



Uso del método de descomposición empírica de modos para eliminar fluorescencia en espectros Raman de tejido biológico.

F. León-Bejarano¹, M. G. Ramírez-Eliás¹, M. O. Méndez¹, G. Dorantes-Méndez¹, A. Alba¹

¹Facultad de Ciencias, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México

Resumen— La espectroscopía Raman ha sido utilizada para diversas aplicaciones biomédicas con éxito. Sin embargo, un problema con los espectros Raman de tejido biológico es la presencia de ruido asociado a diversos factores entre ellos el ruido debido a la fluorescencia, generado por la excitación biomolecular. Este ruido algunas veces es mayor a la propia señal Raman por lo que la reducción o eliminación del mismo es fundamental para un correcto análisis de los espectros. En este trabajo se propone el uso del método de descomposición empírica de modos (EMD, por sus siglas en inglés) para eliminar la fluorescencia y el ruido en espectros Raman. EMD es un método de separación de señales adaptativo y libre de parámetros usado para señales no estacionarias. EMD fue probado en tejido biológico mostrando resultados similares a los métodos utilizados actualmente como el algoritmo de Vancouver (VRA, por sus siglas en inglés). Además, se calculó el coeficiente de correlación entre EMD y VRA siendo de 0.92.

Palabras clave—Espectroscopia Raman, autofluorescencia, eliminación de ruido.

I. INTRODUCCIÓN

La espectroscopía Raman es una técnica usada en medicina y biología para caracterizar muestras biológicas de tejido, células y fluidos, dadas las ventajas que ofrece por ser una técnica no invasiva, relativamente rápida y que proporciona información a nivel molecular. Adicionalmente, cada muestra analizada tiene su espectro característico o “huella molecular” [1] [2]. Sin embargo uno de los grandes problemas de los espectros Raman y específicamente de tejido biológico es que la señal va acompañada de ruido generado por el instrumento de medición, por fuentes externas y el ruido debido a la fluorescencia. Este último puede ser de órdenes de magnitud mayor que la propia señal Raman impidiendo obtener información asociada a la composición molecular de la muestra. Por lo tanto, es necesario eliminar el ruido en los espectros antes de la etapa de análisis. El proceso de eliminación de ruido debe llevarse a cabo con especial cuidado para evitar pérdida de información, y adaptarse al tejido que se pretenda analizar.

Para eliminar el ruido, se han empleado dos enfoques diferentes: el experimental y el computacional. Los métodos que usan el enfoque experimental, se basan en ajustes o mejoras a la instrumentación y estos incluyen los de excitación desplazada y sistemas de tiempo limitado [3] [4]. Métodos experimentales como los que se mencionan son un tanto complejos debido a que implican largos tiempos de adquisición que dificultan su uso para aplicaciones en tejido biológico, por estas razones se ha incrementado el uso de los métodos computacionales dada su rapidez, fácil

implementación y bajo costo. Entre algunos de los métodos computacionales que resaltan se encuentran: los de ajuste polinomial, destacando el método de ajuste multipolinomial modificado y el Algoritmo de Vancouver [5] [6]. Debido a su simplicidad y efectividad se han convertido en los métodos más usados siendo el algoritmo de Vancouver el de mayor uso y el cual podría ser considerado como uno de los métodos estándar. Otros métodos computacionales que se han implementado para eliminación de ruido en espectros Raman son: la transformada de Fourier [7], la transformada Wavelet [8], derivadas de primer y segundo orden [9], mínimos cuadrados penalizados y mínimos cuadrados asimétricos [10]. Hablando específicamente de los métodos basados en ajuste polinomial Lieber y colaboradores crearon el ajuste polinomial modificado (modpoly), que trabaja realizando un ajuste polinomial de forma iterativa [5]. En este método el usuario debe seleccionar el grado del polinomio que mejor modele a la fluorescencia de fondo del tejido que se analiza, el cual finalmente será restado a la señal muestreada. Uno de los grandes inconvenientes que presenta esta técnica es que puede llegar a agregar falsas bandas Raman debido a que no considera la contribución del ruido en el espectro antes de remover la línea base. Como una mejora al modpoly, Zhao et al., proponen el algoritmo de Vancouver (VRA) el cual funciona bajo las mismas bases del modpoly pero con mejoras que incluyen la contribución del ruido y la separación de las bandas Raman de mayor intensidad previas a corregir la línea base, así como también adicionalmente introduce la aplicación de un filtro para suavizar los espectros como paso previo a la eliminación de fluorescencia [7]. Sin embargo pese a su simplicidad, VRA presenta inconvenientes al ser un método que requiere de la selección del grado de polinomio a ajustar, que varía dependiendo del tipo de tejido o material que se analiza, como consecuencia requiere de conocimiento a priori por parte del usuario para determinar el valor del grado correcto.

Por otra parte, el método de descomposición empírica de modos (Empirical Mode Decomposition, EMD) ha surgido para el análisis de señales no estacionarias. EMD es un método adaptativo en tiempo-frecuencia no lineal, que ha sido usado en aplicaciones médicas en señales de electrocardiograma, electroencefalograma para la evaluación de ciertas patologías [11].

En este contexto, este trabajo de investigación tiene como objetivo evaluar el uso de EMD para eliminar ruido y corregir la línea base en espectros Raman de tejido biológico. El trabajo queda organizado de la siguiente forma, en la sección II se definirá la información de los datos y métodos utilizados para este trabajo, así como las métricas utilizadas

para evaluar el uso de los métodos planteados. Por otra parte, la sección III da a conocer los resultados obtenidos al aplicar EMD y compararlo con el método estándar, finalmente en la sección IV se da a conocer la discusión y conclusiones.

II. METODOLOGÍA

A. Protocolo

Para el presente reporte, se obtuvieron 20 espectros de piel humana *in vivo*, usando un sistema Raman portátil R3000, con un diodo laser de 785 nm, resolución espectral de 8 cm⁻¹, una potencia laser de 90 mW, rango espectral de 800 cm⁻¹ a 1800 cm⁻¹ y un tiempo de integración de 10s. Los espectros fueron adquiridos el mismo día, por el mismo operador, bajo las mismas condiciones, realizando cuatro mediciones en el antebrazo derecho de un individuo sano en la misma zona de piel y considerando como señal final al espectro promedio. Al final de las mediciones un conjunto de quince espectros Raman fue considerado para cada muestra.

Un espectro Raman puede expresarse como una combinación lineal de tres componentes

$$y = x_{true} + b + n. \quad (1)$$

En donde x_{true} es el espectro Raman libre de ruido y fluorescencia, b es la fluorescencia de fondo, n el ruido en la señal y finalmente y es la señal Raman cruda adquirida con el espectrómetro con fluorescencia y ruido. De tal forma que para obtener una señal libre de ruido es necesario restar la fluorescencia de fondo y el ruido identificados a la señal cruda.

C. Descomposición empírica de modos (EMD)

EMD fue definido por Huang et al. [12], es un método adaptativo para separar señales en un conjunto de componentes que varían en el tiempo, modulados en amplitud y frecuencia. EMD descompone una señal como la suma de funciones de modo intrínseco (IMF) más un residuo. Una IMF es una función con igual número (o puede diferir en uno) de extremos y cruces por cero. EMD se aplica de la siguiente manera:

- 1.- Se localizan los máximos y mínimos de la señal.
- 2.- Todos los máximos se conectan mediante un spline cúbico para obtener la envolvente superior e_u .
- 3.- Los mínimos se conectan mediante un spline cúbico para obtener la envolvente inferior e_i .
- 4.- Se calcula el promedio entre e_u y e_i , con lo que se obtiene un prototipo de la primera IMF (h).
- 5.- El paso número 4 se repite de forma iterativa tomando como señal el prototipo de IMF (h) proporcionado en el paso número 4, este proceso denominado proceso de tamizado se repite hasta que la salida en 4 cumpla con las especificaciones de una IMF (H).

El criterio utilizado para finalizar el proceso de tamizado queda definido como en la siguiente ecuación [10]:

$$SD = \sum_{t=0}^T \frac{(|h_{k-1}(t) - h_k(t)|)^2}{h_{k-1}^2(t)}. \quad (2)$$

En donde k representa el número de iteración durante el proceso de tamizado, h es la IMF prototipo, SD es la diferencia entre la IMF prototipo (h) anterior y la actual. Posteriormente SD es comparada con un umbral de 0.01, el cual determina cuando se obtiene una IMF (H) y por lo tanto finaliza el proceso de tamizado. El procedimiento para obtener las IMF restantes se realiza utilizando la siguiente ecuación:

$$r_n(t) = r_{n-1}(t) - H_n(t). \quad (3)$$

En donde $H_n(t)$ es la n -ésima IMF, $r_n(t)$ es el residuo actual y la nueva entrada al proceso de tamizado, $r_{n-1}(t)$ representa el residuo anterior. Una vez que se encuentra la k -ésima IMF el proceso termina cuando el residuo $r_n(t)$ es una función monótona de la que no se pueden extraer más IMFs. Finalmente la señal original se puede interpretar como [10]:

$$x(t) = \sum_{i=1}^M H_i(t) + r_M(t). \quad (4)$$

Aplicando EMD a espectroscopía Raman $x(t)$ es el espectro Raman crudo con fluorescencia y ruido, $H_i(t)$ son las IMF extraídas del espectro y $r_M(t)$ es el residuo después de obtener las IMFs del espectro Raman. En base a la observación, el residuo $r_M(t)$ más la última IMF $H_M(t)$ modelan a la fluorescencia de fondo, y las primeras cinco IMFs modelan al ruido de alta frecuencia, siendo la relación entre la ecuación (4) y la ecuación (1) de:

$$b = H_M(t) + r_M(t), \quad (5)$$

$$n = \sum_{i=1}^5 H_i(t), \quad (6)$$

finalmente $x_{true} = y - b - n$, será el espectro Raman libre de fluorescencia y ruido.

III. RESULTADOS

Los espectros Raman fueron procesados con EMD, posteriormente se verificó visualmente que las bandas identificadas en un principio en el espectro crudo se conservaran en las señales procesadas sin ruido ni fluorescencia. Además, VRA se utilizó como referencia por ser un método establecido así como uno de los más utilizados actualmente para eliminar ruido y fluorescencia, dada su simplicidad y efectividad. Posteriormente, los espectros de piel fueron sometidos a VRA y los resultados comparados con los espectros procesados con EMD, para lo cual se calculó el coeficiente de correlación coseno como medida de similitud entre ambos resultados siendo de 0.92. La Fig. 1 muestra el espectro original sin filtrar, las líneas grises verticales identifican el máximo valor de las bandas Raman

registradas en la literatura: aproximadamente en 852 cm^{-1} asociado a $\delta(CCH)$, en 1005 cm^{-1} relacionado a: $\nu(CC)$ y anillo de fenilo, aproximadamente en 1030 cm^{-1} asociado a $\nu(CC)$, 1080 cm^{-1} relacionado con ácidos nucleicos y $\nu(CC)$, 1269 cm^{-1} correspondiente con: amida III, $\nu(CN)$ y $\delta(NH)$, 1445 cm^{-1} asociado a proteínas $\delta(CH_2)$, $\delta(CH_3)$ y finalmente aproximadamente en 1655 cm^{-1} relacionado con amida I [13].

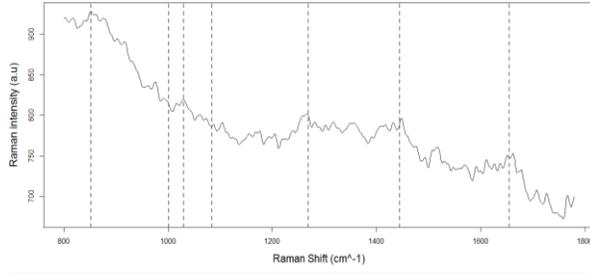


Fig. 1. Espectro Raman de piel humana sana con ruido y fluorescencia.

Por otra parte la Fig. 2, muestra la señal procesada con EMD (línea negra sólida) y VRA (línea negra punteada), en la que las líneas grises verticales identifican a las bandas Raman principales definidas en la literatura y mostradas en la Fig. 1. Los espectros de la Fig.2 muestran como VRA y EMD son semejantes a excepción de algunas regiones en las que existe diferencia de intensidad pero la forma es similar en ambos resultados.

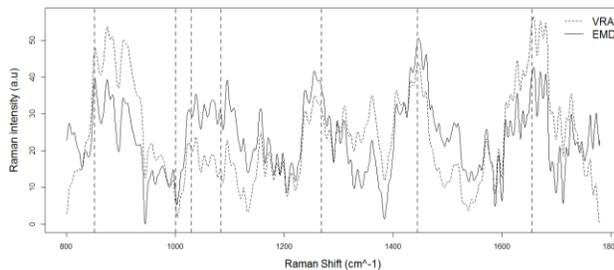


Fig. 2. Espectro Raman de piel humana libre de ruido, procesado con EMD y VRA.

IV. DISCUSIÓN

En los resultados presentados en la sección anterior, se puede observar como EMD proporciona una señal libre de ruido, en la que ya no existe esa curva de fondo que modifica la intensidad de las bandas Raman. Los resultados de EMD muestran similitud con los resultados de aplicar VRA mostrando una correlación de 0.92 entre éstos, lo que sugiere que no se agregaron falsas bandas Raman. EMD permitió identificar los picos Raman. Así como no es sensible al ruido, dado que permite recuperar los espectros Raman conservando la información original, respetando la variabilidad existente en distintos tipos de espectros, en especial en el caso de espectros Raman que provienen de tejido biológico.

V. CONCLUSIÓN

EMD fue capaz de obtener las bandas Raman identificadas en la literatura de forma clara y consistente. EMD permite eliminar la fluorescencia además del ruido generado por el instrumento de medición y por fuentes externas. Con EMD no hay pérdida de información y a nivel usuario tiene la ventaja de que no es necesario definir parámetros ni contar con información *a priori* para poder obtener la señal libre de ruido. Por lo tanto se propone como un método útil para la eliminación de ruido en espectros Raman de muestras biológicas.

RECONOCIMIENTOS

Este trabajo fue soportado por CONACYT, número de beca 414995.

BIBLIOGRAFÍA

- [1] T. Huser and J. Chan, "Raman spectroscopy for physiological investigations of tissues and cells," *Advanced Drug Delivery Reviews*, vol. 89, pp. 57, 2015.
- [2] F.J. González, R. Valdés-Rodríguez, M.G. Ramírez-Eliás, C. Castillo-Martínez, V.M. Saavedra-Alanís and B. Moncada, "Noninvasive detection of filaggrin gene mutations using Raman spectroscopy," *Biomedical Optics Express*, vol. 2, pp. 3363, 2011.
- [3] A.P. Shreve, N.J. Cherepy. and R.A. Mathies, "Effective rejection of fluorescence interference in Raman spectroscopy using a shifted excitation difference technique," *Applied Spectroscopy*, vol. 46, pp. 707, 1992.
- [4] F. Knorr, Z.J. Smith and S. Wachsmann-Hogiu, "Development of a time-gated system for Raman spectroscopy of biological samples," *Optics Express*, vol. 18, pp.20049, 2010.
- [5] J. Zhao, H. Lui, D. McLean and H. Zeng, "Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy," *Applied Spectroscopy*, vol. 61, pp.1225, 2007.
- [6] C.A. Lieber and A. Mahadevan-Jansen, "Automated Method for Subtraction of Fluorescence from Biological Raman Spectra," *Applied Spectroscopy*, vol. 57, pp. 1363, 2003.
- [7] G. Schulze, A. Jirasek, M.L. Yu, A. Lim, R.F.B. Turner, M.W. Blades, "Investigation of Selected Baseline Removal Techniques as Candidates for Automated Implementation," *Applied Spectroscopy*, vol. 59, pp. 545, 2005.
- [8] Z. M. Zhang, S. Chen, Y. Z. Liang and et al., "An intelligent background-correction algorithm for highly fluorescent samples in Raman spectroscopy," *Journal of Raman Spectroscopy*, vol. 41, pp. 659, 2010.
- [9] A. O'Grady, A.C. Dennis, D. Denvir, J.J. McGarvey, S.E. Bell, "Quantitative Raman spectroscopy of highly fluorescent samples using pseudosecond derivatives and multivariate analysis" *Analytical Chemistry*, vol. 73, pp. 2058, 2001.
- [10] P. Candeloro, E. Grande, R. Raimondo, D. Di Mascolo, F. Gentile, M.L. Coluccio, G. Perozziello, N. Malara, M. Francardi and E. Di Fabrizio, "Raman database of amino acids solutions: a critical study of Extended Multiplicative Signal Correction," *Analyst*, vol. 138, pp. 7331, 2013.
- [11] S. Liu, R. Gao, D. John, J. Staudenmayer and P. Freedson, "Tissue Artifact Removal from Respiratory Signals Based on Empirical Mode Decomposition," *Annals of Biomedical Engineering*, vol. 38, pp. 1, 2008.
- [12] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-

stationary time series analysis, in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, eds. Sir Mark Welland FRS, pp. 903, London, 1998.

- [13] Y. Li, R. Chen, H. Zeng, Z. Huang, S. Feng, and S. Xie, "Raman spectroscopy of Chinese human Skin in vivo", Chinese optics letters, vol. 5, pp.105, 2007.