

COMPARACIÓN Y TRANSICIONES ESPECTRALES MEDIANTE EL ALGORITMO FUZZY C-MEANS

PACS 43.75.Zz

Martínez Rodríguez, Brian⁽¹⁾; Liern Carrión, Vicente⁽²⁾

Institución: (1) Universidad Politécnica de Valencia.

Dirección: Camino de Vera, s/n

Población: 46022, Valencia.

País: España.

Teléfono: 963877000.

E-mail: brian.martinez@csmvalencia.es

Institución: (2) Dep. Matemáticas para la Economía y la Empresa. Universidad de Valencia.

Dirección: Av. dels Tarongers, s/n.

Población: 46022, Valencia

País: España.

Teléfono: 963 828 370.

E-mail: vicente.liern@uv.es

Palabras Clave: Comparación, transiciones, espectro, *fuzzy*, *c-means*.

ABSTRACT

In this paper a new mathematical tool is proposed in order to compare several frequency spectra with different number of overtones. To that effect, an average spectral dissimilarity will be defined and calculated by means of fuzzy clustering processes. In addition, meanwhile this algorithm is converging new spectra are being generated, as a kind of intermediate stages that create a transition from an distant initial point to the final one, in which the difference between both spectra has been minimized. Those transitions can be applied in several fields of contemporary composition such electronic music or spectral composition.

RESUMEN

En el presente trabajo se propone una herramienta matemática para la comparación de espectros que contienen un diferente número de sonidos parciales. Para ello se define una disimilitud media espectral que se calcula mediante procesos de *fuzzy clustering*. Además, en el transcurso de la propia convergencia del algoritmo se generan, a modo de estadios intermedios, nuevos espectros que suponen una transición desde un estado inicial hasta otro final. Desde este punto de vista, la convergencia consiste en minimizar la diferencia entre estados. Las transiciones pueden utilizarse en diferentes campos de la composición musical, como la música electroacústica o la composición espectral.

1. INTRODUCCIÓN

Los algoritmos de *clustering* o agrupamiento de datos han sido profusamente utilizados en numerosos campos de la ciencia como la geoestadística, la biofísica, la medicina o la inteligencia artificial en sus disciplinas de visión por ordenador, procesado de imágenes,

clasificación automática o reconocimiento automático de patrones, entre muchas otras¹. Dentro del ámbito de la musicología moderna, el clustering ha proporcionado interesantes resultados en el reconocimiento de patrones y estilos musicales, tal como muestran los estudios realizados por Logrippio y Stepien (1986), o más recientemente, en la determinación de la modalidad musical de géneros relativos a la música antigua (un interesante ejemplo de esta aplicación es el estudio realizado por Tompkins, 2017). Gran parte del éxito de dichos algoritmos radica en su gran versatilidad que permite su aplicación a numerosos problemas, así como en su relativa sencillez de implementación y velocidad de cálculo.

Dado un conjunto inicial de datos $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ donde cada valor es un vector observable de q dimensiones, la idea principal de los algoritmos de *clustering* es establecer una partición de las n observaciones iniciales en un número total de k subconjuntos $S = \{S_1, S_2, \dots, S_k\}$, siendo $k \leq n$. El más conocido de estos procedimientos es el célebre algoritmo *k-means* (Steinhaus, 1956; MacQueen, 1967), que propone realizar dicha partición a través de la minimización del cuadrado de la suma de las distancias existentes entre cada punto de un subconjunto S_j con su respectivo punto medio \mathbf{v}_j (al que denominados *centroide*) tal y como podemos ver en la siguiente expresión:

$$\operatorname{argmin} \sum_{j=1}^k \sum_{\mathbf{x} \in S_j} \|\mathbf{x} - \mathbf{v}_j\|^2$$

La implementación algorítmica estándar de este k-means fue propuesta por Stuart P. Lloyd en 1982 y consiste en una iteración recursiva mediante la cual los centroides, después de ser inicializados, convergen hacia unas posiciones finales que permiten establecer la partición final de los datos. Para una explicación más detallada de dicho proceso véase Lloyd (1982).

2. SOFT CLUSTERING

Una interesante variación del algoritmo anteriormente descrito es el denominado *fuzzy c-means* (en lo sucesivo FCM), propuesto por J. C. Dunn en 1973 y posteriormente desarrollado por Bezdek en 1981. En este algoritmo se realiza la partición de los datos iniciales $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ en un número m de subconjuntos finales $S = \{S_1, S_2, \dots, S_m\}$ bajo la afirmación de que cada elemento del conjunto inicial puede pertenecer a más de un centroide. De esta manera, el resultado final del algoritmo será que cada elemento del conjunto inicial $\mathbf{x}_i \in X$ presenta un coeficiente de pertenencia al centroide $\mathbf{v}_j \in S$, siendo la suma de los coeficientes de pertenencia de un punto \mathbf{x}_i a todos los centroides igual a uno, para cualquier punto de X . Tal y como se explica en Bezdek (1981), para la consecución de esta *soft-partition* será necesario realizar una minimización iterativa de alguna de las funciones descritas por esta expresión general:

$$J_\lambda = \sum_{i=1}^n \sum_{j=1}^m (u_{ij})^\lambda \|\mathbf{x}_i - \mathbf{v}_j\|^2$$

siendo $\lambda \in [1, \infty)$ la ponderación (o grado de *fuzzyness* del proceso) y u_{ij} el coeficiente de pertenencia del elemento \mathbf{x}_i al centroide \mathbf{v}_j . De acuerdo con Bezdek (1981), en cada iteración se debe actualizar el valor de los coeficientes de pertenencia y de los centroides a través de las siguientes expresiones:

$$u_{ij} = \frac{1}{\sum_{k=1}^m \left[\frac{\|\mathbf{x}_i - \mathbf{v}_j\|}{\|\mathbf{x}_i - \mathbf{v}_k\|} \right]^{\frac{2}{\lambda-1}}}, \quad \mathbf{v}_j = \frac{\sum_{i=1}^n u_{ij}^\lambda \cdot \mathbf{x}_i}{\sum_{i=1}^n u_{ij}^\lambda}$$

¹ Véase a modo de ejemplo Honarkhah & Caers (2010); Eisen, Spellman, Brown, & Botstein (1998); Hartigan (1975); Bezdek, Keller, Krisnapuram & Pal (2006); o Bezdek (2013) donde se pueden encontrar interesantes aplicaciones de procedimientos de *data clustering* a distintos ámbitos científicos.

Se exige una condición de finalización y esto es lo que determina el número de iteraciones. Una vez alcanzado considerará que el algoritmo ha convergido. En este momento tenemos una matriz U que representa la partición *fuzzy* de los elementos del conjunto inicial en los m subconjuntos S . Para una mayor explicación del algoritmo así como de su implementación informática véase Bezdek (1981).

3. CÁLCULO DE LA DISIMILITUD ESPECTRAL MEDIA MEDIANTE EL FCM

Consideremos dos espectros de frecuencias, compuestos por un total de n y m sonidos parciales respectivamente. Nos planteamos el parecido o diferencia entre ambos espectros, o dicho de otra manera: ¿Es posible establecer una medida de la disimilitud entre estos dos espectros? Para ello necesitamos previamente precisar el concepto de *sonido parcial*. La definición más sencilla consiste en caracterizar un parcial mediante sus dos propiedades básicas: frecuencia e intensidad. Definiciones más complejas podrían realizarse mediante la incorporación de nuevas propiedades, como por ejemplo la dependencia temporal de cada parcial o la teoría de conjuntos *fuzzy* aplicada a las frecuencias (explicado en profundidad en Liern, 2005; León & Liern, 2012). Esto puede resolverse de gracias a la utilización de un número de dimensiones superior para caracterizar los observables requeridos. A continuación se dan las definiciones necesarias para calcular la disimilitud espectral media.

Definición 1: Un espectro de frecuencias es una secuencia $\mathcal{F} = \{x_i\}_{i=1}^n$, donde cada $x_i \in \mathbb{R}^2$ es un sonido parcial formado por los valores observables de frecuencia e intensidad.

Consideremos por ejemplo el análisis espectral realizado mediante FFT de una nota La_3 (220Hz) producida por un saxo alto en $Mi\flat$, tal y como se muestra en la Figura 1.

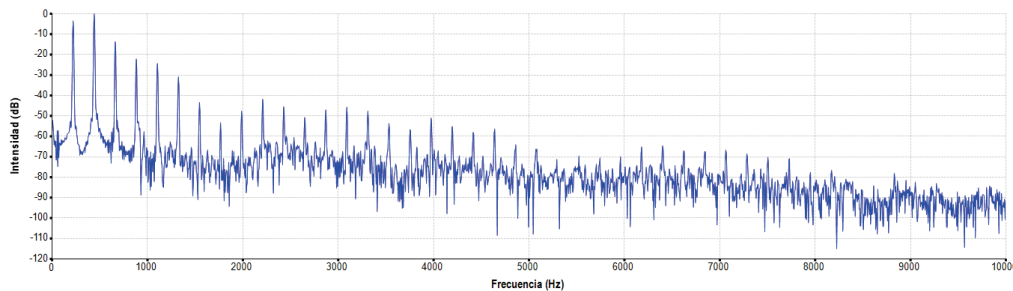


Figura 1: Análisis espectral de la nota La_3 (220Hz) emitida por un saxo alto.

Si filtramos el resultado anterior para obtener los n primeros picos más marcados (lo máximos locales de mayor intensidad²), obtendremos que el espectro de frecuencias \mathcal{F}^{La_3} del saxo puede ser representado como se muestra en la Figura 2. En la Tabla 1 podemos observar cuáles serían los 39 primeros valores numéricos obtenidos para cada parcial del espectro $\mathcal{F}^{La_3} = \{x_i\}_{i=1}^{39}$.

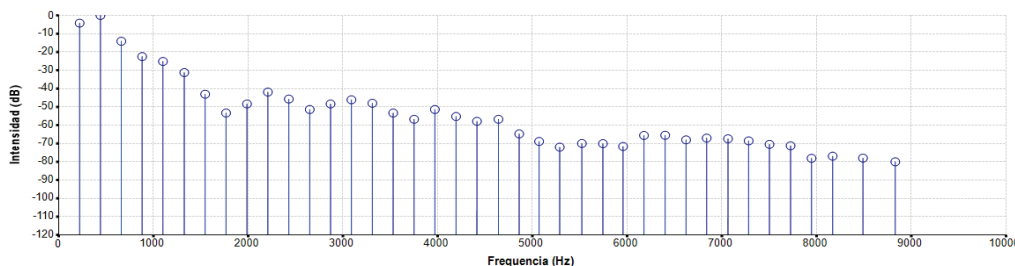


Figura 2: Espectro \mathcal{F}^{La_3} de frecuencias filtrado con los $n = 39$ máximos locales más intensos (en lo sucesivo \mathcal{F}^A).

² Este filtrado puede realizarse mediante diferentes técnicas como el análisis *Cepstrum* o *Harmonic Product Spectrum* (HPS), ambas explicadas concisamente en Park (2009).

Tabla 1: Secuencia de parciales formada por duplas x_i de frecuencia (Hz) e intensidad (dB) del espectro \mathcal{F}^A .

Nº	Hz	dB	Nº	Hz	dB	Nº	Hz	dB
1	220	-4,19	14	3089	-46,09	27	5956	-71,61
2	440	00,00	15	3310	-48,06	28	6177	-65,62
3	660	-14,09	16	3530	-53,33	29	6401	-65,52
4	880	-22,48	17	3751	-56,76	30	6622	-68,02
5	1100	-25,14	18	3970	-51,42	31	6839	-67,04
6	1325	-31,24	19	4194	-55,23	32	7064	-67,42
7	1545	-43,05	20	4414	-57,90	33	7282	-68,57
8	1766	-53,33	21	4641	-56,76	34	7502	-70,47
9	1987	-48,38	22	4860	-64,76	35	7725	-71,23
10	2207	-41,90	23	5070	-68,90	36	7946	-78,09
11	2428	-45,71	24	5288	-72,00	37	8169	-76,95
12	2649	-51,42	25	5522	-70,00	38	8489	-78,00
13	2869	-48,38	26	5744	-70,09	39	8830	-80,00

Consideremos ahora dos espectros de frecuencias \mathcal{F}^A y \mathcal{F}^B , con distinto número de sonidos parciales. De forma análoga a como en Martínez & Liern (2017) fue definida una *función de disimilitud* entre dos melodías con distinto número de notas, se puede proponer una *función de disimilitud* entre dos espectros de frecuencias de distinto número de sonidos parciales.

Definición 2: Sean $\mathcal{F}^A = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \subset \mathbb{R}^2$ y $\mathcal{F}^B = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\} \subset \mathbb{R}^2$ dos espectros de frecuencias, donde $n > m$, $d: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ una función distancia y u_{ij} los coeficientes finales de pertenencia calculados con el algoritmo FCM (tomando como datos iniciales los puntos de \mathcal{F}^A y como *centroides* iniciales los puntos de \mathcal{F}^B). La disimilitud media \mathcal{D} entre \mathcal{F}^A y \mathcal{F}^B se define como:

$$\mathcal{D}(\mathcal{F}^A, \mathcal{F}^B) = \frac{1}{n \cdot m} \sum_{i=1}^n \sum_{j=1}^m u_{ij} \cdot d(\mathbf{x}_i - \mathbf{y}_j)$$

De esta manera la disimilitud media representa una media ponderada de la distancia de cada uno de los puntos del espectro A a cada uno de los puntos del espectro B (*centroides*). Por propia construcción, la ponderación viene dada por los coeficientes de pertenencia, previamente calculados mediante el algoritmo FCM.

4. TRANSICIONES ESPECTRALES

Sin duda, un resultado potencialmente útil de nuestro procedimiento es que los *centroides*, representados en este caso por los sonidos parciales del espectro \mathcal{F}^B , experimentan cambios en su posición durante el transcurso iterativo del algoritmo FCM. En definitiva, se trata de progresivos movimientos en búsqueda de una convergencia final.

Dichos movimientos vienen determinados por el cálculo y actualización de sus coordenadas que se realiza en cada paso del proceso. Suponen estadios intermedios que pueden interpretarse como diferentes etapas de una transición desde su posición inicial hasta la posición final, en la cual la función distancia objetivo ha sido minimizada.

Cada paso intermedio puede asimilarse a nuevos espectros de frecuencias, generados a partir de la progresión del espectro \mathcal{F}^B desde su estado inicial, más alejado del espectro \mathcal{F}^A , hasta el final en el que la distancia entre ambos espectros se ha visto minimizada.

Ejemplo 1: Comparemos el espectro \mathcal{F}^A presentado anteriormente en la Figura 2, con el espectro \mathcal{F}^B de frecuencias (Figura 3) generado de forma que existen evidentes diferencias entre ambos espectros:

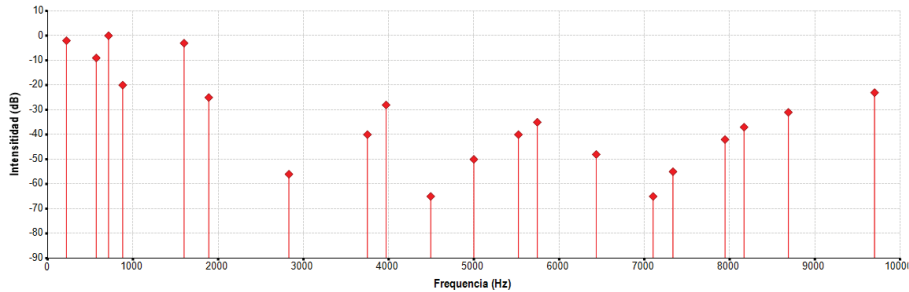


Figura 3: Espectro \mathcal{F}^B de frecuencias generado sintéticamente con $m = 20$ parciales.

Calcularemos la disimilitud media entre \mathcal{F}^A y \mathcal{F}^B mediante el procedimiento descrito en la Definición 2, para un exponente de peso fuzzy $\lambda = 2.5$. En la Figura 4 podemos observar algunos de los diferentes estadios intermedios por los que atraviesa el espectro B en su proceso de convergencia. La disimilitud media calculada es $\mathcal{D}(\mathcal{F}^A, \mathcal{F}^B) = 25.4036132366$.

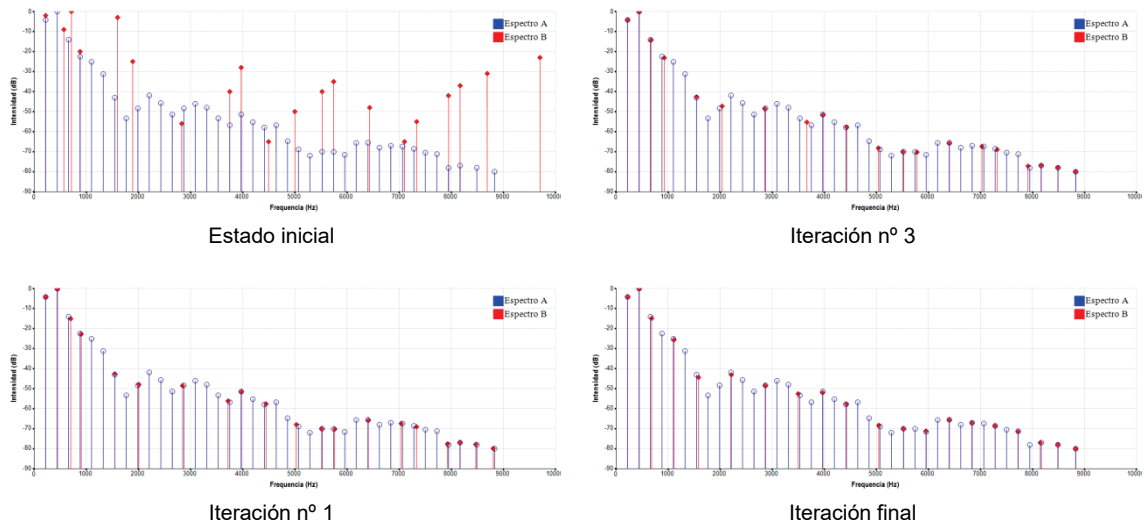


Figura 4: Diferentes estadios intermedios del espectro B durante la aplicación del algoritmo FCM.

Ejemplo 2: Comparemos el espectro \mathcal{F}^A con un espectro \mathcal{F}^C , cercano a \mathcal{F}^A , tal y como se puede observar en la siguiente figura:

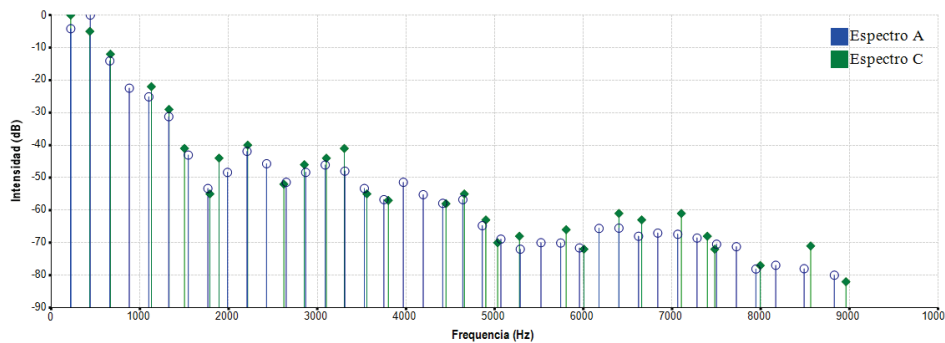


Figura 5: Comparación de espectros \mathcal{F}^A y \mathcal{F}^C .

El resultado final de la aplicación del algoritmo FCM sobre \mathcal{F}^A y \mathcal{F}^C (con el mismo parámetro $\lambda = 2.5$ e idéntico criterio de finalización es) el siguiente:

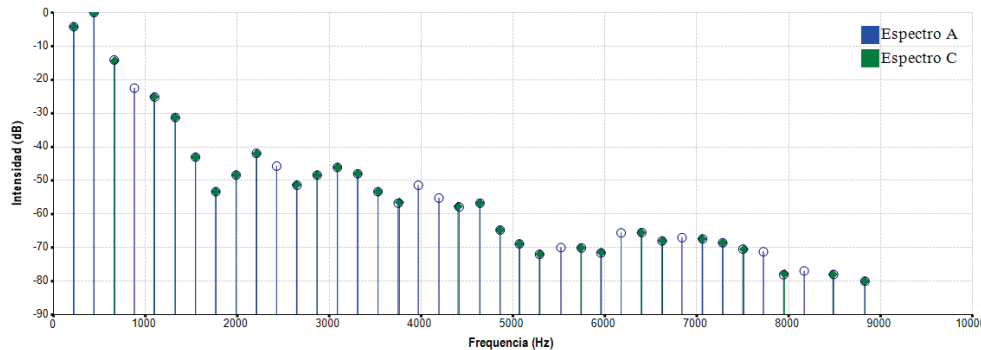


Figura 6: Estado final en el proceso de cálculo de la disimilitud media entre los espectros \mathcal{F}^A y \mathcal{F}^C .

Con los datos del Ejemplo 2, obtenemos un valor de disimilitud media de $\mathcal{D}(\mathcal{F}^A, \mathcal{F}^C) = 2.4933121089$, mucho menor que el obtenido en el Ejemplo 1, hecho que pone de manifiesto la evidente proximidad de ambos espectros.

5. CONCLUSIONES

La disimilitud espectral media propuesta en se muestra como una herramienta útil para estimar el grado de proximidad o lejanía que tienen dos espectros de frecuencias distintos. Esta disimilitud, que se expresa con un número real no negativo, puede ser fácilmente calculada mediante la implementación informática del algoritmo FCM, el cual alcanza la convergencia de forma computacionalmente rápida, incluso para un elevado número de puntos y *centroides*.

Los resultados de disimilitud media obtenidos tanto en los ejemplos como en los casos reales analizados por los autores arrojan una correlación entre la aparente diferencia de los espectros considerados y su valor de disimilitud media.

La transiciones espectrales, obtenidas gracias a los estadios intermedios por los que pasa el espectro B en su camino algorítmico hacia la convergencia, pueden suponer una interesante fuente de creación *tímblica* en ámbitos musicales como la composición electroacústica o la composición espectral, ya que permiten definir procesos intermedios entre dos timbres o estadios iniciales cualesquiera. Esto podría ampliar las técnicas de búsqueda tímbrica propias de la composición contemporánea actual, así como los procesos estructuralistas sobre los cuales la forma musical puede construirse.

Queda como objeto de estudio en futuras investigaciones la utilización de otras funciones distancia en la implementación del FCM para analizar su impacto en los resultados de disimilitud final.

REFERENCIAS BIBLIOGRÁFICAS

- Bezdek, J. C. (2013). *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media.
- Bezdek, J. C., Keller, J., Krisnapuram, R. & Pal, N. (2006). *Fuzzy models and algorithms for pattern recognition and image processing* (Vol. 4). Springer Science & Business Media.
- Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3, 32.

- Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. En *Proceedings of the National Academy of Sciences*, 95(25), 14863-14868.
- Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley & Sons.
- Honarkhah, M. & Caers, J. (2010). Stochastic simulation of patterns using distance-based Pattern modeling. *Mathematical Geosciences*. 42(5), 487-517.
- León, T. & Liern, V. (2012). Mathematics and soft computing in music. En R. Seising, V. Sanz (Ed.), *Soft Computing in Humanities and Social Sciences* (pp. 451-466). Springer.
- Liern, V. (2005). Fuzzy tuning systems: the mathematics of the musicians. *Fuzzy Sets and Systems*, 150(1), 35-52
- Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129-137.
- Logrippo, L. & Stepien, B. (1986). Cluster analysis for the computer-assisted statistical analysis of melodies. *Computers and the Humanities*, 20(1), 19-33.
- MacQueen, J. (1967, Junio). Some methods for classification and analysis of multivariate observations. En *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*. 1(14), 281-297. University of California Press.
- Martínez, B. & Liern, V. (2017, Junio). A fuzzy-clustering based approach for measuring similarity between melodies. En *Proceedings of The Sixth International Conference on Mathematics and Computation in Music. Mexico City, México*.
- Park, T. H. (2009). *Introduction to digital signal processing: Computer musically speaking*. World Scientific.
- Steinhaus, H. (1956). Sur la division des corp materiels en parties. *Bulletin de l'Académie Polonaise des Sciences*, 1, 801-804.
- Tompkins, D. (2017, Junio). A Cluster Analysis for Mode Identification in Early Music Genres. En *Proceedings of The Sixth International Conference on Mathematics and Computation in Music. Mexico City, México*.