding is proven by higher value of S/N produced by the proposed algorithms for each added noise levels. Processing time of the direct form filters are shorter than the proposed filters, but not significantly different. Thus, the proposed filters could be realized for real time applications.

7. Acknowledgement

This study is supported by Universiti Sains Malaysia Research University (RU) grant entitled "Study on capability of Fourier Transform Infrared (FTIR) spectral characteristics for development of intelligent cervical precancerous diagnostic systems".

8. References

[1] J. S. Wang, J. S. Shi, Y. Z. Xu, X. Y. Duan, L. Zhang, J. Wang, L. M. Yang, S. F. Weng, and J. G. Wu, "FT-IR spectroscopic analysis of normal and cancerous tissues of esophagus", *World Journal of Gastroenterology,*9, p. 1897-1899. 2003.

[2] S. G. El-Tawil, R. Adnan, Z. N. Muhamed, and N. H. Othman, "Comparative study between Pap smear cytology and FTIR spectroscopy: A new tool for screening for cervical cancer", *Pathology*, 40, p. 600-603, 2008.

[3] K. Tumer, N. Ramanujam, J. Ghosh, and R. Richards-Kortum, "Ensembles of radial basis function networks for spectroscopic detection of cervical precancer", *IEEE Transactions on Biomedical Engineering*, 45, p. 953-961, 1998.

[4] L. Quintero, C. Matthaus, S. Hunt, and M. Diem, "Denoising of single scan Raman spectroscopy signals", *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, 2010.

[5] S. Hargittai, "Savitzky-Golay least-squares polynomial filters in ECG signal processing", in *Computers in Cardiology*. 2005.

[6] M. Čupek, P. Matajka, and K. Volka, "Noise reduction in Raman spectra: Finite impulse response filtration versus Savitzky-Golay smoothing", *Journal of Raman Spectroscopy*, 38(9), p. 1174-1179, 2007.

[7] J. Luo, and J. Bai, "Savitzky-golay smoothing and differentiation filter of even length: A gram polynomial approach", *Spectroscopy*, 20(11), p. 27-34, 2005.

[8] A. Savitzky, and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures", *Analytical Chemistry*, 36(8), p. 1627-1639, 1964.

[9] C. Battistoni, G. Mattogno, and G. Righini, "Spectral noise removal by new digital smoothing routine", *Journal of Electron Spectroscopy and Related Phenomena*, 74(2), p. 159-166, 1995.

[10] P. Marchand, and L. Marmet, "Binomial smoothing filter: A way to avoid some pitfalls of least-squares polynomial smoothing", *Review of Scientific Instruments*, 54(8), p. 1034-1041, 1983.

[11] J. Kauppinen, and J. Partanen, "Fourier transforms in spectroscopy", *Wiley-VCH*, 2001.

[12] Y. Jusman, S. N. Sulaiman, N. A. Mat-Isa, I. A. Yusoff, R. Adnan, N. H. Othman, and A. Zaki, "Capability of new features from FTIR spectral of cervical cells for cervical precancerous diagnostic system using MLP networks", *Proceedings of TENCON in Singapore*, p. 1-6, 2009.

[13] Y. Jusman and N. A. Mat-Isa, "Denoising-based cascaded algorithms for smoothing of different level additive white gaussian noise-corrupted spectra", *Proceedings of ICCDE in Kuala Lumpur, Malaysia*, p. 925-930. 2011.