

SALSA DATASET: PRIMERA BASE DE CONOCIMIENTO DE MÚSICA SALSA

SALSA DATASET: FIRST SALSA MUSIC KNOWLEDGE BASE

Gerardo M. Sarria M.¹

gsarria@javerianacali.edu.co

Mario Julián Mora²

mjmora@usbcali.edu.co

Carlos Arce-Lopera³

caarce@icesi.edu.co

DOI: [10.17230/ricercare.2016.5.5](https://doi.org/10.17230/ricercare.2016.5.5)

- ¹ Doctor en Ingeniería (Ciencias de la Computación) de la Universidad del Valle, Cali. Profesor asociado de la Pontificia Universidad Javeriana, Cali.
- ² Magíster Executive en Innovación Tecnológica de la Escuela de Negocios EOI de Madrid, España. Decano de la Facultad de Ingeniería de la Universidad San Buenaventura, Cali.
- ³ Doctor en Ingeniería de la Yokohama National University de Japón. Profesor asistente de la Universidad ICESI, Cali.

Resumen

Un género musical muy conocido en nuestra región y que hace parte de nuestra identidad cultural caleña es la salsa. Su origen se remonta a los años 30 del siglo pasado y desde entonces este género ha sido modificado por diferentes artistas en diversas regiones del mundo, cada uno con una visión diferente de él, con experiencias culturales y con aporte de nuevos instrumentos y nueva tecnología. Esto hace que la salsa sea un género intrínsecamente complejo y difícil de definir en términos cualitativos. A pesar de la creciente popularidad del género en el mundo, la salsa no ha sido analizada desde el punto de vista computacional para derivar cuáles son los componentes acústicos que la definen y diferencian de los demás géneros musicales. En este documento abordaremos los resultados del proceso de creación de una base de conocimiento de música salsa que está disponible en forma gratuita para la comunidad científica y que recopila la información acústica de más de 20.000 canciones de dicho género musical. Con esta información, que caracteriza la señal acústica, se pretende modelar diferentes características del género mediante técnicas de inteligencia artificial. Además de hacer accesible esta información a investigadores interesados en la música salsa, otro aporte importante de este proyecto es proporcionar herramientas para que la base de conocimiento crezca con la ayuda de la comunidad científica. Para ello se desarrolló un software que extrae la información acústica pertinente de canciones que tienen los usuarios para ser enviada y adicionada a la base de conocimiento.

Palabras clave: Género Musical Salsa, Dataset, Recuperación de Información Musical.

Abstract

Salsa is a well-known musical genre and part of our cultural identity. Its origins date back to the 30s of the last century and it has grown in popularity since its origins. Ever since, by different artists in different regions of the world have modified the genre, through different visions of what salsa is, experimenting and adding new instruments and technology. Thus, salsa becomes an intrinsically complex and difficult genre in qualitative terms. However, little we know at a computational level, about which are the acoustic characteristics defining this music, to make it different from the rest of musical genres. In this paper we show the results of a process that builds a knowledge base of salsa music freely available to the scientific community. This base gathers acoustic information of over 20.000 salsa songs. We pretend to use this information to model different characteristics of the genre by means of AI techniques. In addition to making information accessible to researchers interested in salsa music, another important contribution of this project is to provide tools to enlarge the knowledge base with the help of the scientific community. To achieve this, we developed a software that extracts the pertinent acoustic information from songs belonging to users and then it includes them into the knowledge base.

Keywords: *Salsa Musical Genre, Dataset, Music Information Retrieval.*

Introducción

La música siempre ha estado presente en la vida del hombre. Los primeros instrumentos musicales fueron, tal vez, huesos de animales o ramas de árboles que se golpeaban contra superficies semielásticas o huecas. Los sonidos resultantes, combinados con los emitidos por ellos mismos, crearon las primeras composiciones musicales (Fitch, 2005). A medida que el tiempo ha transcurrido, la música ha cambiado y evolucionado debido a diversos factores como la religión, la popularidad, la tecnología, la política, la región geográfica, el estilo de vida de las personas y la necesidad de expresar emociones (McDermott, 2008). Se puede apreciar cómo de la música clásica hemos pasado al blues y el jazz, y de allí saltamos al rock. Por otra parte, del jazz pasamos al latín jazz, al son montuno, luego al mambo, el boogaloo, la guaracha y de allí a la salsa. El género salsa tiene raíces tan diversas que ha sido muy difícil encontrar una definición para el término musical. Sin embargo, sus características acústicas nos parecen familiares a la escucha y hacen parte de nuestra identidad cultural.

Un primer problema radica en que dependiendo de la persona a quien se le pregunte sobre la definición del género salsa se obtendrá una respuesta diferente, y además, según el momento histórico y la ubicación geográfica pueden notarse diferencias entre una canción salsa con respecto a otra del mismo tipo. Muchos músicos y orquestas, con sus diferentes instrumentos y culturas, han producido canciones originales con su sello particular. No obstante, todas las mencionadas canciones pueden incluirse en el género salsa. Todos estos factores, que hacen tan difícil una definición concreta del género musical y que dependen de la percepción o de la experiencia de un músico o musicólogo, dificultan un análisis directo basado en encuestas y percepciones humana (Aucouturier y Pachet, 2003). Por lo tanto, es necesario optar por un análisis científico en el que las características acústicas de las canciones del género salsa puedan ser modeladas y analizadas para lograr no solo clasificar las canciones sino también estudiar la evolución del género.

Por otra parte, la salsa se ha popularizado en el mundo y sin embargo poco se sabe acerca de cuáles son las características acústicas que definen el género. La complejidad de sus raíces, al igual que la diversidad de los artistas que la componen, hace de la salsa un género musical muy atractivo para analizar (Morales, 2003; Duany, 1984). Otro problema radica en que las aproximaciones simples que resulten de acudir a expertos musicólogos siempre estarán sesgadas por sus percepciones y experiencia (Aucouturier y Pachet, 2003). Por lo tanto, un análisis científico que utilice técnicas avanzadas de inteligencia artificial, de recuperación de información y de análisis acústico es una idea original y necesaria para poder dar respuesta al problema planteado.

En el presente artículo hablaremos sobre un proyecto que propone realizar un análisis científico que emplea características cuantificables de la música para garantizar que los resultados del modelo computacional sean independientes de cualquier percepción humana. Además, al comparar los resultados de expertos musicólogos con los del modelo computacional se podría modelar y teorizar acerca de la percepción musical humana. El modelo computacional resultante permitirá responder preguntas del estilo siguiente: ¿cuáles son las diferencias en ritmo, melodía, instrumentos, etc., entre salsa de Puerto Rico y salsa de Colombia?, ¿es posible afirmar quién es el autor de una canción desconocida?, ¿es posible decir en cuál año fue escrita una canción particular? Y lo más importante de este modelo es que las respuestas a dichas preguntas se plantearán en términos de características acústicas de fácil medición. Por otro lado, este proyecto acentúa la importancia del género musical estudiado y que nos identifica desde el punto de vista cultural (Padilla, 1989), puesto que sus resultados ayudarían a popularizar y legitimar el género salsa como expresión cultural de nuestra región y, a la vez, permitiría entregarle al mundo una definición clara de que se entiende como salsa.

Marco teórico y estado del arte

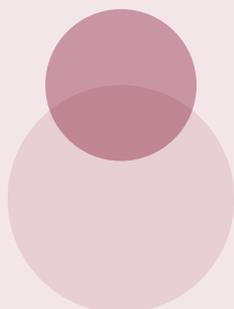
En los últimos 20 años, la forma en que la música se crea, se produce, se distribuye y se adquiere ha cambiado en forma radical (Li y Li, 2012). Las grandes colecciones de música que una persona podía tener antes en discos de vinilo, casetes y discos compactos han cambiado hacia los dispositivos portables de música digital como los *mp3 players*, *iPods* o, incluso, celulares. La tecnología actual permite adquirir y almacenar millones de canciones de manera fácil y rápida, lo que ha ocasionado la aparición de un problema: la organización y análisis eficiente de la información. El problema es muy importante dado que el valor de una colección grande de música está limitada a qué tan eficiente es la forma en la que se puede explorar (Pampalk, 2006).

El área de recuperación de información musical (MIR por las iniciales en inglés de la expresión *music information retrieval*) es un campo interdisciplinario de investigación que busca desarrollar técnicas de búsqueda y recuperación de datos relacionados con la música. La mayor parte de las investigaciones en esta área se ha centrado en la recuperación basada en contenido (Orio, 2006), es decir, la información que se encuentra almacenada está descrita por medio de un conjunto de características que se extraen de modo directo de su contenido mismo. Las técnicas usadas para recuperar información basada en contenido incluyen métodos probabilísticos y estadísticos, aprendizaje de máquina (*machine learning*) y el descubrimiento de conocimiento. El último es un proceso que consiste en una serie de pasos para identificar patrones en datos que sean válidos, nuevos, potencialmente útiles y sobre todo entendibles (Cios, Pedrycz, Swinarski y Kurgan, 2007).

El descubrimiento de conocimiento en música tiene varios tipos de tareas:

- **Manejo de datos**, por ejemplo para indexar bases de datos musicales (Rauber, Pampalk y Merkl, 2002).
- **Asociación**, para detectar correlaciones entre diferentes elementos musicales (Xiao, Tian, Li y Zhou, 2008).
- **Secuenciación**, para encontrar patrones que se presentan en cierto número de datos musicales (Peeters, 2007).
- **Clasificación**, para relacionar objetos musicales con una de muchas categorías musicales predefinidas (Wold, Blum, Keislar y Wheaton, 1996)
- **Conglomeración (clustering)**, para particionar un conjunto de objetos musicales en un espacio multidimensional hacia clases (Li, Ogihara y Zhu, 2006).
- **Búsqueda de similitudes**, para explorar un trabajo musical dada una descripción aproximada de una o más de otros trabajos (Schnitzer, Flexer y Widmer, 2009).
- **Resumen**, para crear una extracción concisa e informativa que de mejor manera resuma una pieza musical (Cooper y Foote, 2003).
- **Visualización de datos**, para concebir el contenido de una pieza musical y representar una colección musical completa (Isaacson, 2005).

Las anteriores tareas pueden realizarse sobre una base de conocimiento que contenga las siguientes características: la onda de sonido, las características acústicas, la letra de las canciones, anotaciones o *reviews* de la música, las etiquetas, las listas de reproducción (*playlists*) y las partituras.

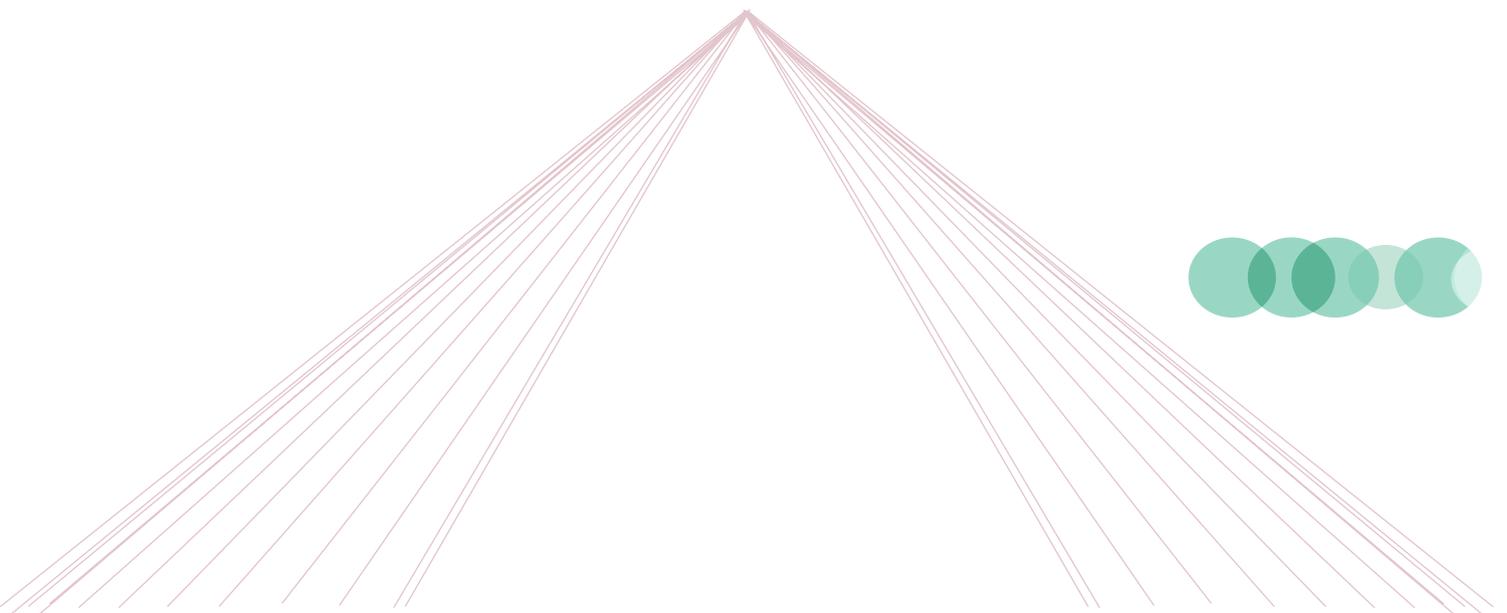


Las características acústicas son:

- **Tonalidad.** Relacionada con la percepción de la frecuencia fundamental de un sonido.
- **Intensidad.** Relacionada con la amplitud de la vibración.
- **Timbre.** Definido como la característica del sonido que les permite a los oyentes percibir como diferentes dos sonidos con la misma tonalidad y la misma intensidad.
- **Tempo.** La velocidad con la que se interpreta una obra musical.
- **Orquestación.** Relacionada con los instrumentos musicales que se emplean para interpretar diferentes voces, acordes y, en general, sonidos.
- **Acústica.** Una especialización de algunas características del timbre.
- **Ritmo.** Relacionado con la repetición periódica de un patrón temporal.
- **Melodía.** Una secuencia de tonos con un timbre similar.
- **Armonía.** La organización de sonidos simultáneos.

En el mundo existen grandes colecciones de música que se han usado para tareas de MIR. Ejemplos de ellas son Echonest (2016), The Million Song Data Set (Bertin-Maheux, Ellis, Whitman y Lamere, 2011), GTZAN (Tzanetakis y Cook, 2002), Codaich (McKay, McEnnis y Fujinaga, 2006), e, incluso el *dataset* de la 5ª Conferencia Internacional en Recuperación de Información Musical (ISMIR, 2004). Estas colecciones han traído consigo metodologías de acceso y recuperación como, por ejemplo, Musiclef (Orío, Rizo, Miotto, Montecchio, Schedl y Lartillot, 2011).

Un esfuerzo para hacer una colección de música latina es *The Latin Music Database* (Silla Jr., Koerich y Kaestner, 2008). Esta base de datos contiene 3.227 canciones de diferentes géneros latinos. Tristemente solo 311 de ellas son de salsa cuyas características pueden ser recuperadas. Por su parte, Proud Music (2016) provee 76 registros de canciones salsa con características como tempo, emoción, carácter, instrumentos, arreglos y compositor.



Construcción de la base de conocimiento

La ausencia de una base de conocimiento con un volumen abultado de canciones del género salsa fue nuestra principal motivación para construir nuestra propia base de conocimiento. Sin embargo, en la etapa de diseño de ella debimos tomar ciertas decisiones con respecto al número de canciones, el formato de audio en el cual estarán codificadas las canciones, las características que se deberían extraer para la futura caracterización y la forma en que sería almacenada la información.

Desde el primer momento pensamos que era necesario tener bastantes canciones para realizar la caracterización. Además, necesitábamos canciones de diferentes intérpretes, compositores, regiones del mundo, año de composición, e, incluso subgéneros de la salsa (por ejemplo: la nueva salsa choke) para hacer diversa la base de conocimientos. Pensamos que cincuenta mil canciones era un buen número; no obstante solo adquirimos 24.108 canciones de colecciones privadas.

Para decidir el formato de audio observamos las bases de datos existentes y encontramos que The Million Song Dataset trabajaba con el format HDF5 (The HDF Group, 2016). Se nos ocurrió, entonces, hacer una extensión a esta base de datos (tener The Million and Twenty Four Thousand A Hundred and Eight Song Dataset), pero en conversaciones con asesores internacionales expertos en recuperación de información acústica nos dimos cuenta de que este formato no era adecuado para el propósito futuro, es decir, la caracterización mediante descubrimiento de conocimiento. En consecuencia, miramos formatos que hubiesen sido usados en sistemas de minería de datos populares como Weka (Holmes, Donkin y Witten, 1994) y RapidMiner (Jungermann, 2009). Constatamos que los formatos XML y ARFF eran adecuados para dicho propósito.

La decisión de cuáles características debíamos extraer de las canciones fue relativamente fácil. La mayoría de sistemas computacionales que hacen análisis de señales de audio (*jAudio*, *Essentia*, *Marsyas*, *Yaafe*, *MIRToolbox*, etc.) extraen las mismas características. Las más conocidas son: espectro de potencia (*power spectrum*), magnitud de espectro (*magnitude spectrum*), variabilidad de espectro (*spectral variability*), centroide espectral (*spectral centroid*), compacidad (*compactness*), fluencia espectral (*spectral flux*), (Coeficientes cepstrales de frecuencia de tipo mel (*mel-frequency cepstral coefficients* o MFCC), raíz cuadrática media (*root mean square* o RMS) y codificación lineal predictiva (*linear predictive coding* o LPC). De esta manera decidimos extraer todas las características que pudiéramos. El lector interesado en conocer dichas características puede referirse a Mitrovic, Zeppelzauer y Breiteneder, 2010).

Por último, con respecto a la forma en que pensamos almacenar la información, dada nuestra experiencia previa con *Oracle Multimedia* (Mora, 2013) inicialmente pensamos usar dicho conocimiento y almacenar la información extraída de las canciones en una base de datos. El problema de esta idea era que la base de datos no podía ser consultada por cualquier persona en el mundo, que es nuestro interés principal, así que decidimos dejar la base de conocimiento en archivos planos que puedan ser abiertos con un editor de texto normal.

Una vez que seleccionamos todos nuestros parámetros iniciales hicimos un análisis preliminar de nuestro conjunto de canciones y encontramos que algunas de ellas estaban repetidas, así que nuestra primera tarea fue encontrar todas las repeticiones y descartarlas. Como el conjunto de canciones fue adquirido de diferentes colecciones privadas, encontramos un poco más de 5.000 canciones repetidas. Cabe anotar que nosotros consideramos una canción repetida como aquella que tiene exactamente el mismo archivo, al contrario de lo que al respecto consideran otros autores que definen que una canción es la repetición de otra si es la misma obra pero (1) cantada por otro

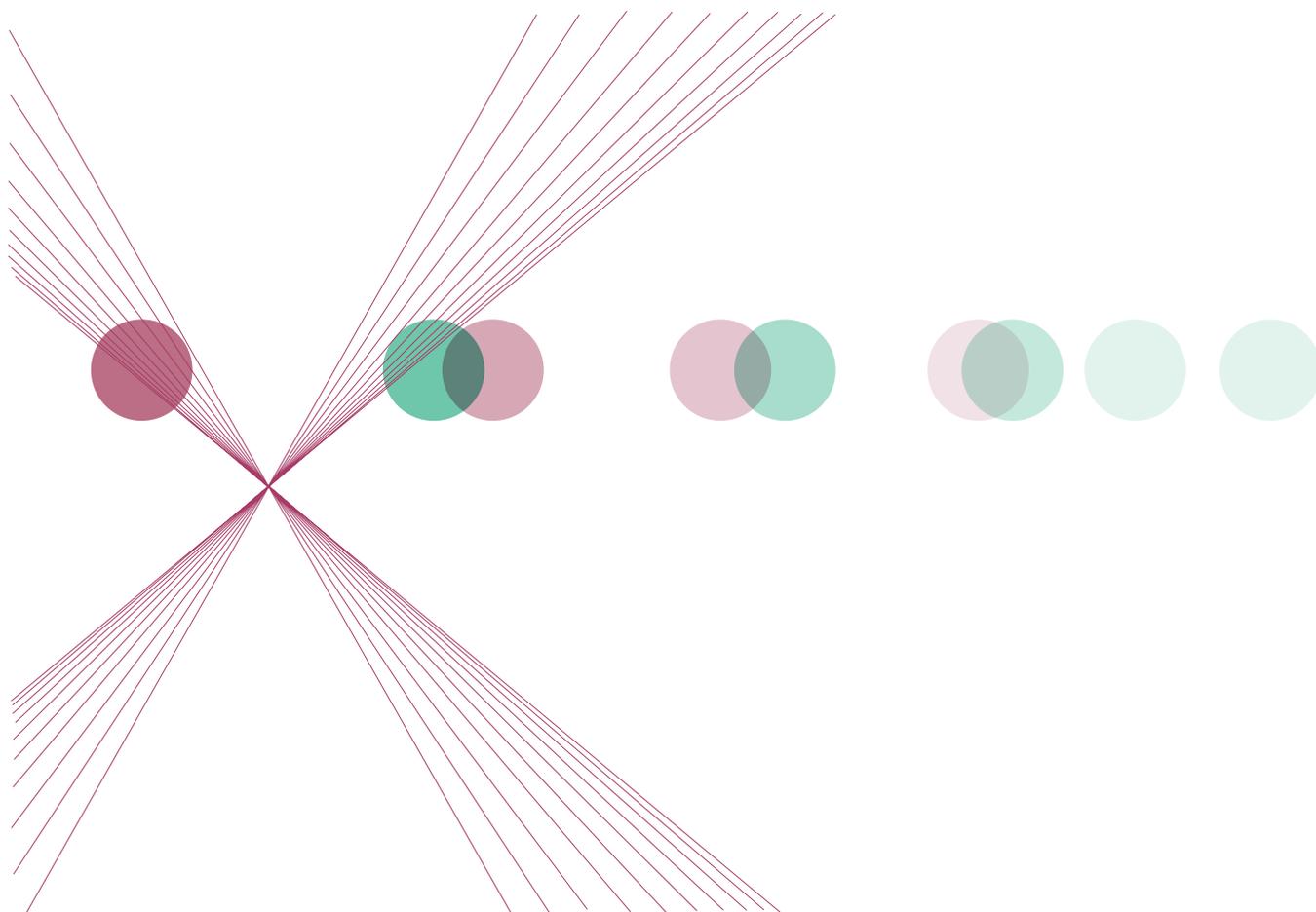
artista, (2) cantada en vivo (o en estudio, dado el caso), o (3) con diferentes arreglos musicales. En beneficio de la futura caracterización no seguimos la definición anterior, y para evitar las mismas fallas que tuvo la base de datos GTZAN, expuestas por Bob L. Sturm (Sturm, 2012), no consideramos canciones repetidas.

Una vez con el conjunto final de canciones por incluir en la base de conocimiento, procedimos a hacer la extracción de características. En este punto tuvimos serios inconvenientes con tamaños de los archivos de audio (i.e. la longitud de las canciones). Cuanto más largas eran las canciones, más prolongado el tiempo de ejecución de los algoritmos de análisis. Consultamos otras bases de datos y encontramos que ninguna considera las canciones completas sino una parte de ellas. The Latin Music Database, por ejemplo, toma 30 segundos de una canción (al principio, en la

mitad y al final). Esta aproximación no nos pareció adecuada puesto que es posible que dichas partes de una canción no sean representativas. Después de analizar muy bien el género salsa encontramos que la parte más representativa es tal vez el coro. Sin embargo, el mismo no está en la misma ubicación (en el sentido cronológico) en todas las canciones, así que había que encontrarlo. El desarrollo del algoritmo de detección de coros de salsa puede verse en Arévalo, Sarria M., Mora y Arce-Lopera, 2015).

Las características de audio se extrajeron a partir de diez segundos del coro de cada canción y con ellas se construyó la base de conocimiento, que cuenta entonces como más de 18.000 canciones y puede consultarse en la siguiente página

web: <http://salsadataset.com/>



Conclusiones y trabajo futuro

Nuestro objetivo general en este proyecto era construir un sistema computacional de descubrimiento de conocimiento para caracterizar el género de música salsa mediante la utilización de técnicas avanzadas de recuperación de información acústica e inteligencia artificial. Para lograrlo dividimos el proyecto en tres etapas.

En la primera se consultó con los asesores internacionales expertos en recuperación de información acústica acerca del estado del arte en caracterización y evolución musical de otros géneros musicales en el mundo. Con esta revisión se buscaba identificar las herramientas tecnológicas y los algoritmos necesarios que deberían implementarse en el proyecto. Quisimos garantizar que las técnicas y herramientas seleccionadas al final de esta etapa fueran relevantes y permitieran la comparación directa con investigaciones actuales en el área científica de recuperación de información musical.

La segunda etapa era fundamental para el proyecto porque determinaba la calidad de sus resultados. La base de conocimiento de las canciones de salsa debía ser de alta calidad puesto que el sistema computacional analiza las señales acústicas para derivar descriptores pertinentes para ser analizados. Además, y como requerimiento para el uso de técnicas de inteligencia artificial, esta base de datos debía ser extensa y con variedad en las canciones en términos de su fecha de producción, de su ubicación geográfica y de su pertenencia a subgéneros de la salsa.

Nos encontramos en este momento en la iniciación de la tercera etapa, que consiste en la implementación de un algoritmo de inteligencia artificial como son, por ejemplo, KNN (*K nearest neighbors*), K-Means, redes de Bayes, EM (*expectation-maximization*), SVM (*support vector machine*) y árboles de decisión, como modelo para el descubrimiento de conocimiento. Se evaluaron diferentes aproximaciones y se escogió la más adecuada, según precisión, carga computacional, simplicidad y expresividad del modelo y relación con otros modelos de percepción musical, entre otras variables. Después se diseñaron las posibles consultas para que los resultados puedan ser analizados e interpretados en forma coherente con el propósito del proyecto. Como por ejemplo, se debió investigar cuál era la consulta más relevante para caracterizar la evolución de la salsa.

Una vez finalizado el desarrollo de este sistema queremos de igual manera hacer comparaciones de los resultados que arroje el sistema con expertos humanos (i.e. *salsómanos*) a manera de prueba de Turing e ir un poco más allá para encontrar relaciones con otras artes, como, por ejemplo, la danza: ¿por qué la salsa se baila de esta manera?, ¿cómo ha sido la evolución del baile de la salsa con respecto a la propia del género?

Referencias

- Arévalo, C., Sarria M., G. M., Mora, M. J., & Arce-Lopera, C. (2015). Towards an efficient algorithm to get the chorus of a salsa song. En *Proceedings of IEEE International Symposium on Multimedia (ISM2015)*, Miami, FL.
- Aucouturier, J.-J., & Pachet, F. (2003). Representing musical genre: a state of the art. *Journal of New Music Research*, 32(1), 83-93.
- Bertin-Maheux, T., Ellis, D. P. W., Whitman, B., & Lamere, P. (2011). The million song dataset. En *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR2011)*, Miami, FL.
- Cios, K. J., Pedrycz, W., Swinarski, R. W., & Kurgan, L. A. (2007). *Data mining: a knowledge discovery approach*. Nueva York: Springer.
- Cooper, M., & Foote, J. (2003). Summarizing popular music via structural similarity analysis. En *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. New Paltz, NY.
- Duany, J. (1984). Popular music in Puerto Rico: toward an anthropology of "salsa". *Revista de Música Latino Americana*, 5(2), 186-216.
- Echonest (2016). *Echonest*. Recuperado el ... de ... de ..., de: <http://developer.echonest.com>
- Fitch, W. T. (2005). The evolution of music in comparative perspective. *Annals of the New York Academy of Sciences*, 1060(1), 29-49.
- Holmes, G., Donkin, A., & Witten, I. H. (1994). *Weka: a machine learning workbench*. Hamilton, New Zealand: University of Waikato, Department of Computer Science, working paper 94/09.
- Isaacson, E. (2005). What you see is what you get: on visualizing music. En *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*. Londres, UK.
- Ismir (2004). *ISMIR2004 audio description contest-genre/artist ID classification and artist similarity*. Recuperado el ... de ... de ..., de: http://ismir2004.ismir.net/genre_contest/
- Jungermann, F. (2009). Information extraction with rapidminer. En *Proceedings of the GSCL Symposium 'Sprachtechnologie und eHumanities'*. Essen, Alemania.
- Li, T., & Li, L. (2012). Music data mining: an introduction. En Li, T., Ogihara, M., & Tzanetakis, G. (ed.). *Music data mining*, pp. ...-... Nueva York: CRC Press.
- Li, T., Ogihara, M., & Zhu, S. (2006). Integrating features from different sources for music information retrieval. En *Proceedings of the 6th International Conference on Data Mining*, Hong Kong, China.
- McDermott, J. (2008). The evolution of music. *Nature*, 453(7193), 287-288.
- McKay, C., McEnnis, D., & Fujinaga, I. (2006). A large publicly accessible prototype audio database for music research. En *7th International Conference on Music Information Retrieval (ISMIR2006)*, Victoria, Canadá.
- Mitrović, D., Zeppelzauer, M. & Breiteneder, C. (2010). Features for content-based audio retrieval. *Advances in Computers*, 78, 71-150.
- Mora, M. J. (2013). *Recuperación basada en contenido de archivos de sonido Mpeg-7 sobre Oracle multimedia*. Cali: Pontificia Universidad Javeriana, trabajo de grado de maestría.
- Morales, E. (2003). *The latin beat: the rhythms and roots of latin music from bossa nova to salsa and beyond*. Boston, MA: Da Capo Press.
- Orio, N. (2006). Music retrieval: a tutorial and review. *Foundations and Trends in Information Retrieval*, 1(1), 1-90.
- Orio, N., Rizo, D., Miotto, R., Montecchio, N., Schedl, M. & Lartillot, O. (2011). Musiclef: a benchmark activity in multimodal music information retrieval.

En *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR2011)*, Miami, FL.

Padilla, F. M. (1989). Salsa music as a cultural expression of Latino consciousness and unity. *Hispanic Journal of Behavioral Sciences*, 11(1), 28-45.

Pampalk, E. (2006). *Computational models of music similarity and their application in music information retrieval*. Viena: Technische Universität Wien, disertación doctoral.

Peeters, G. (2007). Sequence representation of music structure using higher-order similarity matrix and maximum-likelihood approach. En *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, Viena,

Proud Music (2016). *Proud Music*. Recuperado el ... de ... de ..., de: <http://www.proudmusiclibrary.com/en/tag/salsa/>

Rauber, A., Pampalk, E., & Merkl, D. (2002). Content-based music indexing and organization. En *Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 409-410. Nueva York: ACM.

Schnitzer, D., Flexer, A., & Widmer, G. (2009). A filter-and-refine indexing method for fast similarity search in millions of music tracks. En *Proceedings of the 10th International Conference on Music Information Retrieval (ISMIR 2009)*, Kobe, Japan.

Silla Jr., C. N., Koerich, A. L., & Kaestner, C. A. A. (2008). The Latin music database. En *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR 2008)*, Philadelphia, PA.

Sturm, B. L. (2012). An analysis of the gtzan music genre dataset. En *2nd International ACM Workshop on Music Information Retrieval with User-Centered and Multi-modal Strategies (MIRUM2012)*, held as part of the *ACM Multimedia 2012*, Nara, Japón.

The HDF Group (2016). *Welcome to the HDF5 home page!* Recuperado el ... de ... de ..., de: <https://www.hdfgroup.org/HDF5/>

Tzanetakis, G. & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5), 293-302.

Wold, E., Blum, T., Keislar, D., & Wheaton, J. (1996). Content-based classification, search, and retrieval of audio. *IEEE Multimedia*, 3(3), 27-36.

Xiao, L., Tian, A., Li, W., & Zhou, J. (2008). Using a stochastic model to capture the association between timbre and perceived tempo. En *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR 2008)*, Philadelphia, PA.