

UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingeniería y Sistemas de
Telecomunicación



Observatorio virtual sonoro basado en Deep
Learning / Deep Learning Auditory Virtual
Observatory

TESIS DOCTORAL

Presentada para optar al título de Doctor por:

Adrián García Riber
Máster en Investigación Musical

Madrid, 2025



UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingeniería y Sistemas de
Telecomunicación

Doctorado en Música y su Ciencia y Tecnología

Observatorio virtual sonoro basado en Deep Learning / Deep Learning Auditory Virtual Observatory

TESIS DOCTORAL

Presentada para optar al título de Doctor por:

Adrián García Riber
Máster en Investigación Musical

Bajo la dirección de:
Dr. Francisco Serradilla García

Madrid, 2025

Título: Observatorio virtual sonoro basado en Deep Learning / Deep Learning Auditory
Virtual Observatory

Autor: Adrián García Riber

Programa de Doctorado: Música y su Ciencia y Tecnología

Dirección de Tesis:

Dr. Francisco Serradilla García, PhD and Full Professor at ETSISI (Director)

Revisores Externos:

Tribunal de Tesis:

Fecha de Defensa de Tesis:

Actualizando el legado de Pitágoras y Kepler, este trabajo propone un camino para crear música de las esferas. Imitando a Xenakis, se trata de un instrumento para generar sonidos a partir de luz. Parafraseando a Cage, es una herramienta para generar datos espectrales a partir de datos astronómicos.

Agradecimientos

Este trabajo está dedicado a mi familia, lejos y cerca, siempre aquí. A Ruth, por atreverse. A Miquel, por el horizonte de sucesos.

Un agradecimiento especial a Francisco Serradilla y a Rubén García Benito por toda la confianza, apoyo, orientación y contribución a este trabajo, y a Enrique Pérez Montero, Bruce Walker y Scott Flemming por su inestimable ayuda y colaboración. Gracias también a Edith Alonso, Katharina A. Lutz, Enrique Solano, Miriam Cortes-Contreras, Tom Donaldson, Amelia Ortiz, Helena Bas Cañas y Niklas Rönnberg, por estar ahí en momentos clave a lo largo de estos años. Por último, pero no menos importante, gracias a todas las personas voluntarias y anónimas que hicieron posible esta investigación, probando las aplicaciones, participando en las encuestas, y revisando los artículos publicados durante el transcurso de esta tesis.

Esta investigación utilizó datos de los servicios de librerías STELIB, MILES y CaT desarrollados por el Observatorio Virtual Español en el marco del Grupo de Trabajo de la Comisión G5 de la IAU: Bibliotecas Espectrales Estelares.

También utilizó datos recopilados por las misiones TESS y Kepler, obtenidos del archivo de datos MAST del Space Telescope Science Institute (STScI). La financiación para las misiones TESS y Kepler proviene del Programa Explorador y de la Dirección de Misión Científica de la NASA. STScI es operado por la Asociación de Universidades para la Investigación en Astronomía (AURA), Inc., bajo el contrato NAS 5-26555 de la NASA.

Además, se utilizaron las curvas de luz del archivo CoRoT. La misión CoRoT, lanzada el 27 de diciembre de 2006, fue desarrollada y operada por el CNES, con la participación de los Programas Científicos de la ESA, el RSSD de la ESA, Austria, Bélgica, Brasil, Alemania y España.

Finalmente, se utilizaron datos proporcionados por Calar Alto Legacy Integral Field Area (CALIFA) survey (<https://califa.caha.es/>). Basado en observaciones recogidas en el Centro Astronómico Hispano Alemán (CAHA) en Calar Alto, operado conjuntamente por el Max-Planck-Institut für Astronomie y el Instituto de Astrofísica de Andalucía (CSIC).

Abstract

Thanks to the efforts and cooperation of the international community, nowadays it is possible to analyze astronomical data captured by the observatories and telescopes of major space agencies around the world from a personal computer. The development of Virtual Observatory (VO) technology and the standardization of the formats it uses, allow professional and amateur astronomers to access astronomical data and images through the internet with relative ease. Immersed in this context of global accessibility, present and future massive astronomical data archives represent a great opportunity to consolidate the use of sound in the analysis and communication of scientific information. Sonification provides additional dimensions to graphical representations, enhances the accessibility of the archives, offers alternative display strategies to specific case studies, and brings scientific concepts closer to potential audiences. In addition, the use of astronomical data for creative purposes in musical applications and sound art allows the creation of original compositions and sound designs.

This thesis offers a wide collection of sonification and musification strategies implemented to analyze the implications and potential goals of future developments towards a proposal for an Auditory Virtual Observatory. This proposal could be useful not only in educational and outreach activities, but also in scientific research, providing complementary analysis perspectives and improving accessibility for blind and low vision users (BLV). Supported by the implementation of experimental prototypes that allow the multimodal -graphical and auditory- exploration of public access databases and stellar catalogs, this work offers an extended theoretical and practical analysis of the current sonification techniques, as well as the application of deep learning to Sonification and Music composition, using real sky case studies from the Space Telescope Science Institute (STScI), the Spanish Virtual Observatory (SVO), the Barbara A. Mikulski Archive for Space Telescopes (MAST), the NASA exoplanet Archive, the Convection, Rotation, and planetary Transits (CoRoT) archive, and the Calar Alto Legacy Integral Field Area (CALIFA) survey. Mainly based on automation, parameter mapping, and deep learning algorithms, this work provides a collection of multimodal analysis tools ranging from simple mono auditory graphs, to interactive binaural prototypes based on autoencoders, going through complete data-driven music composition systems using neural networks with attention mechanisms.

This research also includes three user studies with qualitative and quantitative feedback from more than 150 participants experienced and non-experienced in Astronomy and Music. The qualitative results show a good acceptance of the proposals, with a general tendency to rate higher the representations using musical forms, than those using abstract representations. Although the quantitative results should be considered indicative, they suggest that all participants could retrieve useful information from the sonifications, pointing to experience in the analysis of sound events as more helpful than previous knowledge of the data, and highlighting the importance of training and attention to detail for the understanding of complex auditory information. All these results are expected to provide potential guidelines to inspire future designs of widely accepted auditory representations for scientific analysis and outreach.

Resumen

Gracias al esfuerzo y la cooperación de la comunidad internacional, hoy en día es posible analizar los datos astronómicos capturados por los observatorios y telescopios de las principales agencias espaciales de todo el mundo desde un ordenador personal. El desarrollo de la tecnología del Observatorio Virtual (VO) y la estandarización de sus formatos, permiten a astrónomos profesionales y aficionados acceder a datos e imágenes astronómicas, a través de internet, con relativa facilidad. Dentro de este contexto de accesibilidad global, los archivos masivos de datos astronómicos presentes y futuros representan una gran oportunidad para consolidar el uso del sonido en el análisis y la comunicación de información científica. La sonificación proporciona dimensiones adicionales a las representaciones gráficas, mejorando la accesibilidad de los archivos, ofreciendo estrategias de visualización alternativas en casos de estudio específicos y acercando la ciencia a todos los públicos. Además, el uso de datos astronómicos con fines creativos en aplicaciones musicales y de arte sonoro permite la creación de composiciones y diseños sonoros originales.

Esta tesis ofrece una amplia colección de estrategias de sonificación y musificación implementadas para analizar las implicaciones y objetivos potenciales de futuros desarrollos hacia una propuesta de Observatorio Virtual Sonoro, que podría ser útil no sólo en actividades educativas y de divulgación, sino también en investigación científica, proporcionando perspectivas de análisis complementarias y mejorando la accesibilidad para usuarios ciegos o con visión reducida (BLV). Apoyada en la implementación de prototipos experimentales que permiten la exploración multimodal –gráfica y sonora– de bases de datos y catálogos estelares de acceso público, se ofrece un amplio análisis teórico y práctico de las técnicas actuales de sonificación, así como de la aplicación del deep learning a la sonificación y a la composición musical, utilizando casos de estudio reales del Space Telescope Science Institute (STScI), del Observatorio Virtual Español (SVO), del Archivo de Telescopios Espaciales Barbara A. Mikulski (MAST), del Exoplanet Archive de la NASA, del archivo Convection, Rotation, and planetary Transits (CoRoT), y del estudio Calar Alto Legacy Integral Field Area (CALIFA). Basadas principalmente en automatización, mapeado de parámetros y algoritmos de deep learning, las herramientas de análisis desarrolladas van desde simples gráficos sonoros mono, hasta prototipos binaurales interactivos basados en autoencoders, pasando por completos sistemas de composición musical formados por redes neuronales con mecanismos de atención.

Esta investigación también incluye tres estudios cualitativos y cuantitativos con más de 150 participantes con y sin experiencia en Astronomía y Música. Los resultados cualitativos muestran una buena aceptación de las propuestas, y una tendencia general a valorar mejor las representaciones que utilizan formas musicales frente a las que utilizan representaciones sonoras abstractas. Aunque los resultados cuantitativos no son concluyentes, sugieren que todos los participantes pudieron recuperar información útil de las sonificaciones, señalando que la experiencia en el análisis de eventos sonoros es más útil que el conocimiento previo de los datos, y destacando la importancia del entrenamiento y la atención al detalle en la comprensión de información sonora compleja. Se espera que estos resultados proporcionen pautas potenciales que puedan inspirar diseños futuros de sonificaciones ampliamente aceptadas para el análisis y la divulgación científica.

Tabla de Contenido

Agradecimientos	v
Abstract	vi
Resumen	vii
Lista de Figuras	xii
Lista de Tablas	xxi
Abreviaturas y Acrónimos	xxv
1 Introducción	1
1.1 Sonificación y musificación	4
1.2 Objetivos y preguntas de investigación	5
1.3 Estructura de la tesis	6
2 Precedentes y estado del arte	9
2.1 Breve historia sobre armonías	9
2.2 Visión general del Observatorio Virtual	12
2.3 Análisis científico y divulgación de catálogos astronómicos	14
2.4 Sonificación de datos astronómicos	15
2.5 Música algorítmica e inteligencia artificial	17
2.6 Tendencias actuales en composición musical automatizada	23
3 Exploración multimodal de catálogos estelares	25
3.1 Curvas de luz y espectros estelares	26
3.2 La librería STELIB	26
3.3 El módulo FITS2OSC	27
3.4 Sonificación de curvas de luz con FITS2OSC	28
3.5 Sonificación de espectros estelares con Astronify	29
4 Paisajes sonoros multivariable	31
4.1 Archivos de validación de datos en series temporales de TESS	31
4.2 Sonificación multivariable secuencial	32
4.3 Explorador DVT	33
4.4 Estética interdisciplinar	35
4.5 Objetos de interés de TESS	35
5 Musificación del análisis de periodograma en curvas de luz	39

5.1	Objetos de interés del catálogo Kepler	40
5.2	Generación de partituras a partir del análisis de periodograma	40
5.3	Síntesis de audio a partir de curvas de luz	42
6	Sonificación simbólica espacializada	45
6.1	El archivo CoRoT	45
6.2	Mapeado simbólico	46
6.3	Representación multimodal a tiempo real con pre-procesado	50
6.4	Espacialización binaural y ambisónica	51
7	Sonificación basada en deep learning	53
7.1	La librería MILES	54
7.2	Autocodificación de espectros estelares	54
7.3	Sonificación del espacio latente generado por sparse autoencoders	56
7.4	Espacialización del espacio latente generado por autoencoders variacionales	59
8	Sistema de composición musical basado en deep learning	61
8.1	Conversión de la librería MILES en acordes musicales	61
8.2	Secuencia de acordes OBAFGKM	63
8.3	Redes LSTM con atención	63
8.4	Sistema de composición musical basado en deep learning	66
8.5	Composiciones no supervisadas a partir de espectros estelares	68
9	Instrumentos musicales para la exploración de catálogos astronómicos	71
9.1	La librería CaT	72
9.2	Comparativa de autoencoders con la librería CaT	72
9.3	Análisis espectral por detección de umbral	74
9.4	Mapeado del espacio latente	75
9.5	Diseño e implementación del sintetizador de sonido	76
9.6	Sistema de acompañamiento para piano a partir de espectros estelares	77
10	Espectroscopía de Campo Integral (IFS) Multimodal Interactiva	81
10.1	Sobre CALIFA	81
10.2	Autocodificación de CALIFA	82
10.3	<i>ViewCube</i> descripción general	84
10.4	Implementación de <i>SoniCube</i>	85
11	Evaluación	89
11.1	Primer estudio. Análisis cuantitativo de sonificaciones de espectroscopía de transmisión y análisis cualitativo de musificaciones	90
11.1.1	Diseño del estudio I	90
11.1.2	Resultados cuantitativos. Reconocimiento sonoro de líneas de transmisión	91
11.1.3	Resultados cualitativos. Musificación	94
11.2	Segundo estudio. Evaluación de la sonificación simbólica y de la sonificación mediante deep learning	95
11.2.1	Diseño del estudio II	96

11.2.2	Resultados cuantitativos. Sonificación simbólica	96
11.2.3	Resultados cuantitativos. Sonificación basada en deep learning	100
11.2.4	Análisis comparativo. Sonificación simbólica - sonificación mediante deep learning	103
11.2.5	Resultados cualitativos. Estudio II	104
11.3	Tercer estudio. Evaluación de ViewCube. IFS multimodal interactiva	105
11.3.1	Diseño del estudio III	105
11.3.2	Resultados generales. Estudio III	107
11.3.3	Resultados cuantitativos por subgrupos. Estudio III	108
11.3.4	Resultados cualitativos. Estudio III	114
12	Discusión	117
12.1	Sobre la estética	117
12.2	Sobre la sonificación de espectros de transmisión	118
12.3	Sobre la sonificación simbólica	119
12.4	Sobre la sonificación basada en deep learning	121
12.5	Sobre el análisis comparativo de metodologías de sonificación	122
12.6	Sobre la espacialización de las sonificaciones	122
12.7	Sobre las composiciones musicales generadas a partir de datos astronómicos	123
12.8	Sobre la utilización de instrumentos musicales en la exploración de catálogos	123
12.9	Sobre las herramientas multimodales interactivas	124
13	Conclusión	127
	Referencias	133
	Anexos	161
.1	Formato de archivos FITS	161
.2	El protocolo OSC	161
.3	Tránsitos planetarios	162
.4	Clasificación de espectros estelares	163
.5	Temperatura efectiva, diagrama HR, metalicidad y gravedad superficial	166
.6	Espectroscopía de transmisión sonificada	167
.7	Mapeado de parámetros formalizado	168
.8	Descripción formal de un autoencoder	169
.8.1	Sparse autoencoders	169
.8.2	Autoencoders variacionales	169
.8.3	Divergencia de Kullback-Leibler	170
.9	Descripción formal de una red LSTM	171
.10	Mecanismo de atención en redes LSTM	173
.11	Descripción formal de un sintetizador	173
.12	Codificación binaural	175
.13	Codificación ambisónica	175
.14	Coficiente de determinación	176
.15	Análisis de correlación	177

.16	Intervalos de confianza de Jeffreys	180
.17	Listado de aplicaciones desarrolladas	180
.18	Composiciones musicales generadas	182
.19	Material Complementario	183
.20	Música relacionada con la astronomía	183
.21	Código fuente	183
.22	Partitura musical	191

Lista de Figuras

1.1	Diagrama de Venn que representa el espacio interdisciplinar en el que se enmarca este trabajo. Intersección entre la sonificación y la aplicación del deep learning al análisis científico y a la creación artística.	2
2.1	Extracto del <i>Libro V</i> de <i>Harmonices Mundi</i> , Johanes Kepler, 1619 (Kepler, 1969, p.207). Representaciones musicales de Saturno, Júpiter, Marte, la Tierra, Venus, Mercurio y la luna.	10
2.2	Luz natural en casa y diagrama de la analogía color-música descrita por Newton en su segundo artículo sobre luz y colores, presentado a la Royal Society en 1675 (Pesic, 2006, p. 294).	11
2.3	Izquierda: Organización de la infraestructura del Observatorio Virtual Nacional (NVO) en 2003 (Quinn y Górski, 2004), proyecto liderado por el investigador principal Dr. Alexander Szalay (Universidad Johns Hopkins) y el co-investigador principal Dr. Roy Williams (Instituto Tecnológico de California). Derecha: Arquitectura IVOA que incluye los estándares en 2010 (Arviset y Gaudet, 2010).	12
2.4	Diagrama HR de Kepler 1661 obtenido con VOSA (SVO).	13
2.5	Espacio circular Ars Informatica-Ars Electronica (P. Vickers y Hogg, 2006).	17
2.6	Neumas. Extracto de <i>Micrologus</i> , Guido d'Arezzo, 1030. Capítulo XVII, <i>Ut queant laxis</i> . (Paris, Bibliothèqne nationale, Lat. 7211, f. 85v. XII ^e siècle, Sud de la France)(D'Arezzo, 1996, p. 81).	18
2.7	Extracto del manuscrito con la matriz de vectores de <i>Achorripsis</i> , 1957 (Xenakis, 1992, pp. 28-135). Las filas representan las clases tímbricas o agrupaciones instrumentales, las columnas representan bloques temporales o secciones de la pieza, que consta de 7 agrupaciones tímbricas intercambiables y 28 columnas o segmentos de tiempo.	20
3.1	Curva de luz de la estrella variable KIC 11013096 con coordenadas RA:18:47:44.69 y DEC:+48:30:29.63. Variaciones de flujo PDCsap (electrones por segundo) en función del tiempo (BKJD = BJD-2454833). STScI.	25
3.2	Espectro de la estrella HD 097916 (RA:11:15:54.23 DEC:+02:05:12.09). Variaciones de flujo normalizadas (ADU) en función de la longitud de onda (Å). SVO.	26
3.3	Diagrama de bloques del módulo de sonificación <i>FITS2OSC</i>	28

3.4	Curvas de luz representando las variaciones de flujo del sistema estelar binario KIC 6504534 con <i>FITS2OSC</i> . Flujo PDCsap (electrones por segundo) en función del tiempo en BKJD (BJD-2454833). Más información en NASA, 2011b , 2016 .	28
3.5	Espectro tipo <i>O</i> de HD 057060 con coordenadas RA: 07:18:40.38 DEC:-24:33:31.32. Representación gráfica de la división en bandas de color utilizada en el módulo de sonificación de espectros de <i>Astronify</i> . Flujo normalizado frente a longitud de onda (Å).	30
4.1	Curva de luz y modelo de tránsito para TIC 220479565, sub-Neptuno excéntrico transitando una estrella enana tipo M2 (Cointepas et al., 2021). Flujo relativo en función de la fase orbital ajustada al momento del tránsito.	32
4.2	Diagrama de bloques del <i>Explorador DVT</i> basado en el módulo <i>FITS2OSC</i> .	33
4.3	Captura de pantalla de la interfaz de usuario del explorador DVT durante la sonificación de la estrella variable de periodo largo TIC 80081010, con coordenadas RA: 22:03:16.76 DEC: -31:26:42.64.	34
4.4	Mapa de cobertura del estudio AstraLux-TESS (Lillo-Box et al., 2024). Software: Aladin, proyección de Aitoff.	36
5.1	Campo de visión completo de 116 grados cuadrados (FOV) del único campo observado durante la misión inaugural de Kepler (izquierda) y captura de imagen durante la misión Kepler (derecha) (NASA, 2020a).	40
5.2	Dos ejemplos del análisis de periodograma Box Least Squares de las curvas de luz del catálogo KOI (Thompson et al., 2018)(STScI). Los gráficos muestran cada curva de luz (flujo SAP en función del tiempo en JD), su análisis de periodograma (Energía en función del periodo en días) y las curvas plegadas alrededor del periodo de mejor ajuste (flujo normalizado en función del tiempo en días), para los objetos Kepler 006891512, que presenta un periodo principal de 1,747 días (izquierda), y Kepler 008292840, con el mejor ajuste en 2,969 días (derecha).	41
5.3	Primeros 16 compases de la pieza musical <i>KOI 400-800</i> , generada a partir del análisis del periodograma BLS de 400 curvas de luz KOI.	42
5.4	Dos capturas de pantalla de la interfaz <i>LC2Music</i> durante la sonificación del catálogo KOI. Izquierda: curva de luz de kplr008758161, falso positivo de las disposiciones de Kepler y NExScI (STScI, 2013) con un periodo principal claro cercano a 1 día. Derecha: curva de luz de kplr008719419 (KOI-3506), candidato a falso positivo del NASA Exoplanet Archive (NASA, 2011a).	42
5.5	Mapa de cobertura de 289 estrellas del KIC albergando al menos un planeta (Bastien, 2014). Software: Aladin, proyección de Aitoff.	43
6.1	Curva de luz CoRoT (flujo blanco en electrones/32s en función del tiempo en JD) y metadatos utilizados en la sonificación de CHR 0223959618. El prototipo extrae las coordenadas, tipo de espectro, luminosidad, clase y valores característicos de cada curva de luz. Además, calcula el periodo principal encontrado en la curva, en este caso de 3,96 días.	46

6.2	Diagrama HR de 452 objetos de la librería CoRoT, ilustrando la conversión simbólica de tipos y clases espectrales propuesta. Temperatura efectiva (K) y rangos de los tipos espectrales frente a Magnitud absoluta y Luminosidad con referencia solar.	47
6.3	VARIABLES extraídas del archivo FITS de cada objeto. El esquema incluye la descripción de cada función (naranja), metadatos y mapeado (azul) y los archivos de salida generados a partir de cada curva de luz.	48
6.4	Diagrama de bloques del sintetizador de sonificación implementado en <i>Csound</i> y <i>Cabbage</i> . Las líneas continuas representan conexiones MIDI, de audio y de imagen. Las líneas discontinuas representan conexiones OSC.	49
6.5	Mapa de cobertura del archivo IAS CoRoT que muestra los objetos (puntos azules) y la posición de referencia (centroide) de las coordenadas de espacialización (cruz roja), ubicada en RA 100.2196 DEC 9.6773 (grados).	50
7.1	Mapa de cobertura de la librería de espectros estelares MILES. Software: Aladin, proyección de Aitoff.	54
7.2	Ilustración de un modelo de autoencoder aplicado a espectros estelares. Para la reducción de los 4367 valores de cada espectro a un tensor del espacio latente de 10 ejes, el codificador utiliza dos capas ocultas con 2.099.350 parámetros. Para la reconstrucción del espectro decodificado, el decodificador utiliza dos capas ocultas con 1.913.175 parámetros.	55
7.3	Comparación gráfica de los resultados obtenidos con un autoencoder variacional de cuatro dimensiones y cuatro capas entrenado con 985 espectros. Espectro de entrada (izquierda) y salida decodificada (derecha) para la estrella HD184406 de tipo <i>K</i> , con coordenadas RA:19:34:05.35 y DEC:+07:22:44.18. Desviación $R^2 = 0.9965$. Librería MILES, IAC.	56
7.4	Comparación de los espectros original y decodificado a la salida de un autoencoder variacional de 4 dimensiones (izquierda) y gráfico de cuantiles (derecha). Coeficiente de determinación R^2 para la codificación de HD 057061 (arriba) y HD 026965 (abajo). Librería MILES, IAC.	57
7.5	Interfaz gráfica del prototipo VAE2Sound. Espectro original (izquierda) y espectro decodificado (derecha) para la estrella HD 040657 con coordenadas RA:06:00:03.35 y DEC:-03:04:27.33. Representación adicional mediante círculos de los altavoces activos (dimensiones del espacio latente) que generan la sonificación.	60
8.1	Arquitectura simplificada del codificador variacional utilizado para extraer las características latentes de cada espectro estelar de la librería MILES. Los números indican la dimensión de cada capa <i>dense</i> utilizada, y el espacio latente final de cuatro dimensiones. El acorde de muestra G4 G4 D5 A#4 se genera a partir del espectro de la estrella BD+15 1305, con coordenadas (J2000) RA 06:44:42.98 DEC 14:54:35.97.	62
8.2	Resultados del autoencoder para la estrella HD 003369 tipo B5V y para la HD 114606 G1V. Espectro original (izquierda) y salida decodificada del autoencoder variacional 4D (derecha) en azul. Error residual en naranja.	64

8.3	Diagrama de red que incluye el tipo y la dimensión de cada capa para la arquitectura LSTM dual con atención utilizada en el sistema de composición musical.	64
8.4	Partituras resultantes y función de pérdida (sparse categorical cross-entropy) para la alturas de nota o pitch (azul) y para las duraciones de nota (rojo) de cuatro experimentos con diferentes conjuntos de entrenamiento: un sólo motete de Lasso (arriba a la izquierda), la misma pieza repetida 100 veces (arriba a la derecha), 50 piezas musicales de Des Prez y Lasso (abajo a la izquierda) y 1318 piezas de Palestrina (abajo a la derecha). Vale la pena mencionar cómo la función de pérdida de altura de nota (pitch) refleja una tasa de aprendizaje inesperadamente rápida para el corpus de Palestrina (el que tiene más piezas), probablemente motivada por las características musicales altamente deterministas del compositor.	65
8.5	Sistema de composición musical de cuatro voces a partir de datos astronómicos basado en deep learning. Diagrama de bloques de la arquitectura completa basada en VAE-LSTM con atención y Pitch Class Set Theory. Los espectros estelares se convierten en acordes de cuatro notas (Partitura 1) utilizando un autoencoder variacional. En un proceso paralelo, una red neuronal LSTM con atención entrenada con el corpus MAESTRO genera la Partitura 2, activada por una pieza musical de entrada (<i>In me transierunt</i>). Finalmente, el algoritmo de análisis basado en la Pitch Class Set Theory analiza las semejanzas entre los acordes de la Partitura 1 y la Partitura 2 y genera la pieza sonora final con los acordes coincidentes extraídos de los espectros estelares.	67
9.1	Espectros decodificados de HD 124186 generados por los tres modelos de autoencoder analizados. Flujo de brillo (ADU) en función de la longitud de onda (Å). La curva negra corresponde al espectro original. La línea punteada azul a la salida (arriba) y al error residual (abajo) del autoencoder de dos capas. La línea naranja muestra el error residual del VAE de cuatro capas y la roja la salida del autoencoder de 6 capas. Como puede observarse, el modelo de seis capas genera una salida completamente decorrelada con la entrada para este espectro.	73
9.2	Rendimiento del autoencoder de dos capas, R^2 para los 706 espectros codificados de la librería CaT y porcentaje de espectros con R^2 superior a 0,9 para el autoencoder de dos capas (verde-izquierda), el VAE de cuatro capas (azul-centro) y el autoencoder de seis capas (rojo-derecha).	74
9.3	Interfaz de usuario mostrando la imagen de salida generada en el bloque de pre-procesado para el espectro de la estrella tipo K0 III, HD 216228. En azul, flujo normalizado (ADU) en función de la longitud de onda (Å). Las líneas y puntos rojos representan las amplitudes de las líneas de absorción detectadas en el análisis secundario.	75
9.4	Diagrama de bloques del módulo de sonificación implementado en <i>Csound</i> y <i>Cabbage</i> . Las líneas sólidas representan rutas MIDI, de audio e imagen. Las líneas discontinuas representan mensajes OSC.	77

9.5	Diagrama de bloques del sintetizador incluyendo el pre-procesado. El sistema captura a “tiempo real” la nota MIDI tocada por el usuario, compara la nota con la nota fundamental de los acordes estelares de la base de datos, y proporciona una representación multimodal del espectro con el primer acorde coincidente encontrado en la base de datos pre-procesada.	78
10.1	Comparativa de autoencoders para la galaxia espiral (Scd) NGC 5406. Auto-encoder de seis capas y seis dimensiones (negro, continuo) vs VAE de cuatro capas y seis dimensiones (naranja, punteado). Espectros reconstruidos y error residual del espectro original para el spaxel (34,34). Sparse autoencoder: $R^2 = 0,99$ (espectro), $R^2 = 0,98$ (cubo de datos), 39,12% de los espectros con $R^2 > 0,9$, 100 epochs, una hora por cubo. VAE: $R^2 = 0,97$ (espectro), $R^2 = 0,98$ (cubo de datos), 4,92% de los espectros con $R^2 > 0,9$, 291 epochs, 5h 30' por cubo. Flujo normalizado (ADU) vs longitud de onda (\AA).	83
10.2	Resultados del autoencoder de seis capas y seis dimensiones. Espectro reconstruido (línea discontinua negra) y original (línea sólida azul) con error residual (línea punteada naranja) del spaxel (35,35) de la galaxia espiral (Sbc) NGC 5784 (izquierda) y la galaxia elíptica (E4) NGC 5682 (derecha). Dos ejemplos de una galaxia antigua y una región de formación estelar del estudio CALIFA. Respectivamente, $R^2 = 0,98$ y $R^2 = 0,95$. Flujo normalizado (ADU) en función de la longitud de onda (\AA).	84
10.3	Curvas de aprendizaje que muestran el error cuadrático medio (MSE) por epoch durante los procesos de entrenamiento y validación para la galaxia espiral (SAb) NGC7047 y la galaxia elíptica (E5) UGC03960. Estos cubos de datos corresponden, respectivamente, al mejor y al peor resultado proporcionado por el autoencoder. Nótese el cambio de escala en el eje de abscisas de las representaciones.	85
10.4	Interfaz de usuario de <i>ViewCube</i> mostrando el cubo de datos de la galaxia espiral (Sbc) NGC 5732. Ventana de imagen 2D (izquierda) y representación multimodal (ventana espectral arriba y espectro sonoro de la sonificación abajo) para el spaxel (35,35) (derecha).	86
10.5	Diagrama de bloques de <i>SoniCube</i> . Pre-procesado y cálculos “a tiempo real”, incluyendo flujos de datos, OSC y señales de audio.	87
11.1	Resultados del reconocimiento sonoro de las líneas de absorción y emisión en galaxias del Large Early Galaxy Census (STScI). Tasa de acierto promedio, desviación estándar e intervalos de confianza de Jeffreys para los participantes con experiencia (0,35) y sin experiencia (0,11). La línea discontinua indica la tasa promedio de acierto aleatorio (0,23).	91
11.2	Resultados del reconocimiento sonoro de líneas de absorción y emisión en galaxias del Large Early Galaxy Census (STScI). Tasa de acierto promedio, desviación estándar e intervalos de confianza de Jeffreys para las subcategorías extraídas del grupo con experiencia. Astrónomos músicos (0,47), músicos (0,35), astrónomos (0,37) y 4 muestras aleatorias de participantes sin experiencia (0,27). La línea discontinua indica la tasa promedio de acierto aleatorio (0,23).	93

11.3	Análisis cualitativo por grupos de experiencia en Música y Astronomía. Valores pico, valores atípicos, promedio y barras de error para las respuestas registradas de los grupos con y sin experiencia. También se incluye una subclasificación adicional no equilibrada del grupo de participantes con experiencia para mostrar las preferencias de astrónomos profesionales y aficionados, músicos profesionales y aficionados, y astrónomos músicos profesionales y aficionados. Las preguntas referenciadas en el eje x están disponibles en: https://forms.office.com/e/iCFyXfwsdT	94
11.4	Evaluación del público para dos composiciones generadas por el sistema de composición autónomo descrito en el Capítulo 8. La red fue entrenada usando los trabajos corales de Giovanni Pierluigi da Palestrina. Se pidió a los participantes que calificaran las composiciones de 0 a 10.	95
11.5	Sonificación de mapeado simbólico. Tasa de acierto promedio para 16 participantes con experiencia (0,78, izquierda) y 16 sin experiencia (0,48, derecha), para tareas de clasificación del tipo espectral. La línea discontinua muestra la tasa de acierto aleatorio (0,14).	97
11.6	Sonificación de mapeado simbólico. Análisis de subgrupos de cuatro muestras. Tasa de acierto promedio para astrónomos músicos (0,81, izquierda), músicos (0,94, centro-izquierda), astrónomos (0,58, centro-derecha) y participantes sin experiencia (0,5, derecha). La línea discontinua muestra la tasa de acierto aleatorio (0,14).	98
11.7	Sonificación mediante deep learning. Tasa de acierto promedio para 16 participantes con experiencia (0,48, izquierda) y 16 sin experiencia (0,39, derecha), para tareas de clasificación del tipo espectral. La línea discontinua muestra la tasa de acierto aleatorio (0,14).	101
11.8	Sonificación mediante deep learning. Análisis de subgrupos de cuatro muestras. Tasa de acierto promedio para participantes con experiencia en Astronomía y Música (0,5, izquierda), sólo en música (0,69, centro-izquierda), sólo en astronomía (0,08, centro-derecha) y participantes sin experiencia (0,44, derecha). La línea discontinua muestra la tasa de acierto aleatorio (0,14).	102
11.9	Gráfico comparativo de las tasas promedio de acierto para cada grupo de estudio, obtenidas para sonificaciones generadas mediante mapeado simbólico y mediante deep learning. La línea discontinua representa la tasa de acierto aleatorio (0,14).	103
11.10	Evaluación cualitativa de 40 participantes, el 90 % de la muestra completa consideró que la propuesta podría ser útil en actividades de divulgación para introducir los tipos de espectros estelares (izquierda), y el 72,5 % declaró su interés en escuchar más sonificaciones astronómicas después de completar la encuesta (derecha).	104

11.11	Ejemplos de edad/tipo de galaxia presentados en los vídeos de entrenamiento. Espectro de una región de formación estelar en la galaxia espiral (S0) NGC 3395 (izquierda), espectro cerca del centro de la galaxia espiral de edad intermedia (Sd) NGC 2347 (centro), y espectro de una región cerca del centro de la galaxia espiral retirada (Sb) NGC 6125 (derecha). Las figuras superiores muestran los mapas de continuo producidos con el filtro de banda estrecha indicado por las curvas verdes representadas en los espectros de las figuras inferiores. Los espectros corresponden a spaxels específicos, resaltados por cuadrados rojos en la imagen superior. Los ejes de las figuras superiores representan desplazamientos en segundos de arco relativos al centro de la galaxia. La barra de colores representa el flujo del cubo de datos convolucionado con el filtro, en unidades de 10^{-16} erg cm ⁻² s ⁻¹ . Los espectros en las figuras inferiores se representan con longitudes de onda en Angstroms.	106
11.12	Tasas de acierto promedio para 65 participantes en las preguntas simples (izquierda, magenta, 0,515), para 29 astrónomos profesionales (centro, azul, 0,554) y para 31 no astrónomos (derecha, verde, 0,481), referidas al porcentaje promediado por pregunta de acierto aleatorio 0,34 (línea discontinua).	107
11.13	Tasas de acierto promedio en preguntas combinadas, resultados globales y por subgrupos.	109
11.14	Tasas de acierto promedio en preguntas simples por campo de experiencia. De izquierda a derecha: Astrónomos músicos (0,54, 10 participantes), Astrónomos no músicos (0,58, 12 participantes), Músicos no astrónomos (0,48, 11 participantes) y submuestra de participantes sin experiencia (0,55, 12 participantes).	109
11.15	Porcentaje de acierto por pregunta y grupo de experiencia. Resultados para Astrónomos y No astrónomos. Preguntas de ubicación del sonido referenciadas como “Loc”, preguntas de distancia al centro de la galaxia referenciadas como “Dist”, preguntas de tipo de galaxia referenciadas como “Tipo” y preguntas combinadas (acierto = todas las opciones de opción múltiple correctas) referenciadas como “Comb”. Las líneas de puntos representan la tasa de referencia de elección aleatoria para cada pregunta.	110
11.16	Porcentaje de acierto por pregunta y grupo de experiencia. Resultados para Astrónomos músicos, Astrónomos no músicos, Músicos no astrónomos y participantes sin experiencia. Preguntas de ubicación del sonido referenciadas como “Loc”, preguntas de distancia al centro de la galaxia referenciadas como “Dist”, preguntas de tipo de galaxia referenciadas como “Tipo” y preguntas combinadas (acierto = todas las opciones de opción múltiple correctas) referenciadas como “Comb”. Las líneas de puntos representan la tasa de acierto aleatorio para cada pregunta. Merece la pena destacar los resultados de elección aleatoria para los participantes sin experiencia en las preguntas Loc-4, Dist-7 y Comb-13.	111

11.17	Tasa de acierto promedio de las dos preguntas combinadas por bloques (Ubicación, Distancia y Tipo/Edad). De izquierda a derecha, tasa de acierto para Astrónomos, No astrónomos y tasas de subgrupos para Astrónomos músicos, Astrónomos no músicos, Músicos no astrónomos y participantes sin experiencia. La línea de puntos indica la tasa promedio de acierto aleatorio (0,33), considerando tres posibles respuestas por bloque (correcta, incorrecta y no respondida), ya que 19 participantes no introdujeron ninguna respuesta en algunos bloques.	112
11.18	Tasas de acierto promedio en preguntas simples y combinadas por grupos de edad. La línea discontinua representa la tasa promedio de acierto aleatorio (valor promedio de las preguntas simples 0,34 y 0,2 en las combinadas). Preguntas simples (izquierda): Participantes <21 = 0,25, entre 21 y 30 = 0,55, entre 31 y 40 = 0,50, entre 41 y 50 = 0,56, entre 51 y 60 = 0,44 y >60 = 0,45. Preguntas combinadas (derecha): Participantes <21 = 0, entre 21 y 30 = 0,08, entre 31 y 40 = 0,17, entre 41 y 50 = 0,24, entre 51 y 60 = 0,33 y >60 = 0,05.	112
11.19	Tasa de acierto en preguntas simples para participantes entrenados sólo con vídeos (V) frente a los participantes que probaron la aplicación en vivo (L). De izquierda a derecha, resultados globales, astrónomos frente no astrónomos y subgrupos de experiencia. La línea de puntos muestra la tasa de elección aleatoria promedio (0,34) de preguntas con 2, 3 y 4 posibles respuestas. . . .	113
11.20	Evaluación cualitativa de <i>ViewCube</i> . Interactividad: resultados de los 42 participantes que probaron la aplicación en persona. Utilidad y estética: muestra completa, 67 participantes. El 81 % declaró que la aplicación tiene una “buena interactividad”, el 79,1 % la encontró “útil” y el 58,2 % con “buen sonido”. . .	115
1	Caracterización del tránsito planetario (Seager y Mallen-Ornelas, 2003). . . .	162
2	Clasificación de espectros estelares (IAU, 2020).	164
3	Comparación de las curvas de respuesta de Kepler, CoRoT y Johnson (Jason Rowe, extraído de 2009IAUS..253..121R) (NASA, 2020a). Funciones de transmisión escaladas para alcanzar el 100 % de transmisión. Espectro de la estrella A2V alcanzando su punto máximo en la banda UV (cian), y espectro de la estrella M2V alcanzando su punto máximo en el infrarrojo (naranja).	165
4	Dos ejemplos de diagramas HR utilizados respectivamente para la clasificación y evolución estelar (Harvard-Smithsonian, 2015). Clasificaciones espectrales de Harvard y Morgan-Keenan sobre un diagrama HR (izquierda), y la trayectoria evolutiva del sol (derecha).	166
5	Ejemplo de espectro LEGA-C pre-procesado. Espectro restado del continuo y líneas de emisión marcadas mediante el módulo <code>find_lines_threshold</code> para la galaxia COSMOS M01 con coordenadas RA: 150.220555, DEC: 2.41104. . .	167
6	Secuencia de cuatro pasos de una celda de red LSTM (Sherstinsky, 2020), con \vec{x} representando el vector de entrada a la celda, \vec{s} representando la señal de estado de la celda, y \vec{v} representando la señal de valor observable de la celda para propósitos externos.	171

7	Celda de red LSTM simple con las etapas del sistema delineadas por rectángulos discontinuos (Sherstinsky, 2020), donde \vec{x} es el vector de entrada, \vec{s} es el estado de la celda, y \vec{v} es el valor observable de la celda para propósitos externos. $\vec{a}_{cu}[n]$ es el nodo de acumulación, con parámetros $W_{x_{cu}}, W_{s_{cu}}, W_{v_{cu}}$, y \vec{b}_{cu} . $\vec{a}_{cs}[n]$ es el nodo de acumulación de la puerta que controla la fracción de la señal de estado de la celda ($\vec{s}[n - 1]$), con parámetros $W_{x_{cs}}, W_{s_{cs}}, W_{v_{cs}}$, y \vec{b}_{cs} . $\vec{a}_{cr}[n]$ es el nodo de acumulación de la puerta que controla la fracción de la señal estimada de lectura ($\vec{r}[n]$) utilizada para liberar la señal observable de la celda, con parámetros $W_{x_{cr}}, W_{s_{cr}}, W_{v_{cr}}$, y \vec{b}_{cr} . $\vec{a}_{du}[n]$ es el nodo de acumulación para la función de transformación de datos, que actualiza la señal estimada ($\vec{u}[n]$), con parámetros $W_{x_{du}}, W_{v_{du}}$, y \vec{b}_{du}	172
8	Curva de nivel de transferencia utilizada para el control de valores de flujo salientes que pueden causar daños auditivos al representar cubos de datos inexplorados.	174
9	Resultados individuales de R^2 mostrando la peor codificación (arriba) y la mejor (abajo) para la librería STELIB, correspondiendo respectivamente a las estrellas HD 092809 y HD 184960. Vale la pena observar cómo incluso para $R^2 = 0,62213$, el modelo representa las principales características espectrales del espectro de entrada alrededor de 4700 y 5900 Å. Sin embargo, genera una línea H α falsa alrededor de 6600 Å.	178
10	Distribución de los resultados de R^2 para la librería MILES utilizando un VAE (azul/arriba) y un sparse autoencoder (verde/abajo), entrenados en el conjunto de datos original (izquierda) y en el conjunto de datos aumentado x20 (derecha). R^2 para cada espectro analizado.	179
11	Correlación entre los seis valores latentes generados por el autoencoder y los anchos equivalentes (EW) de H α de NGC3395 (izquierda, EW alto) y NGC2347 (derecha, EW intermedio). Ambos ejemplos muestran la tendencia de correlación general más alta compartida entre los valores latentes L3, L1 y L4. CALIFA survey.	180
12	EW de H α por spaxel (verde), vector latente por spaxel (azul) y EW de H α vs vector latente (negro) para NGC3395 y NGC2347.	181
13	<i>INT</i> . Composición polifónica no supervisada, generada a partir de espectros estelares.	193

Lista de Tablas

1.1	Musificación y Sonificación: ventajas, desventajas, aplicaciones principales y funciones fundamentales. El asterisco corresponde a las funciones principales del lenguaje natural. Basado en Schubert, 2009.	4
4.1	Mapeado de variables estelares para opcodes de <i>CSound</i>	34
4.2	Ejemplos de TOI sonificados. Coordenadas y referencia de publicación	36
6.1	Estrategia simbólica de mapeado que asigna cada tipo espectral a una nota musical, propuesta para la sonificación de los siete tipos de espectros estelares de la clasificación de Harvard con valores de frecuencia para la octava 4 (conversiones de pitch en <i>CSound</i> : https://csound.com/manual/misc/pitch/).	47
6.2	Ejemplo de mapeado para el objeto con CoRoT ID 0223959618.	48
7.1	Influencia de la dimensión del espacio latente en los resultados de la codificación. R^2 para el conjunto completo (varianza ponderada: el valor proporcionado es el valor promedio ponderado con las varianzas de cada coeficiente individual calculado para cada espectro).	56
7.2	Ejemplo de mapeado para la estrella HD 040657, con coordenadas RA:06:00:03.35 y DEC:-03:04:27.33. Valor latente, frecuencia y valor de frecuencia fundamental final para un factor de multiplicación interactivo 10x.	58
7.3	Ejemplo de mapeado para la estrella HD 040657, con coordenadas RA:06:00:03.35 y DEC:-03:04:27.33.	59
8.1	Espectros estelares utilizados para generar la secuencia OBAFGKM. Resumiendo las clases de luminosidad de Morgan-Keenan (MK), Ia y Ib corresponden a supergigantes luminosas, II a gigantes brillantes, III son gigantes normales, IV son subgigantes, V corresponde a estrellas enanas de la secuencia principal, VI a subenanas y D a enanas blancas. El número permite escalar cada tipo de 0 (el más caliente) a 9 (el más frío).	63
9.1	Ejemplos de las variables y mapeados utilizados en la sonificación de los espectros de HD168775, M71Q, HD4307, HD57061 y M5 II-53. Valor MIDI de la nota de entrada seleccionada por el usuario, objeto coincidente, vector latente y acorde equivalente obtenido del análisis con <i>Music 21</i> , R^2 entre el espectro original y el representado, y número de líneas de transmisión detectadas.	76

11.1	Resumen de las tasas de acierto por grupos de participantes para el reconocimiento sonoro de líneas de transmisión. Desviación estándar e intervalos de confianza.	93
11.2	Sonificación de mapeado simbólico. Resumen de las tasas de acierto promedio por grupos de participantes, desviación estándar e intervalos de confianza. Los subgrupos han sido extraídos del grupo global y han sido equilibrados para permitir su comparación.	99
11.3	Sonificación mediante deep learning. Resumen de las tasas de acierto por grupos de participantes. Desviación estándar e intervalos de confianza. Subgrupos extraídos del grupo global y equilibrados con fines comparativos.	102
11.4	Evaluación cuantitativa. Resultados para secciones de preguntas simples y combinadas mostradas por grupo de experiencia y edad.	114
11.5	Evaluación cualitativa. Porcentajes por grupo de experiencia. Los valores con asterisco corresponden a la muestra de participantes que probaron la aplicación en persona.	114
1	Longitudes de onda electromagnéticas para líneas espectrales comunes (Survey, 2003).	163
2	Listado de aplicaciones desarrolladas	180

Abreviaturas y Acrónimos

- ADU** Analog Digital Unit
- ADSR** Attack/Decay/Sustain/Release
- AI** Artificial Intelligence
- ALMA** Atacama Large Milimeter/submillimeter Array
- ANSI** American National Standards Institute
- AURA** Association of Universities for Research in Astronomy
- AVO** Astrophysical Virtual Observatory
- BJD** Barycentric Julian Date
- BKJD** Barycentric Kepler Julian Date
- BLS** Box-fitting Least-Squares
- BLV** Blind or low vision
- CaT** Calcium triplet
- CAHA** Centro Astronómico Hispano en Andalucía
- CALIFA** Calar Alto Legacy Integral Field Area
- CARMENES** Calar Alto high-Resolution search for M dwarfs with Exoearths with Near-infrared and optical Echelle Spectrographs
- CCD** Charge Coupled Device
- CDS** Centre de Données Stellaires
- CEMAMu** Centre d'Etudes de Mathématique et Automatique Musicales
- CGM** Circumgalactic Medium
- CHEOPS** Characterizing Exoplanets Satellite
- CMB** Cosmic Microwave Background
- CNES** Centre national d'études spatiales
- CoRoT** Convection, Rotation, and planetary Transit
- CSIC** Consejo Superior de Investigaciones Científicas

CSIR Commonwealth Scientific and Industrial Research
CTIO Cerro Tololo Inter-American Observatory
CNN Convolutional Neural Networks
DAW Digital Audio Workstation
DEC Declination
 DR Data Release
DVT Data Validation Time-series
ELBO Evidence Lower Bound Optimization
 ESA European Space Agency
 ESO European Southern Observatory
 EUV Extreme Ultraviolet
FITS Flexible Image Transport System
 FOA First Order Ambisonics
 FOV Field Of View
FWHM Full Width at Half Maximum
 GAN Generative Adversarial Networks
GLAS Geneva Lake Astrophysics and STEAM
 GPU Graphics Processing Unit
HARPS High Accuracy Radial Velocity Planet Searcher
 HDU Header and Data Unit
 HPF High Pass Filter
 HR Hertzsprung-Russel diagram
 IAA Instituto de Astrofísica de Andalucía
 IAC Instituto de Astrofísica de Canarias
 IAS Institut d'Astrophysique Spatiale
 IAU International Astronomical Union
ICAD International Conference on Auditory Display

IFS Integral Field Spectroscopy
IFU Integral Field Unit
IGM Intergalactic Medium
INT Isaac Newton Telescope
IRCAM Institut de Recherche et Coordination Acoustique/Musique
IVOA International Virtual Observatory Alliance
JKT Jacobus Kaptein Telescope
KIC Kepler Input Catalog
KL Kullback–Leibler divergence
KOI Kepler Object of Interest
LAMOST Large Sky Area Multi-Object Fiber Spectroscopic Telescope
LCO Las Cumbres Observatory
LFO Low Frequency Oscillator
LPF Low Pass Filter
LIGO Laser Interferometer Gravitational- Wave Observatory
LISA Laser Interferometer Space Antenna
LLM Large Language Model
LSTM Long Short Term Memory
 M_{\odot} Mass of the Sun
MAST Mikulski Archive for Space Telescopes
MIDI Musical Instrument Digital Interface
MIT Massachusetts Institute of Technology
MK Morgan-Keenan
MSE Mean Square Error
NASA National Aeronautics and Space Administration
NN Neural Network
NVO National Virtual Observatory

OSC Open Sound Control

PDCsap Pre-search Data Conditioning SAP

R_* Radium of a star

RA Right Ascension

RNN Recurrent Neural Network

RSSD Research and Scientific Support Department

SAP Simple Aperture Photometry

SDSS Sloan Digital Sky Survey

SED Spectral Energy Distribution

STScI Space Telescope Science Institute

STEM Science Technology, Engineering, and Mathematics

SVO Spanish Virtual Observatory

TCE Threshold Crossing Events

TESS Transiting Exoplanet Survey Satellite

TIC TESS Input Catalog

TOI TESS Object of Interest

UDO User Defined Opcode

UI User Interface

UNOOSA United Nations Office for Outer Space Affairs

UPIC Unité Polyagogique Informatique du CEMAMu

UV Ultraviolet

VAE Variational Autoencoder

VO Virtual Observatory

VOSA Virtual Observatory SED Analyzer

WHT William Herschel Telescope

Capítulo 1

Introducción

En las últimas décadas, la forma en que se explora el universo ha cambiado significativamente gracias a la digitalización y al desarrollo de la tecnología del Observatorio Virtual (VO) (Quinn, 2003). En la era actual de la Astronomía basada en datos, este entorno estandarizado permite la conexión entre recursos de instituciones internacionales, misiones de ciencia espacial, telescopios y equipos de investigación (Arviset et al., 2012) de todo el mundo, haciendo posible explorar el universo a través de imágenes y datos capturados por las principales misiones internacionales con un ordenador personal.

De manera similar, el mundo de la acústica, del diseño sonoro, la composición musical o incluso la investigación musical, también se han visto fuertemente modificados por el aumento exponencial del poder computacional de los ordenadores personales actuales. Procesos altamente complejos como la generación de paisajes sonoros ambisónicos, la síntesis de instrumentos musicales y orquestas sinfónicas con gran realismo, o la grabación y postproducción de extensos proyectos multipista, pueden realizarse hoy en día con una sola máquina.

Bajo este paradigma común de digitalización global, esta tesis analiza las intersecciones entre Música y Astronomía utilizando Inteligencia Artificial y Sonificación con la intención de proporcionar una perspectiva objetiva sobre la representación de datos astronómicos a través del sonido. Tal y como ilustra el diagrama de Venn de la Figura 1.1, el trabajo está orientado a demostrar la utilidad de este enfoque interdisciplinar, no sólo en actividades creativas, educativas y de divulgación, sino también en la investigación científica, liberando el potencial que tiene la Sonificación para ayudar en la exploración de datos astronómicos (Tucker Brown et al., 2022).

El uso de representaciones sonoras se basa en la capacidad del sistema auditivo humano para reconocer cambios y patrones temporales, su rápida respuesta temporal y la capacidad de monitorizar y procesar múltiples conjuntos de sonidos en paralelo (Walker y Nees, 2011). El diseño de una herramienta de sonificación requiere la definición de un caso de estudio, el análisis del fenómeno y la creación de una experiencia sonora (Lenzi y Ciuccarelli, 2024), involucrando desafíos tanto técnicos como perceptuales (Worrall, 2019).

Alineado con los múltiples objetivos de Bardelli et al., 2021, y la motivación de de Campo, 2009 para sonificar datos científicos, el uso de la sonificación en representaciones multimodales para



Figura 1.1: Diagrama de Venn que representa el espacio interdisciplinar en el que se enmarca este trabajo. Intersección entre la sonificación y la aplicación del deep learning al análisis científico y a la creación artística.

la exploración de datos astronómicos tiene el potencial de proporcionar representaciones sonoras precisas (Bly, 1982), y construir conexiones interdisciplinarias naturales entre las propiedades de ambos campos, ampliando las posibilidades en la representación y perceptualización de la información científica (Malikova. et al., 2017). Además, puede hacer que los catálogos y bases de datos estelares sean más accesibles para usuarios ciegos o con baja visión (BLV) (Noel-Storr y Willebrands, 2022), permitiendo a los investigadores sumergirse en sus casos de estudio para identificar patrones, tendencias y valores atípicos en los conjuntos de datos (K. W. Hall et al., 2019), y navegar por los archivos de datos masivos generados por telescopios espaciales (K. Arcand et al., 2020) o misiones de observación terrestre.

Considerando la propuesta de Barrass, 2012, para integrar estética y funcionalidad como puente entre los métodos científicos y artísticos, y entendiendo el “sonido como material maleable” (Franinovic y Serafin, 2013), la sonificación y musificación de datos astronómicos también abre nuevos horizontes para la creatividad aplicada a la música y al arte sonoro, que además proporcionan un valor añadido en comunicaciones científicas, actividades de divulgación (Metallinou, 2022) y narración transmedia de comunicaciones científicas (Ortega-Alonso et al., 2024).

El ecosistema de proyectos dedicados a la sonificación de datos astronómicos ha sido recientemente representado por estudios como Zanella, Harrison et al., 2022, e iniciativas como el informe de Sonificación de la Oficina de las Naciones Unidas (UNOOSA, 2022), materializando el interés de la comunidad internacional en el uso del sonido como una forma de contribuir a un futuro inclusivo y accesible de las Ciencias Espaciales y la Astronomía (Zanella, Noel-Storr et al., 2022).

En concreto, en el contexto de la comunidad edificada alrededor de la International Conference

on Auditory Display (ICAD), la temática de los proyectos de sonificación de datos astronómicos se ha centrado principalmente en la generación de representaciones del sol, de los planetas y del sistema solar, haciendo esperable la expansión de la disciplina hacia diversos aspectos de la Astronomía y la Astrofísica (García-Benito, 2023).

La sonificación se ha utilizado con éxito para estudiar la estructura estelar y planetaria (Lynch, 2017), para comunicar información a personas ciegas y con baja visión (BLV) (Merced y Gastrow, 2018), y para fomentar la participación pública mediante la representación de varias décadas de datos astronómicos (K. K. Arcand et al., 2024). Proyectos como Audible Universe (Harrison et al., 2022; Misdariis et al., 2022) o Astro Accesible (Pérez-Montero et al., 2017), herramientas de análisis como xSonify (Candey et al., 2006; Diaz-Merced et al., 2011), Astronify (Brosseur et al., 2020) o Highcharts Sonification Studio (Cantrell et al., 2021), y redes de colaboración como el Sonification World Chat (GLAS-Education, 2020), son sólo una pequeña muestra de los esfuerzos invertidos por investigadores internacionales para la aceptación y el establecimiento de las representaciones sonoras en contextos científicos.

Enmarcada en este entorno de acceso universal, esta investigación tiene como objetivo explorar el potencial del deep learning para proporcionar representaciones sonoras no supervisadas que podrían utilizarse en el análisis de datos científicos, en metodologías de exploración complementarias inclusivas y en aplicaciones artísticas y de divulgación pública. Este trabajo ofrece una colección abierta de implementaciones prácticas de sonificación y musificación aplicadas a catálogos reales de exoplanetas, estrellas y galaxias, que utilizan algoritmos de aprendizaje profundo y tecnología ambisónica y binaural para la creación de representaciones sonoras inmersivas.

Para ello, hemos desarrollado dos líneas de investigación paralelas destinadas a fomentar el uso de representaciones sonoras en Astronomía y Astrofísica, a mejorar la accesibilidad de los archivos astronómicos y a estudiar el potencial de la tecnología de sonificación basada en redes neuronales. La primera línea analiza las posibilidades de la sonificación para la representación de datos astronómicos en contextos científicos. La segunda se centra en la generación de musificaciones y composiciones musicales basadas en datos astronómicos. Ambas líneas de investigación utilizan cuatro metodologías: el análisis bibliográfico multidisciplinar, la implementación práctica de los diseños y el estudio cuantitativo y cualitativo de las propuestas.

El propósito general de esta tesis es proporcionar un conjunto de recursos que podrían ser útiles como punto de partida para futuras implementaciones hacia un Observatorio Virtual Sonoro Interoperable. Los posibles usuarios finales de las propuestas que presentamos incluyen a científicos y científicas con interés en complementar las representaciones gráficas de sus casos de estudio con sonificaciones, personas dedicadas a la comunicación científica, a la educación o a la promoción de las disciplinas de Ciencia, Tecnología, Ingeniería y Matemáticas (STEM), personal docente e investigador con interés en involucrar a sus estudiantes en conceptos astronómicos o astrofísicos a través del sonido, y artistas de la creación y de la composición musical con la intención de acercar la ciencia al público en general utilizando piezas musicales basadas en datos astronómicos.

1.1 Sonificación y musificación

Tal y como sugieren Zanella, Harrison et al., 2022, “el gran potencial teórico latente en el uso del sonido en la investigación de conjuntos de datos complejos puede ser de especial interés en la identificación de señales o variaciones características y en la exploración preliminar de datos en Astronomía y Ciencia Espacial”.

Con este argumento como punto de partida, resulta necesario diferenciar entre sonificación y musificación o música basada en datos, como principales disciplinas con el potencial de proporcionar representaciones sonoras capaces de transmitir la información de dichos conjuntos de datos a una determinada audiencia.

Por una parte, Hermann, 2008 propone que, “una técnica que utiliza datos como entrada y genera señales de sonido (eventualmente en respuesta a una excitación o activación adicional opcional), puede llamarse sonificación, si y sólo si:

- (C1) El sonido refleja propiedades o relaciones objetivas en los datos de entrada.
- (C2) La transformación es sistemática. Esto significa que se proporciona una definición precisa de cómo los datos (y las interacciones opcionales) hacen que el sonido cambie.
- (C3) La sonificación es reproducible: dados los mismos datos e interacciones idénticas (o activaciones), el sonido resultante tiene que ser estructuralmente idéntico.
- (C4) El sistema puede usarse intencionalmente con diferentes datos, y también usarse repetidamente con los mismos datos”.

Tabla 1.1: Musificación y Sonificación: ventajas, desventajas, aplicaciones principales y funciones fundamentales. El asterisco corresponde a las funciones principales del lenguaje natural. Basado en Schubert, 2009.

	Musificación	Sonificación
Ventajas	Representación general Participación pública	Representación objetiva Precisión
Desventajas	Falta de precisión Pérdida de información	Abstracción Curva de aprendizaje
Aplicaciones	Divulgación científica Banda sonora original	Análisis científico Accesibilidad e Inclusión
Funciones	*Comunicación Representación simbólica Respuesta emocional Disfrute Entretenimiento	*Comunicación Representación simbólica *Representación del conocimiento

Por otra parte, el término musificación se usa para describir cualquier representación musical de un conjunto de datos. Aunque música y sonificación comparten numerosos aspectos, especialmente aquellos relacionados con la percepción auditiva, la principal diferencia entre sonificación y musificación radica en su objetivo final, algo que determina las estrategias de diseño e implementación de las representaciones sonoras. En este sentido, Scaletti (Dean,

2018) menciona cómo “quizás la distinción más importante entre sonificación y música es la diferencia en la intención. El objetivo y propósito de la sonificación de datos es ayudar a comprender, explorar, interpretar, comunicar y razonar sobre un fenómeno, un experimento o un modelo, mientras que en el arte sonoro, el objetivo es hacer pensar a una audiencia creando un flujo de experiencia para ellos”.

Esta tesis describe, tanto estrategias de sonificación para el análisis científico, como métodos de musificación de datos astronómicos, todo ello con el fin de proporcionar un conjunto de herramientas para la exploración sonora de librerías, catálogos y bases de datos. En la Tabla 1.1 se muestran las principales características de la dicotomía musificación-sonificación con el objetivo de ilustrar los dos bloques conceptuales en los que se basan los capítulos de este trabajo.

1.2 Objetivos y preguntas de investigación

Esta investigación nació con la doble intención de permitir la exploración de los archivos de las agencias espaciales internacionales a través del sonido y de promover el uso de información astronómica en aplicaciones creativas y de composición musical. Con la creación de un marco experimental en torno al concepto de un Observatorio Virtual Sonoro, la tesis tiene como objetivo analizar en profundidad la naturaleza de las relaciones entre Tecnología, Sonido, Música, Astronomía y Astrofísica, respondiendo a tres preguntas fundamentales:

Q1. ¿Podría la conversión de datos astronómicos en sonido ser útil para el análisis científico en Astronomía y Astrofísica?

Q2. ¿Es posible extraer información esencial de catálogos estelares utilizando inteligencia artificial, que pueda entenderse mediante sonificaciones basadas en criterios no arbitrarios o subjetivos?

Q3. ¿Pueden las técnicas actuales de deep learning, basadas en criterios objetivables, generar música a partir de datos astronómicos para ofrecer una revisión empírica de la idea clásica de la *Música de las Esferas*?

Los siguientes objetivos específicos se establecieron para responder a estas preguntas a lo largo de los capítulos del trabajo:

- Estudiar las relaciones históricas entre Música y Astronomía.
- Analizar el desarrollo de las observaciones astronómicas y las posibilidades científicas del Observatorio Virtual (VO).
- Identificar soluciones de sonificación aplicables a casos prácticos de análisis en Astronomía y Astrofísica.
- Revisar las arquitecturas actuales de deep learning para su aplicación en el análisis de datos astronómicos y en composición musical.
- Proponer y evaluar metodologías de exploración complementarias inclusivas que puedan hacer que los datos astronómicos sean accesibles para cualquier persona a través del

sonido.

- Proporcionar un método para la generación de representaciones sonoras no supervisadas orientado al análisis de datos astronómicos.
- Generar representaciones multimodales de datos astronómicos para usar en actividades educativas y de divulgación.
- Desarrollar sintetizadores de sonido basados en datos astronómicos para promover el interés en las disciplinas STEM.
- Explorar el potencial del deep learning para proporcionar una revisión no supervisada del concepto clásico de la *Música de las esferas*.
- Crear composiciones autónomas originales a partir de datos astronómicos que puedan usarse en la creación de bandas sonoras originales para películas, televisión y contenido audiovisual.

1.3 Estructura de la tesis

Esta tesis ha sido organizada por bloques temáticos en un esfuerzo por cubrir, de la manera más amplia posible, el abanico de posibilidades de sonificación aplicables a la gran variedad de archivos astronómicos de acceso público disponibles.

Los capítulos están basados en artículos presentados en conferencias internacionales que se encuentran referenciados a lo largo de texto principal. Cada capítulo ofrece una descripción de los diferentes catálogos y tipos de datos utilizados, así como una prueba de concepto orientada a la demostración, evaluación y potencial validación de las diferentes propuestas.

De este modo, este trabajo analiza la capacidad de representación sonora en función del tipo y número de variables representadas, desarrolla el concepto de sonificación basada en deep learning, implementa un sistema de composición musical no supervisada a partir de espectros estelares, propone el uso de instrumentos musicales digitales para la exploración de archivos astronómicos y estudia el potencial de la espacialización sonora para proporcionar experiencias de análisis inmersivas y para mejorar la comprensión de las sonificaciones.

Capítulo 1. Introducción

Descripción inicial del contexto, objetivos y preguntas de investigación. Este capítulo ofrece una visión general de sus líneas de trabajo, contenidos principales, resultados potenciales, desarrollos y propuestas.

Capítulo 2. Precedentes y estado del arte

Análisis de los antecedentes históricos, estado del arte, conceptos y tecnología relacionados con los temas involucrados en esta investigación: Música, Sonido, Sonificación, Inteligencia Artificial, Astronomía y Astrofísica. Este capítulo ofrece un recorrido desde las primeras referencias de la música de las esferas en la antigua Grecia hasta los actuales desarrollos generativos globales.

Capítulo 3. Exploración multimodal de catálogos estelares

Descripción de la implementación de *FITS2OSC*, el módulo fundacional utilizado a lo largo de diferentes etapas de esta investigación para la sonificación de flujos de datos. Este capítulo ofrece el análisis sonoro de dos casos de estudio: la exploración secuencial de la librería estelar STELIB y el análisis de las curvas de luz de KIC 6504534, de la misión Kepler.

El capítulo incluye una sección dedicada a la descripción de un módulo experimental de sonificación espectral para *Astronify*, desarrollado en colaboración con el equipo del Mikulski Archive for Space Telescopes (MAST) del Space Telescope Science Institute (STScI).

Capítulo 4. Paisajes sonoros multivariable

Análisis de las posibilidades de la representación sonora de múltiples variables simultáneas, sobre el caso de estudio de los archivos de series temporales de validación de datos (DVT) de la misión Transiting Exoplanet Survey Satellite (TESS). Este capítulo propone un enfoque de sonificación interactivo, multicanal y multivariable, capaz de generar paisajes sonoros complejos, a la vez que proporciona servicios de VO para los objetos representados. Como prueba de concepto, se incluye la representación sonora secuencial de 132 objetos de interés (TOI) que albergan, al menos, un planeta confirmado durante la exploración de 48 sectores de la misión TESS.

Capítulo 5. Musificación del análisis de periodogramas en curvas de luz

Descripción de las estrategias de desarrollo propuestas para la generación de partituras musicales a partir del análisis de periodicidad en series temporales. Basado en los métodos *Lomb Scargle* y *Box Least Squares*, este capítulo proporciona resultados musicales obtenidos a partir de la conversión en notas musicales de los periodos encontrados en las curvas de luz del catálogo de objetos de interés de la misión Kepler (KOI).

Capítulo 6. Sonificación simbólica espacializada

Descripción del modelo simbólico para la representación multimodal espacializada interactiva de las curvas de luz del archivo de Convección, Rotación y Tránsitos planetarios (CoRoT) del Institut d'Astrophysique Spatiale (IAS). Este capítulo proporciona representaciones sonoras en configuración ambisónica o binaural seleccionable por el usuario y basada en las coordenadas reales de los objetos. La sonificación se realiza mediante una estrategia de mapeado simbólico que convierte cada tipo espectral en una nota musical, ofreciendo una representación de estética musical, que además permite la diferenciación de espectros estelares a través del sonido.

Capítulo 7. Sonificación basada en deep learning

Descripción de la metodología de sonificación basada en deep learning que convierte en sonido los vectores latentes extraídos de los espectros estelares mediante autoencoders. Este capítulo explora los resultados obtenidos utilizando sparse autoencoders y autoencoders variacionales. Se incluyen pruebas de concepto con prototipos que utilizan síntesis de sonido aditiva para generar sonificaciones no supervisadas en formato estéreo, envolventes multicanal, binaurales y ambisónicas, de las librerías de espectros estelares MILES (IAC) y STELIB (SVO).

Capítulo 8. Sistema de composición musical basado en deep learning

Diseño e implementación de un sistema de composición autónomo completo basado en deep learning que genera piezas originales a partir de espectros estelares. Este capítulo proporciona una visión completa del sistema de composición basado en redes neuronales, desarrollado para la creación de obras musicales polifónicas no supervisadas a partir de los espectros estelares de la librería MILES.

Capítulo 9. Instrumentos musicales para la exploración de catálogos astronómicos

Diseño e implementación de un sintetizador que permite la exploración multimodal de la librería de espectros estelares del triplete de Ca II (CaT) del Observatorio Virtual Español (SVO). Este prototipo reacciona “a tiempo real” proporcionando un acompañamiento armónico para cada nota del teclado del piano, generado a partir de datos de objetos reales mediante deep learning.

Capítulo 10. IFS multimodal interactiva

Este capítulo explora el potencial de la sonificación en el campo de la espectroscopía de campo integral (IFS) a través de la implementación y validación del módulo *SoniCube*. Este módulo de sonificación ha sido incorporado en la herramienta de análisis IFS *ViewCube* desarrollada por Rubén García-Benito del Instituto de Astrofísica de Andalucía (IAA-CSIC). La aplicación multimodal final proporciona visualizaciones y sonificaciones de la información espectral contenida en cubos de datos a través de un entorno interactivo de dos dimensiones.

Capítulo 11. Evaluación

Análisis y resultados de los tres estudios realizados con participantes especializados y no especializados en Astronomía y Música. Este capítulo proporciona información cuantitativa y cualitativa de más de 150 participantes, evaluando las propuestas y proporcionando orientaciones que pueden ser de utilidad en futuros desarrollos.

Capítulo 12. Discusión

Discusión de los principales hallazgos y resultados de la investigación, incluyendo posibles líneas de investigación y prospectiva.

Capítulo 13. Conclusión

Contribuciones y limitaciones de la investigación en su esfuerzo por demostrar el valor añadido de la sonificación en la representación de datos astronómicos.

Capítulo 2

Precedentes y estado del arte

El Universo, entendido no sólo como fuente de inspiración sino también como fuente de armonía musical, ha ocupado la mente de matemáticos, músicos y astrónomos desde los tiempos de la antigua Grecia (ver Anexo.20). Actualizando este concepto a la tecnología actualmente disponible, la conexión de diferentes flujos de datos astronómicos con la generación y control de variables de sonido, abre una ventana de posibilidades para la exploración y comunicación científica, así como para el diseño sonoro y la composición musical.

2.1 Breve historia sobre armonías

Parece acordado mencionar a Pitágoras (siglo VI a.C.) y su *Música de las esferas* como la primera obra de referencia práctica y teórica en el campo interdisciplinar que relaciona Música y Astronomía. Según el etnomusicólogo Mark Ballora, el uso en la escuela pitagórica del monocordio para representar las relaciones entre intervalos de quintas perfectas, y las distancias a la tierra de los cuerpos que en ese momento se creía que orbitaban a su alrededor, también puede verse como la primera evidencia de sonificación de datos astronómicos (Ballora, 2014). Esta idea de relacionar los intervalos musicales con las distancias y velocidades orbitales de los planetas trasciende desde la antigüedad al Renacimiento de la mano de Platón y Aristóteles. En cuanto a Platón, a través de dos fuentes principales: *El mito de Er* y *El pasaje del Timeo*, ambos pertenecientes a su obra de diez volúmenes *La República* (Stephenson, 1994). En cuanto a Aristóteles, a través de su clara descripción de la teoría de la armonía de las esferas en *De caelo* 290 b12 y en 291 a8, en la que afirma que este concepto es pitagórico (Pabón, 2009). Sin embargo, y a pesar de que poco se sabe con certeza sobre las doctrinas pitagóricas, autores especializados en teorías griegas de la armonía de las esferas como Von Jan (Jan, V. 1893, citado en Stephenson, 1994, p.23), afirmaron que las referencias astronómicas de la escuela pitagórica deben interpretarse como simples analogías y no como una teoría astronómica debido a su limitado conocimiento sobre el tema.

En el siglo II d.C., Ptolomeo también esbozó los conceptos de armonías de los cuerpos celestes en su libro *Armónicos*, describiéndolos como meras conexiones racionales que obedecen a las leyes generales del movimiento. Para Ptolomeo, el movimiento ordenado, tanto en las estrellas como en la música, sigue ciertos patrones, de modo que el estudio de estos patrones en un

campo puede ayudar a la comprensión de otros campos (Stephenson, 1994).

En cuanto a la cultura occidental, los escritos de Boecio (480-524 d.C.) sobre la lógica aristotélica y el *Quadrivium*, tuvieron un papel muy influyente en la difusión de las teorías musicales pitagóricas durante la Edad Media. Desde principios del siglo VI, la Aritmética, la Geometría, la Astronomía y la Música, entendida como la ciencia de los números que describen el sonido, representaron los cuatro campos principales de la ciencia cuantitativa. El término *Quadrivium* se utilizó para agrupar estas disciplinas, consideradas como las cuatro ramas de las Matemáticas, capaces de describir el conocimiento del mundo natural. Boecio estableció una clasificación tripartita de la música -Mundana, Humana e Instrumentalis-, que implicaba la aceptación de las relaciones planetarias con los intervalos musicales desde una posición aparentemente continuista. *Música mundana*, refiriéndose a la armonía de las esferas del cielo, *Música humana* tratando la influencia de la música en el alma humana y *Música instrumental*, lo que actualmente conocemos como música (Stephenson, 1994). La sinfonía No. 41 *Júpiter* de Mozart podría haber sido inspirada por la definición de música mundana de Boecio (D. Zhang, 2023).

Las ideas pitagóricas y neopitagóricas también influyeron en el mundo musulmán. En el siglo X, la relación entre los fenómenos astronómicos y la música se presentó en la *Risala fi'l-Musiqā* (Epístola sobre la Música) de los Ikhwan as-Safa (Shiloah, 1978). Este tratado describe cómo las proporciones entre los diámetros de la Tierra, la esfera de aire (en referencia al concepto antiguo de atmósfera), la Luna, Venus, el Sol, Júpiter y las estrellas fijas correspondían a las proporciones de intervalos musicales armónicos perfectos, mientras que las proporciones de los intervalos de Mercurio, Marte y Saturno no lo hacían (Pacholczyk, 1996).

Volviendo a occidente, durante el siglo XV, humanistas como Coluccio Salutati o músicos como Johannes Tinctoris rechazaron la idea de la existencia de la música mundana o Mundane, mientras que teóricos de la música como Franchino Gaffurio afirmaron que sólo podía ser escuchada por personas verdaderamente virtuosas. A finales del siglo XVI, en un intento por describir la música en su totalidad, Gioseffo Zarlino publicó el tratado *Istitutioni harmoniche* (1558), que difundió la teoría armónica pitagórica como punto de referencia para la teoría musical de la época.



Figura 2.1: Extracto del *Libro V* de *Harmonices Mundi*, Johannes Kepler, 1619 (Kepler, 1969, p.207). Representaciones musicales de Saturno, Júpiter, Marte, la Tierra, Venus, Mercurio y la luna.

Lejos del misticismo de los antiguos griegos, los enfoques de Ptolomeo junto con los descubrimientos de las bases empíricas de las consonancias musicales de Vincenzo Galilei -padre de

Galileo- en su *Dialogo della musica antica, et della moderna* (1581), habrían inspirado la obra *Harmonices Mundi*, en la que Johannes Kepler expone cómo los planetas se mueven describiendo una órbita elíptica alrededor del sol. El libro fue publicado en 1619, y es considerado la obra maestra del pensamiento interdisciplinario astronómico-musical. En el *Libro V* de este tratado, Kepler tradujo los parámetros de movimiento y distancia de los planetas del sistema solar en intervalos musicales, algo que aparentemente le llevó a formular las ecuaciones que le permitieron sentar las bases de la Astronomía.

El extracto de la Figura 2.1 muestra las representaciones musicales del sistema solar realizadas por Kepler. Como puede apreciarse, la órbita de cada objeto fué representada por una escala musical, obtenida mediante la asociación de las distancias de afelio y perihelio con las relaciones de los intervalos musicales, haciendo coincidir sus frecuencias con las velocidades angulares de los objetos celestes. Cabe destacar que esta asociación se anticipó a la identificación de los conceptos de frecuencia y altura de un sonido (pitch), realizada por Mersenne en su ley de *L'harmonie universelle* de 1636 (García Martín et al., 2009). En palabras del propio Kepler, “La astronomía y la música son varias nacionalidades de una patria común, la geometría” (Kepler, 1619, citado en Smirnov, 1999).

Encontrando un equivalente óptico a la obra de Kepler, Isaac Newton propuso en su obra *Opticks* (1704), una analogía color-música para definir los siete colores espectrales asociados a los siete tonos de la escala diatónica (Gouk, 1986). Como puede verse en la Figura 2.2, impuso una octava musical a los colores del espectro prismático, y observó cómo los anillos de color en cuerpos transparentes delgados parecían abarcar una sexta mayor, proporcionando una evidencia significativa que ya apuntaba hacia la teoría ondulatoria de la luz (Pesic, 2006).

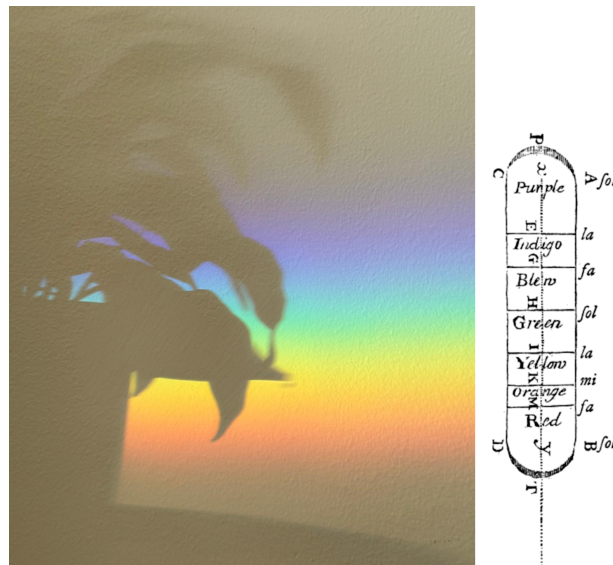


Figura 2.2: Luz natural en casa y diagrama de la analogía color-música descrita por Newton en su segundo artículo sobre luz y colores, presentado a la Royal Society en 1675 (Pesic, 2006, p. 294).

2.2 Visión general del Observatorio Virtual

La *Carte du Ciel*, lanzada en el Observatorio de París en 1887, fue el proyecto más extenso de cooperación internacional en el campo de la Astronomía que puede considerarse como un precedente de la Unión Astronómica Internacional (IAU). Su objetivo era producir una carta y un catálogo de todas las estrellas visibles, lo que llevó al establecimiento de la cooperación mundial promovida con la fundación de la IAU en 1919 en Bruselas, Bélgica (Montmerle y Fauque, 2022).

En 1973, se creó el Centre de Données Stellaires (o Centro de Datos Estelares, CDS) en Estrasburgo, Francia, para combinar los catálogos estelares acumulados en los observatorios europeos. Los años siguientes se caracterizaron por el creciente volumen de datos astronómicos y la digitalización de los catálogos. La necesidad de una estructura a escala global llevó a la creación en 2000 del Observatorio Virtual Internacional (IVO), con el foco puesto en los aspectos técnicos del intercambio de datos y en la interoperabilidad. En 2002, se creó la Alianza Internacional del Observatorio Virtual (IVOA), para el desarrollo e implementación de estándares de intercambio de información entre los participantes de todo el mundo (Montmerle y Fauque, 2022).

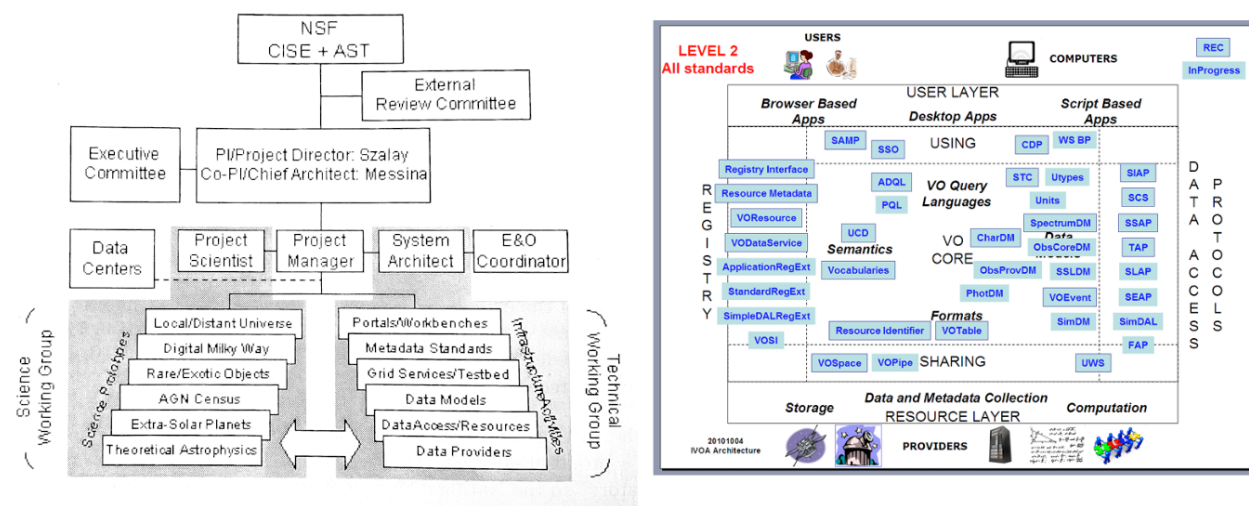


Figura 2.3: Izquierda: Organización de la infraestructura del Observatorio Virtual Nacional (NVO) en 2003 (Quinn y Górski, 2004), proyecto liderado por el investigador principal Dr. Alexander Szalay (Universidad Johns Hopkins) y el co-investigador principal Dr. Roy Williams (Instituto Tecnológico de California). Derecha: Arquitectura IVOA que incluye los estándares en 2010 (Arviset y Gaudet, 2010).

Actualmente, la tecnología del Observatorio Virtual (VO) (Djorgovski y Williams, 2005) permite la exploración del universo con un grado de accesibilidad sin precedentes, contribuyendo a aumentar la comprensión y el conocimiento global del universo sobre un paradigma basado en Ciencia Abierta (Molinari et al., 2020). Esta iniciativa combina la experiencia de observatorios astronómicos, centros de archivos de datos, departamentos de astronomía universitaria y grupos de informática y tecnología de la información en diecisiete organizaciones diferentes

(Quinn y Górski, 2004). La gran mayoría de los datos que se utilizan en el Observatorio Virtual son de dominio público, lo que significa que cualquiera en el planeta puede leerlos, pero sólo el personal del archivo puede modificar los archivos físicos (Rixon et al., 2004).

La Figura 2.3 muestra la infraestructura organizativa del Observatorio Virtual Nacional de EE. UU. (NVO), uno de los socios fundadores de la Alianza Virtual Internacional (IVOA) junto con el Observatorio Virtual Astrofísico del Observatorio Europeo Austral (AVO, ESO) y Astrogrid del Reino Unido (Hanisch et al., 2015). La figura también proporciona una visión general de los estándares y la arquitectura de IVOA en 2010. Las últimas actualizaciones de los estándares y la arquitectura técnica del IVOA pueden consultarse en Dowler et al., 2021.

Object	Model	T _{eff}	LogL		Age		Mass			
KIC6504534	BHAC15	4800	(4750,4850)	-0.5195	(-0.5348,-0.5048)	0.0428	..(0.0397,0.0500)	0.8271	(0.8048,0.8595)	[1]

[1] The distance to one of the closer curves has been estimated as the one to the closest point in the curve

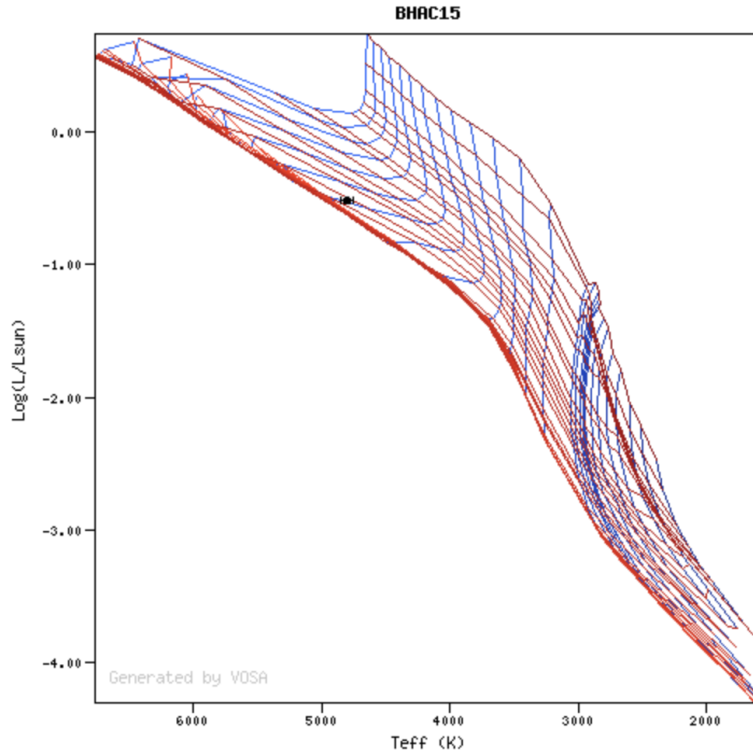


Figura 2.4: Diagrama HR de Kepler 1661 obtenido con VOSA (SVO).

Los recursos interoperables basados en estándares VO que los astrónomos utilizan en el análisis de sus casos de estudio se incluyen en herramientas como el atlas celeste *ALADIN* (Boch y Fernique, 2014; Bonnarel et al., 2000; CDS, 2000a), que proporciona acceso directo a servidores astronómicos activos en todo el mundo; *Simbad* (CDS, 2000b; Wenger et al., 2000), que ofrece información sobre los objetos astronómicos estudiados en artículos científicos; *Vizier* (CDS, 2000c), que proporciona la librería completa de catálogos astronómicos publicados; *TOPCAT* (M. Taylor, 2005; M. B. Taylor, 2005), que permite la edición y visualización gráfica

interactiva de datos en formato tabular; o el analizador de distribución de energía espectral (SED) del SVO *VOSA* (Bayo et al., 2008; SVO, 2008), que calcula estimaciones de masa y edad de las fuentes estudiadas.

Como ejemplo de las posibilidades de este tipo de herramientas, la Figura 2.4 muestra el diagrama de Hertzsprung-Russel (HR) (ver Anexo.5) de KIC 6504534 (Kepler 1661) obtenido con *VOSA*, herramienta de análisis desarrollada por el SVO. KIC 6504534 fue descubierto y catalogado como un sistema binario de 28,2 días en la segunda revisión del Eclipsing Binary Kepler catalog (Prša et al., 2011, Slawson et al., 2011). El sistema alberga un planeta circumbinario en tránsito del tamaño de Neptuno ($R_p = 3,87 \pm 0,06 R_{\text{Tierra}}$), llamado Kepler-1661 b (Socia et al., 2020b), que fue encontrado en la fotometría de la misión Kepler. El planeta tiene un periodo cercano a 175 días y su órbita precede con un periodo de sólo 35 años. El catálogo KIC proporciona las siguientes estimaciones para los parámetros estelares: $Kepmag = 14,216$, $T_{\text{eff}} = 4748 \text{ K}$, $\log g = 4,46$, metalicidad = -0,10 y una contaminación = 0,00 para las cuatro temporadas de Kepler. Como puede observarse, la estimación de la temperatura efectiva con *VOSA* ($T_{\text{eff}} = 4800 \text{ K}$) es consistente con la literatura.

2.3 Análisis científico y divulgación de catálogos astronómicos

Tal y como mencionamos en la Sección 2.2, actualmente la mayoría de archivos, catálogos y librerías de objetos astronómicos tales como MAST (STScI, 2008), SDSS(NSF, 2025) o GAIA (ESA, 2025), son de dominio público. No obstante, el uso y la interpretación de su información varía notablemente en función de los niveles de experiencia y objetivos de los usuarios.

Las personas dedicadas a la investigación científica en Astronomía y Astrofísica, generalmente trabajan con grandes conjuntos de datos de observaciones capturadas mediante telescopios como Kepler (M. STScI, 2017), Hubble (NASA, 2025a), JWST (NASA, 2025b), Nancy Grace Roman (NASA, 2025c), ALMA (ESO, 2025), CAHA (CAHA, 2025) o Vera C. Rubin (NSF, 2023).

A menudo, estas observaciones pasan por procesos de reducción de datos (Tody, 1986, 1993) y calibración para permitir su estudio (Bristow et al., 2008; Freudling et al., 2013), que habitualmente se basa en análisis estadísticos (Feigelson et al., 2021), técnicas de machine learning (Bufano et al., 2022) y simulaciones (Villaescusa-Navarro et al., 2021) que permiten caracterizar mediante comparación los fenómenos observados.

Además, las observaciones suelen organizarse en librerías de objetos seleccionados según el campo de estudio y los objetivos del proyecto (SVO, 2019), lo que proporciona versiones procesadas, etiquetadas y simplificadas de los grandes conjuntos de datos (Guillochon et al., 2017), que pueden ser utilizadas por los estudiantes a lo largo de su carrera formativa.

En cuanto al uso de los archivos por parte del público en general, además de los portales institucionales, existen plataformas interactivas como Stellarium (D. Brown, 2015; Chéreau et al., 2025; Zotti y Neubauer, 2024), o WW Telescope (Carver, 2023), que proporcionan información en forma de planisferios virtuales, así como portales educativos (IAA, 2025;

IAC, 2025) con numerosas actividades participativas que han contribuido activamente a la proliferación de proyectos de ciencia ciudadana (Zooniverse, 2025), divulgación científica (ICCUB, 2025) y educación (AURA, 2016; Caltech, 2020; CfA, 2025; ESO, 2010; NASA, 2023).

2.4 Sonificación de datos astronómicos

La sonificación se ha utilizado ampliamente para mostrar información en numerosos campos de investigación, especialmente desde el auge de la tecnología de computación digital a mediados de los noventa (Walker, 2023). Su aplicación en Astronomía ofrece ejemplos anteriores de análisis de datos a través del sonido, como el estudio del “ruido estelar” de Grote Reber, quien confirmó en 1938 la existencia de emisiones de radio no térmicas en la Vía Láctea (descubiertas previamente por Jansky, 1933), lo que llevó a notables descubrimientos posteriores en el campo de la radioastronomía (Kellermann y Bouton, 2023).

Ejemplos más recientes incluyen la sonificación utilizada en el observatorio de ondas gravitacionales del interferómetro láser LIGO para convertir señales de ondas gravitacionales en sonidos audibles (“LIGO Laboratory, Gravitational Wave Open Science Center”, 2021), la sonificación basada en piano para la misión espacial del interferómetro láser LISA (Valle y Korol, 2022), las aplicaciones de Helioseismología que representan números de manchas solares e irradiaciones espectrales ultravioleta extremas (EUV) (Pesnell et al., 2021), y las sonificaciones de datos de series temporales fotométricas de Kepler para V344 Lyra y V1504 Cyg (Tutchton et al., 2012).

Proyectos institucionales como *Astronify* del STScI (Brasseur et al., 2020), *Sound from Around the Milky Way* (NASA y the Harvard-Smithsonian Center for Astrophysics, 2019) y *A Universe of Sound* (K. Arcand et al., 2020) del centro de rayos X Chandra de la NASA, coexisten con equipos de investigación internacionales que desarrollan composiciones, estudios y herramientas que contribuyen a establecer el estado del arte en este campo.

Obras como *Sonification of Dark Matter* (Bonet et al., 2016), *Sonification of the zCOSMOS Galaxy Dataset* (Bardelli et al., 2021), *Sonification of Planetary Orbits in Asteroid Belts* (Quinton et al., 2021), proyectos como *CosMonic* (García-Benito, 2019), y herramientas como *x-sonify* (García et al., 2019) o *Sonification Sandbox* (Walker y Cothran, 2003), actualizada a *Highcharts Sonification Studio* (Cantrell et al., 2021), son ejemplos destacados del marco internacional interdisciplinar en el que se ubica esta tesis.

La combinación de representaciones visuales y sonoras puede ofrecer una mejor comprensión de un fenómeno (Enge et al., 2024), lo que hace que el uso del sonido para la representación de variables físicas sea un área de investigación establecida (Dubus y Bresin, 2013). También puede permitir una interpretación holística de los datos para descubrir relaciones previamente invisibles (Cooke et al., 2017), o facilitar tareas analíticas individuales, que incluyen estimación y comparación de puntos, identificación de tendencias y análisis de las estructuras de datos (Walker y Nees, 2011).

Además, tal y como describieron Alexander et al., 2014, la sonificación puede ser útil para detectar anomalías técnicas, como el ruido artificial inducido por instrumentación en series

temporales de magnetómetros de alta resolución, y para identificar la actividad ondulatoria incrustada en el viento solar turbulento en datos Wind-MAG de 11Hz.

La sonificación ha demostrado ser una herramienta muy útil para ayudar en el análisis interactivo de datos hiperspectrales, generando series temporales sónicas relacionadas con el contenido espacial y espectral relativo a las posiciones del ratón seleccionadas por los usuarios (Bernhardt et al., 2007). También puede mejorar la percepción de la densidad en la visualización de datos complejos presentados en coordenadas paralelas y diagramas de dispersión (Rönnberg y Jimmy, 2016), y garantizar un acceso simple a la información para personas ciegas y no ciegas, mejorando la accesibilidad de los datos astronómicos y el trabajo del personal científico con interés en metodologías de exploración complementarias (Casado et al., 2024).

Además, la mejora de la información visual mediante la incorporación de sonificaciones permite a las comunidades videntes y ciegas o con baja visión (BLV) tener experiencias astronómicas a niveles similares (K. K. Arcand et al., 2024). Incluyendo a personas ciegas desde el inicio del proceso de diseño, Casado y García, 2024 propusieron el caso de estudio de las galaxias del *Sloan Digital Sky Survey* (SDSS) para mostrar las posibilidades de *sonoUno*. Su aplicación multimodal, destinada a mostrar sonido e imágenes de cualquier conjunto de datos, permitió el descubrimiento de la estrella variable UCAC4 459-09273 por estudiantes BLV utilizando la sonificación. Como investigador ciego, Foran et al., 2022 confirmó del uso de *StarSound* en el trabajo de galaxias de alto corrimiento al rojo 1D, para la verificación y análisis inicial del marco en reposo de los espectros ultravioleta (UV) de galaxias distantes. También desarrolló la herramienta de sonificación táctil *VoxMagellan* para analizar imágenes 2D y conjuntos de datos multidimensionales.

En el campo del análisis fotométrico y espectroscópico, Trayford et al., 2023 propusieron la audificación de cubos de datos espectrales (conversión directa de datos en frecuencias audibles), para demostrar que la información física se puede extraer directamente del sonido con STRAUSS (Trayford y Harrison, 2023). Utilizando *Star Sounder*, Huppenkothen et al., 2023 proporcionaron una sonificación interactiva del diagrama de Hertzsprung-Russell basada en la coincidencia cruzada entre la tabla estelar de Kepler y la versión de datos DR2 de Gaia. La introducción de una perspectiva sónica también fue valorada en el análisis espectroscópico de quásares por Hansen et al., 2020, concluyendo que, en espectros de candidatos a sistemas de medio intergaláctico/circungaláctico (IGM/CGM), la sonificación puede permitir un descubrimiento e identificación de manera más rápida que el escaneo visual.

Incluyendo la interactividad multimodal, Starks, 2018 exploró los cubos de datos de la radioimagen de las *Galaxias Antena* del Atacama Large Millimeter/submillimeter Array (ALMA), utilizando *Galaxy player* dentro del proyecto *Soniverse*. Además, el uso de sonificaciones espacializadas fue destacado por Quinton et al., 2020, 2021, como una estrategia de mapeado de parámetros eficaz y con el potencial de detectar cambios repentinos entre múltiples fuentes dentro del campo de la búsqueda de planetas exosolares. La espacialización sonora también se implementó en representaciones inmersivas de datos de astronomía antártica (West et al., 2018), en sonificaciones de datos simulados de fusión de agujeros negros (St George et al., 2018), en el proyecto *Kepler Concordia* (Snook et al., 2020), y en la experiencia multi-participante performativa *INSTRUMENT* (West et al., 2018).

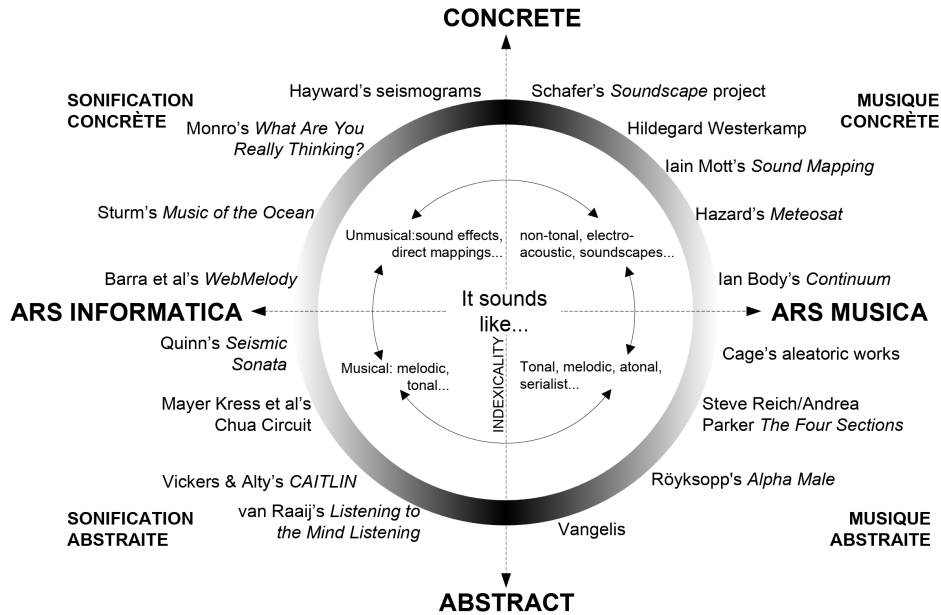


Figura 2.5: Espacio circular Ars Informatica-Ars Electronica (P. Vickers y Hogg, 2006).

El potencial de la sonificación para la expresión artística también ha sido reivindicado como un aspecto importante a considerar por pioneros en el campo (Gresham-Lancaster, 2012) o grupos de investigación como RADICAL (Boehringer et al., 2022). Se ha incluido en los premios anuales de sonificación (Lenzi et al., 2024), y se ha materializado en el contexto de interpretaciones musicales de datos astronómicos con obras como *CMB (Cosmic Microwave Background)* (Ducharme, 2023), o *When parallel lines converge* (K. Arcand et al., 2020).

Las representaciones de “Ars Informatica-Ars Musica” de la Figura 2.5, definidas por P. Vickers y Hogg, 2006, dibujaron el espacio en el que la música y la sonificación pueden tener lugar, ofreciendo, al mismo tiempo, una visión panorámica de las intersecciones entre ambos fenómenos sonoros. Además, Bonet Filella, 2019, propuso el concepto de “musificación como sonificación organizada”, compartiendo datos complejos con todo tipo de audiencias.

2.5 Música algorítmica e inteligencia artificial

Una de las primeras expresiones de creación de un modelo de composición musical sistematizado se encuentra en el tratado *Micrologus* (1030) de Guido d’Arezzo, que describió un mecanismo de generación automática de melodías a partir de texto en el que letras y sílabas fueron mapeadas a alturas y grupos de alturas de notas, denominadas neumas. Tal y como puede verse en la Figura 2.6, “La correspondencia entre el discurso (oratio) y el canto (cantus), estaba hecha” (Sullivan, 1989).

Atendiendo a la introducción de factores aleatorios en la composición musical occidental, destacó el *Würfelwalzer* o *Vals de Dados*. Esta forma de composición introdujo elementos de azar mediante el lanzamiento de un dado para determinar qué frases musicales de una

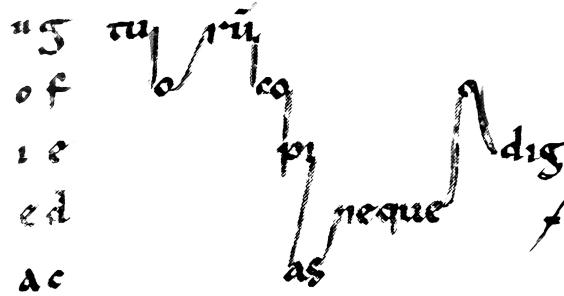


Figura 2.6: Neumas. Extracto de *Micrologus*, Guido d'Arezzo, 1030. Capítulo XVII, *Ut queant laxis*. (Paris, Bibliothèque nationale, Lat. 7211, f. 85v. XII^e siècle, Sud de la France)(D'Arezzo, 1996, p. 81).

composición preparada se tocaban. Desde 1757 hasta 1812, se publicaron al menos veinte juegos de dados en Europa (Albini, 2018), volviéndose bastante populares con el *Musikalisches Würfelspiel* o *Juego Musical de Dados* de Wolfgang Amadeus Mozart, publicado póstumamente en 1792 (Hedges, 1978).

Respecto a los primeros instrumentos musicales basados en principios electrostáticos simples, de los cuales se tiene constancia, cabe destacar el *Clavecin Électrique* diseñado por Jean-Baptiste Delaborde en Francia en 1759. Aunque se desarrollaron varios prototipos de instrumentos musicales electrónicos a finales del siglo XIX por ingenieros eléctricos tales como Elisha Gray, William Duddell o Ernst Lorenz, Thaddeus Cahill no completó el primer modelo de su *Telharmonium* hasta 1900 (D. Dunn, 1992).

Durante las décadas siguientes, la popularidad de los desarrollos tecnológicos aumentó el número de nuevas expresiones e instrumentos musicales electrónicos. En 1913, Luigi Russolo y Ugo Piatti construyeron la orquesta de instrumentos de ruido electromecánico *Intonarumori*. En Alemania, Jörg Mager desarrolló el instrumento de teclado basado en el oscilador de radiofrecuencia *Sphärophon* (1926). En Francia, Maurice Martenot introdujo sus *Ondes Martenot* (1928), y Leo Theremin presentó en el mismo año el *Theremin*. En 1929, en Estados Unidos, Laurens Hammond construyó su primer órgano electrónico y, en los años siguientes, patentó la técnica de reverberación de muelles, siendo el primero en lograr la producción en masa de instrumentos musicales electrónicos (D. Dunn, 1992).

En Estados Unidos, en 1939, Homer Dudley creó el vocoder para aplicaciones no musicales asociadas con el análisis del habla. En España, Juan García de Castillejo publicó el libro *La telegrafía rápida, el triteclado y la música eléctrica* (1944), que describe el electrocompositor automático que utilizaba desde 1933 (Noé, 2022). A finales de los años cuarenta, Pierre Schaeffer adoptó en Francia el uso de grabadoras de cinta y experimentó con la espacialización del sonido desarrollando su *Musique Concrète* (Schaeffer, 2012).

En 1951, el CSIR Mk1 se utilizó en la Conferencia de Computación Australiana para reproducir música (Doornbusch, 2004). Programado por Geoff Hill para reproducir melodías, fue probablemente la primera interpretación musical pública de una computadora, y el comienzo de la era de la informática musical. No obstante, parece que algunas notas musicales fueron generadas por un ordenador con un altavoz en el laboratorio de computación de Alan Turing

en Manchester en 1948 (Copeland y Long, 2016).

En 1956, Alan Turing planteó la pregunta fundacional “¿Pueden pensar las máquinas?” y John McCarthy propuso el término Inteligencia Artificial (IA) en el proyecto de verano de Dartmouth sobre Inteligencia Artificial (IA) (Turing, 2009). Un año después, Lejaren Hiller y Leonard Isaacson estrenaron *Illiatic Suite*, la primera composición musical construida por un sistema de IA e interpretada por un cuarteto de cuerdas, que introdujo varios paradigmas de la informática musical que aún permanecen vigentes hoy en día (Miranda, 2021).

Desde esta primera composición en 1957 hasta el momento actual de revolución tecnológica, la IA se ha utilizado ampliamente en composición musical (Hernandez-Olivan y Beltran, 2022), interpretación musical (De Mantaras y Arcos, 2002), diseño sonoro (Barahona-Ríos, 2023), producción musical y procesado de señal (Moysis et al., 2023, Purwins et al., 2019), en análisis musical (Temperley, 2004, Pons et al., 2017), en recomendación musical (Schedl, 2019) y en educación (Holland, 2013; Li y Wang, 2024).

Podemos distinguir tres periodos principales de tendencias y desarrollos en el campo de las herramientas musicales de IA, destacando algunos precedentes como el trabajo de Conlon Nancarrow para la integración de la automatización musical en pianolas, o los *schematas Rhythmicon* y *Musamaton*, contruidos por Leo Theremin para la composición de nuevas obras con máquinas (Cope, 1991).

Durante lo que puede considerarse como un periodo de exploración temprana (décadas de 1950-1970), marcado por la introducción de algoritmos informáticos para la composición musical y la generación de sonido, Max Mathews desarrolló los primeros programas de síntesis de sonido en los Bell Telephone Laboratories (Mathews, 1963); John Cage colaboró con Lejaren Hiller para crear HPSCHD (abreviatura informática de clavecín en inglés), una de las primeras interpretaciones multimedia concebidas como un collage aleatorio de capas superpuestas (Husarik, 1983); y Iannis Xenakis formalizó modelos matemáticos para componer Música Estocástica como *Achorripsis* (1957), cuya matriz de composición se muestra en la Figura 2.7, y puede interpretarse como una sonificación de distribuciones de probabilidad (Childs, 2002).

En 1969, John Edward Rothgeb completó su tesis sobre composición musical informatizada en la Universidad de Yale, en la que resolvió el problema de la armonización del bajo no cifrado (De Mantaras y Arcos, 2002); y en 1970 John Horton Conway propuso el modelo fundamental del autómatas celular a través del *Juego de Conway*, una secuencia de células que evolucionan determinísticamente en pasos de tiempo discretos según reglas definidas, involucrando los valores de sus vecinos más cercanos (Conway, 1978). En 1975, Laurie Spiegel compuso la versión de la música de las esferas que iniciaba el disco de oro enviado al espacio en la sonda Voyager (J. A. Caballero, 2024). A finales de los años setenta, el campo estaba en plena efervescencia, tal y como puede verse en los informes de las conferencias internacionales de música por computadora de 1977 y 1978 (“Papers from the 1977 International Computer Music Conference”, 1978; Roads, 1978).

En las décadas de 1980 y 1990, el uso de computadoras se extendió rápidamente gracias al desarrollo de la Interfaz Digital de Instrumentos Musicales (MIDI) (Loy, 1985), lo que permitió a las computadoras comunicarse con instrumentos digitales para mejorar su interactividad. En 1983, Stephan Wolfram ya había detallado los modelos matemáticos de los autómatas celulares,

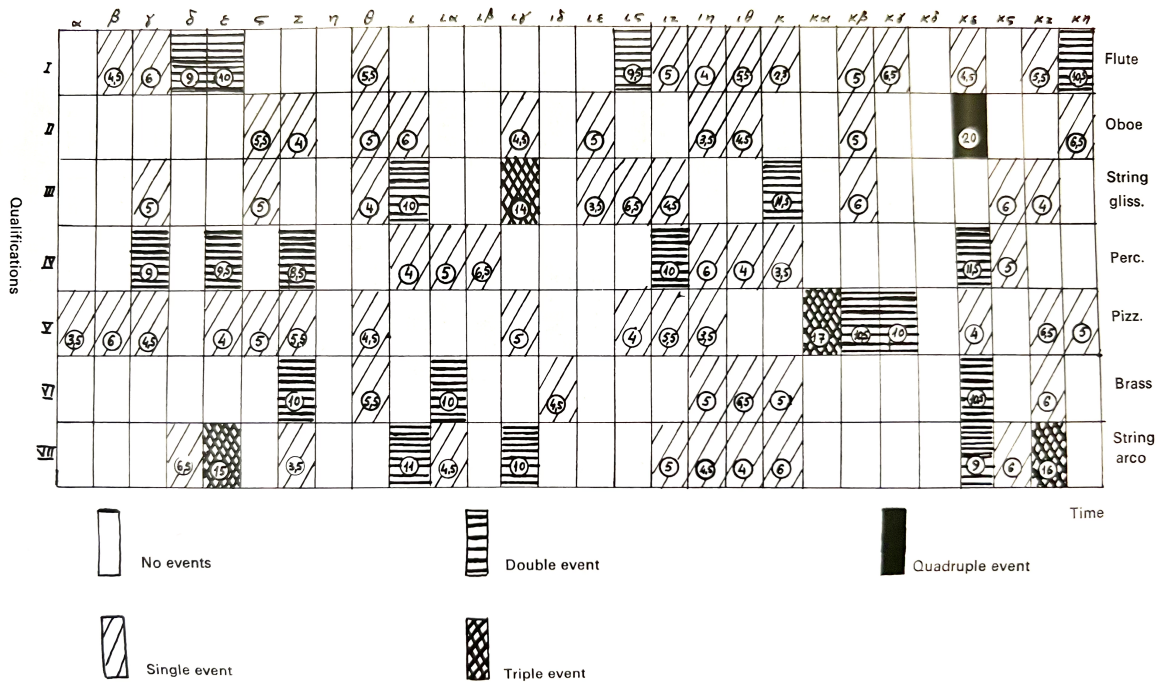


Figura 2.7: Extracto del manuscrito con la matriz de vectores de *Achorripsis*, 1957 (Xenakis, 1992, pp. 28-135). Las filas representan las clases tímbricas o agrupaciones instrumentales, las columnas representan bloques temporales o secciones de la pieza, que consta de 7 agrupaciones tímbricas intercambiables y 28 columnas o segmentos de tiempo.

demonstrando cómo a partir de un número reducido de reglas sencillas pueden llegar a generarse patrones complejos (Wolfram, 1983). Iannis Xenakis fue uno de los primeros compositores en aplicar el enfoque de los autómatas celulares para producir progresiones armónicas y nuevas combinaciones de timbres en su pieza *Horos* (Solomos, 2005). También fomentó la creación de nuevas formas de interacción musical a través del desarrollo de prototipos experimentales como el UPIC (Unité Polyagogique Informatique du CEMAMu), una máquina dedicada a la composición interactiva de partituras musicales a través del dibujo a tiempo real (Xenakis, 1992).

Basado en los lenguajes MUSIC-N desarrollados por Max Mathews, Barry Vercoe creó *CSound* en el MIT Media Lab (Vercoe et al., 1986), una poderosa herramienta abierta para la síntesis de sonido, la composición algorítmica y la música generativa, que marcó el comienzo del sonido generado por computadora a través de código. Por otro lado, David Zicarelli comercializó *M* (Zicarelli, 1987), uno de los primeros sistemas de composición interactivos que utilizaba MIDI para combinar técnicas algorítmicas y señales de entrada del usuario. En 1987, Pauline Oliveros estrenó en público el trabajo *Echoes from the Moon*, en el que utilizaba transmisiones de radio para proyectar sonido hacia la luna y escuchar sus reflexiones (Barrett, 2021). *Experiments in Musical Intelligence* (EMI) de David Cope (Cope, 1989) produjo algunos de los primeros proyectos de transferencia de estilo basados en IA, generando música clásica al estilo de Bach, Mozart o Chopin, a finales de los ochenta.

Con el objetivo de proporcionar un control a tiempo real sobre la música, orientado a las

interpretaciones en vivo, Miller Puckette desarrolló MAX (Puckette, 1991) en el Institut de Recherche et Coordination Acoustique/Musique (IRCAM) de París. Este software de programación gráfica proporcionaba un entorno unificado para describir los flujos de control y de señal de audio. Cinco años después, David Zicarelli lanzó una versión comercial del software y Puckette desarrolló un sucesor natural de código abierto, llamado Pure data (Puckette et al., 1996), que simplificaba su estructura de datos y permitía la representación de gráficos. En aquellos años, GenJam (J. Biles et al., 1994) ya estaba en el escenario, proporcionando una poderosa herramienta basada en algoritmos genéticos, y buscando espacios melódicos restringidos para construir solos de jazz “ajustados” por un mentor humano o por la audiencia al completo (J. A. Biles y Eign, 1995). Las primeras especificaciones del protocolo Open Sound Control (OSC), sucesor del MIDI y desarrollado en el Center for New Music and Associated Technologies (CNMAT) de la Universidad de California, fueron presentadas en la International Computer Music Conference (ICMC) de 1997, proporcionando una herramienta abierta para el control de dispositivos electrónicos basada en red, ampliamente utilizada en la actualidad.

El comienzo del siglo XXI se caracterizó por el desarrollo de sistemas de transmisión y recomendación basados en el auge de la recuperación de información musical (MIR). Este campo permitió el análisis de datos de audio, simbólicos y bibliográficos para extraer información musical del tono, el tempo, el timbre y el género (Downie, 2003), así como de la emoción (Baum, 2006), la tensión (Barchet et al., 2024) y la segmentación musical (McCallum, 2019). En el terreno artístico, Kaija Saariaho estrenaba *Asteroid 4179: Toutatis*, inspirada en el complejo movimiento de precesión del asteroide (Saariaho, 2005). Unos años antes, el desarrollo de MusicXML (Good, 2012), un sistema de codificación de texto basado en XML y diseñado para compartir archivos de partituras entre aplicaciones musicales, había ampliado las posibilidades del análisis simbólico. Una de estas aplicaciones, *music21* (Cuthbert y Ariza, 2010), integró con éxito potentes herramientas de software con conocimientos musicales sofisticados, para músicos con poca experiencia en programación y para programadores con habilidades modestas en teoría musical.

Desde la perspectiva de la interpretación en vivo, *Wekinator* (Fiebrink y Cook, 2010) permitió el uso del aprendizaje automático como meta instrumento, enfatizando la interacción humano-computadora para desarrollar rápidamente prototipos aplicados a diferentes campos, como por ejemplo la visualización basada en timbres o los motores de melodías improvisadas por el ordenador influenciadas por el tono; y McCartney, 2002 lanzó *Supercollider*, una plataforma para la síntesis de audio y la composición algorítmica que contribuyó a impulsar la escena del live coding (Collins et al., 2003). Conceptualizado en los años noventa y desarrollado hasta la fecha en Georgia Tech, Shimon, el músico robótico impulsado por IA, improvisa, compone y toca instrumentos musicales (Weinberg et al., 2008, 2020).

Tal y como analizan Civit et al., 2022, en la última década, el uso de modelos de deep learning ha fomentado el desarrollo comercial y la adopción generalizada¹ de herramientas de inteligencia artificial en música (Alom et al., 2018). Estudios como Briot et al., 2020, proporcionan un análisis detallado de numerosos sistemas basados en redes neuronales profundas para la composición y la interpretación musical. Algunos ejemplos pueden incluir modelos generativos

¹Herramientas comerciales de IA para la composición musical: Aiva, Suno, Jukebox, ACE-Step o Staccato.

de audio sin procesar como WaveNet (Oord et al., 2016), que opera directamente a nivel de forma de onda y se utiliza en el modelado de audio musical y el reconocimiento de voz; modelos polifónicos como DeepBach (Hadjeres et al., 2017), capaz de generar corales muy convincentes imitando el estilo de Bach; sistemas simbólicos como MidiNet, que utiliza una combinación de redes neuronales convolucionales y redes generativas adversarias (CNN-GAN) para la generación de MIDI; o modelos generativos como Music VAE (Roberts et al., 2018), que utiliza un decodificador jerárquico en una arquitectura de autoencoder variacional para mejorar las estructuras musicales a largo plazo en un modelo de trío formado por batería, bajo y melodía.

No obstante, fueron los modelos basados en transformers (Vaswani et al., 2017) los que marcaron el punto de inflexión en la aplicación masiva de la inteligencia artificial generativa a la creación musical. Y. Ma et al., 2024, ofrecieron una revisión de los modelos fundacionales aplicados a este campo, entre los que podemos destacar el music transformer (C.-Z. A. Huang et al., 2018) desarrollado por *Magenta* (Dubreuil, 2020), un departamento de investigación de IA de Google Brain que proporciona conjuntos abiertos de herramientas para generar secuencias y melodías monofónicas y polifónicas en MIDI y audio sin procesar (raw); y los sistemas basados en Large Language Models (LLM) (Raiaan et al., 2024) como Jukebox (Dhariwal et al., 2020) de OpenAI o ACE-Step (Gong et al., 2025) de StepFun/ACE Studio, que generan música en el dominio del audio sin procesar, condicionado por el artista y el género para dirigir el estilo musical y vocal, y que permiten incluso editar música y letra mediante descripciones de texto (Zhao et al., 2025).

La producción musical también es un campo prolífico para la inteligencia artificial. Actualmente se utiliza en iniciativas abiertas² y soluciones comerciales³ para la mezcla y masterización automáticas. Explorando los beneficios de la creación de subgrupos para mejorar la claridad en la mezcla, Ronan, 2019, propuso un sistema inteligente para minimizar los efectos de enmascaramiento en sesiones multipista. Basado en un ajuste de mínimos cuadrados para cualquier respuesta de frecuencia deseada, Z. Ma et al., 2013, implementaron un ecualizador inteligente para proporcionar curvas de ecualización apropiadas según el contexto cultural, de género y social de una pieza elegida.

Además, el uso de redes neuronales convolucionales (CNN), autoencoders, autoencoders variacionales (VAE) y redes generativas adversarias (GAN) permitió la mejora de la calidad de audio y la reducción de ruido, superando potencialmente las técnicas tradicionales de procesado de señal (Lohani et al., 2024). Estos enfoques también permitieron la separación de las señales deseadas del ruido de fondo con una eficacia sin precedentes⁴, incluso permitiendo la separación de instrumentos mezclados en una banda sonora para generar pistas aisladas en una sesión multipista⁵.

²Herramientas abiertas de IA: <https://github.com/sony/FxNorm-automix>, <https://github.com/ai-mastering>

³Herramientas comerciales de IA para mezcla y masterización: Roex, Landr, Cryo Mix o Neuron.

⁴Herramientas comerciales de IA para la eliminación de ruido: ElevenLabs o Voice AI

⁵Herramientas comerciales de IA para la separación de pistas de audio: Splitter AI o Stem Splitter

2.6 Tendencias actuales en composición musical automatizada

Tal y como comentamos en la Sección 2.5, la composición musical automatizada ha avanzado significativamente en los últimos años gracias al desarrollo interdisciplinar de modelos basados deep learning, lenguaje natural y síntesis de audio generativa (Z. Wang et al., 2024).

Los actuales sistemas basados en LLMs son capaces de generar audio raw (Kim et al., 2018; Oord et al., 2016), música simbólica (MIDI) y partituras (Deng et al., 2024; Ji et al., 2023; Shih et al., 2022; Xu et al., 2024), y de sintetizar audio a partir de ruido aleatorio mediante modelos diffusion (Q. Huang et al., 2023; Kong et al., 2020; Mittal et al., 2021) y GANs (Donahue et al., 2018; Engel et al., 2019; Kumar et al., 2019). La interacción con estos modelos se basa en *prompts* o descripciones introducidas por los usuarios (Xu et al., 2024; Y. Zhang et al., 2025) para controlar el modelo, que se usan para definir el estilo musical, la duración y estructura de la composición, la instrumentación, e incluso la letra de las pistas vocales, que son sintetizadas con absoluto realismo (Agostinelli et al., 2023; Borsos et al., 2023; Copet et al., 2023; Wei et al., 2024). Además, también permiten la creación de composiciones multimodales (Wu et al., 2024) que combinan música con imagen para generar bandas sonoras de películas (Haseeb et al., 2024; Luo et al., 2023; Xie et al., 2025) y contenidos audiovisuales, a partir de vídeos (Ji et al., 2025; L. Zhang y Fuentes, 2025) o fotogramas (Vasireddy et al., 2024).

En cuanto al potencial creativo de estos sistemas autónomos, cabe mencionar la asistencia a la composición tradicional (Laidlow, 2024), la mezcla experimental de estilos musicales (Jing et al., 2019) y la posibilidad de generar sonidos originales (Herremans et al., 2017). Del mismo modo, cabe destacar la posibilidad de crear hilos musicales continuos “a tiempo real”, personalizados y emocionalmente adaptativos (Gao et al., 2024), que permiten la generación automática de música en juegos (Marra y Ferreira, 2024), video juegos y aplicaciones de realidad virtual (Hutchings y McCormack, 2019; Kaae, 2017), espectáculos con narrativa o interpretaciones en vivo (Dubnov et al., 2022).

Capítulo 3

Exploración multimodal de catálogos estelares

La incorporación de representaciones sonoras a las herramientas de análisis gráfico puede conducir a una comprensión más rica de las imágenes astronómicas, los espectros y los datos numéricos utilizados en el análisis científico. Este capítulo describe el módulo fundamental que hemos desarrollado para sonificar flujos de datos, y que utilizamos en diferentes pruebas de concepto a lo largo de este trabajo de investigación. Como resultado, proporcionamos un método de generación de representaciones multimodales automáticas de datos unidimensionales, en forma de gráficos secuenciales visuales y sonoros, que permite la exploración de espectros estelares y catálogos de curvas de luz.

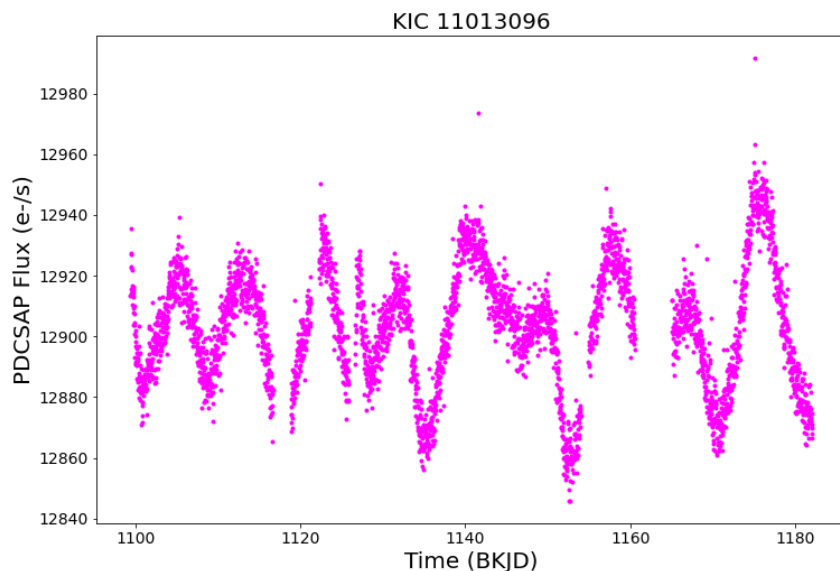


Figura 3.1: Curva de luz de la estrella variable KIC 11013096 con coordenadas RA:18:47:44.69 y DEC:+48:30:29.63. Variaciones de flujo PDCsap (electrones por segundo) en función del tiempo (BKJD = BJD-2454833). STScI.

3.1 Curvas de luz y espectros estelares

Las curvas de luz son representaciones gráficas bidimensionales de las variaciones del flujo de brillo, observadas en un objeto de investigación, a lo largo del tiempo. El análisis de estas series temporales, centrado principalmente en la detección y evaluación de la periodicidad, se utiliza comúnmente para la clasificación de sistemas binarios, estrellas variables y supernovas, así como para el descubrimiento de planetas extrasolares basados en el método de detección de tránsito (paso de un objeto entre una fuente de luz y el observador, ver Anexo.3 para una descripción detallada). La Figura 3.1 muestra un ejemplo de este tipo de series temporales correspondiente a la estrella KIC 11013096 .

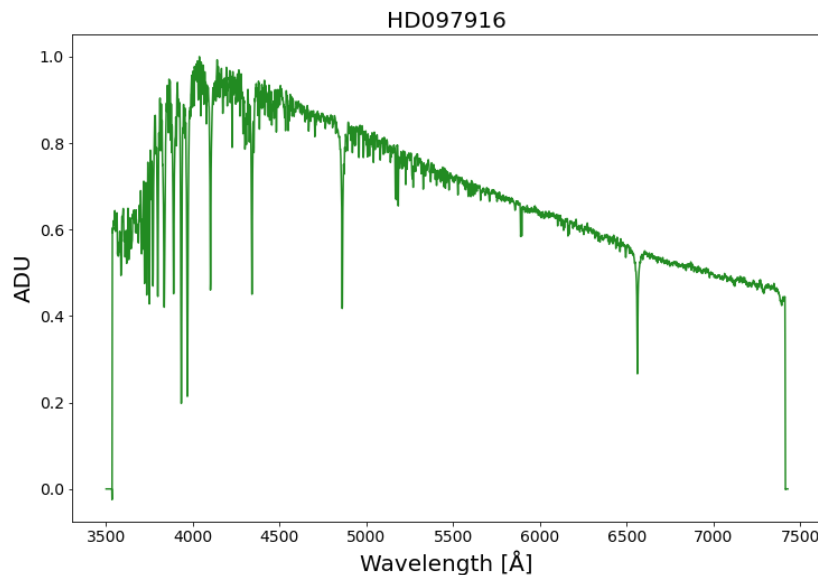


Figura 3.2: Espectro de la estrella HD 097916 (RA:11:15:54.23 DEC:+02:05:12.09). Variaciones de flujo normalizadas (ADU) en función de la longitud de onda (Å). SVO.

Los espectros estelares son representaciones gráficas de las variaciones del flujo de luz de un objeto de investigación en función de la longitud de onda. Su análisis se utiliza para determinar la composición química, la temperatura, la masa y la densidad de estrellas, galaxias y planetas dentro y fuera del sistema solar (ver Anexo.4). La Figura 3.1 muestra, a modo de ejemplo, el espectro estelar de la estrella HD 97916.

3.2 La librería STELIB

STELIB (SVO, 2003) es una librería estelar espectroscópica creada para ser utilizada en la síntesis de poblaciones y modelos evolutivos de galaxias. Incluye 249 espectros de estrellas en el rango visible (3200 a 9500 Å), incluyendo la mayoría de tipos espectrales y clases de luminosidad, y cubriendo un rango relativamente amplio en metalicidad. La librería proporciona un conjunto homogéneo de espectros estelares con una resolución espectral relativamente alta (alrededor de 3 Å), basada en observaciones recogidas en el telescopio Jacobus Kaptein de 1 m (JKT), en el Observatorio Roque de los Muchachos, La Palma, Islas

Canarias, España, y en el telescopio de 2,3 m de la Universidad Nacional Australiana en Siding Spring (SSO), Australia (Le Borgne et al., 2003). La librería está disponible públicamente en archivos FITS (IAU, 2018), proporcionando metadatos de encabezado y 6700 valores correspondientes a la variación de flujo de luz por espectro (ver Anexo.1 para la descripción del formato FITS).

3.3 El módulo FITS2OSC

Con el objetivo de proporcionar una primera aproximación a la exploración sonora de catálogos astronómicos, implementamos el módulo *FITS2OSC* (García Riber, 2022), que convierte las variaciones de flujo en el brillo de los objetos estudiados en mensajes Open Sound Control (OSC) (ver Anexo.2). El módulo envía valores numéricos “a tiempo real” al sintetizador, que realiza un mapeado paramétrico (Grond y Berger, 2011) sin restricciones de nota o escala, proporcionando un puente perfecto para la transducción directa entre variables astronómicas y parámetros de sonido. Concebido como una herramienta para escanear librerías estelares, y basado en la filosofía del gráfico sonoro (Walker y Mauney, 2010), este prototipo realiza un mapeado simple del valor de x al eje de tiempo, y del valor de y a la frecuencia que controla un oscilador sinusoidal (ver Anexo.7). De este modo, se pueden generar indistintamente representaciones sonoras de datos unidimensionales, tales como curvas de luz y espectros estelares.

El siguiente vídeo muestra el prototipo *FITS2OSC* durante la exploración de 256 espectros de la librería estelar STELIB (SVO). Como se puede observar, las líneas de absorción y emisión en los espectros se identifican fácilmente como valles y picos profundos que modifican bruscamente la altura del sonido (pitch), a lo largo de la línea de tiempo. La representación también permite comprender la distribución de energía del espectro en función de la longitud de onda, algo que podría ser útil para la clasificación sonora de espectros.

<https://vimeo.com/911990140>

En cuanto al diagrama de bloques del módulo, mostrado en la Figura 3.3, el pre-procesado de datos, la visualización y la generación de mensajes OSC se desarrollan en *Python*. La librería *astropy* (Price-Whelan et al., 2018; T. P. Robitaille et al., 2013) se utiliza para abrir el archivo FITS. El pre-procesado se realiza utilizando *numpy* (Harris et al., 2020; Van Der Walt et al., 2011), y la visualización gráfica se implementa con *matplotlib* (Hunter, 2007). Finalmente, se utiliza *python-osc* (“Python-OSC”, 2013) para generar los mensajes OSC a partir de los valores de flujo y las marcas de tiempo que se necesitan en el motor de síntesis de sonido. Para este último propósito, proponemos el entorno *CSound*, ampliamente documentado (Boulangier, 2000), ejecutado sobre el frontend de audio *Cabbage* (Walsh et al., 2008). Esta combinación proporciona un entorno extremadamente eficiente para diseñar, probar y comparar prototipos multimodales con audio de alta calidad. No obstante, el módulo permite el uso de cualquier sintetizador de sonido alternativo que admita mensajes de entrada *OSC*, como por ejemplo *Max* (Puckette, 1985), *Pure Data* (Puckette et al., 1996) o *Supercollider* (McCartney, 1996). El prototipo está disponible en código abierto en el siguiente enlace:

<https://github.com/AuditoryVO/AutoFITS2Sound>

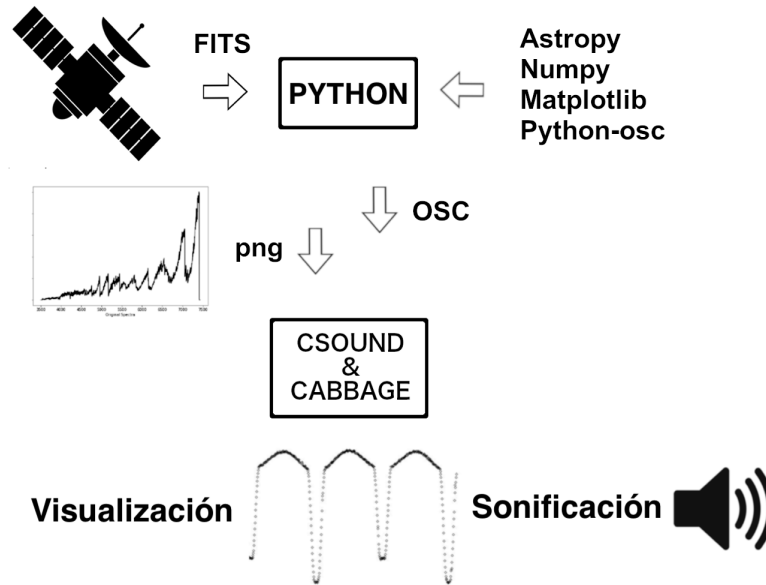


Figura 3.3: Diagrama de bloques del módulo de sonificación *FITS2OSC*.

3.4 Sonificación de curvas de luz con *FITS2OSC*

Tal y como describe Socia et al., [2020a](#), KIC 6504534 es un sistema binario de aproximadamente 28,2 días descubierto en la segunda revisión del catálogo de estrellas binarias de la misión Kepler. Alberga un exoplaneta en órbita, detectado y confirmado recientemente mediante el método de detección de tránsito fotométrico. Este objeto de interés llamado KIC6504534b o Kepler 1661-b es un planeta circumbinario en tránsito del tamaño de Neptuno, en el límite de la zona habitable, con un periodo de alrededor de 175 días. Se pueden encontrar detalles adicionales del planeta en Hanno Rein, [2020](#); NASA, [2020b](#).

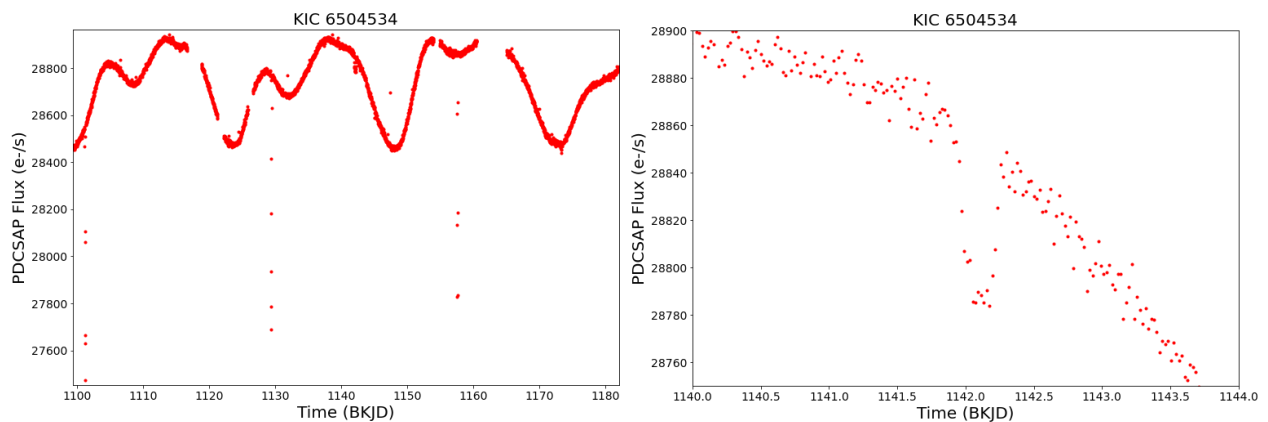


Figura 3.4: Curvas de luz representando las variaciones de flujo del sistema estelar binario KIC 6504534 con *FITS2OSC*. Flujo PDCsap (electrones por segundo) en función del tiempo en BKJD (BJD-2454833). Más información en NASA, [2011b](#), [2016](#).

Observando la primera curva de luz de la Figura 3.4 (izquierda), es posible apreciar una pequeña disminución del flujo de luz alrededor de 1142,1 BKJD, correspondiendo al segundo tránsito descrito en la caracterización del planeta, que se utilizó para permitir una estimación aproximada del periodo orbital del candidato. La segunda curva de luz de la Figura 3.4 (derecha), ofrece una vista ampliada de este tránsito planetario obtenida con las opciones interactivas del gráfico generado. Los cálculos utilizados para la descripción de los sistemas planetarios detectados con el método de tránsito se describen en el Anexo.3. El siguiente vídeo proporciona una demostración de la representación multimodal de la curva de luz de KIC 6504534 / Kepler 1661-b:

<https://vimeo.com/1008778114>

3.5 Sonificación de espectros estelares con *Astronify*

Astronify (Brasseur et al., 2020) es una aplicación implementada en *Python* para la sonificación de datos astronómicos, desarrollada por el Space Telescope Science Institute (STScI, 1981). Basada en el motor de generación de sonido Pyo (Belanger, 2016), inicialmente, la aplicación fue diseñada para sonificar curvas de luz. En 2022, en colaboración con el personal del Mikulski Archive for Space Telescopes (MAST), desarrollamos un módulo experimental en *Astronify* para la sonificación de espectros estelares. El siguiente vídeo muestra diferentes ejemplos de espectros sonificados (ver Anexo.4) utilizando este complemento experimental, que proporciona representaciones de tipo “vista previa” que podrían ser útiles para permitir la exploración preliminar de archivos espectrales a través del sonido.

<https://vimeo.com/787754481/786c33a8d0>

El módulo genera una representación gráfica esquemática del espectro, dividiendo el rango visible en bandas de longitud de onda asociadas a colores representativos tal y como muestra la Figura 3.5. En el prototipo inicial, se asociaron cinco ondas sinusoidales de diferentes frecuencias fundamentales a estas bandas para ser mostradas secuencialmente. Al final de la representación, todas las ondas sinusoidales se reproducían juntas aproximando un acorde A6/9 (azul = 110 Hz, verde = 165 Hz, amarillo = 277 Hz, naranja = 415 Hz y rojo = 698 Hz). Las frecuencias se modificaron finalmente a 300, 400, 500, 600 y 700 Hz para permitir la reproducción y diferenciación utilizando los altavoces genéricos de un ordenador portátil.

La amplitud del sonido de cada banda se calculó como el área bajo la curva de flujo utilizando la función `trapz` de *numpy* (Numpy, 2008). Esta estrategia permite la diferenciación sonora entre distribuciones de flujo, como puede apreciarse con claridad al comparar muestras espectrales de tipo O y M .

Además, el módulo ofrece una estimación de la presencia y abundancia de líneas de absorción y emisión para cada banda, representada por un efecto de trémolo. La frecuencia de oscilación del trémolo se basa en la desviación estándar de cada banda, normalizada por la desviación estándar máxima en todo el espectro completo.

Las versiones experimentales adicionales incorporan tres modos de sonificación seleccionables por el usuario, incluido un enfoque de piano panoramizado. Todas las versiones están disponi-

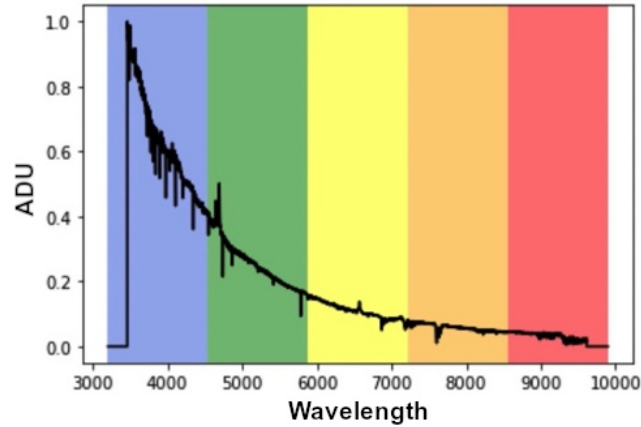


Figura 3.5: Espectro tipo *O* de HD 057060 con coordenadas RA: 07:18:40.38 DEC:-24:33:31.32. Representación gráfica de la división en bandas de color utilizada en el módulo de sonificación de espectros de *Astronify*. Flujo normalizado frente a longitud de onda (Å).

bles en el GitHub del STScI y se pueden utilizar en la evaluación de estrategias apropiadas para la generación de sonificaciones comprensibles, de manera análoga a la propuesta desarrollada en el Anexo .6 y evaluada en la Sección 11.1.1.

<https://github.com/spacetelescope/astronify/tree/main>

https://github.com/spacetelescope/astronify/tree/piano_mode

Capítulo 4

Paisajes sonoros multivariable

Los productos de las misiones de telescopios espaciales se pueden dividir en dos amplias categorías: datos sin procesar (raw) y datos procesados listos para el análisis científico. La primera categoría incluye la información original grabada, requerida en investigaciones en profundidad y aplicaciones específicas (Jaschek, 1989). Los espectros y curvas de luz que se describen en el Capítulo 3 pertenecen a esta categoría. A través de procesos de reducción y calibración de datos, los objetos astronómicos y sus atributos se extraen de los datos en bruto para proporcionar productos etiquetados listos para su uso científico (Szalay et al., 2001). Estos productos procesados incluyen información como el telescopio, la fecha y hora de la observación, las coordenadas y el tipo espectral de los objetos, o el nivel RMS de la señal registrada. Comúnmente están disponibles para la consulta instantánea a través de bases de datos y sistemas de información tales como Exo-Dat (Deleuil et al., 2009).

Este capítulo analiza las posibilidades de la representación sonora de múltiples variables simultáneas utilizando datos procesados. Aplicado al caso de estudio de los archivos de validación de datos en series temporales (DVT) de la misión Transiting Exoplanet Survey Satellite (TESS), se propone un enfoque de sonificación multivariable multicanal interactivo. La implementación práctica de la propuesta que hemos desarrollado proporciona paisajes sonoros complejos a partir de la temperatura efectiva, la metalicidad, la gravedad superficial y el radio estelar de las estrellas analizadas (ver Anexo.5), así como el periodo orbital, la duración del tránsito y la profundidad de los tránsitos planetarios (ver Anexo.3) incluidos en el catálogo TOI, de objetos de interés de la misión TESS (García Riber y Serradilla, 2022).

4.1 Archivos de validación de datos en series temporales de TESS

Los archivos DVT de TESS son el resultado de un análisis preliminar de posibles señales periódicas similares a tránsitos, encontradas en las curvas de luz de TESS, que se utilizan en el proceso de confirmación de tránsitos planetarios. En este proceso, las variaciones del flujo de brillo registradas por el telescopio y almacenadas en curvas de luz con cadencia de dos minutos, son analizadas por un módulo de búsqueda de planetas en tránsito (TPS). El TPS

crea eventos de cruce de umbral (TCE), que podrían ser “consistentes con planetas en tránsito, pero también con estrellas binarias, estrellas variables o incluso ruido” (Mullally y Fleming, 2018). Los archivos DVT archivan las series temporales relacionadas con la validación de datos para la búsqueda de tránsitos y el modelado de TCE, proporcionando una HDU primaria con metadatos y una extensión FITS HDU para cada TCE encontrado en la curva de luz del objeto observado, así como una extensión adicional con estadísticas sobre la búsqueda.

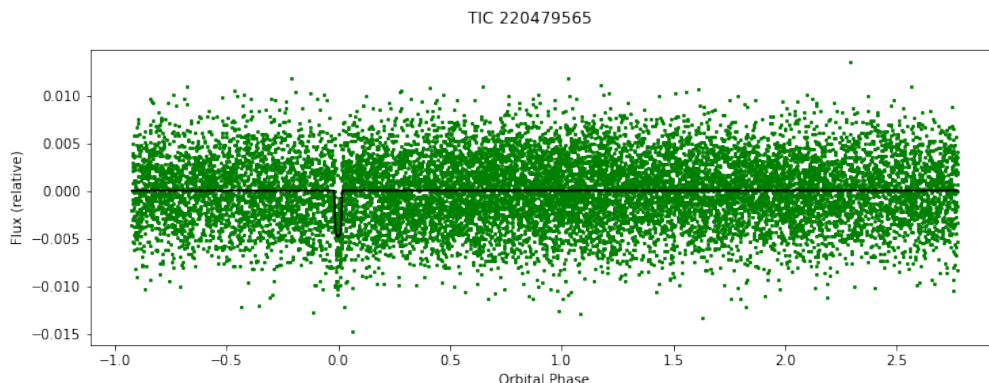


Figura 4.1: Curva de luz y modelo de tránsito para TIC 220479565, sub-Neptuno excéntrico transitando una estrella enana tipo M2 (Cointepas et al., 2021). Flujo relativo en función de la fase orbital ajustada al momento del tránsito.

Los archivos DVT incluyen pre-procesado relacionado con el ajuste de nivel, la eliminación de armónicos, la normalización y la unión entre sectores. Todos los detalles sobre la estructura de datos se describen en la documentación de productos de datos científicos de TESS (NASA, 2015; Twicken et al., 2020).

La Figura 4.1 muestra la curva de luz y el modelo de tránsito para TIC 220479565, confirmado como un planeta tipo “sub-Neptuno (entre super-Tierras y sub-Neptunos) transitando la estrella enana M2 TOI-269 (TIC 220479565, $V = 14,4$ mag, $J = 10,9$ mag, $R_{\star} = 0,40 R_{\odot}$, $M_{\star} = 0,39 M_{\odot}$, $d = 57$ pc). El candidato a exoplaneta ha sido identificado en múltiples sectores de TESS y validado mediante espectroscopía de alta precisión (HARPS) y seguimiento fotométrico terrestre (ExTrA y Observatorio Las Cumbres - Observatorio Interamericano Cerro Tololo, LCO-CTIO)” (Cointepas et al., 2021).

4.2 Sonificación multivariable secuencial

Hoy en día, la grabación multipista es una técnica estándar en la industria musical, incluso en los entornos de música clásica (Bartlett et al., 2013). Una vez finalizada la grabación, todas las pistas pasan por un mezclador para elaborar una mezcla final en la que se ajustan todos los aspectos necesarios en cada una de las fuentes que componen el programa sonoro (Bartlett y Bartlett, 2016). Este método permite controlar el equilibrio entre fuentes *a posteriori*, minimizando la cantidad de sonido de un instrumento que es captada por el micrófono de otro instrumento y realzando pequeños detalles en cada sonido grabado (Huber y Runstein, 2013).

Proponemos utilizar esta optimización consolidada del proceso de grabación y mezcla como

paradigma de diseño para la sonificación de representaciones sonoras multivariables, basadas en mapeados de frecuencia, y mostradas interactivamente en paisajes sonoros multicanal. De acuerdo con las sugerencias de Flowers, 2005, relacionadas con comparaciones secuenciales de datos sonificados, la adición de un motor de exploración secuencial automático a estos paisajes sonoros multipista de archivos DVT, convierte la propuesta en una herramienta de monitorización continua, que permite la identificación de tendencias y la percepción de correlaciones armónicas e inarmónicas entre variables y entre objetos en estudio.

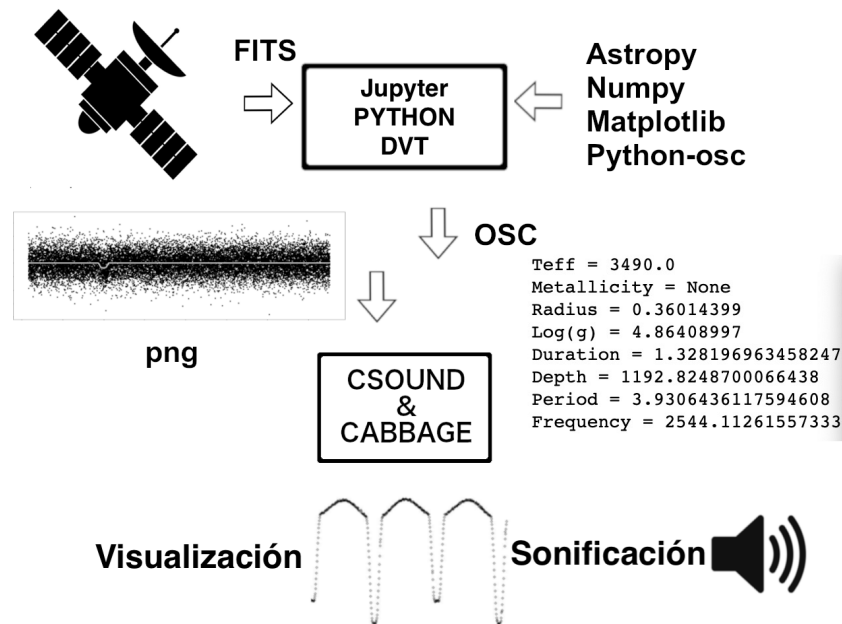


Figura 4.2: Diagrama de bloques del *Explorador DVT* basado en el módulo *FITS2OSC*.

4.3 Explorador DVT

A continuación presentamos dos implementaciones prácticas para probar el enfoque de paisaje sonoro multivariable introducido en la Sección 4.2. La primera, proporciona una exploración secuencial del catálogo TOI. Como puede verse en el primer vídeo de demostración, la duración y la profundidad de los tránsitos controlan, respectivamente, la duración y la amplitud de los sonidos, ofreciendo una representación sonora que puede recordar a la corriente musical espectralista (Harvey, 2000). El segundo diseño permite a los usuarios analizar objetos individuales con herramientas VO a través de consultas a *Aladin Lite* y *Vizier*. Las dos versiones del prototipo proporcionan control de sonido interactivo y utilizan el módulo *FITS2OSC* para conectar el motor de procesamiento de *Python* con *Cabbage-CSound*, que se encarga de sintetizar el sonido y proporcionar el entorno gráfico.

<https://vimeo.com/728364932>

<https://vimeo.com/723815913>

Como puede verse en la Figura 4.2, en la representación multimodal de cada objeto se utilizan las curvas de luz almacenadas en las extensiones FITS para cada TCE . En este sentido, la

Figura 4.1 es también un ejemplo de los archivos .png (Portable Network Graphics), generados con *matplotlib* e incorporados en la interfaz de usuario del prototipo.

En cuanto a la implementación de audio, los valores de cada variable de análisis se extraen de los archivos DVT utilizando las librerías *astropy* y *numpy*, y se convierten en mensajes OSC con *python-osc*, que son enviados a *CSound*, donde cada instrumento independiente genera la sonificación de una variable, conformando el paisaje sonoro final.

Tabla 4.1: Mapeado de variables estelares para opcodes de *CSound*.

Variable	Frec. factor	CSound opcode	Amp. factor
Teff	1/2	oscil	1/20
Metallicity	+/- 10000	dust2	1/2
Radius	100	fmb3	1
Logg	100	buzz	1/4
TPeriod	1/10000	fmvoice	1/2

La Tabla 4.1 enumera las variables objeto de análisis, los opcodes de *CSound* (Boulangier, 2000) y los factores de corrección utilizados en el proceso de mapeado de frecuencia y amplitud. La duración y la profundidad de cada tránsito detectado equivalen, respectivamente, a la duración y la amplitud de cada paisaje sonoro sintetizado, generando secuencias de sonificaciones con carácter musical. La estrategia de mapeado utilizada en este estudio persigue la conversión “más directa posible”, tratando de reducir el proceso únicamente a la introducción de factores de multiplicación que ajusten el dominio de variables astronómicas al dominio de variables sonoras. Por tanto, no utiliza modelos mentales (Walker y Kramer, 2005) (como los aplicados comúnmente en la sonificación de la cantidad de masa), con el fin de mantener el enfoque numérico directo y reducir las inferencias necesarias en la comprensión del paisaje sonoro completo.

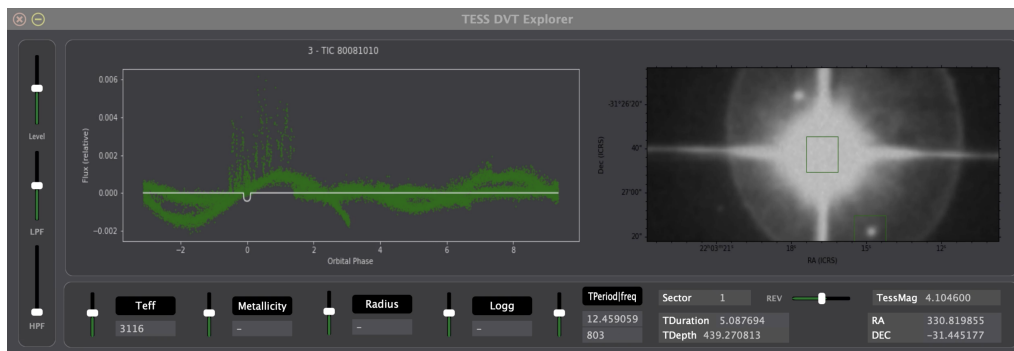


Figura 4.3: Captura de pantalla de la interfaz de usuario del explorador DVT durante la sonificación de la estrella variable de periodo largo TIC 80081010, con coordenadas RA: 22:03:16.76 DEC: -31:26:42.64.

Tal y como muestra la Figura 4.3, además de la representación de la curva de luz y una imagen de la estrella generada mediante el módulo de Python *APLpy* (T. Robitaille y Bressert, 2012; T. Robitaille, 2019), la interfaz de usuario proporciona controles de nivel para cada variable, filtros paso alto y paso bajo ajustables en frecuencia y un módulo de reverberación destinado

a generar sonificaciones de inspiración musical. El prototipo está disponible en código abierto en: <https://github.com/AuditoryVO/DVT-Explorer>

4.4 Estética interdisciplinar

Abordar los desafíos de un proyecto de sonificación, implica de forma natural la consideración de aspectos multidisciplinares que afectan fuertemente a la representación final y a su uso potencial. Como destacó Walker y Lane, 2001, muchas preguntas críticas deben resolverse para generar sonificaciones útiles. La traducción de datos a música puede llevar a la pérdida de información y la distracción del objeto de estudio, por lo que es necesario establecer algunos puntos de referencia aceptables que permitan a los receptores comprender la información contenida en un mensaje sonificado.

En este sentido, las convenciones y el conocimiento musical compartido por una audiencia han sido señalados como punto de partida potencial, que puede ayudar en la identificación de correlaciones entre los sonidos y los datos, aumentando el nivel de interés de la audiencia por el tema representado (Middleton et al., 2018). El conocimiento implícito de las estructuras musicales (Tillmann et al., 2000), y la existencia de redes neuronales específicas de la música en el cerebro humano (Peretz y Zatorre, 2005), son áreas de investigación actualmente activas. Perani et al., 2010 sugiere que “la arquitectura neuronal subyacente al procesamiento de la música en los recién nacidos es sensible a los cambios en la clave tonal, así como a las diferencias en consonancia y disonancia”, lo que puede llevar a demostrar que la mayoría de las expresiones musicales básicas ya son entendidas por el ser humano en los primeros días de vida.

Considerando todos estos factores, proponemos la búsqueda de un equilibrio entre las formas musicales abstractas y las clásicas, orientadas a la generación de sonificaciones con un cierto grado de abstracción, pero lo suficientemente accesibles como para facilitar su comprensión por la audiencia. Este enfoque podría reducir los errores en la interpretación (P. Vickers, 2006) y permitir conexiones entre los sonidos percibidos y los datos subyacentes. Moviéndonos alrededor del espacio de perspectiva estética descrito por Vickers (Vickers, 2013), que se presentó en la Figura 2.5, sugerimos la sonificación inspirada en la corriente musical espectral como solución estética con el potencial de preservar con precisión las relaciones de datos en el mapeado, produciendo sonificaciones con un nivel moderado de abstracción. Este enfoque invita al oyente a centrarse en el objeto de estudio, destaca el mensaje sobre la forma y reduce la distracción y la fatiga inducidas por exposiciones prolongadas.

4.5 Objetos de interés de TESS

El siguiente enlace proporciona un vídeo que muestra la representación sonora secuencial de 132 objetos de interés de TESS (*TOI*), que albergan al menos un planeta confirmado durante la exploración de los primeros 48 sectores de la misión. La Tabla 4.2 presenta las coordenadas y enlaces a las publicaciones de referencia de algunos de estos objetos.

<https://vimeo.com/702520208>

Habitualmente, la validación y confirmación de los objetos de interés se realiza mediante técnicas de observación combinadas en misiones de seguimiento. A modo de ejemplo, la Figura 4.4 muestra el mapa de cobertura del reciente estudio AstraLux-TESS (Lillo-Box et al., 2024), que proporciona imágenes de alta resolución espacial capturadas por el observatorio de Calar Alto (Almería) para la validación y confirmación fotométrica de señales de tránsito de planetas de pequeño tamaño ($R < 4 R_{\text{Tierra}}$) en TESS.

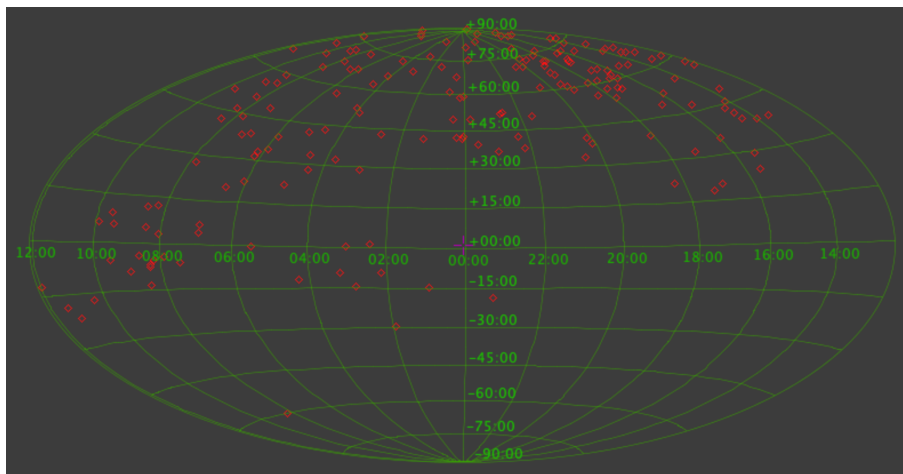


Figura 4.4: Mapa de cobertura del estudio AstraLux-TESS (Lillo-Box et al., 2024). Software: Aladin, proyección de Aitoff.

Tabla 4.2: Ejemplos de TOI sonificados. Coordenadas y referencia de publicación

Nombre	RA(2000)	DEC(2000)	Ref.
TIC 172900988	08:34:38.80737	+31:33:14.6804	Kostov et al., 2021
TIC 166527623	13:50:06.27970	-40:50:08.8811	Rizzuto et al., 2020
TIC 257060897	15:10:07.66974	+72:42:37.2466	Montalto et al., 2022
TIC 439867639	02:53:15.81819	+00:03:08.7844	Morello et al., 2023
TIC 142937186	03:29:09.83348	+31:21:46.9614	Giacalone et al., 2022
TIC 212957629	00:35:13.21911	-03:22:14.2880	Wells et al., 2021
TIC 230001847	19:22:28.79651	+60:51:13.9011	Grunblatt et al., 2022
TIC 257060897	12:58:57.68144	+77:39:41.6896	Schanche et al., 2022
TIC 441420236	20:45:09.53250	-31:20:27.2379	Wittrock et al., 2022
TIC 358107516	03:24:54.86176	-73:57:27.2632	Trifonov et al., 2021
TIC 176956893	06:43:19.94322	-66:56:51.6331	Saunders et al., 2022
TIC 298663873	18:31:46.49083	+56:39:03.0671	Dalba et al., 2022
TIC 237913194	01:29:46.95650	-60:44:23.8365	Schlecker et al., 2020
TIC 27491137	14:29:34.24261	+39:47:25.5433	Osborn et al., 2022
TIC 188589164	15:58:18.79764	+35:24:24.2798	Kemmer et al., 2022

TIC 172900988 es un sistema binario con un periodo cercano a 19,7 días y una excentricidad cercana a 0,45, que alberga un planeta en tránsito detectado desde el sector 21. El planeta transitó la estrella primaria y cinco días después transitó la estrella secundaria (Kostov et al., 2021).

HIP 67522 b es un planeta tipo Júpiter caliente en tránsito confirmado, que orbita TIC 166527623 ($T_{\text{eff}} = 5650$, $M = 1,2M_{\text{Tierra}}$) en la asociación Scorpius-Centaurus (Sco-Cen OB) de 10 a 20 millones de años de antigüedad (Rizzuto et al., 2020).

TIC 257060897 b es un sistema inflado que tiene una de las densidades más pequeñas conocidas hasta el momento (Montalto et al., 2022).

TOI-2445 b es un exoplaneta tipo súper Tierra que orbita una estrella de tipo M. Su masa es de 2,1 Tierras y tarda 0,4 días en completar una órbita alrededor de TIC 439867639 (Morello et al., 2023).

TOI-2427.01 es un candidato a planeta de $1,80 R_{\text{Tierra}}$ con un periodo orbital de 1,31 días que orbita una enana tipo K (TIC 142937186) que está a 28,5 pc de distancia y tiene una magnitud V (luz amarillo-verde, 545 nm con un ancho de banda de 84 nm) de 10,30 (Giacalone et al., 2022).

TOI-2406 b es un planeta tipo sub-Neptuno que orbita la estrella enana de tipo M TIC 212957629. La baja metalicidad de la estrella, el tamaño relativamente grande y el periodo corto del planeta, hacen de TOI-2406 b un resultado inusual de la formación de planetas y su caracterización proporciona una importante restricción observacional para los modelos de formación (Wells et al., 2021).

TOI-2337 b es un planeta masivo no inflado ($1,6 M_{\text{Júpiter}}$, $0,9 R_{\text{Júpiter}}$) con la órbita de periodo más corto ($P = 2,9943$ días) jamás observada alrededor de una estrella gigante roja ($3,2 R_{\odot}$, $1,4 M_{\odot}$) (Grunblatt et al., 2022).

TOI-2257 b (TIC 198485881) es un planeta tipo sub-Neptuno de periodo largo (35 días) que orbita una estrella tipo M3 a 57,8 pc. Su profundidad de tránsito es lo suficientemente grande (0,4%) como para ser detectada con telescopios terrestres de tamaño mediano (Schanche et al., 2022).

AU Mic (TIC 441420236) alberga un joven sistema exoplanetario cercano que sirve de laboratorio para sondear y caracterizar sistemas exoplanetarios jóvenes que podrían presentar planetas adicionales no detectables mediante el método de tránsito y que necesitan ser confirmados. El modelo parece ser consistente con una cadena multiplanetaria resonante compacta de periodo 4:6:9 (Wittrock et al., 2022).

TOI-2202 es una estrella de tipo K con una masa de $0,82 M_{\odot}$, un radio de $0,79 R_{\odot}$ y metalicidad similar a la solar. Este sistema es muy interesante porque presenta dos planetas cálidos de masa joviana que orbitan cerca de la resonancia de movimiento medio 2:1, que es una configuración poco habitual (Trifonov et al., 2021).

TIC 176956893 es una estrella sub-gigante evolucionada masiva ($M = 1,53 M_{\text{Júpiter}}$, $R = 2,90 R_{\text{Júpiter}}$) que alberga un planeta tipo Júpiter caliente que fue marcado inicialmente como falso positivo por el *TESS Quick-Look Pipeline*. Combinando análisis fotométricos de TESS junto con fotometría y mediciones de velocidad radial desde tierra, TOI2184 b finalmente ha sido reportado como un planeta en órbita (Saunders et al., 2022).

TOI-2180 b es un planeta gigante de $2,8 M_{\text{Júpiter}}$ que orbita la estrella G5 ligeramente evolucionada TIC 298663873 (Dalba et al., 2022).

TIC 237913194 b, con una masa de $M = 1,942 M_{\text{Júpiter}}$ y un radio de $R = 1,117 R_{\text{Júpiter}}$ (lo que implica una densidad bruta similar a la de Neptuno), orbita una estrella de tipo G ($M = 1,026 M_{\odot}$, $V = 12,1$ mag) con un periodo de 15,17 días en una de las órbitas más excéntricas de todos los gigantes cálidos conocidos ($e \approx 0,58$) (Schlecker et al., 2020).

TOI-2076 es un sistema de tres planetas en tránsito de tipo sub-Neptuno que orbitan la estrella joven (340 ± 80 Myr) de tipo K brillante ($G = 8,9$ mag) TIC 27491137. Aunque es un sistema planetario validado, las órbitas de los dos planetas exteriores no estaban restringidas, ya que sólo se vieron dos tránsitos no consecutivos en la fotometría de TESS. Esto dejó 11 y 7 posibles periodos alias para cada planeta. Para revelar las verdaderas órbitas de estos dos planetas de periodo largo, se requirió un seguimiento fotométrico basado en observaciones desde el espacio y terrestres de TOI-2076 c y d con los telescopios CHEOPS, SAINT-EX y LCO (Osborn et al., 2022).

GJ 3929 b, es un planeta caliente del tamaño de la Tierra que orbita la estrella enana cercana M 3.5 V, GJ 3929 (TIC 188589164, TOI-2013). El modelado conjunto de observaciones fotométricas de los sectores 24 y 25 de TESS, junto con 73 observaciones espectroscópicas de CARMENES y observaciones de tránsito de seguimiento de SAINT-EX, LCOGT y OSN, proporcionaron un radio planetario de $R_b = 1,150 R_{\text{Tierra}}$, una masa de $M_b = 1,21 M_{\text{Tierra}}$ y un periodo orbital de $P_b = 2,6162745$ días. La densidad resultante es compatible con la densidad media de la Tierra, aproximadamente $5,5 \text{ g/cm}^3$ (Kemmer et al., 2022).

Capítulo 5

Musificación del análisis de periodograma en curvas de luz

El análisis de periodograma *Lomb-Scargle* es uno de los métodos estadísticos estándar para buscar periodicidad en datos muestreados irregularmente cuando no se puede aplicar el análisis de Fourier (Ivezić et al., 2019). Se utiliza principalmente en la búsqueda de estrellas variables y en el análisis de datos de velocidad radial. El periodograma muestra una estimación de la potencia de la transformada de Fourier en función del periodo, a partir de la cual podemos leer los periodos detectados en la señal. Básicamente, consiste en ajustar un modelo sinusoidal a los datos para cada frecuencia, obteniendo mayor potencia cuanto más se ajusta el modelo a la señal de análisis.

Todos los detalles matemáticos e históricos relacionados con este método pueden encontrarse en VanderPlas, 2018, también una buena referencia de técnicas complementarias para el análisis periódico de observaciones puntuales tales como los enfoques bayesianos o los métodos de *Fourier*, *Phase-folding* y *Box-fitting Least-Squares* (BLS). Este último tipo de análisis BLS, ajusta las series temporales de entrada a funciones periódicas en forma de caja, las cuales representan mejor que los senos y cosenos el comportamiento de una curva de luz durante un tránsito, resultando de utilidad en la búsqueda de exoplanetas y estrellas binarias (Kovács et al., 2002).

Desde la perspectiva práctica del análisis del periodograma de una curva de luz, la librería *Astropy*, permite la implementación de los periodogramas estándar y generalizados de *Lomb-Scargle* y *Box-fitting Least-Squares*, manejando adecuadamente las magnitudes con sus unidades adjuntas. El módulo de *herramientas de astroestadística de Astropy* (Astropy, 2011-2022b) proporciona numerosas funciones, algoritmos y herramientas estadísticas que permiten incluso la normalización del flujo (sigma-clipping). Toda la documentación, funcionalidad, herramientas comunes y procesos estadísticos necesarios para realizar Astronomía y Astrofísica en *Python* están disponibles en *astropy:docs* (Astropy, 2011-2022a).

Este capítulo tiene como objetivo complementar mediante musificaciones las representaciones gráficas de los análisis de periodograma en curvas de luz, utilizando casos de estudio extraídos del catálogo de entrada Kepler (KIC).

5.1 Objetos de interés del catálogo Kepler

Un objeto de interés del catálogo Kepler (KOI), es una estrella observada por el telescopio espacial Kepler que potencialmente alberga uno o más planetas en tránsito. Se trata de un catálogo derivado del catálogo de entrada Kepler (KIC) que contiene 150.000 estrellas observadas desde 2009 hasta 2013 (misión Kepler) y desde 2013 hasta 2018 (misión K2). El proyecto Kepler identifica objetos periódicos similares a tránsitos examinados en sus curvas de luz para identificar candidatos KOI. Contiene alrededor de 10.000 objetos astronómicos hasta la fecha. Como se muestra en la Figura 5.1, el fotómetro de Kepler consta de 21 módulos, cada uno con dos CCD de 2200x1024 píxeles. Cada módulo cubre 5 grados cuadrados de cielo (NASA, 2020a).

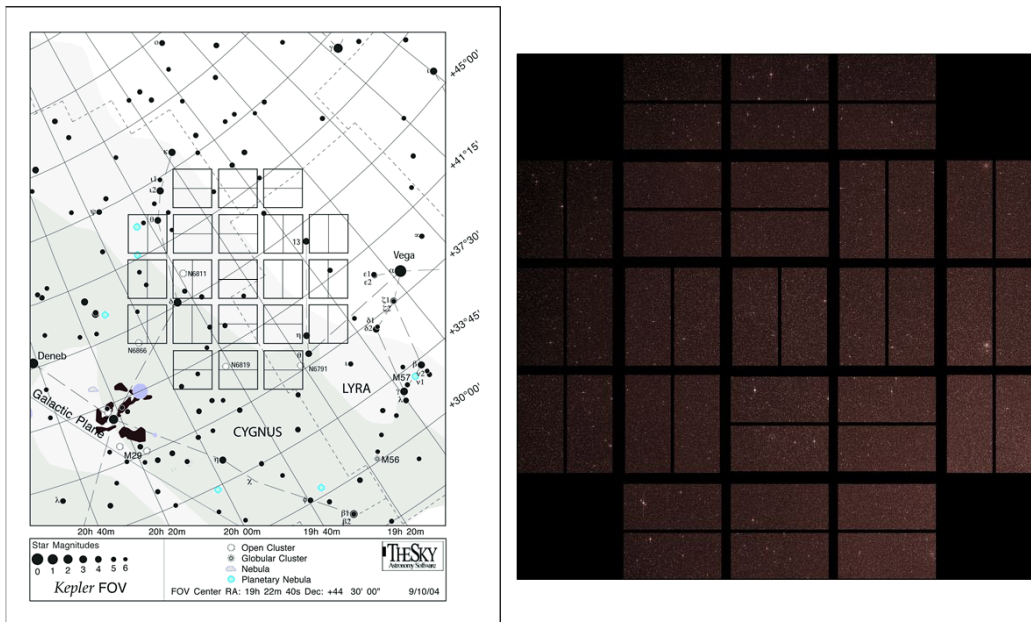


Figura 5.1: Campo de visión completo de 116 grados cuadrados (FOV) del único campo observado durante la misión inaugural de Kepler (izquierda) y captura de imagen durante la misión Kepler (derecha) (NASA, 2020a).

5.2 Generación de partituras a partir del análisis de periodograma

Esta sección presenta *LC2Score*, el prototipo exploratorio que hemos implementado para la generación automática de partituras a partir del análisis de periodogramas en curvas de luz. La aplicación, desarrollada como prueba de concepto en *Python*, utiliza las librerías *astropy*, *numpy* y *matplotlib* para proporcionar una representación de cada curva de luz, su análisis de periodograma y la curva plegada alrededor de su periodo principal, que es traducido a una nota musical. El conjunto de herramientas para el análisis de musicología asistido por computadora *Music 21* (Cuthbert y Ariza, 2010) se utiliza para generar la partitura, finalmente renderizada en *MuseScore3* (Schweer, 2002).

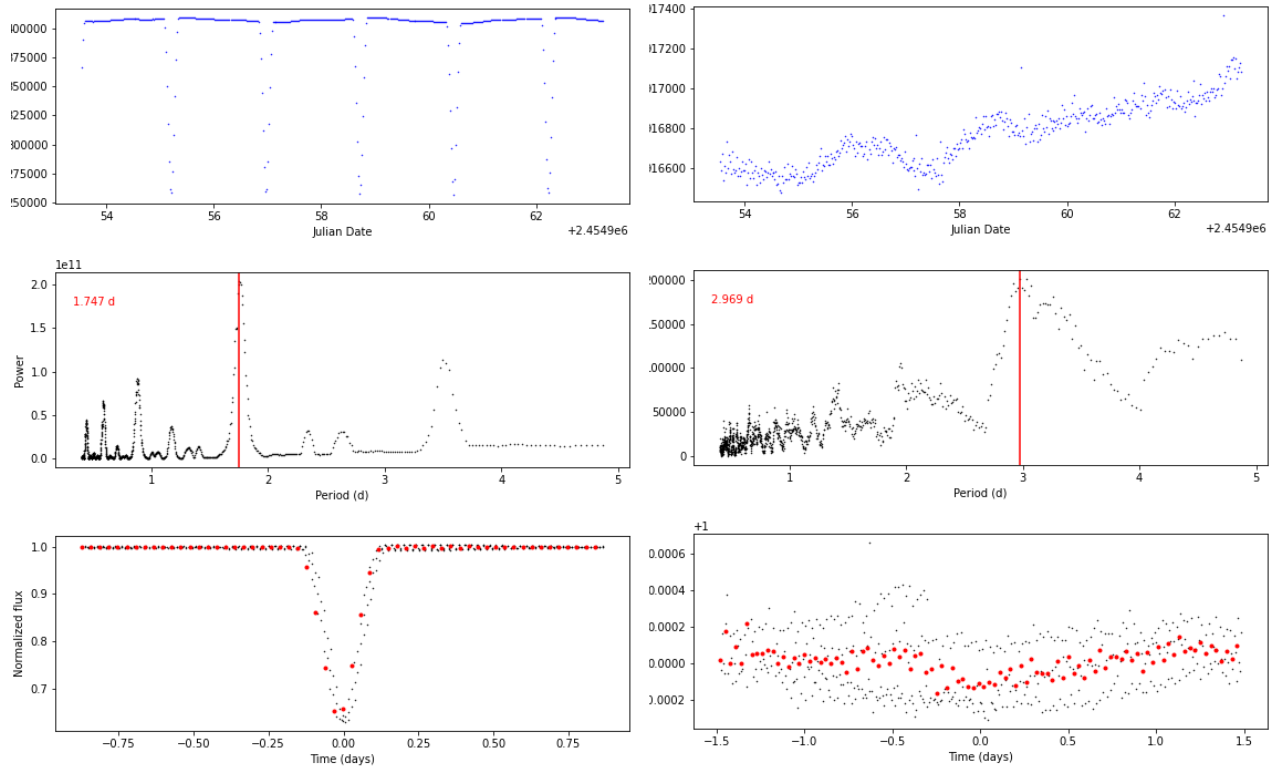


Figura 5.2: Dos ejemplos del análisis de periodograma Box Least Squares de las curvas de luz del catálogo KOI (Thompson et al., 2018)(STScI). Los gráficos muestran cada curva de luz (flujo SAP en función del tiempo en JD), su análisis de periodograma (Energía en función del periodo en días) y las curvas plegadas alrededor del periodo de mejor ajuste (flujo normalizado en función del tiempo en días), para los objetos Kepler 006891512, que presenta un periodo principal de 1,747 días (izquierda), y Kepler 008292840, con el mejor ajuste en 2,969 días (derecha).

La Figura 5.2 proporciona un ejemplo del análisis de periodograma utilizando el caso de estudio de las curvas de luz del catálogo KOI (M. STScI, 2017).

La asignación de notas musicales se calcula directamente a partir del periodo de mejor ajuste obtenido con el análisis del periodograma de cada curva, siguiendo una filosofía de conversión directa “periodo orbital-periodo de nota musical”, multiplicado por un factor de 100 para ajustarlo a los rangos audibles. Los resultados se aproximan a la frecuencia fundamental de la nota más cercana sobre una base de escala cromática de 12 notas. La Figura 5.3 muestra los 16 compases iniciales de una pieza generada a partir de 400 curvas de luz KOI.

La intensidad sonora de cada nota se obtiene de la energía del periodo de mejor ajuste y se escala logarítmicamente para que coincida con las 128 etapas de velocidad del estándar MIDI (Anderton, 1987). La duración de cada nota depende de la media de energía obtenida en los resultados del periodograma para cada curva de luz. Las duraciones se agrupan en cuatro categorías que van desde figuras musicales de *corchea* hasta *redonda* para generar la partitura final. La aplicación de software abierto está disponible en el siguiente enlace:



Figura 5.3: Primeros 16 compases de la pieza musical *KOI 400-800*, generada a partir del análisis del periodograma BLS de 400 curvas de luz KOI.

<https://github.com/AuditoryVO/LC2Score>

El siguiente vídeo muestra una versión de arpa sintetizada de la pieza *KOI 100*, generada a partir de las 100 primeras curvas de luz del catálogo KOI. Resulta interesante destacar la abundancia de periodos principales alrededor de 4,8 días (nota Sol sostenido), así como las reminiscencias orientales en el carácter de la partitura automática generada con esta propuesta.

<https://vimeo.com/678882792>

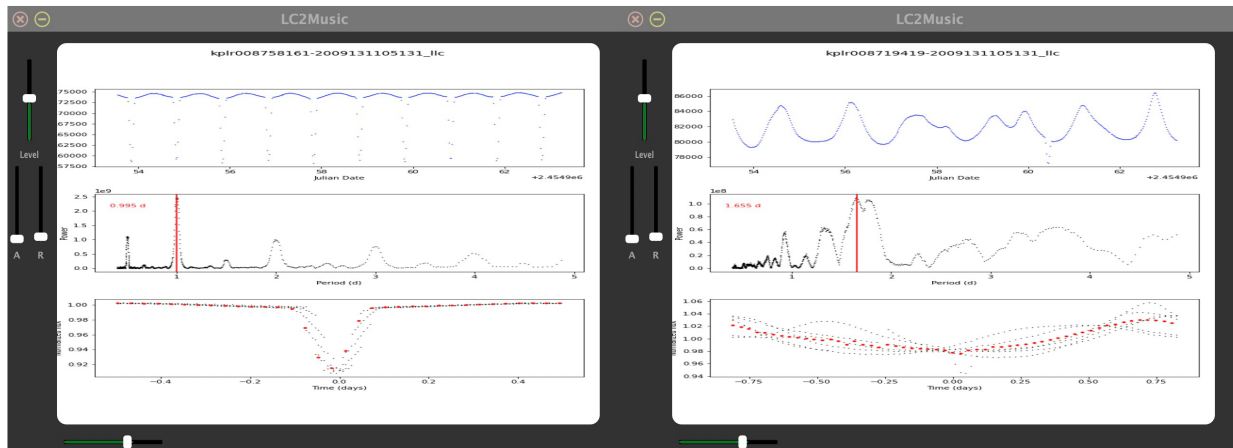


Figura 5.4: Dos capturas de pantalla de la interfaz *LC2Music* durante la sonificación del catálogo KOI. Izquierda: curva de luz de kplr008758161, falso positivo de las disposiciones de Kepler y NExScI (STScI, 2013) con un periodo principal claro cercano a 1 día. Derecha: curva de luz de kplr008719419 (KOI-3506), candidato a falso positivo del NASA Exoplanet Archive (NASA, 2011a).

5.3 Síntesis de audio a partir de curvas de luz

La resolución en frecuencia del enfoque descrito en la sección anterior (basado en MIDI) presenta la limitación asociada a la cuantización de notas. Esta cuantización puede enmascarar los resultados del análisis del periodograma con una diferencia en frecuencia inferior a $2^{(1/12)}$ (equivalente a la separación en frecuencia entre notas). Una posible solución es obtener la

frecuencia fundamental del sonido sintetizado directamente del periodo principal del análisis del periodograma de cada curva ($f=1/T$), y escalarlo al rango audible. En comparación con el enfoque de generación de partituras, esta alternativa proporciona resultados microtonales, lo que aumenta la precisión de las representaciones para cambios suaves de periodo o frecuencia.

La Figura 5.4 muestra la interfaz de usuario del sintetizador *LC2Music*, que utiliza este último enfoque ($f=1/T$), durante la sonificación de kplr 008758161 y kplr 008719419. El prototipo genera notas de flauta sintetizadas “a tiempo real” a partir del análisis de periodogramas BLS, e incluye controles de nivel, tiempo de ataque y relajación. Los periodos de mejor ajuste encontrados en las curvas de luz son multiplicados por un factor de 100 para llevar las frecuencias fundamentales al rango de frecuencias audibles.

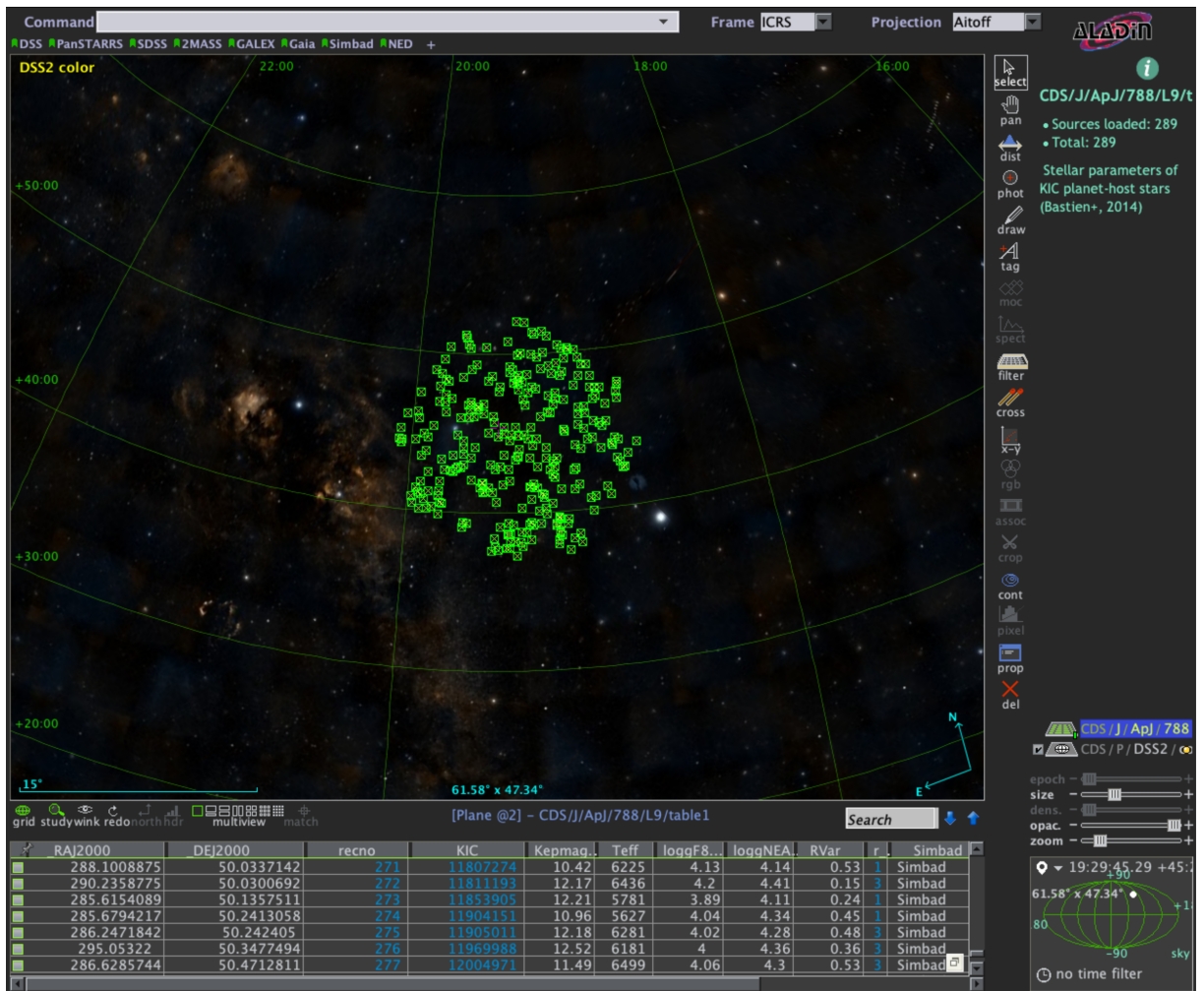


Figura 5.5: Mapa de cobertura de 289 estrellas del KIC albergando al menos un planeta (Bastien, 2014). Software: Aladin, proyección de Aitoff.

Sincronizados con la generación de sonido, la aplicación proporciona tres gráficos que muestran cada curva de luz, su distribución de periodograma con el “mejor periodo” resaltado y las curvas plegadas a su alrededor. La duración de cada representación es constante, proporcionando un escaneo secuencial de los objetos que permite a los usuarios centrarse en los resultados

sonoros del periodograma.

Con el objetivo de permitir un análisis comparativo de las propuestas, utilizamos en la implementación de *LC2Music* las curvas de luz del catálogo KOI, utilizadas para la generación de partituras en la Sección 5.2. El siguiente vídeo muestra las posibilidades de la propuesta durante la exploración de la región del cielo representada en la Figura 5.5.

<https://vimeo.com/641213364>

El prototipo de software abierto está disponible en el siguiente enlace:

<https://github.com/AuditoryVO/LC2Music>

Capítulo 6

Sonificación simbólica espacializada

Centrado en el análisis del potencial de las sonificaciones espacializadas, este capítulo presenta la exploración de la información etiquetada y de fácil acceso disponible en los encabezados de las curvas de luz N2 del archivo Convection, Rotation and planetary Transits (CoRoT) del Institut d’Astrophysique Spatiale (IAS) (Baudin et al., 2006). Para ello, proponemos el uso de la estrategia de sonificación de mapeado simbólico (García Riber y Serradilla, 2024b), que convierte los tipos espectrales en notas musicales. Basada en las pautas para mejorar la combinación de visualización y sonificación de información para la comprensión de programas establecidas por Hussein et al., 2009, la propuesta podría implementarse en archivos masivos para proporcionar información sonora del tipo de estrella y de la periodicidad detectada en sus curvas de luz. El objetivo final es ofrecer una metodología de representación inclusiva complementaria a la visualización, que pueda hacer que los datos procesados para el análisis científico del archivo sean accesibles para cualquier persona.

La prueba de concepto que presentamos en las siguientes secciones incluye espacialización ambisónica (Malham y Myatt, 1995) de primer orden y binaural, conmutables por el usuario. Las representaciones están vinculadas a las coordenadas reales de los objetos representados, mejorando la conexión con la audiencia a través de audio inmersivo y proporcionando sonificaciones espacializadas binaurales orientadas a la escucha con auriculares y representaciones sonoras ambisónicas que pueden ser útiles en comunicaciones públicas en vivo. El prototipo que hemos implementado genera sonificaciones preliminares multivariable “a tiempo real” para ayudar en la búsqueda de casos de estudio, que se pueden analizar en profundidad *a posteriori* utilizando metodologías más precisas como las descritas en Tutchton et al., 2012 y Tucker Brown et al., 2022, o con herramientas como Sonouno (De La Vega et al., 2022), Highcharts Sonification Studio (Cantrell et al., 2021), STRAUSS (Trayford y Harrison, 2023), Astronify (Brasseur et al., 2020) o Sonifigrapher (García Riber, 2019). Los resultados de su evaluación pueden encontrarse en la Sección 11.2.2.

6.1 El archivo CoRoT

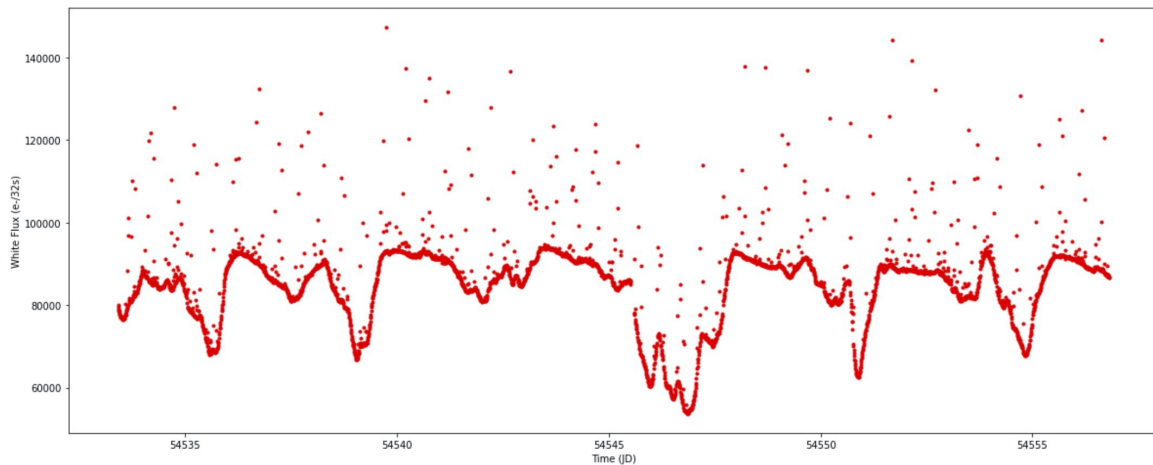
La misión espacial CoRoT fue dirigida por el Centre national d’études spatiales (CNES) con la participación de la Agencia Espacial Europea (ESA), entre diciembre de 2006 y junio de 2014.

Orientada al análisis de la física estelar y a la búsqueda de exoplanetas, CoRoT fue pionera en la detección de planetas rocosos fuera del Sistema Solar. Este trabajo utiliza 474 productos EN2_STAR_MON (anteriormente llamados canal de exoplanetas) (Auvergne et al., 2009), que contienen la información monocromática de todos los píxeles de cada objetivo sumados para proporcionar curvas de luz individuales como la que se muestra en la Figura 6.1.

Los archivos están disponibles públicamente en el archivo CoRoT (IAS, 2016) y en el Centro de Datos Astronómicos de Estrasburgo (CDS, 1983) a través de Vizier (CDS, 2000c). Las curvas de luz se archivan utilizando la nomenclatura mostrada en el siguiente ejemplo para la estrella CHR 0223959618.

<COROTID>_<START_DATE>_<END_DATE>.fits.

EM2_STAR_CHR_0223959618_20080307T215033_20080331T073549.fits



```
MIDI NOTE: 77
-----Lightcurve 18 -----
Spectral type: F5
Luminosity class: III
LC RMS: 85976.2421875
LC Mean: 85357.375
RA: 99.922804
DEC: 9.772137
Bes fit period: 3.957 days
```

Figura 6.1: Curva de luz CoRoT (flujo blanco en electrones/32s en función del tiempo en JD) y metadatos utilizados en la sonificación de CHR 0223959618. El prototipo extrae las coordenadas, tipo de espectro, luminosidad, clase y valores característicos de cada curva de luz. Además, calcula el periodo principal encontrado en la curva, en este caso de 3,96 días.

6.2 Mapeado simbólico

Inspirada en la analogía de color musical de Newton, ilustrada en la Figura 2.2, esta sección propone el uso de la coincidencia simbólica interdisciplinar entre las letras utilizadas en la clasificación espectral estelar de Harvard y las utilizadas en las clases de altura de nota en notación musical, para generar representaciones sonoras rápidas de conjuntos de datos pre-procesados en curvas de luz.

La Tabla 6.1 muestra la estrategia de mapeado simbólico sugerida para convertir intuitivamente en sonido las clases espectrales de los objetos del archivo, etiquetados utilizando la clasificación de Harvard y Morgan-Keenan (ver Anexo.4). La conversión consiste en una correspondencia casi directa entre las siete clases espectrales y las siete notas musicales de la escala de Do mayor.

Tabla 6.1: Estrategia simbólica de mapeado que asigna cada tipo espectral a una nota musical, propuesta para la sonificación de los siete tipos de espectros estelares de la clasificación de Harvard con valores de frecuencia para la octava 4 (conversiones de pitch en *CSound*: <https://csound.com/manual/misc/pitch/>).

Teff	>25.10 ³ K	10.10 ³ -25.10 ³ K	7,5.10 ³ -10.10 ³ K	6.10 ³ -7,5.10 ³ K	5.10 ³ -6.10 ³ K	3,5.10 ³ -510 ³ K	<3,5.10 ³ K
Tipo espectral	O	B	A	F	G	K	M
Nota musical	C	B	A	F	G	D	E
Notación latina	Do	Si	La	Fa	Sol	Re	Mi
Octava 4 (Hz)	261,63	493,88	440	349,23	391,99	293,66	329,63

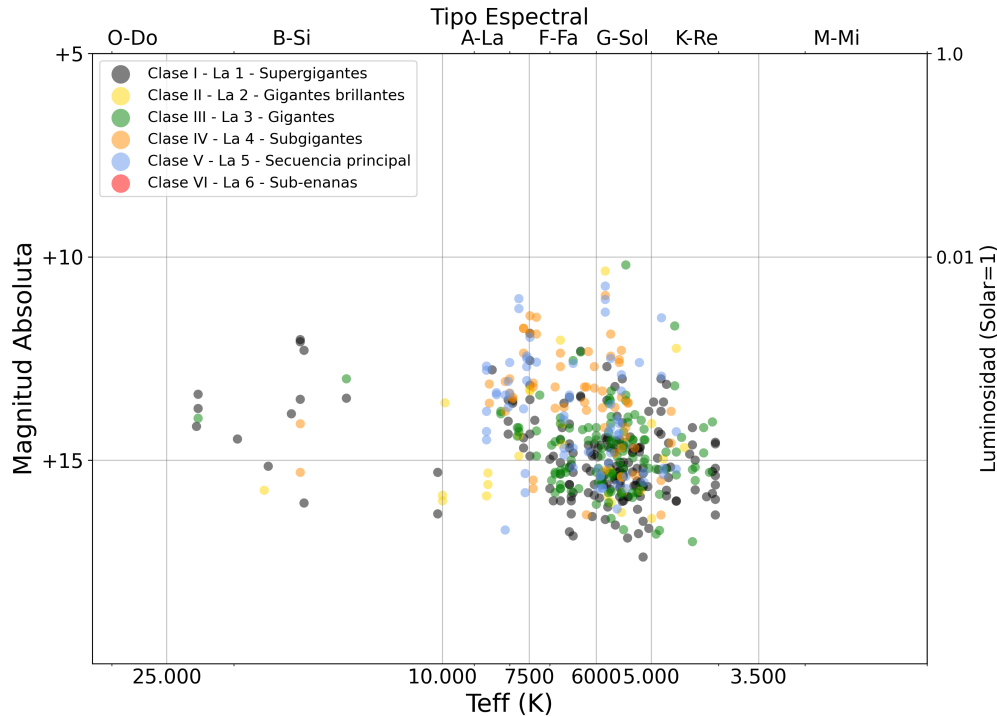


Figura 6.2: Diagrama HR de 452 objetos de la librería CoRoT, ilustrando la conversión simbólica de tipos y clases espectrales propuesta. Temperatura efectiva (K) y rangos de los tipos espectrales frente a Magnitud absoluta y Luminosidad con referencia solar.

La clasificación de Harvard utiliza números de subdivisión para representar la profundidad de las líneas de absorción. Esta numeración se utilizó para controlar la octava de cada nota, asignando los tonos más bajos a las estrellas más calientes (número 0) y los tonos más altos a las más frías (número 9). Un número romano adicional expresa la clase de luminosidad de Morgan-Keenan (desde *I* correspondiendo a estrellas supergigantes, hasta *VII* correspondiendo

a estrellas enanas). Estas clases se representaron con una nota constante (A), cuya octava se hace corresponder con la clase M-K del objeto.

De este modo, una estrella de tipo espectral G8-V se representa como un Sol-8 (tipo G, fría) con La-5 (secuencia principal) y una estrella F0-I, como un Fa-0 (tipo F, caliente) con La-1 (supergigante). Como resultado, la exploración secuencial del catálogo se asemeja a una composición musical en Do mayor cuyas notas orbitan alrededor del sexto grado (la nota A que representa la clase de luminosidad). La Figura 6.2 ilustra esta conversión sobre el diagrama HR (ver Anexo.5) de la librería CoRoT.

Tabla 6.2: Ejemplo de mapeado para el objeto con CoRoT ID 0223959618.

	Tipo espectral	Clase	Flujo medio	Análisis Periodo	RA	DEC
Curva de luz	[SPECTYPE]	[LUMCLASS]	[LC MEAN]	[WHITE FLUX]	[ALPHA]	[DELTA]
Valor	A5	IV	1217305 (e ⁻)	3,96 (días)	100,22	9,51
Sonificación	Nota	Octava equivalente (La)	Amplitud (Log)	Frecuencia LFO corregida	Azimuth	Elevación
Valor	La 5 (880Hz)	La 4 (440Hz)	6,08	3,47 (Hz)	100,22	0 (No 3D)

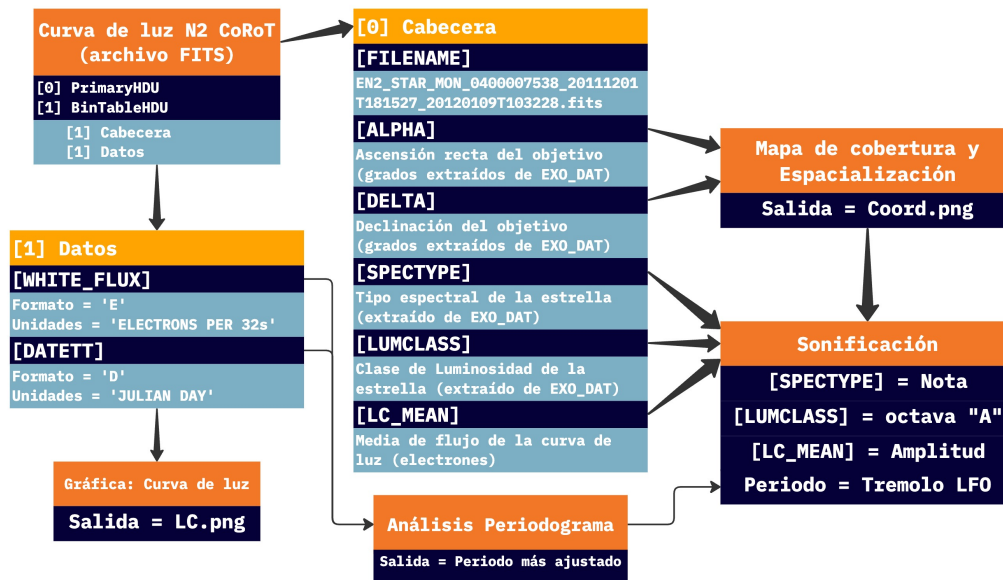


Figura 6.3: Variables extraídas del archivo FITS de cada objeto. El esquema incluye la descripción de cada función (naranja), metadatos y mapeado (azul) y los archivos de salida generados a partir de cada curva de luz.

Por otra parte, analizando el periodograma BLS, se calcula el periodo de mejor ajuste detectado en cada curva de luz, que se hace corresponder con la frecuencia de oscilación de un efecto de trémolo. Adicionalmente, se introdujeron controles interactivos para las sonificaciones de la clase de luminosidad, del periodo detectado y de las amplitudes, a través de una envolvente ADSR (ataque, caída, sostenimiento y relajación). Finalmente, la intensidad del sonido generado se modula en función del flujo medio de cada curva de luz para permitir el análisis de las intensidades de flujo relativas entre objetos. La Figura 6.3 muestra las variables y los archivos de salida utilizados para la sonificación de cada curva de luz. En la Tabla 6.2 se puede ver un ejemplo ilustrativo del mapeado utilizado.

A continuación, se proporciona un vídeo demostrativo de la exploración binaural del archivo CoRoT utilizando el mapeado simbólico propuesto:

<https://vimeo.com/921824459>

Este enfoque fue diseñado con el objetivo de proporcionar la representación sonora más sencilla posible del archivo CoRoT, para comprobar si el uso de conceptos musicales elementales (reconocimiento de las notas de la escala mayor de Do) puede permitir la clasificación sonora de tipos espectrales sin necesidad de disponer de conocimientos previos en Astronomía. Tal y como muestra la sección 11.2.4, esta metodología se usó para establecer un marco comparativo con la metodología de sonificación basada en deep learning descrita en el Capítulo 7.

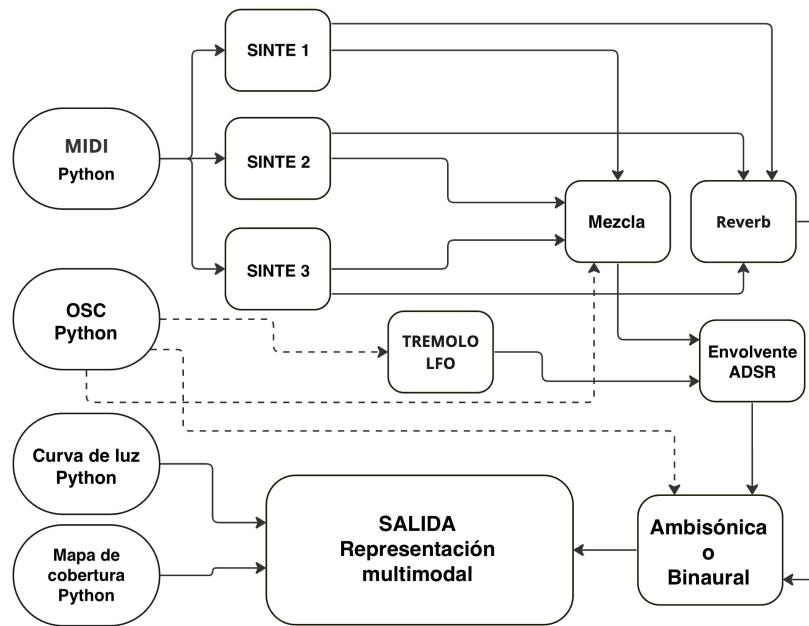


Figura 6.4: Diagrama de bloques del sintetizador de sonificación implementado en *CSound* y *Cabbage*. Las líneas continuas representan conexiones MIDI, de audio y de imagen. Las líneas discontinuas representan conexiones OSC.

Aunque no se ha implementado en este trabajo, una sencilla ampliación del prototipo podría permitir diferenciar las sonificaciones de objetos con idéntico tipo espectral y distinta temperatura. Para esta potencial expansión o para implementaciones futuras de precisión basadas en MIDI, proponemos la sustitución del mapeado simbólico por un mapeado equivalente a través de la temperatura efectiva de cada objeto. De esta forma, pueden utilizarse las centésimas de tono (MIDI cents) para establecer variaciones alrededor de la frecuencia fundamental de la nota, mapeada a la temperatura efectiva central del rango de temperaturas de cada tipo espectral. A modo de ejemplo, para dos curvas de luz con idéntico tipo espectral G4 y con temperaturas efectivas 5600 K y 5100 K, se realizaría un mapeado a Sol, octava 4 en ambos casos, del mismo modo que en el modelo simbólico. Sin embargo, se podrían generar las respectivas desviaciones de +10 cents y -40 cents (respecto a la nota MIDI Sol mapeada a la temperatura central del rango de tipo espectral, 5500 K), que se obtienen como resultado de mapear los límites de temperatura del tipo espectral G (6000 y 5000 K) a +/- 50 cents.

6.3 Representación multimodal a tiempo real con pre-procesado

El prototipo de sonificación que proponemos en este capítulo consta de dos módulos implementados en *Jupyter notebook* (IPython, 2014), que pre-procesan y lanzan la exploración del archivo, y una interfaz de usuario multimodal implementada en *CSound/Cabbage* que permite el control de las sonificaciones “a tiempo real”. El primer módulo de pre-procesado lee los datos de cada archivo FITS de la librería CoRoT, calcula el periodo de mejor ajuste encontrado en cada curva de luz y guarda los datos en matrices usando *numpy*. El módulo de exploración lee las matrices pre-procesadas, genera las representaciones gráficas de la curva de luz y envía notas MIDI y mensajes OSC “a tiempo real” al sintetizador de sonido.

Tal y como muestra la Figura 6.4, el módulo de generación de sonido consta de tres sintetizadores independientes que permiten cubrir el rango espectral equivalente a 9 octavas, desde la nota D0 (que representa el tipo espectral K0), hasta la nota G8 (que representa el tipo espectral G8). Los opcodes de *CSound fmb3* y *vco2* (Lazzarini et al., 2016) cubren los rangos de frecuencia baja y media. Para las notas más altas, se agregó el opcode *oscil3* (CSound, 2003). Todos ellos reciben las notas MIDI de activación enviadas por el módulo de sonificación de *Python*, que también envía a través de OSC las variables que controlan la frecuencia del trémolo asociado al análisis del periodograma, la proporción y las amplitudes de los generadores de sonido asociados al flujo medio de la curva de luz, las coordenadas de cada objeto para la espacialización y la frecuencia de actualización que controla la representación secuencial.

El sistema completo está disponible en código abierto en:

https://github.com/AuditoryVO/CoRoT_Explorer

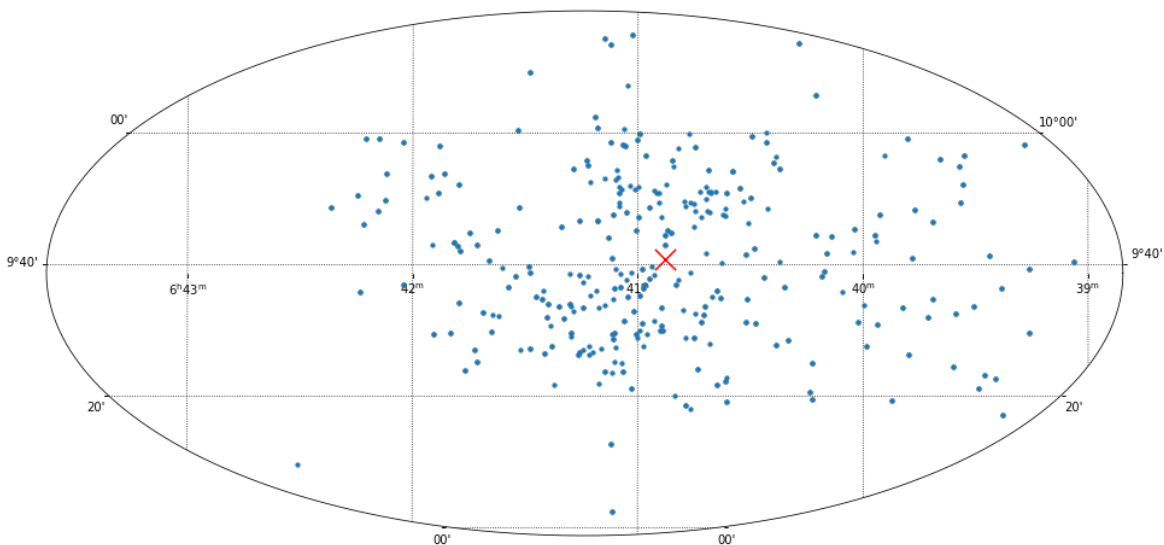


Figura 6.5: Mapa de cobertura del archivo IAS CoRoT que muestra los objetos (puntos azules) y la posición de referencia (centroide) de las coordenadas de espacialización (cruz roja), ubicada en RA 100.2196 DEC 9.6773 (grados).

6.4 Espacialización binaural y ambisónica

Con el objetivo de proporcionar una experiencia de sonificación inmersiva para la exploración del archivo CoRoT, el prototipo proporciona sonificaciones ambisónicas o binaurales de manera conmutable por el usuario.

La configuración ambisónica es adecuada para actividades educativas o de divulgación pública, pero no admite la reproducción con auriculares. El enfoque binaural complementario proporciona espacialización para la monitorización mediante auriculares. La Figura 6.5 muestra el mapa del archivo CoRoT completo y la posición de referencia de las coordenadas de espacialización (centroide) establecida en RA 100.2196 DEC 9.6773 grados, que proporciona una perspectiva inmersiva centrada en las dos configuraciones (ambisónica y binaural). Las representaciones están correlacionadas con una proyección de los objetos del archivo en pantalla plana, manteniendo la compatibilidad entre los dos formatos de audio y la coherencia entre el flujo de datos gráficos y sonoros.

Las ecuaciones que rigen los procesos de codificación y decodificación binaural se pueden encontrar en el Anexo.12. El proceso de codificación y decodificación ambisónica también se describe en el Anexo.13. En este enfoque, utilizamos representaciones ambisónicas de primer orden (FOA), que requieren la codificación en formato-B con tres dimensiones básicas y su posterior decodificación a ocho señales de audio (que alimentan ocho altavoces dispuestos en configuración circular). El modelo fue implementado en *CSound* mediante el desarrollo de un opcode definido por el usuario (UDO) (Neukom, 2014).

Capítulo 7

Sonificación basada en deep learning

El uso de técnicas de deep learning para la extracción automática de características específicas en datos astronómicos ha generado resultados interesantes en la estimación de parámetros atmosféricos a partir de espectros estelares (Yang y Li, 2015). La aplicación de técnicas de reducción basadas en autoencoders para grandes conjuntos de datos astronómicos, tiene además el potencial de mejorar la clasificación y exploración de la complejidad inherente de las galaxias (Portillo et al., 2020), permitiendo la captura de características espectrales intrínsecas, independientemente de los corrimientos al rojo (estiramiento del espectro debido al aumento de la longitud de onda que se genera por el aumento de la distancia al objeto observado), el ruido y las señales parásitas o artefactos presentes en los datos (Melchior et al., 2022).

Los autoencoders (Bourlard y Kamp, 1988; Le Cun, 1986) están formados por un codificador y un decodificador. Ambos son redes neuronales con un número variable de capas ocultas entrenadas para generar una copia imperfecta de la entrada, reduciendo su dimensión a una representación comprimida llamada espacio latente (Goodfellow et al., 2016). Los autoencoders variacionales (Kingma, 2013; Kingma, Welling et al., 2019; Rezende et al., 2014) son una subcategoría de autoencoders en los que los vectores latentes, obtenidos como resultado de la codificación, están restringidos a una distribución de probabilidad continua. Esto permite generar salidas sintéticas realistas a través de la exploración del espacio latente formado por dichos vectores, algo que es de utilidad en la generación de curvas de luz sintéticas, utilizadas en la caracterización de exoplanetas y estrellas distantes (Woodward et al., 2019), y en el aumento de conjuntos de entrenamiento de redes neuronales.

Este capítulo explora el potencial de ambas arquitecturas para la sonificación de espectros estelares, con el objetivo principal de minimizar la subjetividad y la arbitrariedad en el diseño de sonificaciones mediante procesos no supervisados. Se incluyen diversas pruebas de concepto basadas en el caso de estudio de la librería espectral MILES (García Riber y Serradilla, 2024c), cuyo análisis se encuentra ampliado en el Anexo.15.

7.1 La librería MILES

La librería de espectros estelares MILES consta de 985 estrellas que cubren un amplio rango de parámetros atmosféricos. Desarrollada en el Instituto de Astrofísica de Canarias (IAC), los espectros se obtuvieron con el telescopio Isaac Newton de 2,5 m en el Observatorio del Roque de los Muchachos, cubriendo de 3525 a 7500 Å, con una resolución espectral de 2,3 Å (FWHM) (Falcón-Barroso et al., 2011; Sánchez-Blázquez et al., 2006). La Figura 7.1 muestra el mapa de cobertura de la librería completa obtenido con *Aladin*, utilizando la proyección de Aitoff.

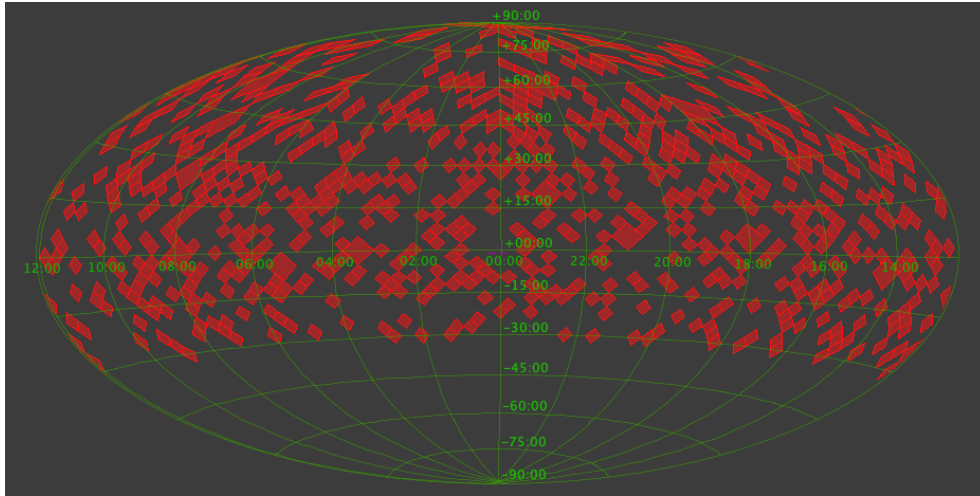


Figura 7.1: Mapa de cobertura de la librería de espectros estelares MILES. Software: Aladin, proyección de Aitoff.

7.2 Autocodificación de espectros estelares

El deep learning es un subconjunto del aprendizaje automático o machine learning, en el que las redes neuronales multicapa aprenden de un conjunto representativo de datos de población. Dentro de esta categoría, los autoencoders representan uno de los algoritmos de aprendizaje no supervisado utilizados para identificar relaciones dentro de los datos.

Tal y como lo define Goodfellow et al., 2016, un autoencoder es una red neuronal que está entrenada para proporcionar a su salida una copia de su entrada. Se compone de dos módulos, un codificador y un decodificador. Ambos son redes neuronales de alimentación directa con un número variable de capas ocultas. La Figura 7.2 presenta un ejemplo de esta arquitectura en la que el codificador toma la entrada y la reduce a una representación de menor dimensión llamada espacio latente. A partir de esta representación comprimida, el decodificador intenta reconstruir la entrada original. El modelo está diseñado para que no pueda aprender a copiar perfectamente, viéndose obligado a priorizar qué aspectos de la entrada deben copiarse. Esta restricción hace que la arquitectura aprenda propiedades útiles de los datos de entrada. La Figura 7.3 muestra la desviación entre los espectros originales y los decodificados medida a través del coeficiente de determinación R^2 (Cameron y Windmeijer, 1997), que representa la

varianza de la salida decodificada del autoencoder relacionada con la varianza del espectro original, con 1 indicando una copia perfecta del espectro original (Ver Anexo .14).

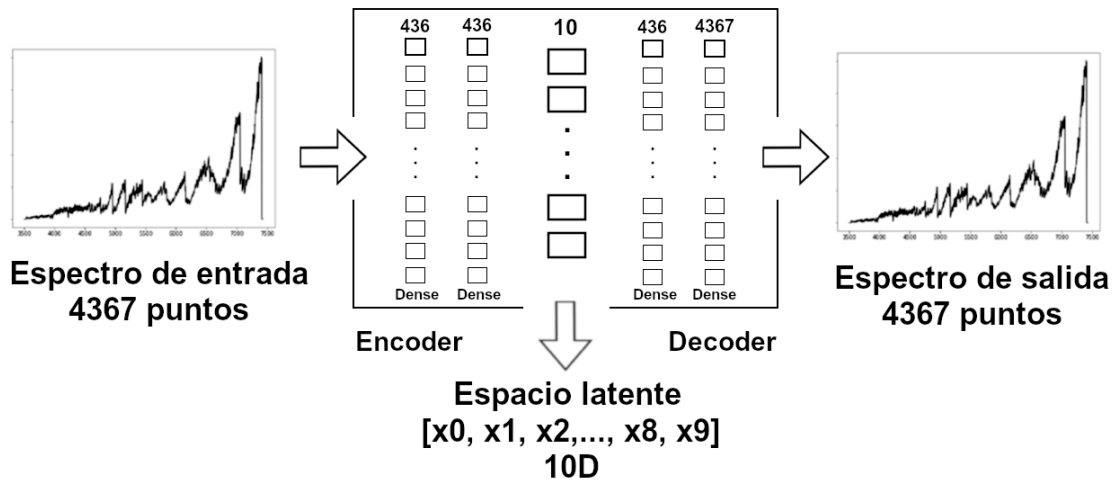


Figura 7.2: Ilustración de un modelo de autoencoder aplicado a espectros estelares. Para la reducción de los 4367 valores de cada espectro a un tensor del espacio latente de 10 ejes, el codificador utiliza dos capas ocultas con 2.099.350 parámetros. Para la reconstrucción del espectro decodificado, el decodificador utiliza dos capas ocultas con 1.913.175 parámetros.

Por otro lado, en los autoencoders variacionales(VAE) se impone una restricción adicional. La representación codificada, es decir, las variables latentes, siguen alguna distribución de probabilidad pre-establecida. Habitualmente, se elige una distribución gaussiana por su generalidad (Briot et al., 2020). En estos modelos VAE, para construir la distribución de variables latentes, el codificador desarrolla una media y una desviación estándar condicionales, re-parametrizadas por un término épsilon (distribuido normalmente).

Además de la función de pérdida de reconstrucción, los autoencoders variacionales incorporan una función de pérdida KL (divergencia de Kullback-Leibler) para mantener la forma de la distribución latente cercana a la normal (ver Anexo.8.3 para una descripción formal). Esta característica se conoce como *manifold learning* o aprendizaje de representación y proporciona una estrategia generativa simple para obtener nuevas entradas sintéticas realistas mediante el muestreo de la distribución normal del espacio latente. Se puede encontrar una descripción formalizada de ambos tipos de autoencoders en el Anexo.8.

Durante el desarrollo de esta investigación, hemos estudiado la influencia de la dimensión del espacio latente y de la configuraciones del autoencoder, realizando numerosas pruebas de reducción de espectros estelares a representaciones de menor dimensión, mediante entrenamientos con diferentes librerías de estrellas y galaxias (ver Anexo11.2.3). Todas las pruebas han sido implementadas utilizando *Jupyter notebook*, *astropy*, *numpy*, *matplotlib*, *tensorflow 2* (Martín Abadi et al., 2015) y *scikit-learn*. La Tabla 7.1 muestra la influencia de la dimensión del espacio latente en la codificación de los 985 espectros de la librería MILES durante 100 epochs. Vale la pena mencionar cómo duplicar el número de dimensiones (de 6 a 12), produce un aumento del R^2 alrededor del 0,0019.

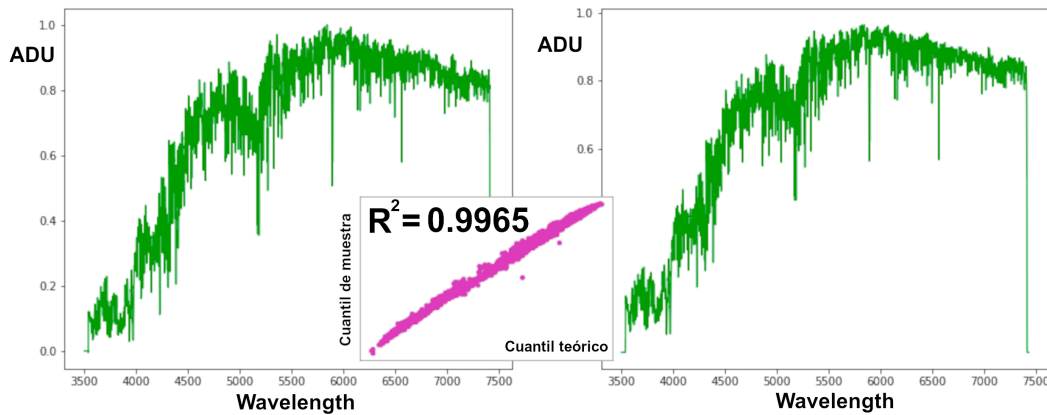


Figura 7.3: Comparación gráfica de los resultados obtenidos con un autoencoder variacional de cuatro dimensiones y cuatro capas entrenado con 985 espectros. Espectro de entrada (izquierda) y salida decodificada (derecha) para la estrella HD184406 de tipo *K*, con coordenadas RA:19:34:05.35 y DEC:+07:22:44.18. Desviación $R^2 = 0.9965$. Librería MILES, IAC.

Tabla 7.1: Influencia de la dimensión del espacio latente en los resultados de la codificación. R^2 para el conjunto completo (varianza ponderada: el valor proporcionado es el valor promedio ponderado con las varianzas de cada coeficiente individual calculado para cada espectro).

Dimensiones	4	6	10	12
R^2	0.9302	0.9504	0.9517	0.9523

Como ejemplo de las configuraciones probadas, la Figura 7.4 muestra los resultados de un autoencoder variacional de cuatro dimensiones y cuatro capas utilizado para la reducción de los espectros de la librería MILES. El codificador entrenó 2.427.624 parámetros durante 100 epochs, proporcionando resultados prometedores a pesar de la simplicidad de la red y del reducido número de curvas utilizadas para el entrenamiento (985 espectros). Los coeficientes R^2 fueron respectivamente 0,8369 y 0,7978 (varianza ponderada: el valor proporcionado es el valor promedio ponderado con las varianzas de cada coeficiente individual) para los conjuntos completos de prueba y entrenamiento (Ver Anexo .14).

7.3 Sonificación del espacio latente generado por sparse autoencoders

Esta sección describe el prototipo software que hemos desarrollado para la representación sonora precisa y objetiva de la información contenida en espectros estelares. La aplicación sintetiza notas y acordes musicales a partir de la sonificación del espacio latente de un sparse autoencoder de dos capas. El modelo es capaz de reducir los 4367 valores de flujo de cada espectro a un vector de 10 dimensiones. Esta dimensión del espacio latente se escogió de entre los modelos puestos a prueba con la intención de proporcionar la conversión “más cercana posible” a los espectros estelares originales (mejores resultados del análisis R^2). Los vectores latentes obtenidos con este modelo, se sonifican utilizando síntesis aditiva de audio y

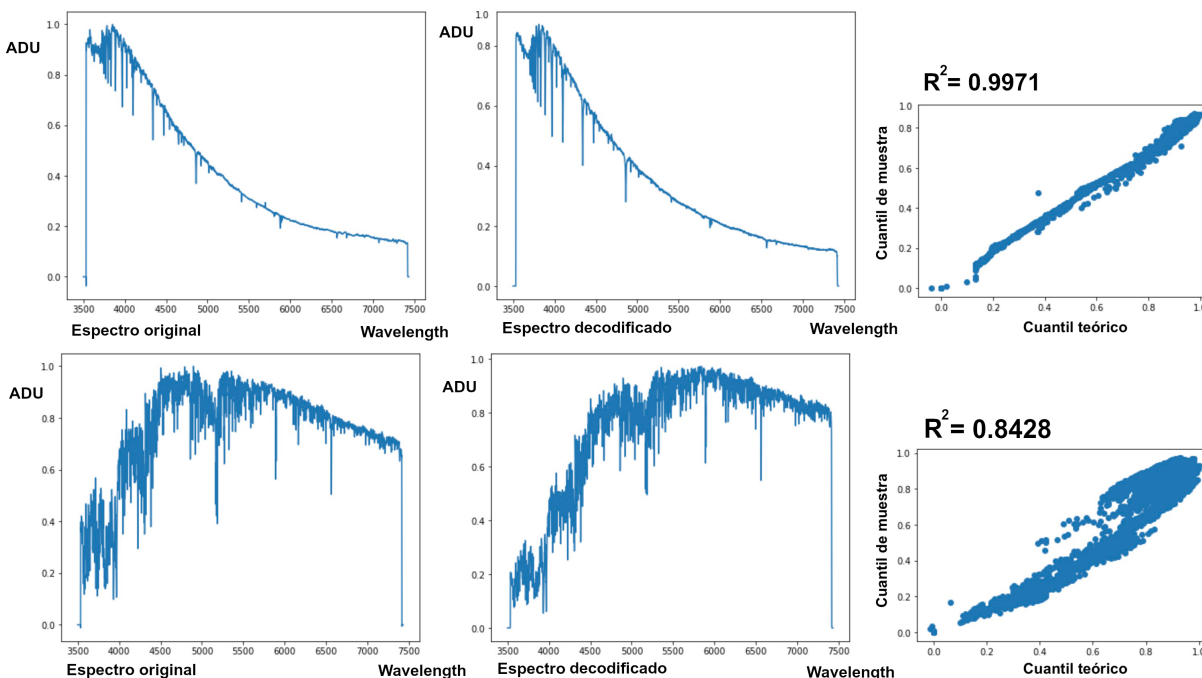


Figura 7.4: Comparación de los espectros original y decodificado a la salida de un autoencoder variacional de 4 dimensiones (izquierda) y gráfico de cuantiles (derecha). Coeficiente de determinación R^2 para la codificación de HD 057061 (arriba) y HD 026965 (abajo). Librería MILES, IAC.

se presentan secuencialmente.

Atendiendo a los detalles técnicos, el modelo de autocodificación ha sido entrenado con 985 espectros de la librería MILES. El codificador utiliza dos capas `dense` intermedias y entrena 2.099.350 parámetros durante 100 epochs. La generación de eventos se implementó en *Python* y se envió a *Cabbage-CSound* a través de OSC. El valor de la función de pérdida final (Error Cuadrático Medio, MSE), fue de 0,0022.

En cuanto al diseño de sonificación, el prototipo puede verse como un sintetizador de audio aditivo de 10 sinusoides (Kleczkowski, 1989; Moorer, 1977) que genera espectros complejos de manera análoga a la desarrollada en un proceso de modelado sinusoidal (Quatieri y McAulay, 1986). Cada dimensión del espacio latente se mapeó directamente a la frecuencia de cada oscilador, únicamente multiplicada por un factor de 10.000 para llevar los valores al rango de frecuencias audibles. Esto permite la monitorización de cada dimensión del autoencoder por separado o la escucha de los vectores multidimensionales del espacio latente.

La elección de osciladores sinusoidales como fuentes sonoras fundamentales se realizó en base a criterios de objetividad, precisión y simplicidad de la representación (Kramer, 2000). Las ondas sinusoidales carecen de armónicos, lo que proporciona una representación biunívoca en el dominio frecuencial y temporal (Zwicker y Fastl, 2013), algo potencialmente útil para el análisis científico de las variables representadas. La percepción humana de estos sonidos ha sido ampliamente estudiada (Fletcher y Munson, 1933; Robinson y Dadson, 1956; Suzuki y Takeshima, 2004),

lo que, tal y como mostramos en el Capítulo 10, permite incorporar diversos modelos de ponderación de la sonoridad en función de la frecuencia (ver Anexo .11) a los diseños de sonificación, o incluso determinar las funciones de escala que relacionan el dominio representado con su representación sonora (Walker y Lane, 2001). Además, los tonos puros ofrecen una representación abstracta y neutra en cuanto al contexto cultural o emotivo, lo que los hace especialmente útiles en representaciones experimentales.

Con la intención de reducir la posible fatiga derivada de un tiempo de exposición largo y de aumentar la musicalidad de las sonificaciones (Barrass y Vickers, 2011; Kramer, 2000; Risset, 2003; Roads, 1996), se introdujo un procesador interactivo de reverberación que suaviza el ataque y las transiciones entre representaciones, generando una sonoridad próxima a la de un órgano.

También se incluyó en el sintetizador un control interactivo de la frecuencia fundamental, para permitir a los usuarios elegir el rango de frecuencias de la representación. Por defecto, las frecuencias generadas entre 20 Hz y 5 KHz son filtradas, aunque las frecuencias de corte de los filtros también se pueden controlar interactivamente. Todas las amplitudes de las sinusoides son iguales (no se implementó ponderación). La interfaz gráfica proporciona una representación del espectro original a la izquierda y del espectro recuperado a partir del vector latente a la derecha. La Tabla 7.2, ofrece el mapeado por defecto realizado para la estrella HD 040657. Como puede apreciarse en el ejemplo, la representación sin actuar sobre el control interactivo de la frecuencia fundamental silenciará las dimensiones L1, L2, L3, L5 y L6, debido a que ofrecen valores por debajo del límite inferior de frecuencia del oído humano (20Hz). Del mismo modo, con el factor de multiplicación interactivo (10x), la dimensión L7 y L8 se silencian debido a que se excede el límite superior (20KHz). Para solucionar este exceso de rango dinámico en futuras implementaciones, se propone una normalización ajustada a los límites de percepción, tal y como realizamos en la Sección 10.

Tabla 7.2: Ejemplo de mapeado para la estrella HD 040657, con coordenadas RA:06:00:03.35 y DEC:-03:04:27.33. Valor latente, frecuencia y valor de frecuencia fundamental final para un factor de multiplicación interactivo 10x.

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
Valor latente	1520094	130328	59053	94887	3383624	166717	3745	76199841	29883587	1386263
Frecuencia (Hz)	152	13	6	9	338	17	0	7.620	2.988	139
F.Final (factor 10x)	1.520	130	60	90	3.380	170	4	76.199	29.883	1.386

Hemos desarrollado dos versiones del sintetizador. La primera, *Autoencoder2Chords*, presenta los 10 osciladores sonando simultáneamente en una representación tipo “acorde”. La segunda, *Autoencoder2Notes*, activa secuencialmente cada oscilador para ofrecer un flujo continuo de notas entre espectros. Los prototipos de software abierto y los vídeos de demostración de ambas versiones están disponibles en:

<https://github.com/AuditoryVO/Autoencoder2Sound>

<https://vimeo.com/902527469>

<https://vimeo.com/902530924>

7.4 Espacialización del espacio latente generado por autoencoders variacionales

Esta sección describe el prototipo que hemos implementado para explorar las posibilidades en la espacialización multicanal de las sonificaciones del espacio latente generado a partir de espectros estelares, descritas en la Sección 7.3. En este caso, se propone la utilización de un autoencoder variacional de 12 dimensiones, que además permite la generación de espectros estelares sintéticos mediante el muestreo del espacio latente, algo que podría ser útil para aumentar conjuntos de datos de reducido tamaño.

El objetivo principal de este diseño es proporcionar una representación envolvente espacializada de los espectros para ser utilizada en presentaciones públicas, conferencias o exhibiciones. En esta ocasión utilizamos 12 dimensiones para hacerlas coincidir con los doce canales de entrada del sistema de audio envolvente 3D, disponible en las instalaciones de la Escuela Superior de Música de Catalunya (ESMUC), donde se realizó una presentación pública en julio de 2023.

En esta implementación, los 12 valores de cada vector del espacio latente se mapean a las frecuencias fundamentales de 12 generadores de ondas sinusoidales independientes. Cada uno de estos osciladores se envía a un altavoz para permitir la espacialización acústica de la representación en un sistema de sonido envolvente formado por 12 altavoces. Un aspecto interesante de esta espacialización es que permite ubicar en el espacio físico los valores de los vectores latentes, proporcionando una forma de monitorizar el comportamiento de la red neuronal.

Respecto al diseño de sonificación, se utilizó la misma estrategia de mapeado paramétrico y generación mediante síntesis aditiva de ondas sinusoidales de la Sección 7.3, incorporando también el procesador de reverberación interactivo para acercar las representaciones al imaginario musical. Como puede verse en la Tabla 7.3, que ofrece el ejemplo de mapeado de la estrella HD 040657, los valores latentes negativos fueron representados mediante la inversión de la señal sinusoidal (cambio de polaridad, fase = 180°).

Tabla 7.3: Ejemplo de mapeado para la estrella HD 040657, con coordenadas RA:06:00:03.35 y DEC:-03:04:27.33.

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
Valor latente	-6192252	14677721	3183953	-3640259	16842898	-2528660	-2562281	8207887	11515150	11849148	13908924	-4402188
Frecuencia (Hz)	619	1.468	318	364	1.684	253	256	821	1.151	1.184	1.391	440
F.Final (factor 10x)	6.192	14.678	3.184	3.640	16.843	2.529	2.562	8.208	11.515	11.849	13.909	4.402
Fase (grados)	180	0	0	180	0	180	180	0	0	0	0	180

La interfaz gráfica del prototipo, mostrada en la Figura 7.5, proporciona el espectro original, el espectro recuperado del vector latente sonificado y una representación de los altavoces activos. Adicionalmente, incorpora controles de silenciado en cada uno de los 12 canales.

El modelo también ha sido entrenado durante 100 epochs en la librería MILES con fines comparativos. El valor de la función de pérdida final (MSE) fue de 0,0158, finalmente reducido a 0,0043 después de una optimización de la tasa de aprendizaje. La estrategia de mapeado utilizada fue idéntica a la del enfoque descrito en la Sección 7.3, convirtiendo los valores latentes en las frecuencias fundamentales de 12 osciladores independientes. Los valores latentes

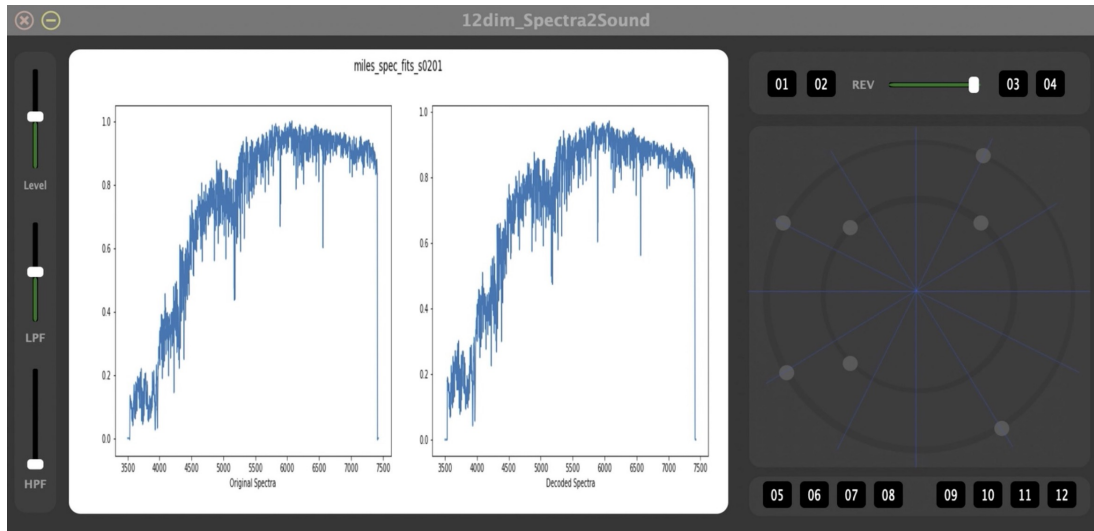


Figura 7.5: Interfaz gráfica del prototipo VAE2Sound. Espectro original (izquierda) y espectro decodificado (derecha) para la estrella HD 040657 con coordenadas RA:06:00:03.35 y DEC:-03:04:27.33. Representación adicional mediante círculos de los altavoces activos (dimensiones del espacio latente) que generan la sonificación.

también se multiplicaron por 10.000. Se utilizaron frecuencias de corte predeterminadas de 20 Hz y 8 kHz, respectivamente, para los controles interactivos de los filtros paso alto (HP) y paso bajo (LP).

Los siguientes enlaces proporcionan un vídeo demostrativo con la exploración exportada en estéreo de 25 espectros estelares de la librería MILES y el prototipo completo en software abierto:

<https://vimeo.com/641221857>

<https://github.com/AuditoryVO/VAE2Sound>

Capítulo 8

Sistema de composición musical basado en deep learning

Este capítulo presenta el sistema de composición musical no supervisado basado en deep learning que hemos desarrollado para generar piezas originales a partir de datos astronómicos. Con el objetivo de ofrecer una revisión automática y objetiva sobre el tema clásico de la armonía de las esferas, el sistema permite la exploración de librerías estelares utilizando arquitecturas de autoencoder. La propuesta se basa en un algoritmo que compara una partitura de acordes estelares creada a partir de su decodificación mediante autoencoders, con la salida de una red neuronal LSTM con atención entrenada con partituras MIDI, utilizando la Pitch Class Set Theory (Forte, 1973). La partitura MIDI final puede ser renderizada por cualquier instrumento virtual o sintetizador.

Este enfoque permite la generación de composiciones completamente originales a partir de espectros estelares y de galaxias, mediante la activación del proceso generativo con una pieza musical de entrada. Además, el sistema que hemos desarrollado proporciona representaciones gráficas de los objetos y de sus espectros, que son las fuentes reales de la composición. Una vez generada la pieza, hemos realizado el proceso (humano) de definición de la instrumentación, diseño sonoro, mezcla y mastering que proporciona el resultado de audio final. El capítulo presenta la pieza *AI-rmonies of the Spheres*, activada por el motete *In me transierunt* de Orlando di Lasso, la música que podría haber inspirado a Johannes Kepler durante la escritura de *Harmonices Mundi* (1619) (García Riber y Serradilla, 2023a).

8.1 Conversión de la librería MILES en acordes musicales

Esta sección describe un método para la conversión de espectros estelares en acordes musicales enfocado a la representación secuencial sonora de los espectros estelares de la librería MILES.

La Figura 8.1 muestra la estructura del autoencoder variacional de cuatro dimensiones utilizado para convertir los espectros en acordes musicales de cuatro notas. Esta estructura se

eligió tomando como referencia del tetracordio, la unidad del sistema armónico en la antigua Grecia (Shirlaw, 1951). Esta estructura musical básica proporciona el equilibrio necesario para la generación de acordes sencillos, pero con un cierto grado de complejidad o interés armónico, lo que favorece la inteligibilidad y mantiene el interés en las piezas generadas.

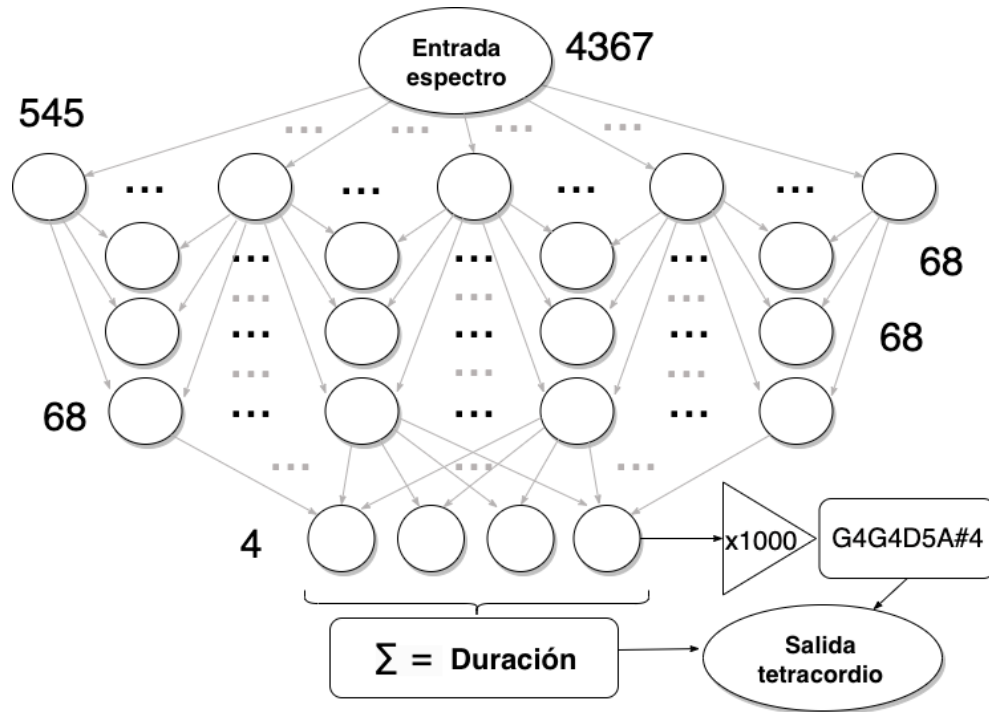


Figura 8.1: Arquitectura simplificada del codificador variacional utilizado para extraer las características latentes de cada espectro estelar de la librería MILES. Los números indican la dimensión de cada capa **dense** utilizada, y el espacio latente final de cuatro dimensiones. El acorde de muestra G4 G4 D5 A#4 se genera a partir del espectro de la estrella BD+15 1305, con coordenadas (J2000) RA 06:44:42.98 DEC 14:54:35.97.

Cada valor latente se multiplica por 1000 para llevar sus valores al rango de frecuencias audibles, y es aproximado a la frecuencia fundamental de la nota más cercana de la escala cromática. La duración de cada acorde se calcula utilizando un mecanismo de ponderación que suma los valores del vector latente, generando los acordes más largos a partir de los vectores con los valores latentes más altos.

El codificador utiliza 4 capas **dense** y entrena 2.427.624 parámetros durante 100 epochs. Los coeficientes de análisis R^2 para los conjuntos de prueba y entrenamiento obtenidos fueron respectivamente 0,8472 y 0,8468 (varianza ponderada). La librería MILES contiene muy pocos espectros de tipo O por lo que el modelo está peor entrenado en este grupo de datos y genera codificaciones más alejadas del original. El conjunto de espectros de entrada se balanceó para compensar este problema mediante la repetición de los tres espectros de tipo O incluidos en la librería (HD 057060, HD 057061 y HD 209975), aumentando el número final de curvas de entrenamiento a 1123 espectros.

En el conjunto balanceado, los coeficientes R^2 para los subconjuntos de prueba y entrenamiento fueron respectivamente 0,8472 y 0,8468 (varianza ponderada). En el conjunto sin balancear, es

Tabla 8.1: Espectros estelares utilizados para generar la secuencia OBAFGKM. Resumiendo las clases de luminosidad de Morgan-Keenan (MK), Ia y Ib corresponden a supergigantes luminosas, II a gigantes brillantes, III son gigantes normales, IV son subgigantes, V corresponde a estrellas enanas de la secuencia principal, VI a subenanas y D a enanas blancas. El número permite escalar cada tipo de 0 (el más caliente) a 9 (el más frío).

Tipo	R ²	MAE	RMSE	Error máx.	Nombre	RA(J2000)	DEC(J2000)
O7Ia	0,988	0,0481	0,0729	0,8733	HD 057060	07:18:40.38	-24:33:31.32
B5V	0,879	0,0259	0,0323	0,1478	HD 003369	00:36:52.80	33:43:09.48
A0V	0,979	0,0196	0,0290	0,1301	HD 031295	04:54:53.69	10:09:02.88
F1V	0,948	0,0598	0,0657	0,2532	HD 222451	23:40:40.56	36:43:14.88
G1V	0,977	0,0337	0,0396	0,2653	HD 114606	13:11:21.36	09:37:33.49
K0V	0,972	0,0523	0,0716	0,2654	HD 233832	11:26:05.52	50:22:32.88
M1V	0,949	0,5578	0,0711	0,1278	HD 036395	05:31:27.41	-03:40:37.99

decir, utilizando directamente los 985 espectros de la librería, fueron 0,8369 y 0,7978 (varianza ponderada). Finalmente, la librería de análisis musical *music 21* se utilizó para generar la partitura con la secuencia de acordes que, finalmente, se renderizó utilizando el software abierto *MuseScore3*.

8.2 Secuencia de acordes OBAFGKM

Para evaluar la propuesta de conversión de espectros estelares en acordes musicales, hemos utilizado una estrella de muestra de cada tipo espectral (ver Anexo.4). Los resultados pueden verse en la Tabla 8.1 (Falcón-Barroso et al., 2011; Prugniel et al., 2007; Sánchez-Blázquez et al., 2006; Valdes et al., 2004), que incluye el valor individual del coeficiente de determinación (R²) obtenido en la codificación, el error absoluto medio (MAE), el error cuadrático medio y el error máximo, obtenidos entre el espectro original y el decodificado. Como puede observarse, el máximo error se produjo para el espectro tipo O, debido a que el entrenamiento se realizó replicando los únicos tres espectros de este tipo existentes en la librería. Con estos ejemplos, hemos generado una secuencia de acordes OBAFGKM que puede escucharse en:

<https://vimeo.com/770510584>

8.3 Redes LSTM con atención

“Una Red Neuronal Recurrente (RNN) es una red neuronal especializada en procesar una secuencia de valores” que “también puede procesar secuencias de longitud variable” a través del intercambio de parámetros a lo largo de todo el modelo. Descrita por primera vez por Rumelhart et al., 1986, y especialmente útil en modelos de tipo secuencia a secuencia, “comparte los mismos pesos a través de varios instantes de tiempo” de modo que “cada elemento a la salida de la red se produce utilizando la misma regla de actualización aplicada a las salidas anteriores. Esta formulación recurrente produce un intercambio de parámetros a través de un gráfico computacional muy profundo” (Goodfellow et al., 2016).

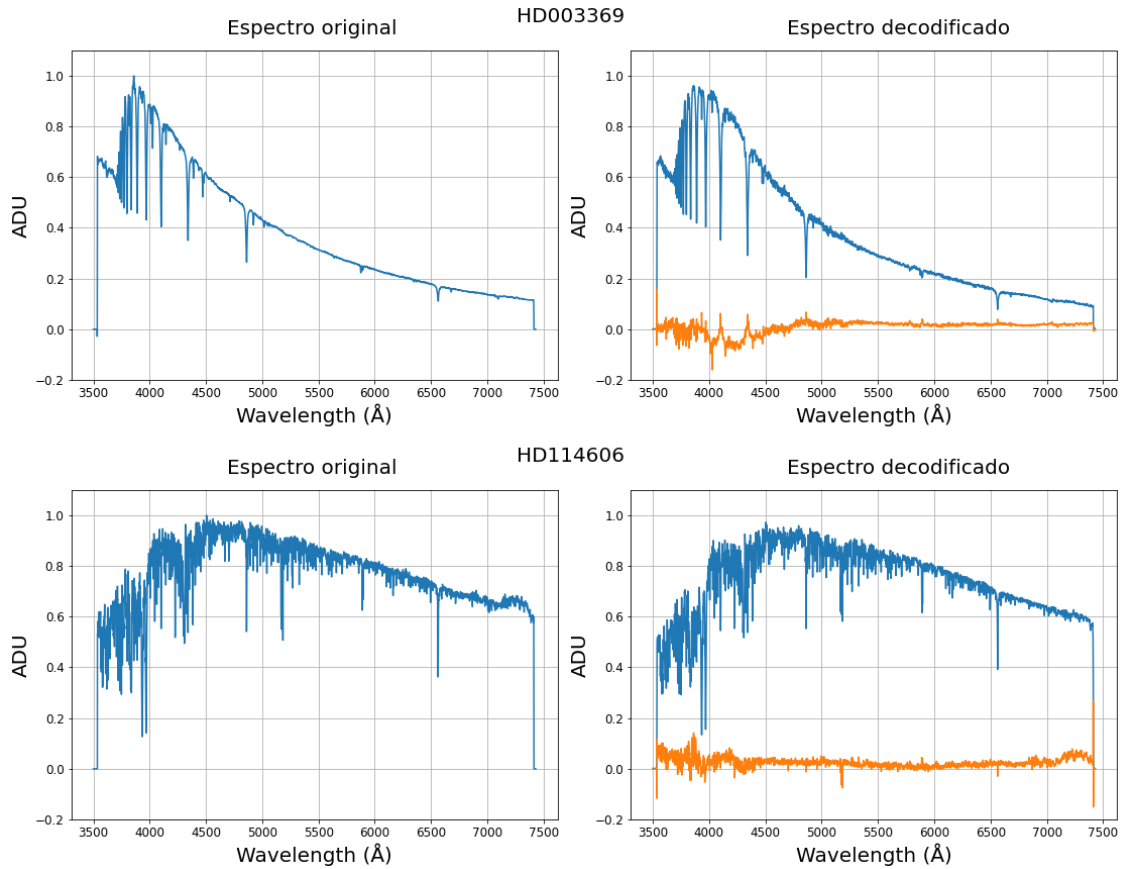


Figura 8.2: Resultados del autoencoder para la estrella HD 003369 tipo B5V y para la HD 114606 G1V. Espectro original (izquierda) y salida decodificada del autoencoder variacional 4D (derecha) en azul. Error residual en naranja.

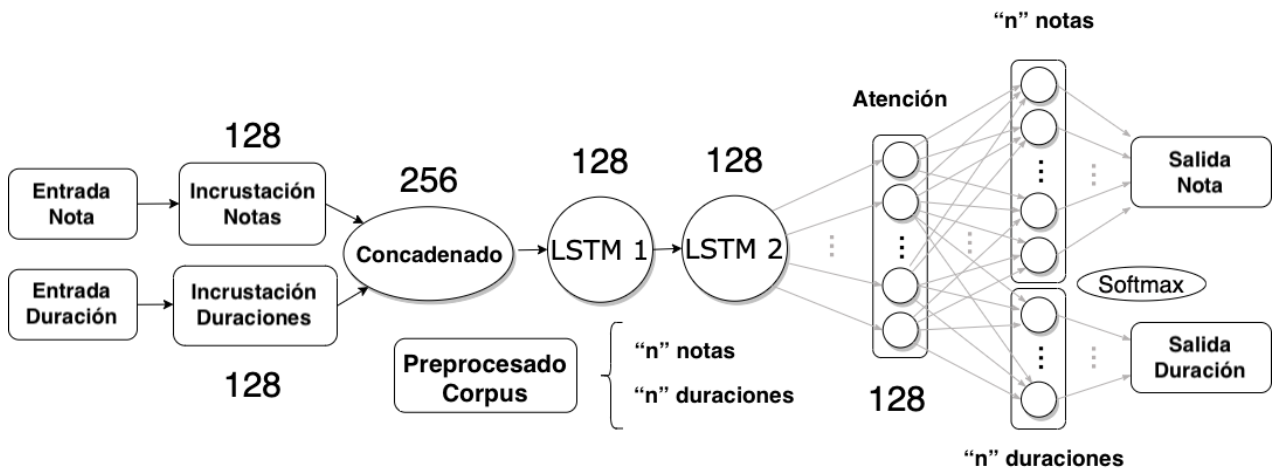


Figura 8.3: Diagrama de red que incluye el tipo y la dimensión de cada capa para la arquitectura LSTM dual con atención utilizada en el sistema de composición musical.

Las arquitecturas de memoria a corto-largo plazo (LSTM), presentadas por primera vez por Graves et al., 2013, se construyen sobre RNN con compuertas, para resolver el problema del desvanecimiento y/o explosión del gradiente. Este problema aparece cuando las RNN intentan aprender dependencias a largo plazo (Torres, 2020), y se resuelve mediante la introducción de bucles automáticos ponderados dependientes del contexto. Las redes LSTM propagan información a través de secuencias largas y permiten que las salidas anteriores se utilicen como entradas a lo largo de las capas ocultas. Estas redes mejoran la capacidad de aprender posibles relaciones entre características a lo largo del tiempo y permiten el mantenimiento en la memoria de características relevantes de los datos de entrada (ver Anexo.9).

Para mejorar la gestión de las dependencias a largo plazo, se puede incorporar un mecanismo de atención adicional (Ver Anexo.10), introducido en el campo de la traducción automática por Bahdanau, 2014. Este enfoque utiliza las partes más relevantes de la secuencia de entrada mediante una combinación ponderada de los vectores de entrada codificados para centrarse “cada instante de tiempo en algunos elementos específicos de la secuencia de entrada” (Briot et al., 2020).

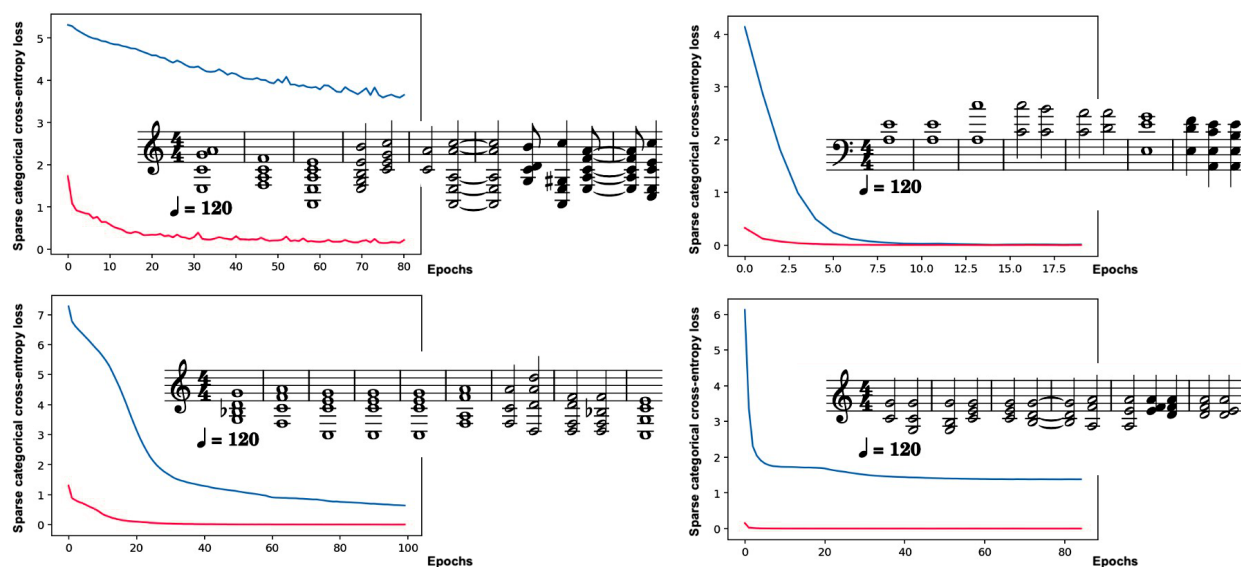


Figura 8.4: Partituras resultantes y función de pérdida (sparse categorical cross-entropy) para la alturas de nota o pitch (azul) y para las duraciones de nota (rojo) de cuatro experimentos con diferentes conjuntos de entrenamiento: un sólo motete de Lasso (arriba a la izquierda), la misma pieza repetida 100 veces (arriba a la derecha), 50 piezas musicales de Des Prez y Lasso (abajo a la izquierda) y 1318 piezas de Palestrina (abajo a la derecha). Vale la pena mencionar cómo la función de pérdida de altura de nota (pitch) refleja una tasa de aprendizaje inesperadamente rápida para el corpus de Palestrina (el que tiene más piezas), probablemente motivada por las características musicales altamente deterministas del compositor.

Basado en estas premisas, realizamos varios experimentos de composición generativa utilizando diferentes obras musicales del siglo XVI para entrenar y probar una red LSTM con atención basada en una de las implementaciones descritas en Babcock y Bali, 2021. La red, representada

en la Figura 8.3, aprende la altura de nota (pitch) y la duración de cada acorde del conjunto de datos gracias a una arquitectura de entrada dual y salida dual, que incluye una estrategia de muestreo basada en un control de “temperatura” que actúa después de introducir las notas en la función de predicción.

La Figura 8.4 resume los experimentos realizados para analizar el comportamiento de la red. La partitura superior izquierda es el resultado de la red cuando se entrena con un sólo archivo MIDI. Los 203 acordes y 7 duraciones de notas diferentes del motete *In me transierunt* de Orlando di Lasso se utilizaron para entrenar una red de 382.803 parámetros. Como era de esperar, los resultados sólo son útiles como punto de partida, proporcionando un error de pérdida de altura de nota (pitch) final (MSE) de 3,6543 después de 81 epochs.

La partitura superior derecha se generó con un corpus aumentado, repitiendo el motete 100 veces. Vale la pena mencionar cómo la red es capaz de imitar la pieza después de menos de 10 epochs. La función de pérdida de altura de nota final (pitch) es de 0,0118 (MSE) después de 20 epochs.

El experimento inferior izquierdo utilizó 50 piezas musicales, 25 de Josquin Des Prez y 25 de Orlando di Lasso. Incluía 3.245 acordes y 46 duraciones diferentes que claramente aumentaron la capacidad de aprendizaje del modelo, en este caso, con 1.174.620 parámetros entrenables. Después de 100 epochs, la función de pérdida de altura de nota final (pitch) fue de 0,6431 (MSE).

La partitura inferior derecha se obtuvo entrenando una red de 3.643.876 parámetros con 12.884 acordes y 15 duraciones diferentes extraídas de 1.318 piezas de Giovanni Pierluigi da Palestrina. La función de pérdida de altura de nota final (pitch) fue de 0,8879 (MSE) después de 41 epochs.

8.4 Sistema de composición musical basado en deep learning

Esta sección describe el sistema de composición completo que hemos implementado para generar piezas musicales a partir de espectros estelares utilizando el caso de estudio de la librería MILES.

Como se muestra en la Figura 8.5, el primer bloque del sistema de composición convierte los espectros de la librería en acordes musicales utilizando el autoencoder variacional de cuatro dimensiones descrito en la Sección 8.1. Cada espectro se convierte en un acorde musical de cuatro notas. La duración de cada acorde depende de los pesos relativos de su vector latente. El siguiente vídeo muestra un extracto de una representación de la librería estelar MILES renderizada con estilo de ciencia ficción (Partitura 1):

<https://vimeo.com/764757244>

El segundo bloque del sistema genera una segunda partitura (Partitura 2) utilizando la red neuronal LSTM con atención de la Sección 8.3. Esta red genera los acordes que proporcionan la estructura musical subyacente en la pieza final, mediante la activación de una pieza musical

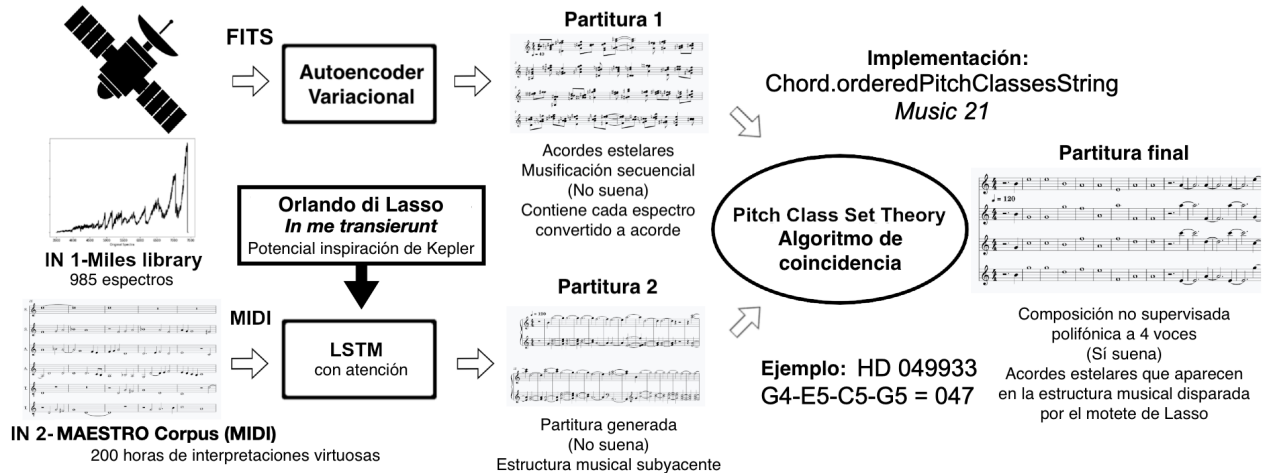


Figura 8.5: Sistema de composición musical de cuatro voces a partir de datos astronómicos basado en deep learning. Diagrama de bloques de la arquitectura completa basada en VAE-LSTM con atención y Pitch Class Set Theory. Los espectros estelares se convierten en acordes de cuatro notas (Partitura 1) utilizando un autoencoder variacional. En un proceso paralelo, una red neuronal LSTM con atención entrenada con el corpus MAESTRO genera la Partitura 2, activada por una pieza musical de entrada (*In me transierunt*). Finalmente, el algoritmo de análisis basado en la Pitch Class Set Theory analiza las semejanzas entre los acordes de la Partitura 1 y la Partitura 2 y genera la pieza sonora final con los acordes coincidentes extraídos de los espectros estelares.

de entrada. El siguiente enlace proporciona un ejemplo de la partitura de salida de la red entrenada con 1318 piezas corales del compositor renacentista Giovanni Pierluigi da Palestrina:

<https://vimeo.com/1043003037>

Finalmente, un algoritmo basado en la Pitch Class Set Theory analiza las semejanzas entre la Partitura 1 y la Partitura 2 para crear la pieza de salida con los acordes estelares coincidentes. El algoritmo realiza la comparación de los acordes de ambas partituras utilizando el método `Chord.orderedPitchClassesString` de la librería *Music 21*.

La Pitch Class Set Theory proporciona uno de los métodos más utilizados para reducir y etiquetar información musical (Forte, 1973). Según los principios de esta teoría, una clase de altura de nota A, es un grupo que contiene todos los registros correspondientes a esa nota con equivalencia de octava (A0, A2, A4, etc), sin distinguir equivalentes enarmónicos (como A3 sostenido y B3 bemol). Traduciendo esas clases a números, con 0 correspondiendo a C y 11 correspondiendo a B, cada acorde generado se reduce a un solo código que permite su comparación para encontrar coincidencias positivas entre corpus. De esta manera, el acorde G4-E5-C5-G5 corresponde al código 047, como ejemplo de un acorde que está presente tanto en la partitura de acordes estelares como en la partitura de acordes musicales y que corresponde a la estrella HD 049933 con coordenadas RA: 06:50:49.8309, DEC: -00:32:27.1675.

Las duraciones resultantes se filtran y reducen a sólo 4 figuras posibles (redonda, blanca, negra y corchea) manteniendo una cadencia lenta en la música que permite la sincronización con la representación gráfica de la estrella fuente “a tiempo real”, generando la representación

multimodal final de los espectros. Al final del proceso, la pieza musical final se renderiza con cualquier sintetizador o DAW.

El sistema se puede descargar de GitHub para reproducir el proceso de composición completo (software abierto) en:

<https://github.com/AuditoryVO/AI-rmonizer>

8.5 Composiciones no supervisadas a partir de espectros estelares

A continuación, presentamos varias piezas que hemos generado con diferentes corpus musicales de entrenamiento construidos a partir de piezas de Josquin Des Prez, Orlando di Lasso y Giovanni Pierluigi da Palestrina, como compositores renacentistas representativos que podrían haber inspirado los pensamientos musicales de Kepler.

Hemos encontrado cierta controversia sobre la influencia de los compositores más relevantes de principios del siglo XVII en la obra de Kepler. Para Ball, 2009, “el mundo musical de Kepler abarcaba la opulencia polifónica de Palestrina y Monteverdi”, mientras que Pesic, 2005 afirma que “él (Kepler) acepta el sistema de Zarlino y se refiere sólo a Lasso y Artusi, nunca a Monteverdi”. De todos modos, parece claro que, aunque Kepler no incluyó ninguna mención explícita en su obra, podría haber estado interesado en lo que hoy conocemos como la escuela franco-flamenca, especialmente centrado en la música de Orlando di Lasso, y más específicamente en su motete *In me transierunt* “como el motete celestial ideal jamás escrito”, que se ajustaba a sus pensamientos astronómico-musicales (Pesic, 2005).

El siguiente vídeo muestra una sincronización de la pieza generada con la partitura y con las representaciones gráficas de cada estrella fuente (espectro e imagen). Esta pieza se generó con el corpus de entrenamiento Lasso-Des Prez y se renderizó en un estilo electroacústico:

<https://vimeo.com/770493178>

Utilizando el sistema con el corpus de entrenamiento de Palestrina, se generó la siguiente pieza para flauta, violín y piano:

<https://vimeo.com/746620075>

Finalmente, para obtener resultados más generales, entrenamos el sistema con el conjunto de datos MAESTRO (Hawthorne et al., 2018). Este corpus ofrece alrededor de 200 horas de interpretaciones de piano virtuosas de obras de diversos compositores importantes, abarcando desde el siglo XVII hasta el siglo XX. Se han utilizado un total de 82.231 notas únicas y 118 duraciones diferentes para entrenar los 21.492.526 parámetros de la red durante 525 epochs. El proceso tardó cuatro horas y media utilizando procesado mediante GPU de alto rendimiento. El siguiente vídeo muestra la composición resultante, *AI-rmonies of the Spheres*, obtenida al alimentar el modelo con las notas y duraciones del motete *In me transierunt* de Orlando di Lasso:

<https://vimeo.com/794718061>

El Anexo.18 proporciona interpretaciones en vivo y composiciones musicales adicionales, que hemos realizado con el sistema de composición autónoma descrito en este capítulo durante el transcurso de la investigación. Finalmente, el Anexo.22 ofrece una versión del sistema que permite generar las representaciones multimodales “a tiempo real”, junto con un ejemplo de las partituras que se obtienen a la salida del sistema, que incluyen la referencia al objeto del cielo real que produce cada acorde.

Capítulo 9

Instrumentos musicales para la exploración de catálogos astronómicos

Este capítulo describe las estrategias de diseño e implementación que hemos utilizado en el desarrollo de un sintetizador que permite la exploración multimodal de la librería de espectros estelares del triplete de Ca II (CaT) del Instituto de Astrofísica de Canarias (IAC). El prototipo reacciona “a tiempo real” para proporcionar un acompañamiento armónico a cada nota del teclado de piano, generado a partir de datos reales de objetos celestes utilizando deep learning.

Una arquitectura de autoencoder de dos capas reduce cada espectro estelar a un acorde MIDI de seis notas. El algoritmo busca la nota fundamental de los acordes estelares resultantes que coinciden con la nota de la tecla del piano presionada por el usuario, y genera la sonificación en función de su vector latente codificado y del número de líneas de absorción detectadas en el espectro. La interfaz de usuario proporciona el nombre del objeto, el espectro y la precisión de la representación sonora a través del R^2 calculado entre el espectro original y el decodificado. Este enfoque genera una representación sonora rápida de las líneas de absorción y de la distribución del flujo de luz en los espectros, que podría implementarse en archivos astronómicos para permitir la exploración, interpretación y clasificación de sus datos a través del sonido.

De este modo, proponemos la utilización de instrumentos musicales digitales y virtuales en la exploración multimodal interactiva de librerías estelares. Aprovechando la popularidad del teclado de piano en la cultura occidental y el potencial de atracción intrínseco en la música y el sonido, se espera que el enfoque pueda ser de utilidad en actividades educativas y de divulgación para acercar los conceptos de Astronomía a todos los públicos, además de ser aplicable en diseño sonoro y composición musical para la creación de timbres y texturas complejas (García Riber y Serradilla, [2024a](#)).

9.1 La librería CaT

El triplete de Ca en el infrarrojo cercano se refiere a tres líneas espectrales de calcio ionizado, alrededor de las longitudes de onda de 8498 Å, 8542 Å y 8662 Å, que se utilizan comúnmente en el estudio de la actividad cromosférica de las estrellas, que son regiones clásicamente vistas como situadas por encima de la fotosfera brillante, y caracterizadas por un gradiente de temperatura positivo y una marcada desviación del equilibrio radiativo (J. C. Hall, 2008). El triplete también es útil en la clasificación espectral estelar, ya que es prominente en estrellas de tipo espectral G, K y M, y en el estudio de cúmulos globulares, galaxias activas y de tipo temprano, y regiones H II extragalácticas.

La librería de calibración empírica del triplete de Ca II (CaT) (Cenarro et al., 2001) contiene 706 estrellas en el rango espectral del infrarrojo cercano, desde 8348 Å hasta 9020 Å, cubriendo una amplia variedad de tipos espectrales, clases de luminosidad, metalicidades y temperaturas efectivas (ver Anexo.5). Los espectros de la librería se obtuvieron durante 21 noches entre 1996 y 1997 con el Telescopio Jacobus Kapetyn (JKT), el Telescopio Isaac Newton (INT) y el Telescopio William Herschel (WHT) en el Observatorio del Roque de los Muchachos (La Palma, España), y con el telescopio de 3,5 m en el Observatorio de Calar Alto (Almería, España). Esta librería fue definida para proponer un nuevo conjunto de índices que permitan cuantificar con precisión la fuerza del triplete de Ca en poblaciones estelares con un amplio rango de edades y metalicidades.

9.2 Comparativa de autoencoders con la librería CaT

Como se describió en capítulos anteriores, los autoencoders se han utilizado ampliamente para aprender representaciones eficientes de los conjuntos de datos. En esta sección, se utilizan de nuevo en la reducción de la dimensionalidad de cada espectro de la librería CaT para su conversión en sonido. Este proceso tiene como objetivo proporcionar una representación sonora objetiva, directa, inteligible y “lo más cercana posible” a los espectros reales.

Después de pruebas intensivas con diferentes modelos y configuraciones de red, proponemos una arquitectura de autoencoder de dos capas **dense** para la reducción de los datos espectrales de la librería CaT. Con un número total de 121.705 parámetros entrenables en el codificador y 105.197 en el decodificador, este modelo permite la reducción de los 790 puntos de cada espectro a un vector latente de seis valores con una pérdida final de $4,37 \times 10^{-04}$ (MSE). La primera capa del codificador reduce la dimensión de cada espectro de 790 a 131 puntos utilizando 103.621 parámetros. Estos 131 puntos se reducen en la segunda capa al vector latente final de seis dimensiones utilizando 17.292 parámetros.

Cada espectro se replicó 50 veces utilizando una mezcla aleatoria en el proceso de división del conjunto para entrenar el modelo durante 100 epochs. Este enfoque fue útil para hacer que el autoencoder convergiera un 30,37% más rápido que con un aumento equivalente en el número de epochs de entrenamiento (1.325 segundos frente a 1.903 segundos en 10 procesos de entrenamiento). Cabe destacar que el objetivo de este proceso es obtener una copia “lo más parecida posible” a cada espectro de la librería para ser sonificado, y que el autoencoder no estará expuesto a datos desconocidos. Utilizando este enfoque, el R^2 para el conjunto de

entrenamiento fue de 0,9296 (28.280 espectros, valor medio en 10 procesos de entrenamiento). Se obtuvieron resultados equivalentes ($R^2 = 0,9337$) con el conjunto de entrenamiento original durante 5.000 epochs (565 espectros, valor medio en 10 procesos de entrenamiento).

Hemos probado y comparado tres tipos y configuraciones de autoencoders utilizando el mismo conjunto de datos (ver Anexo.8 y Shah y Ganatra, 2022 para una visión general de los tipos de autoencoders y sus aplicaciones más comunes). Los siguientes resultados comparan un autoencoder variacional de cuatro capas y dos sparse autoencoders de dos y seis capas.

De nuevo, el coeficiente de determinación (R^2) entre los espectros originales y decodificados se utilizó para evaluar la precisión de la representación ($R^2 = 1$, copia perfecta y $R^2 = 0$, ausencia de correlación). El notebook de análisis utilizado para estudiar las distribuciones de R^2 se puede encontrar en el GitHub del prototipo, proporcionado al final de este capítulo. Todos los modelos se entrenaron utilizando el mismo conjunto de datos aumentado (x50) de la librería CaT, mezclado aleatoriamente en el proceso de división del conjunto.

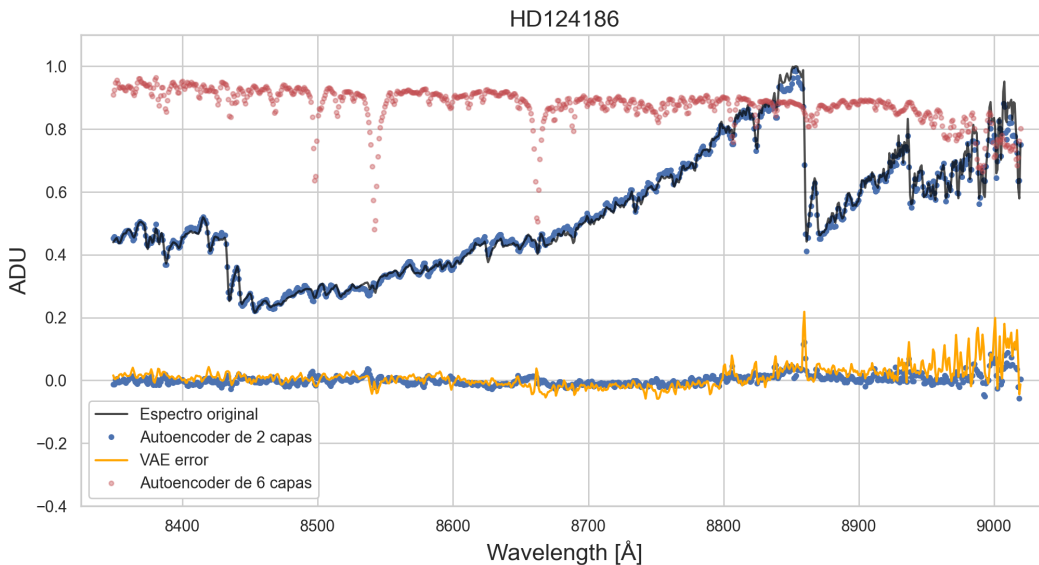


Figura 9.1: Espectros decodificados de HD 124186 generados por los tres modelos de autoencoder analizados. Flujo de brillo (ADU) en función de la longitud de onda (Å). La curva negra corresponde al espectro original. La línea punteada azul a la salida (arriba) y al error residual (abajo) del autoencoder de dos capas. La línea naranja muestra el error residual del VAE de cuatro capas y la roja la salida del autoencoder de 6 capas. Como puede observarse, el modelo de seis capas genera una salida completamente decorrelada con la entrada para este espectro.

En cuanto a la comparativa, los espectros representados en la Figura 9.1 ilustran el comportamiento de los tres modelos al codificar el espectro de la estrella HD 124186. Este objeto muestra explícitamente el impacto negativo observado en los resultados al aumentar el número de capas del modelo. Como puede observarse, la red de seis capas no puede reproducir el espectro original y ofrece una salida muy alejada del espectro de entrada.

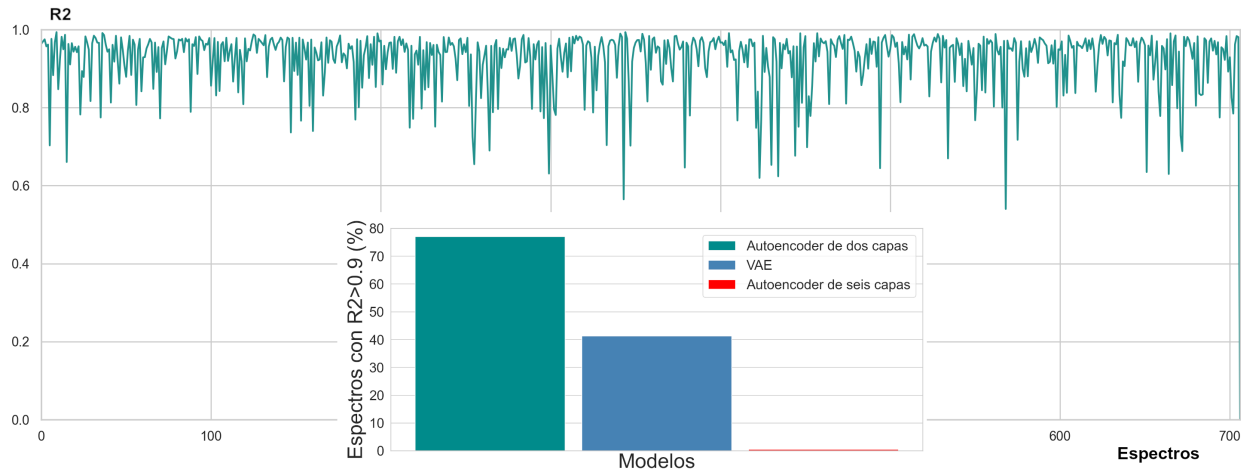


Figura 9.2: Rendimiento del autoencoder de dos capas, R^2 para los 706 espectros codificados de la librería CaT y porcentaje de espectros con R^2 superior a 0,9 para el autoencoder de dos capas (verde-izquierda), el VAE de cuatro capas (azul-centro) y el autoencoder de seis capas (rojo-derecha).

La Figura 9.2 muestra el rendimiento del modelo de dos capas finalmente utilizado en la implementación del sintetizador. Los R^2 individuales entre el espectro original y el decodificado varían de 0,9934, para la mejor codificación, a 0,54 para la peor, con un 77,1% de los espectros por encima de 0,9. Las otras arquitecturas evaluadas mostraron un rendimiento mucho menor, con un 41,3% por encima de 0,9 para el VAE, y un 0,57% para el autoencoder de seis capas. El autoencoder que utiliza sólo dos capas *dense* también proporcionó el menor tiempo de computación, consumiendo 1 segundo por epoch durante el proceso de entrenamiento frente a los 30 segundos por epoch que requiere el VAE, y los 3 segundos por epoch utilizados por el modelo de seis capas.

9.3 Análisis espectral por detección de umbral

Adicionalmente, implementamos en el sintetizador un análisis de los espectros basado en la detección de picos por umbral, con la finalidad de incorporar un mecanismo de variación de timbre controlado por las características del propio espectro y para mejorar las posibilidades de análisis de la representación. La Figura 9.3 muestra un ejemplo de las imágenes generadas en el bloque de pre-procesado que utiliza la función `find peaks` de la librería *SciPy* (“SciPy”, 2001), para detectar las principales líneas de absorción. Los flujos se invirtieron para encontrar mínimos locales separados por al menos 40 muestras, con una prominencia superior a 0,15 ADU y un ancho mayor de una muestra.

El número de líneas espectrales detectadas en cada espectro se mapeó directamente al LFO de un trémolo aplicado a la sonificación completa. Este efecto proporciona una herramienta para la diferenciación sonora de espectros basada en la frecuencia de oscilación inducida en la representación. Aunque las longitudes de onda de cada línea fueron calculadas y se muestran en la salida de *Python*, no se utilizaron en la sonificación. Futuras versiones del prototipo

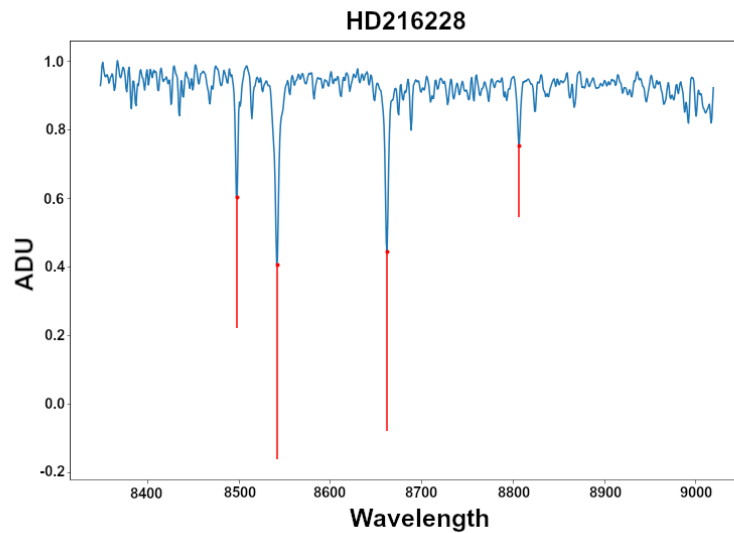


Figura 9.3: Interfaz de usuario mostrando la imagen de salida generada en el bloque de pre-procesado para el espectro de la estrella tipo K0 III, HD 216228. En azul, flujo normalizado (ADU) en función de la longitud de onda (\AA). Las líneas y puntos rojos representan las amplitudes de las líneas de absorción detectadas en el análisis secundario.

podrían beneficiarse de estos cálculos para generar sonidos enarmónicos y permitir la creación de timbres y texturas experimentales.

9.4 Mapeado del espacio latente

La Tabla 9.1 muestra cinco ejemplos de los mapeados utilizados para proporcionar una representación multimodal del espectro estelar asociado a cada nota de piano pulsada. El algoritmo utiliza la librería MIDI *pygame* (“Pygame”, 2000) para obtener los acordes finales de seis notas que controlan la sonificación, convirtiendo a MIDI la nota de la tecla pulsada y comparándola con los valores de cada vector latente (previamente multiplicados por 100 para llevarlos al rango audible), los cuales, a su vez, son interpretados como códigos MIDI de notas ANSI.

Los acordes estelares se almacenan en la memoria temporal del ordenador local y se accede a ellos mediante el algoritmo de coincidencia que genera acompañamientos armónicos “a tiempo real” para cada nota de entrada. Este algoritmo utiliza la librería *Music 21* (Cuthbert y Ariza, 2010) para comparar la nota de entrada con la nota fundamental de los acordes CaT. Encuentra la nota fundamental entre las alturas de las notas del acorde o descarta el acorde, mostrando el primer espectro coincidente encontrado.

Para hacer el diseño lo más intuitivo posible, no se implementó la posibilidad de reemplazar dinámicamente los espectros asignados a cada tecla del piano, algo que permitiría que las repeticiones de nota generasen acompañamientos diferentes en cada pulsación. No obstante, se trata de una característica adicional interesante que puede ser incluida con facilidad en

Tabla 9.1: Ejemplos de las variables y mapeados utilizados en la sonificación de los espectros de HD168775, M71Q, HD4307, HD57061 y M5 II-53. Valor MIDI de la nota de entrada seleccionada por el usuario, objeto coincidente, vector latente y acorde equivalente obtenido del análisis con *Music 21*, R^2 entre el espectro original y el representado, y número de líneas de transmisión detectadas.

MIDI	Objeto	Vector Latente	Notas	Acorde	R^2	Líneas
45	HD168775	76/36/19/98/23/45	E5/C2/G0/D7/B0/A2	C all combinatorial	0,98	7
66	M71 Q	66/68/28/97/42/45	F#4/G#4/E1/C#7/F#2/A5	minor-ninth chord	0,88	7
72	HD4307	68/51/27/98/32/72	G#4/D#3/D#1/D7/G#1/C5	all-interval tetrachord	0,98	4
82	HD57061	37/80/59/97/40/82	C#2/G#5/B3/C#7/E2/A#5	diminished-major 9 th	0,94	2
96	M5 II-53	56/96/44/84/72/36	G#3/C7/G#2/C6/C5/C2	diminished 4 th 8 ^{va}	0,64	13

futuros desarrollos, lo que permitiría la exploración de la librería estelar completa sin la actual restricción a 88 notas o espectros.

9.5 Diseño e implementación del sintetizador de sonido

Tal y como puede verse en la Figura 9.4, la estructura interna del sintetizador consta de tres generadores de sonido que proporcionan una representación inspirada en el modelo determinístico más estocástico (Serra, 2013). Dos sintetizadores paralelos proporcionan la representación determinista (o sinusoidal) de los espectros, mostrando los acordes de la librería CaT como una mezcla continua de la salida del opcode de síntesis de modulación de frecuencia (FM) *fmb3* y del opcode de oscilador de modelado analógico de banda limitada *vco2* (Vercoe et al., 2007). Sus amplitudes respectivas están controladas por el número de líneas espectrales detectadas en cada espectro, haciendo que *fmb3* prevalezca si el número es bajo y viceversa, y proporcionando una variación de timbre basada en las características reales del espectro.

Por otra parte, un generador de ruido aleatorio representa la precisión de la codificación medida a través del coeficiente R^2 , aportando la componente estocástica de la sonificación.

Además, como se mencionó en la Sección 9.3, el módulo de trémolo incluido permite la representación sonora del número de líneas de transmisión presentes en el espectro, ofreciendo una estimación de su abundancia o ausencia que podría ser útil en exploraciones rápidas de catálogos masivos.

De esta manera, cada sonificación posee un timbre diferente basado en las características de los datos representados y en el error de codificación asociado, que se modula conjuntamente en función del número de eventos de interés detectados (las líneas de absorción).

El sintetizador recibe mensajes MIDI de la aplicación de *Python* para activar los sonidos, y datos OSC para activar las imágenes y controlar las variaciones de timbre de la sonificación. Este enfoque abre una amplia ventana de posibilidades para futuros desarrollos que involucren cualquier tipo de instrumento o controlador digital. El módulo incluye envolventes ADSR para permitir la interacción creativa con las sonificaciones y un procesador de reverberación alimentado en paralelo añadido con fines estéticos.

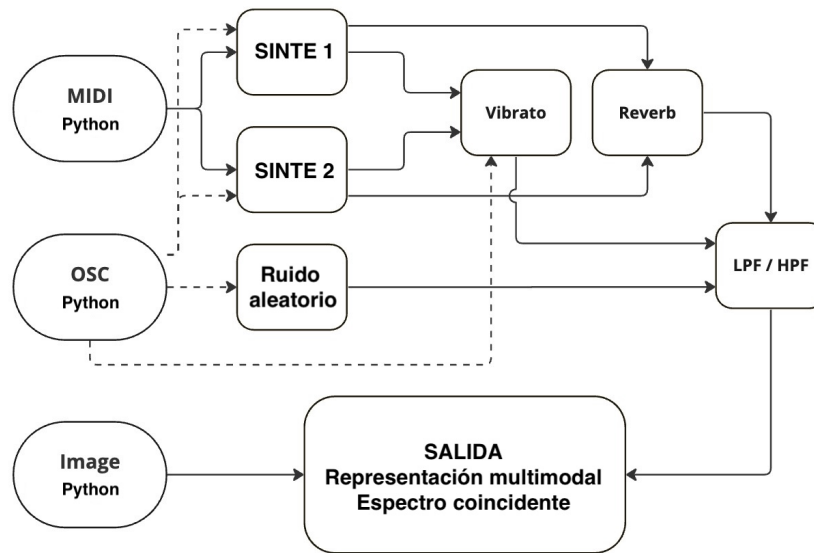


Figura 9.4: Diagrama de bloques del módulo de sonificación implementado en *Csound* y *Cabbage*. Las líneas sólidas representan rutas MIDI, de audio e imagen. Las líneas discontinuas representan mensajes OSC.

El proceso de elección, combinación y mapeado de los módulos de síntesis sonora se basó en criterios estéticos orientados a la generación de nuevas sonoridades dentro del contexto musical de creación de bandas sonoras para películas de ciencia ficción. Este contexto proporciona una metáfora audiovisual que puede considerarse equivalente a la percepción ecológica (Walker y Kramer, 2004) de las sonificaciones de datos astronómicos, ofreciendo asociaciones entre los fenómenos naturales descritos y el imaginario colectivo. En este sentido, los patrones perceptuales introducidos en la cultura audiovisual del género ciencia ficción, pueden servir para construir significado y contexto que ayude a entender la información presentada (Seiça et al., 2023) y a conectar con la audiencia.

9.6 Sistema de acompañamiento para piano a partir de espectros estelares

Esta sección describe el proceso de desarrollo del sistema de representación diseñado para explorar la librería CaT con un teclado de piano. El sistema genera un espectro cuya sonificación está correlacionada armónicamente con la nota de la tecla presionada por el usuario. El diagrama de bloques de la Figura 9.5 proporciona una descripción gráfica del módulo completo. Incluye subprocesos y flujo de datos, desde la entrada de la tecla de piano presionada por el usuario hasta la salida que genera una representación multimodal del espectro correlacionado armónicamente.

El diseño consta de dos notebooks de pre-procesado, una aplicación de generación de eventos “a tiempo real” también construida en *Python*, y una interfaz de usuario multimodal en *CSound/Cabbage* para la implementación del sintetizador. El primer bloque de pre-procesado basado en deep learning entrena un sparse autoencoder con los espectros de la librería CaT,

reduciendo los 790 puntos de cada espectro a un vector latente de seis valores.

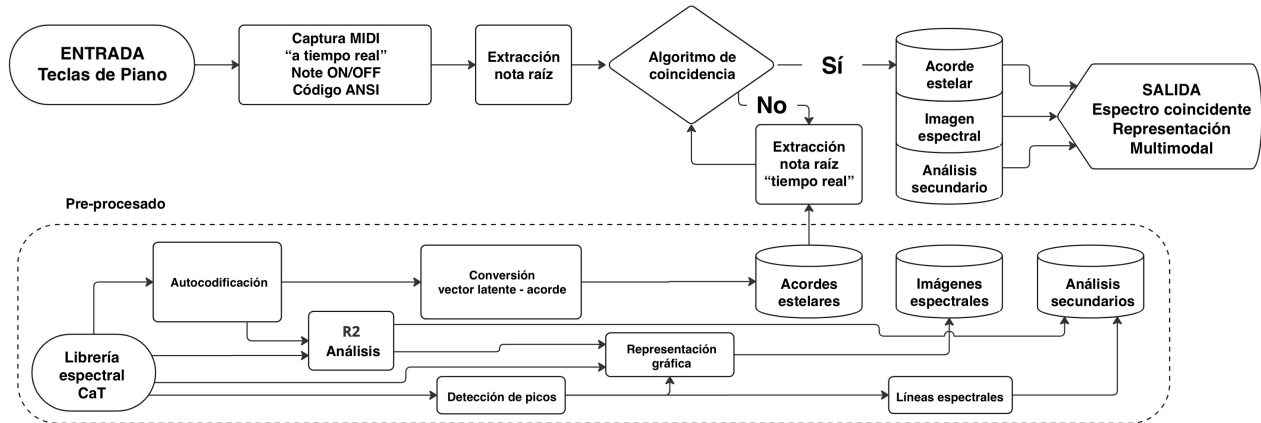


Figura 9.5: Diagrama de bloques del sintetizador incluyendo el pre-procesado. El sistema captura a “tiempo real” la nota MIDI tocada por el usuario, compara la nota con la nota fundamental de los acordes estelares de la base de datos, y proporciona una representación multimodal del espectro con el primer acorde coincidente encontrado en la base de datos pre-procesada.

El segundo bloque de pre-procesado utiliza el autoencoder para convertir los espectros en vectores latentes que se escalan e interpretan como acordes musicales de seis notas. También analiza las líneas de absorción que destacan en cada espectro y genera su representación gráfica. Los resultados del proceso completo se almacenan en una base de datos a la que accede la aplicación de generación de eventos. El tercer notebook captura “a tiempo real” la nota MIDI tocada por el usuario y representa el espectro coincidente, proporcionando un acompañamiento armónico para la nota tocada.

El vídeo enlazado a continuación muestra las posibilidades del prototipo para la generación de composiciones musicales originales y para mostrar en público las representaciones sonoras de espectros estelares.

<https://vimeo.com/913841717>

Las notas de piano que se pueden escuchar en el vídeo se generan con un sampler no implementado en el prototipo, y sólo se utilizan para demostrar la consistencia armónica de la propuesta. El sistema introduce una latencia de 159 ms que no se consideró crítica para este primer enfoque de acompañamiento. La reducción del tiempo de respuesta se puede abordar en futuros desarrollos pre-calculando las notas fundamentales de los acordes estelares y utilizando un script de *Python* en lugar de un *Jupyter Notebook*.

El siguiente vídeo muestra el análisis secundario de las líneas espectrales y la precisión de la sonificación en detalle:

<https://vimeo.com/916886181>

Como puede observarse, la salida de la consola de *Python* muestra la nota tocada, el vector latente extraído del espectro coincidente, el nombre del acorde musical equivalente, el número y las longitudes de onda de las líneas de absorción detectadas, y el resultado del análisis R^2 .

Tanto la precisión como las sonificaciones de la detección de líneas de absorción se pueden silenciar.

El prototipo completo (código abierto) está disponible en:

https://github.com/AuditoryVO/CaT_Sonification_Synthesizer

Capítulo 10

Espectroscopía de Campo Integral (IFS) Multimodal Interactiva

El uso de la sonificación en el análisis de datos científicos proporciona dimensiones adicionales a la visualización, aumentando potencialmente las capacidades analíticas de los investigadores y fomentando la inclusión y la accesibilidad. Este capítulo explora el potencial de la sonificación aplicada a la espectroscopía de campo integral (IFS) a través del diseño e implementación de un módulo de sonificación basado en deep learning llamado *SoniCube*. El módulo se ha integrado dentro de *ViewCube*, la aplicación de análisis IFS desarrollada por Rubén García Benito (IAA-CSIC). La aplicación final genera visualizaciones interactivas y sonificaciones binaurales de la información espectral contenida en cubos de datos, proporcionando una herramienta que incorpora el sonido inmersivo a la espectroscopía 3D. El capítulo incluye una prueba de concepto utilizando los cubos de datos de las galaxias del estudio Calar Alto Integral Field Area (CALIFA) (García Riber et al., 2024), cuya evaluación puede encontrarse en la Sección 11.3.

10.1 Sobre CALIFA

CALIFA (S. F. Sánchez et al., 2016; S. Sánchez et al., 2012; Walcher et al., 2014) es un estudio de más de 600 galaxias con un eje mayor isofotal de banda r entre 45" y 79,2" y un desplazamiento al rojo $0,005 < z < 0,03$, seleccionadas del catálogo fotométrico Sloan Digital Sky Survey (SDSS) DR7. Con el objetivo de ayudar en el estudio de la evolución de las galaxias en el universo local a través del tiempo cósmico, utiliza la espectroscopía de campo integral (Allington-Smith, 2006) para proporcionar un estudio de galaxias IFU (Integral Field Units) de campo amplio que incluye todos los tipos morfológicos, cubriendo masas entre $10^{8,5}$ y $10^{11,5} M_{\odot}$ (S. F. Sánchez et al., 2016).

Las observaciones se obtuvieron con el espectrógrafo de campo integral PMAS/PPak montado en el telescopio de 3,5 m del observatorio de Calar Alto. El rango de longitud de onda entre 3700 y 7500 Å se muestrea utilizando dos configuraciones espectrales diferentes, un modo V500 de baja resolución (3745–7500 Å) con una resolución espectral de 6,0 Å (ancho completo

a la mitad del máximo, FWHM), y un modo V1200 de resolución media (3650–4840 Å) con una resolución espectral de 2,3 Å (García-Benito et al., 2015). La tercera versión de datos de CALIFA (S. F. Sánchez et al., 2016) proporciona 646 objetos en la configuración V500, 484 en la V1200, y la combinación de los cubos de ambas configuraciones (COMBO). Las referencias de morfología proporcionadas en este capítulo fueron extraídas de Walcher et al., 2014.

10.2 Autocodificación de CALIFA

Con la intención de explorar las posibilidades de los autoencoders en la reducción de espectros de galaxias a vectores de baja dimensión, para generar sonificaciones precisas de la información espectral de cubos de datos, esta sección muestra los resultados que hemos obtenido mediante estrategias equivalentes a las descritas con espectros estelares en el Capítulo 7. Basados en el algoritmo de descenso de gradiente, los autoencoders permiten la reducción de la dimensionalidad de los datos mejor que otros enfoques, tales como el análisis de componentes principales (Hinton y Salakhutdinov, 2006). Tal y como se describe en el Capítulo 7 y en el Anexo.8, los autoencoders son modelos de redes neuronales capaces de aprender una aproximación a la función identidad, proporcionando una salida similar a su entrada (Ng et al., 2011). Mediante la reducción de la dimensión de la capa intermedia de la red, el espacio latente, un modelo puede aprender estructuras relevantes de los datos, que también se pueden reconstruir a partir de esta representación intermedia de menor dimensión (Goodfellow et al., 2016).

La dimensión del espacio latente depende del conjunto de datos y la arquitectura utilizada. Con el objetivo de obtener parámetros estelares utilizando redes neuronales convolucionales, Mas-Buitrago et al., 2024 propusieron un autoencoder con espacio latente de 32 dimensiones, mostrado en una matriz de 8x4, para la reducción de espectros CARMENES y espectros sintéticos ACES modelados con “PHOENIX”, 1990. Por otro lado, la reconstrucción a partir de vectores latentes de cuatro dimensiones fue suficiente para que Xiang et al., 2022 analizaran la actividad magnética estelar utilizando autoencoders variacionales en los espectros K2 del Large Sky Multi-Object Fiber Spectroscopic Telescope (LAMOST). Demostrando el potencial de los autoencoders variacionales, Portillo et al., 2020 resumieron la información espectral de las galaxias con sólo seis variables latentes en el Sloan Digital Sky Survey (SDSS). Esta dimensión también funcionó eficazmente para la librería Calcium II Triplet (CaT) codificada mediante sparse autoencoders en el Capítulo 9 de este trabajo.

Diversas pruebas intensivas con diferentes configuraciones de ambas arquitecturas fueron realizadas utilizando los cubos de datos COMBO (V500+V1200) del CALIFA DR3. La Figura 10.1 proporciona un ejemplo comparativo para el cubo de datos de la galaxia espiral (Scd) NGC 5406 con un autoencoder de seis capas y seis dimensiones, y un VAE de cuatro capas y seis dimensiones, ambos implementados utilizando TensorFlow (Abadi et al., 2016).

Para cada cubo de datos codificado, se calculó el coeficiente de determinación (R^2) entre los conjuntos de espectros originales y los reconstruidos, proporcionando una medida de la precisión de la reducción. En cuanto a la duración de los procesos de entrenamiento, el VAE requirió 5,5 veces más tiempo de computación por epoch que el sparse autoencoder, proporcionando finalmente resultados inferiores. Respectivamente, $R^2 = 0,96$ (VAE) vs $R^2 =$

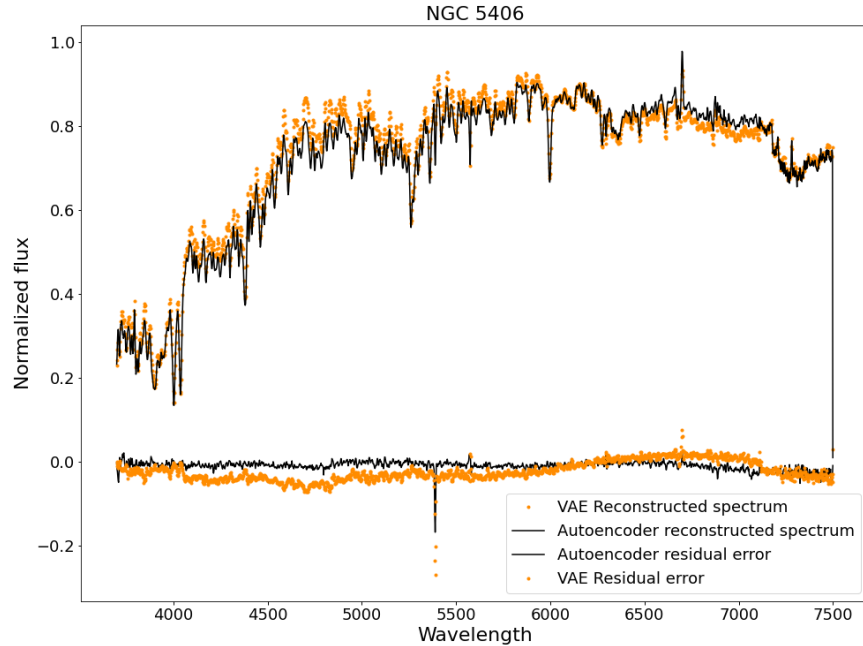


Figura 10.1: Comparativa de autoencoders para la galaxia espiral (Scd) NGC 5406. Autoencoder de seis capas y seis dimensiones (negro, continuo) vs VAE de cuatro capas y seis dimensiones (naranja, punteado). Espectros reconstruidos y error residual del espectro original para el spaxel (34,34). Sparse autoencoder: $R^2 = 0,99$ (espectro), $R^2 = 0,98$ (cubo de datos), 39,12% de los espectros con $R^2 > 0,9$, 100 epochs, una hora por cubo. VAE: $R^2 = 0,97$ (espectro), $R^2 = 0,98$ (cubo de datos), 4,92% de los espectros con $R^2 > 0,9$, 291 epochs, 5h 30' por cubo. Flujo normalizado (ADU) vs longitud de onda (Å).

0,98 (autoencoder) para el cubo de datos completo, con 4,92% (VAE) vs 39,12% (sparse) de los espectros con $R^2 > 0,9$, y $R^2 = 0,96$ (VAE) vs $R^2 = 0,99$ (autoencoder) para el spaxel representado (34,34) en la Figura 10.1.

Basado en estas pruebas, incluimos un módulo de autoencoder de seis capas y seis dimensiones en *SoniCube* para representar “a tiempo real” la información espectral de los cubos de datos con vectores de baja dimensión. Esta arquitectura permitió la reducción de cada espectro de entrada $X_i \in \mathbb{R}^{1901}$ (cada cubo de datos contiene alrededor de 5540 espectros con 1901 valores de flujo por espectro), a una representación de seis dimensiones $Z_i \in \mathbb{R}^6$, y la reconstrucción de X_i a partir del vector latente Z_i , $\hat{X}_i \in \mathbb{R}^{1901}$.

El modelo fue entrenado en cada cubo de datos de forma independiente. La Figura 10.2 proporciona dos ejemplos de los espectros originales y reconstruidos de los cubos de datos de NGC 5784 (Sbc) y NGC 5682 (E4). Los archivos codificados para 446 galaxias de los cubos de datos COMBO DR3 del estudio CALIFA, pueden descargarse desde:

<https://zenodo.org/records/10570065>

La Figura 10.3 muestra tres ejemplos de las curvas de aprendizaje obtenidas durante el proceso de entrenamiento con los espectros de los cubos de datos de NGC7047 (Sab), UGC10331 (E1)

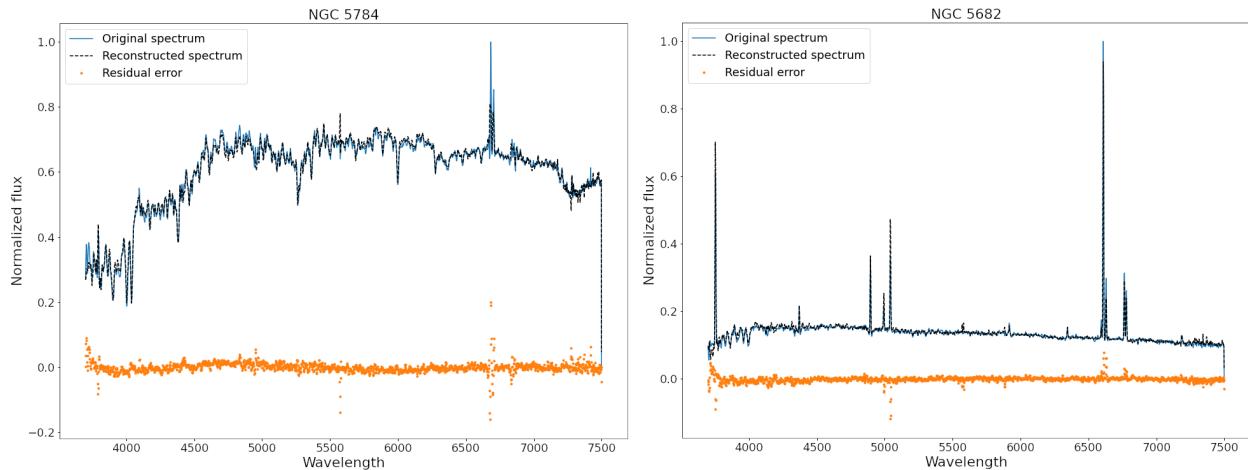


Figura 10.2: Resultados del autoencoder de seis capas y seis dimensiones. Espectro reconstruido (línea discontinua negra) y original (línea sólida azul) con error residual (línea punteada naranja) del spaxel (35,35) de la galaxia espiral (Sbc) NGC 5784 (izquierda) y la galaxia elíptica (E4) NGC 5682 (derecha). Dos ejemplos de una galaxia antigua y una región de formación estelar del estudio CALIFA. Respectivamente, $R^2 = 0,98$ y $R^2 = 0,95$. Flujo normalizado (ADU) en función de la longitud de onda (\AA).

y UGC03960 (E5). Estas galaxias ilustran el rendimiento del autoencoder, correspondiendo respectivamente al mejor resultado de codificación, a un resultado intermedio y al peor. Los coeficientes obtenidos oscilaron entre 0,998 y 0,882 a lo largo del conjunto de datos completo, con un 49 % de los cubos de datos presentando un R^2 superior a 0,96, y un 4,78 % presentando un R^2 inferior a 0,92.

El siguiente vídeo muestra un prototipo preliminar desarrollado para explorar secuencialmente los espectros centrales de las galaxias de CALIFA, proporcionando el coeficiente de determinación de cada codificación, los valores de cada vector latente escalado al margen audible y la ubicación de cada galaxia sobre el mapa de cobertura.

<https://vimeo.com/843617421>

10.3 *ViewCube* descripción general

ViewCube es una aplicación de software abierto escrita en *Python* por Rubén García Benito (IAA-CSIC). Esta herramienta fue desarrollada originalmente para la evaluación rápida de la calidad y las características físicas de los cubos de datos de CALIFA, y para la exploración rápida de productos de datos de alto nivel generados por el pipeline PyCASSO (de Amorim et al., 2017). *ViewCube* dispone de un lector de archivos FITS general y flexible y admite cubos de datos de cualquier procedencia. La aplicación también permite renderizar formatos Raw Stacked Spectra (RSS), siempre que esté disponible un archivo que mapee las posiciones de las fibras.

El objetivo principal de *ViewCube* es facilitar una inspección rápida y eficaz de los cubos de datos, ya sea para una evaluación de la calidad o para un examen centrado de sus

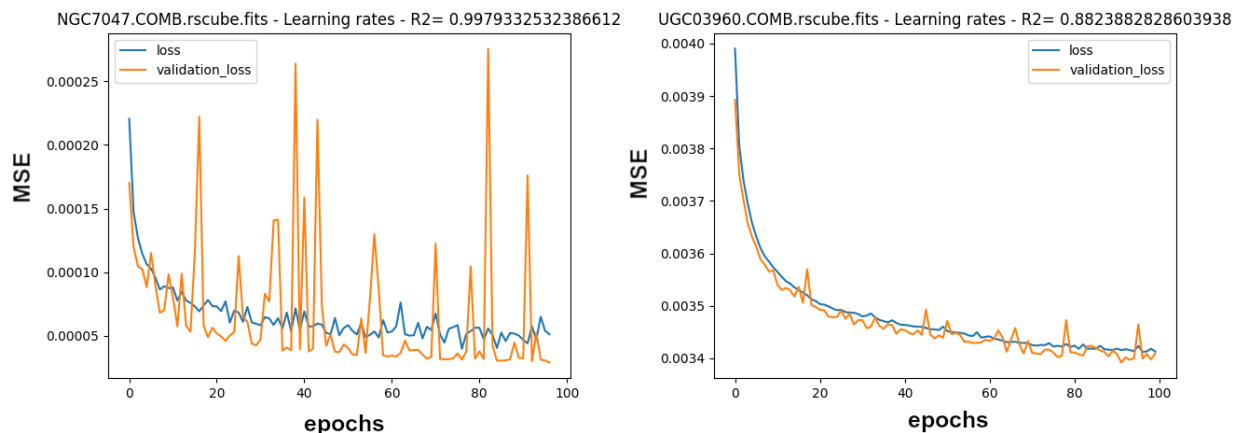


Figura 10.3: Curvas de aprendizaje que muestran el error cuadrático medio (MSE) por epoch durante los procesos de entrenamiento y validación para la galaxia espiral (SAb) NGC7047 y la galaxia elíptica (E5) UGC03960. Estos cubos de datos corresponden, respectivamente, al mejor y al peor resultado proporcionado por el autoencoder. Nótese el cambio de escala en el eje de abscisas de las representaciones.

características. Respecto a los detalles técnicos, todas las visualizaciones que proporciona la aplicación se renderizan utilizando el módulo *matplotlib*.

La Figura 10.4 muestra la interfaz de usuario que dispone de dos ventanas principales: una ventana de imagen, que presenta un mapa 2D del cubo de datos convolucionado a través de una fibra o banda de paso elegida, y una ventana espectral, que muestra el espectro correspondiente a la ubicación seleccionada por el usuario con el cursor. La aplicación permite seleccionar diferentes spaxels o fibras para su comparación, generar un espectro integrado y guardar tanto los espectros individuales como los integrados. Además, los usuarios pueden modificar el filtro utilizado en la convolución del cubo de datos para generar la ventana de imagen, y ajustar la longitud de onda central del filtro arrastrando y soltando la banda de paso del filtro de la ventana espectral.

10.4 Implementación de *SoniCube*

Esta sección describe el módulo de sonido que hemos implementado dentro de *ViewCube* para permitir la sonificación de los espectros asociados con cada elemento espacial de un cubo de datos. El módulo, llamado *SoniCube*, complementa la representación gráfica de la aplicación proporcionando una herramienta completa, multimodal, abierta, y de propósito general para el análisis de IFS. El objetivo final de la interfaz *SoniCube* es permitir la sonificación de cubos de datos mediante diferentes metodologías, adaptadas a las necesidades de análisis específicas de los casos de estudio o a las preferencias de los usuarios. Esta sección presenta el primer método de sonificación implementado, que proporciona una representación general basada en deep learning del cubo de datos.

SoniCube está implementado en *Python* y controla, a través de OSC un sintetizador de sonido desarrollado en *CSound*, mediante el módulo nativo python-osc y la interfaz *ctcsound*

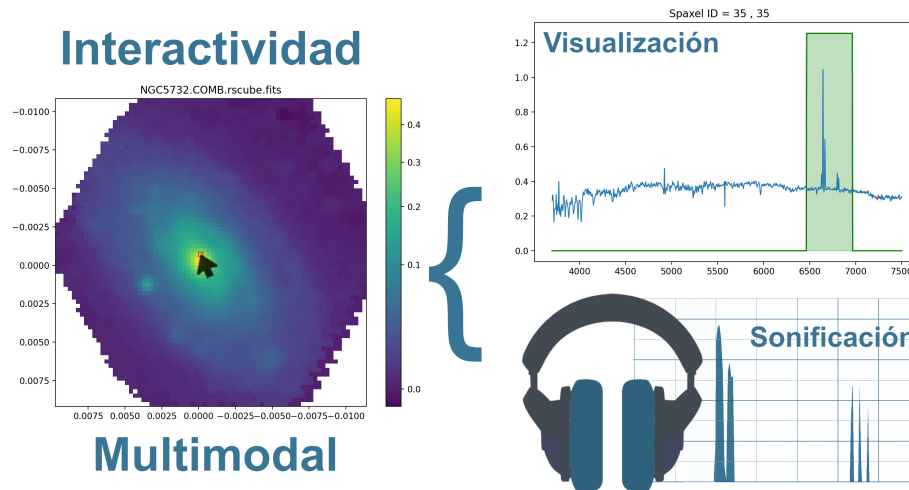


Figura 10.4: Interfaz de usuario de *ViewCube* mostrando el cubo de datos de la galaxia espiral (Sbc) NGC 5732. Ventana de imagen 2D (izquierda) y representación multimodal (ventana espectral arriba y espectro sonoro de la sonificación abajo) para el spaxel (35,35) (derecha).

(Ctcsound, 2022). El módulo proporciona una sonificación interactiva “a tiempo real” del espectro asociado a cada spaxel seleccionado por el usuario.

En esta primera implementación de sonificación, cada espectro se convierte en sonido utilizando el enfoque basado en deep learning descrito en la Sección 10.2. Este proceso proporciona una representación sonora única no supervisada que convierte la información de cada espectro en un solo evento sonoro. Cada evento se genera con un sintetizador aditivo utilizando seis osciladores independientes, alimentados por un vector latente de seis dimensiones que es generado por un autoencoder (Baldi, 2012). El Anexo.21 ofrece una muestra del código fuente desarrollado para ser incluido en *ViewCube*.

El módulo genera un vector latente de seis dimensiones “a tiempo real” a partir de cada espectro seleccionado por el usuario. Los componentes de este vector se interpretan como frecuencias fundamentales para los seis osciladores que sintetizan el sonido. Las seis componentes se multiplican por 10.000 para escalar los valores latentes al rango de frecuencias audibles, generando sonificaciones comprensibles y precisas. Para una descripción formalizada del sintetizador, consultar el Anexo.11.

Además, el módulo calcula el azimut y la distancia radial desde los spaxels seleccionados por el usuario hasta el spaxel de referencia, que se hizo corresponder con el centro de la galaxia en cada cubo de datos. El azimut se utiliza para localizar la representación sonora del espectro dentro del paisaje sonoro binaural (Møller, 1992) generado para cada cubo de datos, proporcionando una representación inmersiva de sus espectros con el oyente situado en el centro de la galaxia. Para obtener más información sobre la codificación binaural, consultar el Anexo.12.

La distancia del spaxel seleccionado al centro de la galaxia se utiliza para controlar la relación

energía directa - energía reverberante (parámetro dry/wet) de un emulador de reverberación (Gardner, 1998), proporcionando la sensación cognitiva asociada al campo sonoro que puede encontrarse en recintos de grandes dimensiones. Este efecto genera de manera virtual la percepción de distancia, como función de la mencionada relación entre energía directa y reverberante (Lu y Cooke, 2010).

La amplitud de cada sonificación se calcula a partir de los flujos absolutos del espectro representado. Para resolver el amplio y variable rango dinámico de densidad de flujo que se encuentra comúnmente en observaciones reales, *SoniCube* incluye dos modos de funcionamiento para representar amplitudes de sonido relativas. En el modo sensible al flujo (predeterminado), la amplitud de cada sonificación se calcula a partir de la mediana logarítmica del flujo absoluto, normalizada a todo el conjunto. De esta forma, se preserva la relación aparente de flujos dentro del cubo de datos.

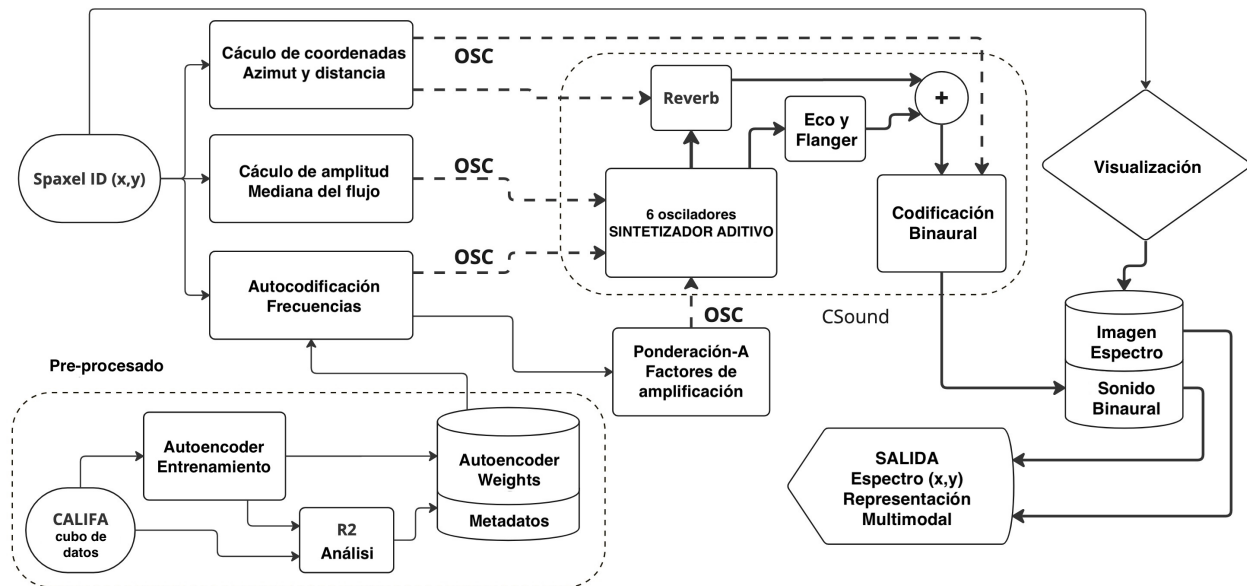


Figura 10.5: Diagrama de bloques de *SoniCube*. Pre-procesado y cálculos “a tiempo real”, incluyendo flujos de datos, OSC y señales de audio.

Por otro lado, si el modo sensible está desactivado, todos los espectros del cubo de datos se representan con la misma amplitud, lo que permite la apreciación de regiones con flujos absolutos relativamente bajos. Adicionalmente, en ambos modos de funcionamiento se implementó un limitador/compresor de rango dinámico de banda ancha de dos etapas (Kates, 2005), que mantiene controlados los valores extremadamente salientes, evitando posibles daños auditivos al analizar cubos de datos no explorados con anterioridad.

Para compensar la respuesta no lineal del oído humano en los dominios de frecuencia y amplitud, se aplicó una curva ponderada A a la matriz de amplitudes utilizando el módulo *librosa* (McFee et al., 2015). La función `A_weighting` incluida en este módulo calcula los factores de amplificación normalizados que proporcionan la respuesta de igual sonoridad según la curva de ponderación A, para cada una de las representaciones. Los resultados obtenidos con esta función fueron satisfactorios para la implementación de este primer módulo

de sonificación evaluado con el estudio CALIFA. Sin embargo, se propone la exploración de contornos de sonoridad alternativos para futuras implementaciones en las que las limitaciones de esta ponderación, descritas en Charbonneau et al., 2012, puedan ser críticas.

Utilizando el protocolo OSC, el módulo envía a *CSound* todas las variables necesarias para sintetizar la representación sonora en un paisaje binaural. El diagrama de bloques de la Figura 10.5 resume el proceso de generación de sonido. Merece la pena observar el uso del azimut y la distancia del spaxel seleccionado en los bloques binaural y de reverberación, así como la mediana de flujo normalizado, y las frecuencias con factores ponderados A del espectro correspondiente en el sintetizador aditivo.

Por último, se agregaron algunos aspectos estéticos adicionales basados en efectos de eco y flanger a la sonificación. Este procesado facilita la localización binaural de la representación y suaviza las transiciones rápidas entre espectros. El siguiente vídeo proporciona una demostración de la aplicación completa.

<https://vimeo.com/1005208084>

La aplicación multimodal completa *ViewCube* está disponible como código abierto en el GitHub de Rubén García Benito:

<https://github.com/rgbIAA/viewcube>

Capítulo 11

Evaluación

Este capítulo ofrece una descripción detallada de tres estudios cuantitativos y cualitativos que hemos realizado para evaluar las ventajas y desventajas asociadas a la sonificación de datos astronómicos. Utilizando los prototipos descritos en capítulos anteriores, e incluyendo casos de estudio específicos basados en objetos reales, el análisis proporciona una visión general de las implicaciones, desafíos y potenciales asociados a la implementación práctica de las propuestas. Se espera que los resultados sean útiles, tanto en actividades de divulgación como en análisis científico, como marco exploratorio para futuros desarrollos hacia una propuesta de observatorio virtual sonoro.

Los tres estudios incluyen una evaluación cuantitativa y cualitativa realizada por participantes especializados y no especializados en Música y Astronomía. Para el análisis estadístico se utilizó la librería de Python *scipy* (“SciPy”, 2001). El primer estudio se centra en la identificación de estrategias de sonificación eficaces para revelar el número de líneas de transmisión superpuestas en representaciones espectroscópicas de galaxias, y en la identificación de las preferencias potenciales de la audiencia para futuros diseños de representaciones sonoras ampliamente aceptadas. El segundo estudio explora el potencial de la sonificación simbólica y de la sonificación basada en deep learning para la representación de espectros estelares. El análisis de resultados incluye una comparativa de ambos enfoques para justificar su uso como complemento a la representación gráfica de datos espectrales. El tercer estudio evalúa la implementación del modelo de sonificación mediante deep learning dentro de *ViewCube* con la intención de valorar su potencial para el análisis multimodal interactivo de cubos de datos de galaxias. El análisis estadístico completo de los tres estudios puede consultarse en:

<https://github.com/AuditoryVO/Statistics.git>

11.1 Primer estudio. Análisis cuantitativo de sonificaciones de espectroscopía de transmisión y análisis cualitativo de musicaciones

Con el objetivo de proporcionar una evaluación cuantitativa y cualitativa de los diseños descritos en capítulos anteriores, se distribuyó un enlace con un cuestionario online a posibles participantes voluntarios entre los contactos personales y profesionales de los autores, desde el 7 de diciembre de 2022 hasta el 25 de enero de 2023. Se aconsejó a los participantes que utilizaran auriculares, sin imponer un límite de tiempo de respuesta. El cuestionario, disponible en el siguiente enlace, incluía 25 preguntas separadas en dos categorías y utilizaba vídeos para proporcionar las representaciones.

<https://forms.office.com/e/iCFyXfwsdT>

11.1.1 Diseño del estudio I

La primera parte de la encuesta estuvo formada por 8 preguntas cuantitativas destinadas a identificar la estrategia de sonificación más eficaz para revelar el número de líneas de emisión y absorción superpuestas en espectros de galaxias, utilizando el Large Early Galaxy Census (LEGA-C) (Van der Wel et al., 2021). El prototipo de sonificación descrito en el Anexo.6 se implementó *ad hoc* para la evaluación de esta tarea. También se introdujeron dos preguntas adicionales con ejemplos de la librería estelar STELIB, sonificadas con el prototipo *FITS2OSC* descrito en la Sección.3, para analizar su utilidad en el reconocimiento de líneas de absorción en espectros estelares.

La segunda categoría de preguntas tuvo como objetivo obtener una evaluación cualitativa de las diferentes estrategias de sonificación y musicación propuestas en los capítulos anteriores. Con la intención de hacer accesible la encuesta a los participantes con menos experiencia, no se establecieron criterios especiales para estas preguntas, tratando de capturar las primeras impresiones de los participantes, que podrían ser útiles en el diseño de representaciones sonoras intuitivas ampliamente aceptadas para aplicaciones de divulgación. El cuestionario también incluía preguntas categóricas adicionales relacionadas con las limitaciones visuales y auditivas, así como con el nivel de experiencia de los participantes en Astronomía y Música. Todas las preguntas se presentaron en orden aleatorio.

Se obtuvieron un total de 46 respuestas de participantes con edades comprendidas entre los 18 y los 65 años, la mayoría de ellos residentes en España, con algunas excepciones de Australia y EE.UU. Los participantes se clasificaron por grupos según su experiencia previa en Astronomía y Música, entendida como su familiaridad con la naturaleza de los datos astronómicos o con su capacidad para prestar atención a los eventos sonoros. La muestra también incluyó dos participantes ciegos o con baja visión (BLV), cuyo análisis se realizó en paralelo con una muestra aleatoria de participantes no BLV para permitir el estudio comparativo de sus resultados. De este modo, se analizaron las respuestas de dos grupos de 15 participantes, con y sin experiencia previa en Música y Astronomía. Además, el grupo de participantes con experiencia se dividió en subgrupos para analizar las respuestas de 3 personas aficionadas o profesionales de la Astronomía, 4 personas aficionadas o profesionales de la Música y 3

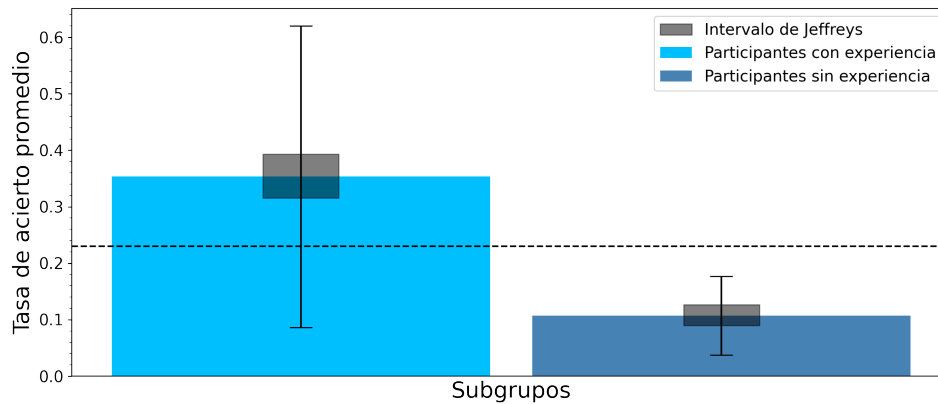


Figura 11.1: Resultados del reconocimiento sonoro de las líneas de absorción y emisión en galaxias del Large Early Galaxy Census (STScI). Tasa de acierto promedio, desviación estándar e intervalos de confianza de Jeffreys para los participantes con experiencia (0,35) y sin experiencia (0,11). La línea discontinua indica la tasa promedio de acierto aleatorio (0,23).

personas pertenecientes a ambos grupos, permitiendo la comparación con los resultados de un subgrupo aleatorio formado por 4 participantes sin experiencia. Las tendencias presentadas a continuación podrían ser útiles como guía preliminar para futuros desarrollos.

11.1.2 Resultados cuantitativos. Reconocimiento sonoro de líneas de transmisión

Esta sección describe el análisis de las preguntas relacionadas con la identificación sonora de líneas de transmisión en espectros estelares STELIB y de galaxias LEGA-C.

En primer lugar se realizaron los recuentos generales con el método del rango intercuartil (IQR) para detectar potenciales valores atípicos sin resultados destacables (Dekking, 2005). A continuación se realizaron los subgrupos, dividiendo la muestra en participantes con y sin experiencia. Tanto los grupos como los subgrupos requirieron la extracción aleatoria de muestras para garantizar el equilibrio de participantes y permitir las comparaciones entre grupos. Una vez equilibrados, se realizó el test de normalidad Shapiro-Wilk (Shapiro y Wilk, 1965) sobre las distribuciones obtenidas. Con ello se comprobó la normalidad de la muestra completa y de los subgrupos formados por 15 participantes con experiencia, 15 sin experiencia, 3 astrónomos músicos y el subgrupo muestreado aleatoriamente con fines comparativos de 4 participantes sin experiencia.

Analizando la muestra completa se obtuvo una tasa promedio general de acierto de 0,26, desviación estándar 0,17, con un intervalo de confianza de *Jeffreys* (Jeffreys, 1946) de (0,24, 0,28) y un 68,3 % de significación. El resultado fue comparado con la tasa promedio de acierto aleatorio (0,23) mediante la prueba t-test (Student, 1908), confirmando que la diferencia del promedio observado con la probabilidad de acierto aleatorio era estadísticamente significativa,

$t(43) = 4,78, p < 0,001$.

El análisis de acierto por participante de las diez preguntas, ofreció una media de 3,89, mediana 3,5 y valor más frecuente 3 (30 % de acierto), obtenido 10 veces (22,7 % de los participantes).

La Figura 11.1, muestra las tasas de acierto obtenidas por los participantes con y sin experiencia. Los participantes con experiencia obtuvieron una tasa promedio de acierto de 0,35, desviación estándar 0,27, con intervalo de Jeffreys (0,31, 0,39) y un 68,3 % de significación. La diferencia con la tasa promedio de acierto aleatorio resultó estadísticamente significativa, $t(14) = 4,01, p = 0,003$. Los participantes sin experiencia obtuvieron una tasa promedio de acierto de 0,1, desviación estándar 0,07, con intervalo de Jeffreys (0,09, 0,13) y un 68,3 % de significación. La diferencia con la tasa promedio de acierto aleatorio resultó estadísticamente significativa, $t(14) = 4,01, p = 0,003$.

Respecto al análisis de acierto por participante, el grupo con experiencia obtuvo una media de 4,8, mediana 4,0 y valor más frecuente 7 (70 % de acierto), obtenido 4 veces (26,6 % de los participantes). El grupo sin experiencia obtuvo una media de 3,26, mediana 3 y valor más frecuente 3 (30 % de acierto), obtenido también 4 veces (26,6 % de los participantes).

Para evaluar si existían diferencias significativas entre los promedios de los grupos de participantes con y sin experiencia se utilizó la prueba ANOVA unidireccional (Moore et al., 2009), $F(1, 28) = 0,241, p = 0,14$. Este resultado confirmó que las diferencias observadas no eran estadísticamente significativas y deben considerarse indicativas.

Las bajas tasas de acierto registradas sugieren que la tarea resultó complicada para todos los participantes, a pesar de que los resultados obtenidos por los participantes con experiencia son 3,3 veces superiores a los obtenidos por los participantes sin experiencia. Esta tendencia observada concuerda con las conclusiones reportadas por Walker y Nees, 2011, Tucker Brown et al., 2022 y Harrison et al., 2022, señalando la importancia del entrenamiento y conocimiento previo del campo de estudio para conseguir una integración útil de la sonificación en este tipo de análisis científicos altamente especializados. El diseño de este bloque de preguntas partió de la hipótesis de que incluso los participantes con menos experiencia serían capaces de contar eventos sonoros, algo que los resultados desmintieron, confirmando que incluso el recuento de eventos sonoros sencillos requiere entrenamiento y concentración.

Atendiendo a los resultados por subgrupos, tal y como puede verse de la Tabla 11.1, los astrónomos músicos obtuvieron una tasa de acierto promedio de 0,46, los astrónomos de 0,37 y los músicos de 0,35. Los astrónomos músicos obtuvieron, por tanto, resultados un 27,3 % mejores que los astrónomos, un 33,3 % mejores que los músicos y un 69,45 % mejores que el grupo aleatorio de 4 participantes sin experiencia. Se utilizó de nuevo el t-test para comprobar que la diferencia de cada promedio respecto a la tasa de acierto aleatorio (0,14) era estadísticamente significativa (Astrónomos músicos $t(2) = 3,82, p = 0,004$, músicos: $t(3) = 2,74, p = 0,022$, astrónomos: $t(2) = 3,14, p = 0,01$ y 4 participantes sin experiencia: $t(3) = 2,77, p = 0,021$).

A continuación, se utilizó el método de Kruskal-Wallis (Kruskal y Wallis, 1952) para comprobar si las diferencias en los promedios de los subgrupos observados eran estadísticamente significativas, puesto que el grupo de astrónomos no presentó una distribución normal. El resultado,

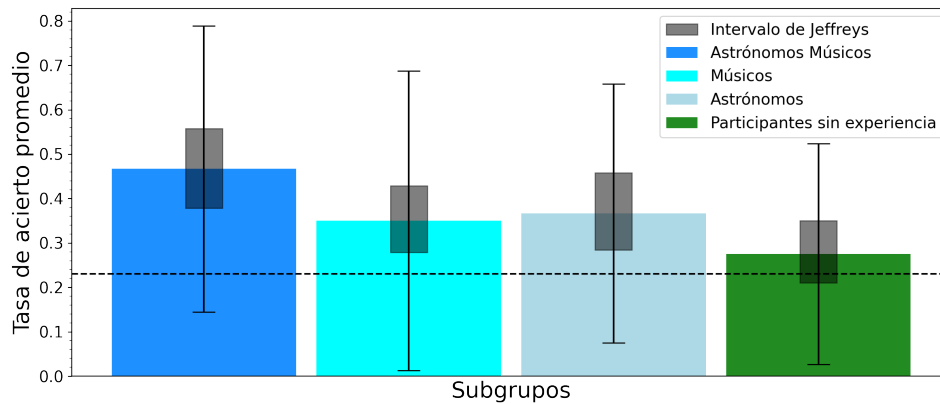


Figura 11.2: Resultados del reconocimiento sonoro de líneas de absorción y emisión en galaxias del Large Early Galaxy Census (STScI). Tasa de acierto promedio, desviación estándar e intervalos de confianza de Jeffreys para las subcategorías extraídas del grupo con experiencia. Astrónomos músicos (0,47), músicos (0,35), astrónomos (0,37) y 4 muestras aleatorias de participantes sin experiencia (0,27). La línea discontinua indica la tasa promedio de acierto aleatorio (0,23).

Tabla 11.1: Resumen de las tasas de acierto por grupos de participantes para el reconocimiento sonoro de líneas de transmisión. Desviación estándar e intervalos de confianza.

Grupos	Participantes	Tasas de acierto	Conf-bajo	Conf-alto	Std
Completo	46	0,2608	0,2409	0,2818	0,1690
Global	44	0,2636	0,2431	0,2851	0,1708
Con experiencia	15	0,3533	0,3154	0,3932	0,2667
Sin experiencia	15	0,1069	0,0901	0,1263	0,0698
AstroMus	3	0,4667	0,3782	0,5573	0,3220
Músicos	4	0,3500	0,2791	0,4282	0,3374
Astrónomos	3	0,3667	0,2844	0,4576	0,2919
Sin experiencia	4	0,275	0,2106	0,3504	0,2486
BLV	2	0,2	0,1264	0,3027	0,2486
No BLV	2	0,3	0,2094	0,4099	0,2581

$H(3) = 8,78$, $p = 0,83$, mostró que las diferencias en los promedios no eran estadísticamente significativas y deben considerarse indicativas.

Respecto al análisis por pregunta, las tasas de acierto extremadamente bajas obtenidas en la pregunta 6, que presentaba un único evento sonoro a modo de acorde generado con osciladores sinusoidales (enlace al cuestionario: <https://forms.office.com/e/iCFyXfwsdT>), sugieren que las representaciones secuenciales podrían ser mucho más útiles que los enfoques basados en acordes para el reconocimiento sonoro de múltiples líneas de transmisión en espectros de galaxias. Las tasas de acierto más altas se obtuvieron en la pregunta 7, que utilizaba un enfoque secuencial, sugiriendo que las representaciones mediante osciladores sinusoidales

mostradas secuencialmente son potencialmente más eficaces que el mapeado de instrumentos musicales para esta tarea, en la que las líneas se presentaban superpuestas en la representación gráfica.

Los participantes BLV obtuvieron una tasa promedio de acierto de 0,2, desviación estándar 0,35. La diferencia del promedio observado con el valor de la tasa promedio de acierto aleatorio no resultó estadísticamente significativa, $t(1) = 0,77$, $p = 0,46$.

11.1.3 Resultados cualitativos. Musificación

En la Figura 11.3 se presenta una síntesis de la evaluación cualitativa registrada en el cuestionario online. Las respuestas a las diez preguntas sobre estrategias de desarrollo y composiciones musicales no supervisadas se clasificaron en cuatro grupos según el nivel de experiencia de los participantes (subgrupos sin equilibrar). Cabe mencionar que las respuestas de los participantes con experiencia en Astronomía presentaron una polarización más suave que el resto de los subgrupos, lo que podría estar relacionado con las expectativas estéticas de los participantes con experiencia en Música y de los participantes sin experiencia.

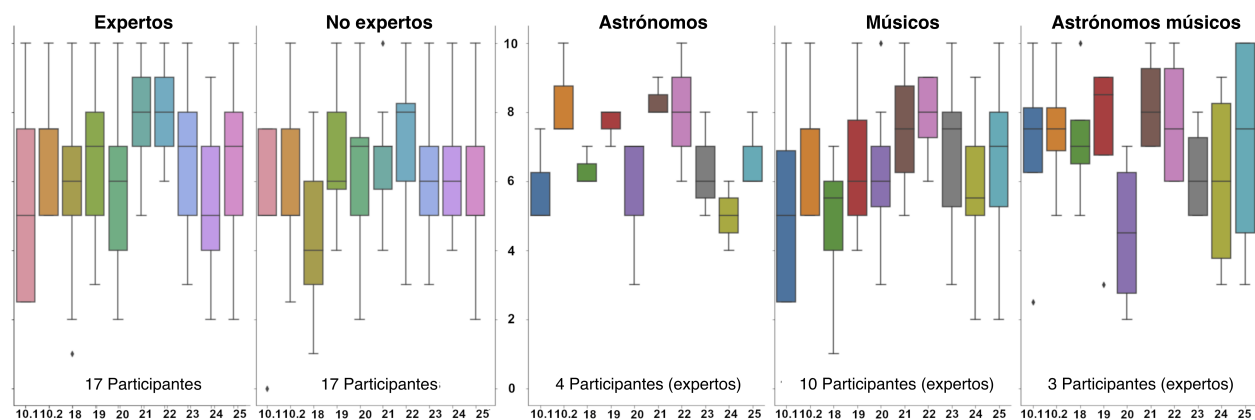


Figura 11.3: Análisis cualitativo por grupos de experiencia en Música y Astronomía. Valores pico, valores atípicos, promedio y barras de error para las respuestas registradas de los grupos con y sin experiencia. También se incluye una subclasificación adicional no equilibrada del grupo de participantes con experiencia para mostrar las preferencias de astrónomos profesionales y aficionados, músicos profesionales y aficionados, y astrónomos músicos profesionales y aficionados. Las preguntas referenciadas en el eje x están disponibles en: <https://forms.office.com/e/iCFyXfwsdT>

En la pregunta número 10, que proporciona una evaluación del prototipo de autoencoder para la representación de espectros estelares, las personas con experiencia en Música y Astronomía encontraron la sonificación mediante ondas sinusoidales más interesante que la representación de piano basada en acordes de la secuencia de tipos espectrales. Esto podría sugerir, en el caso de las personas con experiencia en Astronomía, una clara preferencia por las representaciones científicas abstractas, mientras que en el caso de las personas con experiencia en Música, podría explicarse por las expectativas asociadas al instrumento. Esta cuestión puede representar una línea adicional de investigación para futuros estudios.

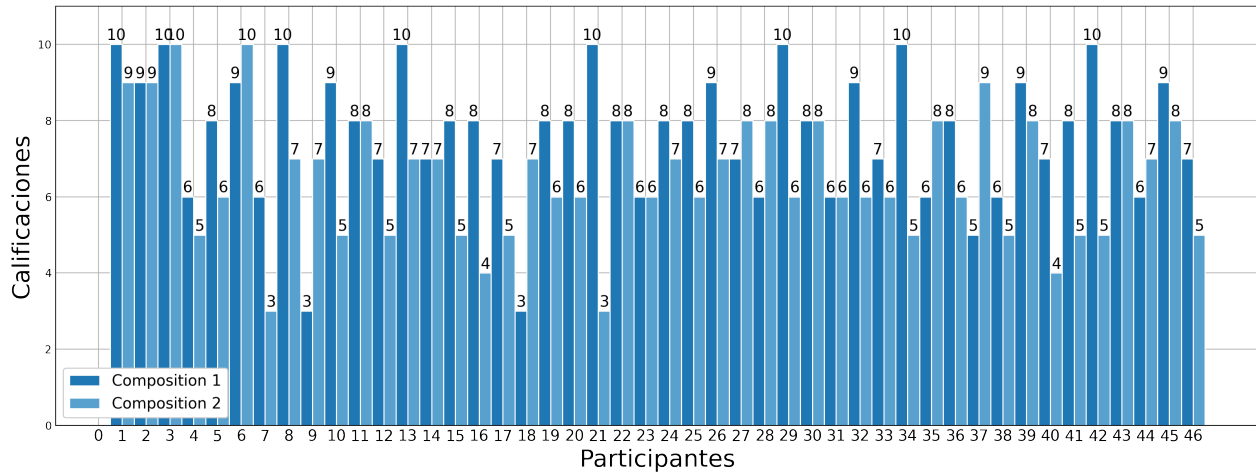


Figura 11.4: Evaluación del público para dos composiciones generadas por el sistema de composición autónomo descrito en el Capítulo 8. La red fue entrenada usando los trabajos corales de Giovanni Pierluigi da Palestrina. Se pidió a los participantes que calificaran las composiciones de 0 a 10.

En el resto del cuestionario, se observó una tendencia generalizada a proporcionar calificaciones más altas a las formas de musicales clásicas frente a representaciones sonoras abstractas. Esta tendencia se puede apreciar con la comparación de los resultados de las preguntas 18, 19 y 20 y en la evaluación subjetiva de las composiciones musicales no supervisadas de las preguntas 21 y 22. El resultado promedio para todas las preguntas cualitativas es de 6,35 y la mediana es de 6,0, lo que sugiere una buena aceptación de las propuestas.

Las siguientes piezas fueron calificadas de 0 a 10 por los 46 participantes, proporcionando una estimación cualitativa de la aceptación del sistema de composición autónomo descrito en el Capítulo 8. La composición 1 obtuvo una calificación media de 7,71, mediana 8 y resultado más frecuente 8, con 13 repeticiones. La composición 2 obtuvo una calificación media de 6,5, mediana 6 y resultado más frecuente 5, con 10 repeticiones. Las calificaciones detalladas pueden verse en la Figura 11.4.

Composición 1: <https://vimeo.com/770493178>

Composición 2: <https://vimeo.com/746620075>

11.2 Segundo estudio. Evaluación de la sonificación simbólica y de la sonificación mediante deep learning

Con el objetivo de evaluar si las estrategias de sonificación propuestas en los Capítulos 6 y 7 pueden ser útiles para la clasificación sonora de espectros estelares, se realizó un estudio cuantitativo y cualitativo basado en un cuestionario online con 16 preguntas y dos vídeos de autoentrenamiento con ejemplos de sonificación. El cuestionario se distribuyó durante los meses de marzo, abril y mayo de 2024 a participantes voluntarios, incluyendo estudiantes de cursos universitarios de astronomía y ciencia de datos, músicos, astrónomos y contactos

personales de los autores.

11.2.1 Diseño del estudio II

El cuestionario formuló 8 preguntas centradas en la clasificación espectral sonora a través de vídeos. Se presentaron cuatro sonificaciones utilizando el mapeado simbólico descrito en el Capítulo 6 y cuatro utilizando el enfoque de sonificación basada en deep learning descrito en el Capítulo 7. Además, los participantes respondieron a 6 preguntas de control demográfico y a 2 preguntas cualitativas adicionales sobre la utilidad de la propuesta y el interés de los participantes en escuchar más sonificaciones una vez realizado el cuestionario. Todas las preguntas se presentaron en orden aleatorio y se recomendó el uso de auriculares, aunque no fueran esenciales. No se requirió ningún conocimiento previo para participar en el estudio. Los vídeos de autoentrenamiento utilizaron un enfoque cercano al concepto Infosonics (Holloway et al., 2022), complementando las sonificaciones con una locución para proporcionar contexto y puntos de referencia (Walker y Kramer, 2004). El cuestionario está disponible en:

<https://forms.office.com/e/veisXfB6fw>

En total se obtuvieron 40 respuestas, principalmente de participantes residentes en España, con algunos participantes de Alemania, Chile y Estados Unidos. La muestra observada contiene 18 participantes que declararon tener experiencia en Astronomía y/o Música, 22 participantes que declararon no tener conocimientos previos en estos campos y dos participantes ciegos o con baja visión (BLV) que fueron analizados en paralelo.

Los resultados y conclusiones finales se obtuvieron analizando las respuestas de un conjunto global de 38 participantes, de dos subconjuntos de 16 participantes con y sin experiencia (muestreados aleatoriamente), y del subconjunto de 2 participantes BLV. Adicionalmente se realizó un análisis por subgrupos que permite la comparación de astrónomos, astrónomos músicos, músicos y participantes sin experiencia. Finalmente, se extrajeron aleatoriamente dos participantes del conjunto global para comparar los resultados de los dos participantes BLV.

Las respuestas registradas se analizaron en dos sentidos, contabilizando los aciertos por pregunta y contabilizando los aciertos por participante. En líneas generales, las diferencias encontradas entre subgrupos de participantes deben considerarse indicativas debido a que la muestra es demasiado pequeña para ofrecer los niveles habituales de confianza estadística. No obstante, los resultados ofrecen un punto de partida para futuros trabajos que examinen estas cuestiones con una muestra de mayor tamaño.

11.2.2 Resultados cuantitativos. Sonificación simbólica

El siguiente análisis corresponde a las cuatro primeras preguntas del cuestionario, centradas en la clasificación espectral utilizando la estrategia de sonificación de mapeado simbólico. Una vez separadas las respuestas de los participantes BLV, los recuentos no presentaron valores atípicos estadísticos según el método del rango intercuartil (IQR). Se utilizó el test de normalidad Shapiro-Wilk sobre las distribuciones obtenidas, obteniendo un resultado de normalidad para la muestra completa y para los subgrupos formados por 16 participantes con y sin experiencia, 3 astrónomos y 4 astrónomos músicos.

La tasa general de acierto promedio de los 38 participantes para las cuatro preguntas fue de 0,57, desviación estándar 0.15, con un intervalo de confianza de Jeffreys de (0,53, 0,61) y un 68,3% de significación. Se realizó un t-test para comparar la media con la tasa de acierto aleatoria (0,14), obteniendo una diferencia estadísticamente significativa, $t(37) = 7,77$, $p = 0,004$, que sugiere que el enfoque podría ser útil para la tarea de clasificación espectral propuesta.

El análisis de acierto por participante de las cuatro preguntas, ofreció una media de 2,29, mediana 2,0 y valor más frecuente 4 (100% de acierto), obtenido 11 veces (28,9% de los participantes).

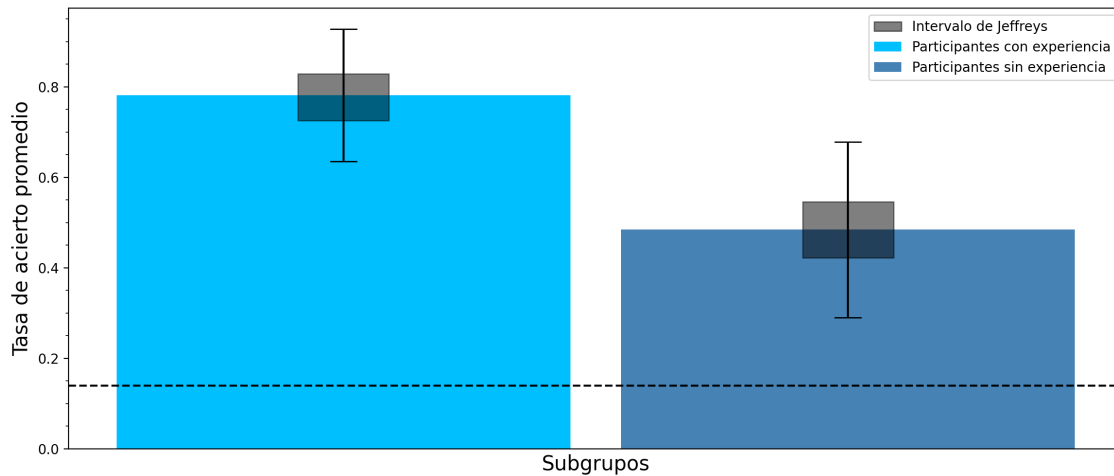


Figura 11.5: Sonificación de mapeado simbólico. Tasa de acierto promedio para 16 participantes con experiencia (0,78, izquierda) y 16 sin experiencia (0,48, derecha), para tareas de clasificación del tipo espectral. La línea discontinua muestra la tasa de acierto aleatorio (0,14).

Las respuestas se dividieron en dos grupos en función de su experiencia previa. El grupo de participantes con experiencia, entendido como el formado por los participantes familiarizados con los datos o por aquellos con entrenamiento en la extracción de información a partir del sonido y la música, se formó con 16 personas aficionadas y profesionales en Astronomía o en Música. El grupo de participantes sin experiencia se formó con otros 16 participantes muestreados aleatoriamente del grupo inicial de 22, que declararon no tener ningún conocimiento previo relacionado con estos campos.

Tal y como muestra la Figura 11.5, el grupo de participantes con experiencia obtuvo una tasa de acierto promedio de 0,78, desviación estándar 0,19, con intervalo de confianza de Jeffreys de (0,72, 0,83) para un 68,3% de significación. La comparación con la tasa de acierto aleatorio de 0,14 mediante un t-test confirmó que la diferencia es estadísticamente significativa, $t(15) = 7,95$, $p = 0,004$.

Este resultado fue 1,62 veces superior a la tasa de acierto promedio del grupo de participantes sin experiencia, que obtuvo un 0,48 de tasa promedio de acierto, desviación estándar 0,11, con intervalo de confianza de Jeffreys (0,41, 0,53), para el mismo porcentaje de significación. La comparación de este resultado con la tasa de acierto aleatorio de 0,14 también ofreció

significación estadística, $t(15) = 8,91$, $p = 0,003$.

Para evaluar si existían diferencias significativas en las medias obtenidas entre participantes con y sin experiencia se utilizó un test ANOVA unidireccional y un t-test. El resultado de ambas pruebas sugirió que las diferencias en los promedios observados son estadísticamente significativas, $F(1,30) = 7,12$, $p = 0,036$, $t(30) = 2,68$, $p = 0,047$. No obstante, el tamaño del efecto, medido con el estadístico de Cohen (Cohen, 2013), fue pequeño ($d = 0,118$), por lo que se puede concluir que aunque las diferencias en las tasas promedio obtenidas por los participantes con y sin experiencia presentan una diferencia real, su magnitud es baja y probablemente tenga una relevancia pequeña en términos prácticos.

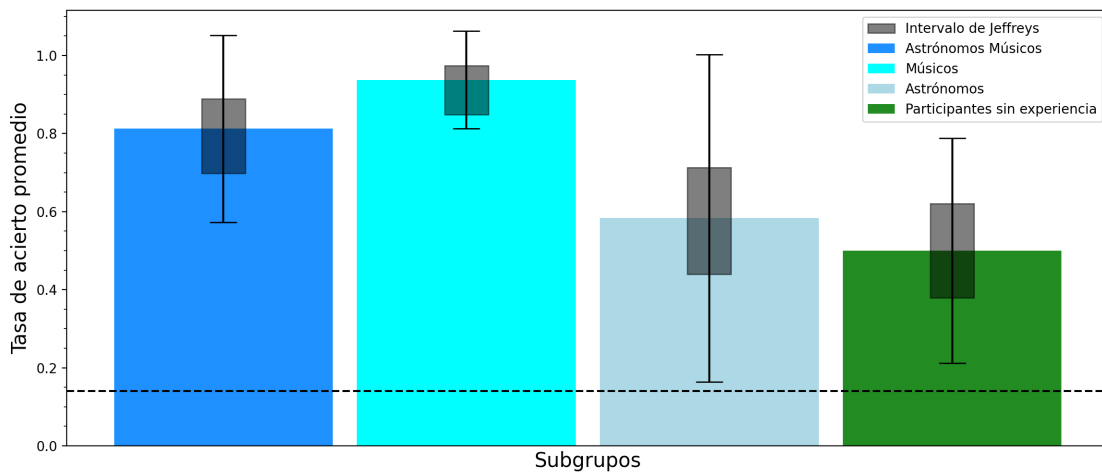


Figura 11.6: Sonificación de mapeado simbólico. Análisis de subgrupos de cuatro muestras. Tasa de acierto promedio para astrónomos músicos (0,81, izquierda), músicos (0,94, centro-izquierda), astrónomos (0,58, centro-derecha) y participantes sin experiencia (0,5, derecha). La línea discontinua muestra la tasa de acierto aleatorio (0,14).

Atendiendo a las diferencias entre preguntas, merece la pena observar las tasas promedio de acierto de la pregunta 1, 73,68 % para la muestra completa, 93,7 % para los participantes con experiencia y 62,5 % para los participantes sin experiencia. Esta pregunta se diseñó como indicador de control con una representación sonora potencialmente sencilla que permitiera comprobar si la mayoría de los participantes habían entendido la propuesta.

Respecto al análisis de acierto por participante, el grupo con experiencia obtuvo una media de 3,12, mediana 3,0 y valor más frecuente 4 (100 % de acierto), obtenido 6 veces (37,5 % de los participantes). El grupo sin experiencia obtuvo una media de 1,93, mediana 1,5 y valor más frecuente 4 (100 % de acierto), obtenido 5 veces (31,2 % de los participantes).

También se estudiaron las diferencias potenciales dentro del grupo de participantes con experiencia mediante el análisis de las respuestas de cuatro subgrupos formados por astrónomos, músicos, astrónomos músicos y participantes sin experiencia. La muestra original incluía 3 astrónomos y 4 astrónomos músicos, lo que obligó a reducir aleatoriamente el grupo de músicos (de 10 a 4 muestras) y el de participantes sin experiencia (de 22 a 4 muestras).

La Figura 11.6 muestra cómo los músicos tuvieron los mejores resultados con una tasa de

acierto promedio de 0,94, frente al 0,81 obtenido por los astrónomos músicos, el 0,58 obtenido por los astrónomos y el 0,5 obtenido por el subgrupo de 4 participantes sin experiencia. Los intervalos de confianza y desviaciones estándar de cada grupo pueden encontrarse resumidos en la Tabla 11.2. El análisis mediante t-test permitió comprobar que las diferencias de los promedios respecto a la tasa de acierto aleatorio (0,14) son estadísticamente significativas, con la excepción del subgrupo de astrónomos. Astrónomos $t(2) = 2,56$, $p = 0,083$, músicos: $t(3) = 14,44$, $p = 0,007$, astrónomos músicos: $t(3) = 6,49$, $p = 0,007$ y 4 participantes sin experiencia: $t(3) = 3,22$, $p = 0,48$.

Mediante el método de Kruskal-Wallis se comprobó que las diferencias en las medias de los subgrupos eran estadísticamente significativas, $H(3) = 7,91$, $p = 0,048$. Posteriormente, se usó una prueba Post Hoc de Dunn (O. J. Dunn, 1961) para comparar las parejas de grupos, obteniendo que la tasa promedio del grupo de músicos es significativamente diferente de la del grupo de astrónomos, $p = 0,017$ y de la del grupo de participantes sin experiencia, $p = 0,031$. Respecto al tamaño del efecto, analizado con el estadístico de Cohen, en ambos casos se obtuvo un efecto medio, con $d = 0,52$ en la comparación con los astrónomos y $d = 0,49$ en la comparación con los participantes sin experiencia.

No obstante, cabe mencionar que realizando la prueba de Dunn con corrección de Bonferroni los subgrupos no presentan diferencias estadísticamente significativas, lo que sugiere, para futuras investigaciones, la necesidad de una muestra de mayor tamaño.

Los resultados notablemente superiores obtenidos por los músicos, confirman que el entrenamiento en percepción auditiva aumenta la comprensión de las sonificaciones (Walker y Nees, 2005, 2011). Del mismo modo, los resultados aceptables obtenidos por el grupo de participantes sin experiencia sugieren que incluso careciendo de conocimientos previos en los campos relacionados, y a pesar de la dificultad para diferenciar sonidos declarada por algunos participantes en el apartado de observaciones del cuestionario, los vídeos de autoentrenamiento facilitaron la comprensión de la propuesta, algo que puede ser de utilidad en proyectos de divulgación y ciencia ciudadana.

Tabla 11.2: Sonificación de mapeado simbólico. Resumen de las tasas de acierto promedio por grupos de participantes, desviación estándar e intervalos de confianza. Los subgrupos han sido extraídos del grupo global y han sido equilibrados para permitir su comparación.

Grupo	Participantes	Tasa de acierto	Conf-bajo	Conf-alto	Std
Completo	40	0,5562	0,5167	0,5950	0,1519
Global	38	0,5723	0,5318	0,6118	0,1463
Con-experiencia	16	0,7812	0,7254	0,8283	0,1943
Sin-experiencia	16	0,4843	0,4226	0,5466	0,1067
AstroMus	4	0,8125	0,6977	0,8895	0,2239
Músicos	4	0,9375	0,8481	0,9739	0,125
Astrónomos	3	0,5833	0,4400	0,7135	0,4194
4-Sin-exp	4	0,5	0,3788	0,6212	0,2886
BLV	2	0,25	0,1328	0,4251	0,2887
2-No BLV	2	0,6250	0,4490	0,7721	0,25

Finalmente, se analizaron las respuestas de los participantes BLV y se compararon con las de un grupo de dos participantes escogidos aleatoriamente de la muestra completa. El grupo de participantes BLV obtuvo una tasa de acierto promedio de 0,25, desviación estándar 0,29, con intervalo de confianza de Jeffreys de (0,13, 0,42) para un 68,3% de significación. La comparación del promedio con la tasa de acierto aleatorio de 0,14 no ofreció significación estadística, $t(1) = 1,25$, $p = 0,3$.

El análisis de acierto por participante BLV de las cuatro preguntas ofreció una media de 1,0 (25% de acierto) y mediana 1,0, con un participante sin ningún acierto (50% de los participantes).

11.2.3 Resultados cuantitativos. Sonificación basada en deep learning

Esta sección analiza el segundo grupo de cuatro preguntas centradas en la clasificación espectral utilizando la estrategia de sonificación mediante deep learning. El análisis se realizó siguiendo el mismo esquema y metodología descrita en la sección anterior con la intención de proporcionar una comparativa entre ambos métodos de sonificación. Los recuentos tampoco presentaron valores atípicos estadísticos según el método del rango intercuartil (IQR), una vez separadas las respuestas de los participantes BLV. Se utilizó el test de normalidad Shapiro-Wilk sobre las distribuciones obtenidas, obteniendo un resultado de normalidad para toda la muestra y para los subgrupos de 16 participantes con y sin experiencia, el de 3 astrónomos y el de 4 participantes sin experiencia.

La tasa de acierto promedio de los 38 participantes para las cuatro preguntas fue de 0,41, desviación estándar 0,034, intervalo de confianza de Jeffreys de (0,37, 0,45) y un 68,3% de significación, lo que sugiere que la tarea de clasificación sonora propuesta no resultó sencilla para los participantes. Mediante un t-test se comparó la media obtenida con la tasa de acierto aleatoria (0,14), obteniendo una diferencia estadísticamente significativa, $t(37) = 23,79$, $p < 0,001$, que sugiere que el enfoque podría ser útil para la tarea de clasificación espectral propuesta.

En el análisis de acierto por participante de las cuatro preguntas, se obtuvo una media de 1,63, mediana 2,0 y valor más frecuente 2 (50% de acierto), obtenido 13 veces (34,2% de los participantes).

Atendiendo a los resultados por subgrupos de participantes con y sin experiencia, tal y como muestra en la Figura 11.7, los participantes con experiencia obtuvieron una tasa de acierto promedio de 0,48, desviación estándar 0,11, con un intervalo de confianza de Jeffreys de (0,42, 0,55) y un 68,3% de significación. La comparación de la media con la tasa de acierto aleatorio (0,14) mediante un t-test confirmó una diferencia estadísticamente significativa, $t(15) = 8,91$, $p = 0,003$.

Este resultado es 1,24 veces mayor que la tasa de acierto promedio del grupo de participantes sin experiencia, que obtuvo un promedio de acierto de 0,39, desviación estándar 0,078, con intervalo de confianza de Jeffreys (0,32, 0,43), para el mismo porcentaje de significación. La comparación de este resultado con la tasa de acierto aleatorio de 0,14 también ofreció

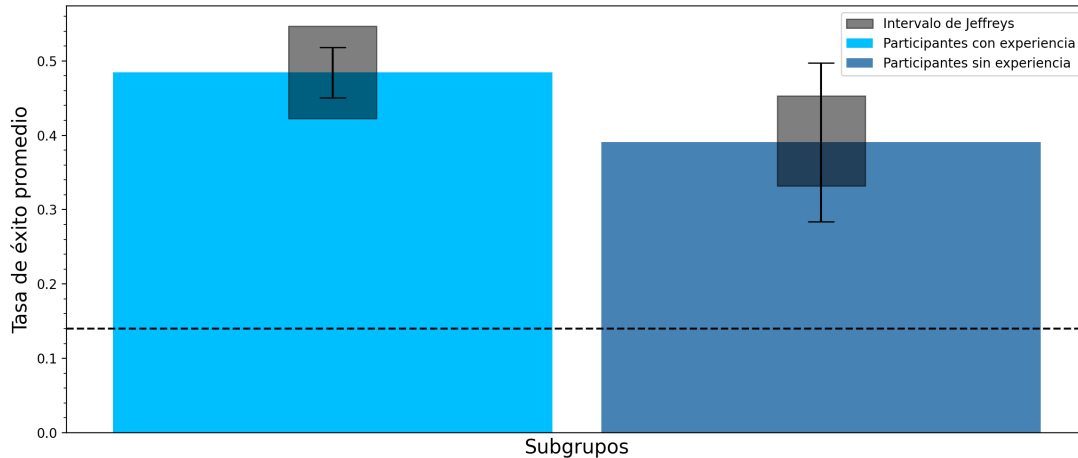


Figura 11.7: Sonificación mediante deep learning. Tasa de acierto promedio para 16 participantes con experiencia (0,48, izquierda) y 16 sin experiencia (0,39, derecha), para tareas de clasificación del tipo espectral. La línea discontinua muestra la tasa de acierto aleatorio (0,14).

significación estadística, $t(15) = 9,71$, $p = 0,002$.

De nuevo se utilizó un test ANOVA unidireccional y un t-test para comprobar si las diferencias en los promedios obtenidos eran estadísticamente significativas, respectivamente $F(1, 30) = 2$, $p = 0,2$ y $t(30) = 1,41$, $p = 0,21$. Ambos tests confirmaron que la diferencia carece de significación estadística por lo que los resultados deben considerarse indicativos.

Respecto al análisis de acierto por participante, el grupo con experiencia obtuvo una media de 1,94, mediana 2,0 y valor más frecuente 2 (50% de acierto), obtenido 8 veces (50% de los participantes). El grupo sin experiencia obtuvo una media de 1,56, mediana 1,0 y valor más frecuente 1 (25% de acierto), obtenido 7 veces (43,7% de los participantes).

De nuevo se estudiaron las diferencias potenciales dentro del grupo de participantes con experiencia. La Figura 11.8 muestra cómo los músicos tuvieron un resultado relativamente superior al del resto de participantes, presentando una tasa de acierto promedio de 0,69 frente al 0,5 obtenido por los astrónomos músicos, el 0,44 obtenido por los participantes sin experiencia y el 0,083 obtenido por los astrónomos.

Los intervalos de confianza y desviaciones estándar de cada grupo pueden encontrarse resumidos en la Tabla 11.3. El análisis mediante t-test permitió comprobar que las diferencias de los promedios respecto a la tasa de acierto aleatorio (0,14) son estadísticamente significativas, con la excepción del subgrupo de astrónomos. Astrónomos $t(2) = 0,44$, $p = 0,69$, músicos: $t(3) = 5,45$, $p = 0,012$, astrónomos músicos: $t(3) = 4,55$, $p = 0,02$ y 4 participantes sin experiencia: $t(3) = 6,44$, $p = 0,008$.

Mediante el estadístico de Kruskal-Wallis se comprobó que las diferencias en los promedios de los subgrupos son estadísticamente significativas, $H(3) = 10,15$, $p = 0,017$. Posteriormente, se usó una prueba Post Hoc de Dunn para comparar las parejas de grupos, obteniendo que la tasa promedio del grupo de astrónomos es significativamente diferente de la del grupo

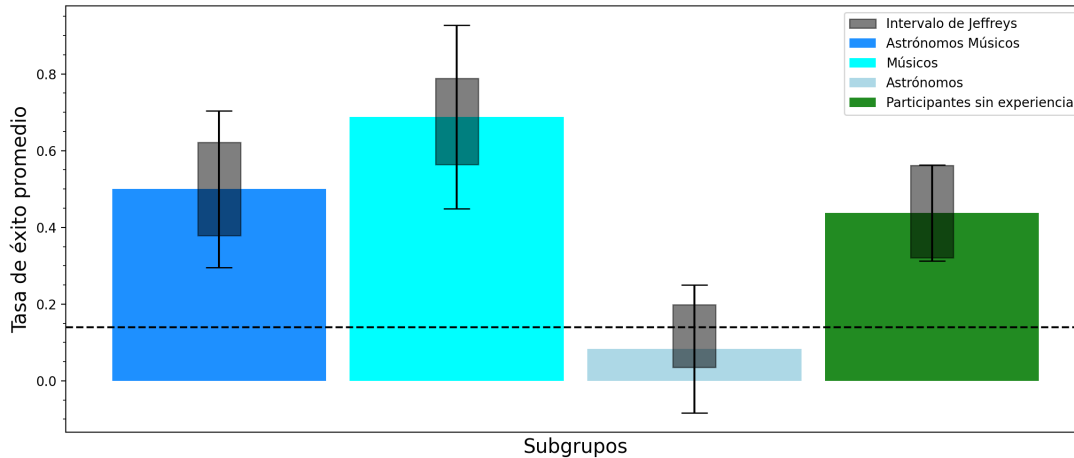


Figura 11.8: Sonificación mediante deep learning. Análisis de subgrupos de cuatro muestras. Tasa de acierto promedio para participantes con experiencia en Astronomía y Música (0,5, izquierda), sólo en música (0,69, centro-izquierda), sólo en astronomía (0,08, centro-derecha) y participantes sin experiencia (0,44, derecha). La línea discontinua muestra la tasa de acierto aleatorio (0,14).

Tabla 11.3: Sonificación mediante deep learning. Resumen de las tasas de acierto por grupos de participantes. Desviación estándar e intervalos de confianza. Subgrupos extraídos del grupo global y equilibrados con fines comparativos.

Grupos	Participantes	Tasas de acierto	Conf-bajo	Conf-alto	Std
Completo	40	0,40	0,3620	0,4392	0,0456
Global	38	0,4078	0,3687	0,4482	0,0339
Con experiencia	16	0,4844	0,4226	0,5466	0,1067
Sin experiencia	16	0,3906	0,3317	0,4528	0,0786
AstroMus	4	0,5000	0,3788	0,6212	0,2041
Músicos	4	0,6875	0,5634	0,7889	0,2393
Astrónomos	3	0,0833	0,0348	0,1981	0,1667
4-Sin-exp	4	0,4375	0,3209	0,5616	0,1250
BLV	2	0,25	0,1328	0,4251	0,2887
2-No BLV	2	0,50	0,3336	0,6663	0,25

de astrónomos músicos, $p = 0,026$ y de la del de músicos, $p = 0,002$. En cuanto al tamaño del efecto, el estadístico de Cohen mostró que la diferencia observada en el promedio de los astrónomos respecto a los músicos es grande, $d = 0,82$ y mediana en el caso de los astrónomos músicos $d = 0,57$.

No obstante, aplicando la corrección de Bonferroni, las diferencias en los promedios de los astrónomos sólo resultaron estadísticamente significativas con el subgrupo de músicos, $p = 0,012$.

Respecto a la comparación de las respuestas de los participantes BLV con las de un grupo de dos participantes escogidos aleatoriamente de la muestra completa, el grupo de participantes

BLV obtuvo una tasa de acierto promedio de 0,25, desviación estándar 0,29, con intervalo de confianza de Jeffreys de (0,13, 0,42) para un 68,3 % de significación. La comparación del promedio con la tasa de acierto aleatorio de 0,14 no ofreció significación estadística, $t(1) = 1,25$, $p = 0,3$.

El análisis de acierto por participante BLV de las cuatro preguntas ofreció una media de 1,0 (25 % de acierto) y mediana 1,0, con un participante sin ningún acierto (50% de los participantes).

11.2.4 Análisis comparativo. Sonificación simbólica - sonificación mediante deep learning

La Figura 11.9 permite comparar los resultados obtenidos con los métodos de sonificación simbólica y mediante deep learning para la tarea de clasificación sonora de tipos espectrales propuesta. Los promedios representados, extraídos a su vez de las Tablas 11.3 y 11.2, permiten comprobar cómo las tasa de acierto de las preguntas de sonificación basada en deep learning fueron más bajas en todos los grupos de participantes, confirmando que el enfoque de sonificación mediante deep learning requiere más entrenamiento que el enfoque simbólico para la tarea de clasificación sonora propuesta. Mediante un t-test se comprobó que las diferencias en las medias observadas en cada grupo con ambas metodologías de sonificación son estadísticamente significativas, $t(74) = 2,2278$, $p = 0,041$. Además, el tamaño del efecto calculado mediante la medida de Cohen indica que estas diferencias son relevantes en términos prácticos, $d = 1,05$.

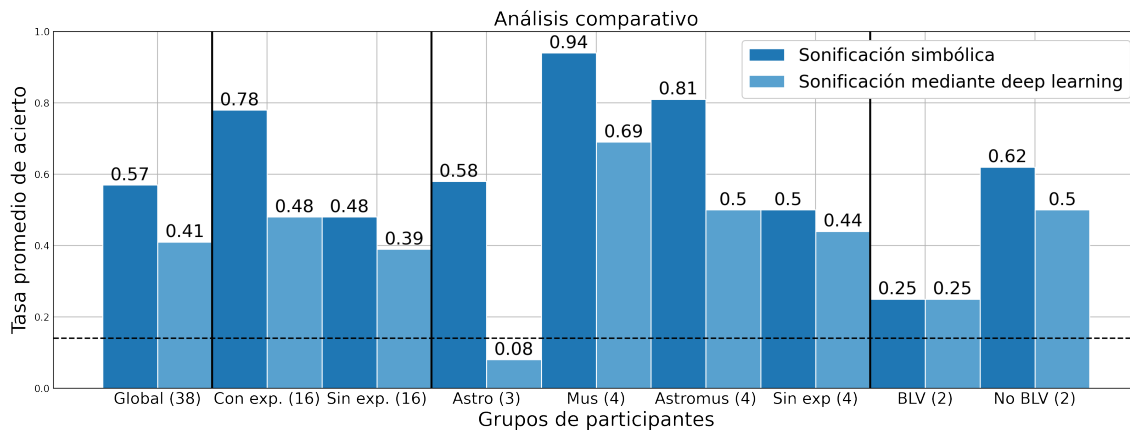


Figura 11.9: Gráfico comparativo de las tasas promedio de acierto para cada grupo de estudio, obtenidas para sonificaciones generadas mediante mapeado simbólico y mediante deep learning. La línea discontinua representa la tasa de acierto aleatorio (0,14).

Por otra parte, la comparación de subgrupos, permite observar cómo las personas con experiencia en Música pudieron clasificar eficazmente los espectros estelares mediante las sonificaciones utilizando los dos enfoques de sonificación. Este resultado sugiere que, con suficiente entrenamiento, el enfoque de sonificación basado en deep learning no introduciría dificultades de interpretación adicionales. Los resultados también confirman que el entrenamiento previo en el análisis de fuentes sonoras, característico del subgrupo de músicos, resultó más útil para

las tareas de clasificación propuestas que el conocimiento previo de los datos representados, característico del subgrupo de astrónomos. Los malos resultados de este último grupo en las preguntas de sonificación mediante deep learning sugieren que el conocimiento previo de los datos pudo producir un sesgo que dificultó la comprensión de las sonificaciones.

Los conocimientos musicales previos de los participantes (reconocimiento básico de las notas de la escala de Do mayor), resultaron de utilidad, incluso en los grupos de participantes sin experiencia, para el reconocimiento de los tipos espectrales en las preguntas de mapeado simbólico. Por otra parte, la abstracción de las sonificaciones generadas mediante deep learning redujo la diferencia entre los aciertos de los participantes con y sin experiencia, pasando de un factor de 1,62 en el mapeado simbólico, a un factor de 1,24 con el uso de deep learning. Este hecho se debió probablemente a que las representaciones resultaron completamente nuevas para todos los participantes.

11.2.5 Resultados cualitativos. Estudio II

Atendiendo a los resultados de las preguntas cualitativas mostradas en la Figura 11.10, el 93,9% de los 40 participantes consideró que las sonificaciones eran útiles en actividades de divulgación para introducir los tipos de espectros estelares, y el 75,8% de los participantes declaró su interés en escuchar más sonificaciones de datos astronómicos después de completar la encuesta.

En cuanto a los subgrupos, en el grupo de participantes con experiencia, el 88,88% de los participantes encontró útiles las sonificaciones en actividades de divulgación para introducir los tipos de espectros estelares, y el 83,33% declaró interés en escuchar más sonificaciones después de realizar el cuestionario. En el grupo sin experiencia, el 88,88% de los participantes encontró útiles las sonificaciones y el 66,66% declaró interés en escuchar más sonificaciones.

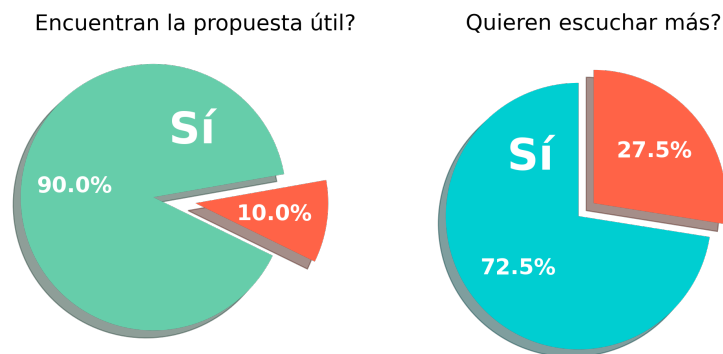


Figura 11.10: Evaluación cualitativa de 40 participantes, el 90% de la muestra completa consideró que la propuesta podría ser útil en actividades de divulgación para introducir los tipos de espectros estelares (izquierda), y el 72,5% declaró su interés en escuchar más sonificaciones astronómicas después de completar la encuesta (derecha).

11.3 Tercer estudio. Evaluación de ViewCube. IFS multimodal interactiva

Esta sección analiza el estudio que se realizó con la intención de evaluar la utilidad potencial de *ViewCube*, la aplicación descrita en el Capítulo 10, que permite el análisis sonoro de cubos de datos. La evaluación consistió en un cuestionario anónimo online y en pruebas presenciales utilizando los espectros de galaxias de CALIFA. Tanto el cuestionario, que incluía vídeos de entrenamiento, como las demostraciones interactivas en persona se dirigieron a participantes especializados y no especializados en Astronomía y Música. Todos los participantes recibieron el mismo formulario en línea, donde indicaron si habían probado la aplicación en persona. El formulario está disponible como referencia en:

<https://forms.office.com/e/RbHVp8Vbtt>.

11.3.1 Diseño del estudio III

El cuestionario online se envió a participantes voluntarios entre el 15 de abril y el 31 de julio de 2024. Incluía cinco vídeos de entrenamiento que podían reproducirse tantas veces como fuera necesario, el primero proporcionando una visión general de la aplicación y el resto como complemento al inicio de cada sección. La encuesta constaba de cuatro secciones específicas orientadas a analizar diversos aspectos de la propuesta. Además, se incluyeron 12 preguntas para recopilar información demográfica, los niveles de experiencia de los participantes en Astronomía y Música, y tres evaluaciones cualitativas sobre la interactividad de la aplicación, su utilidad y la estética de los sonidos empleados en la sonificación.

Se aconsejó a todos los participantes que utilizaran auriculares y que verificaran su correcta colocación en los oídos izquierdo y derecho. Los participantes podían comparar las preguntas con los ejemplos presentados en los vídeos de entrenamiento sin límite de tiempo. A continuación, se describen las secciones y preguntas incluidas. Cada pregunta proporcionaba varias sonificaciones sin gráficos, generadas a partir de los espectros de las galaxias NGC 5784 (Sbc), NGC 5732 (Sbc), NGC 5682 (E4), NGC 6060 (S0a), NGC 7562 (Sbc), NGC 7671 (S0), NGC7800 (Ir), NGC 2638 (Sb) y UGC 00148 (Sb).

Sección 1. Ubicación del sonido. Esta primera sección constaba de cuatro preguntas diseñadas para analizar las posibilidades de la aplicación en la estimación de la ubicación del sonido. Dentro del paisaje sonoro binaural virtual proporcionado, el oyente se coloca virtualmente en el centro de la galaxia, mirando hacia la posición superior. Los vídeos de entrenamiento de esta sección proporcionaron ejemplos utilizando la galaxia espiral (Sa) NGC 7549.

Sección 2. Distancia al centro de la galaxia. Esta sección también incluía cuatro preguntas destinadas a estudiar las posibilidades de la aplicación para proporcionar información sonora sobre la distancia del cursor en movimiento al centro de la galaxia, dentro del paisaje sonoro espectral virtual. Los vídeos de entrenamiento de esta sección proporcionaron ejemplos de la galaxia espiral (Sbc) NGC 5732.

Sección 3. Edad/Tipo de galaxia.

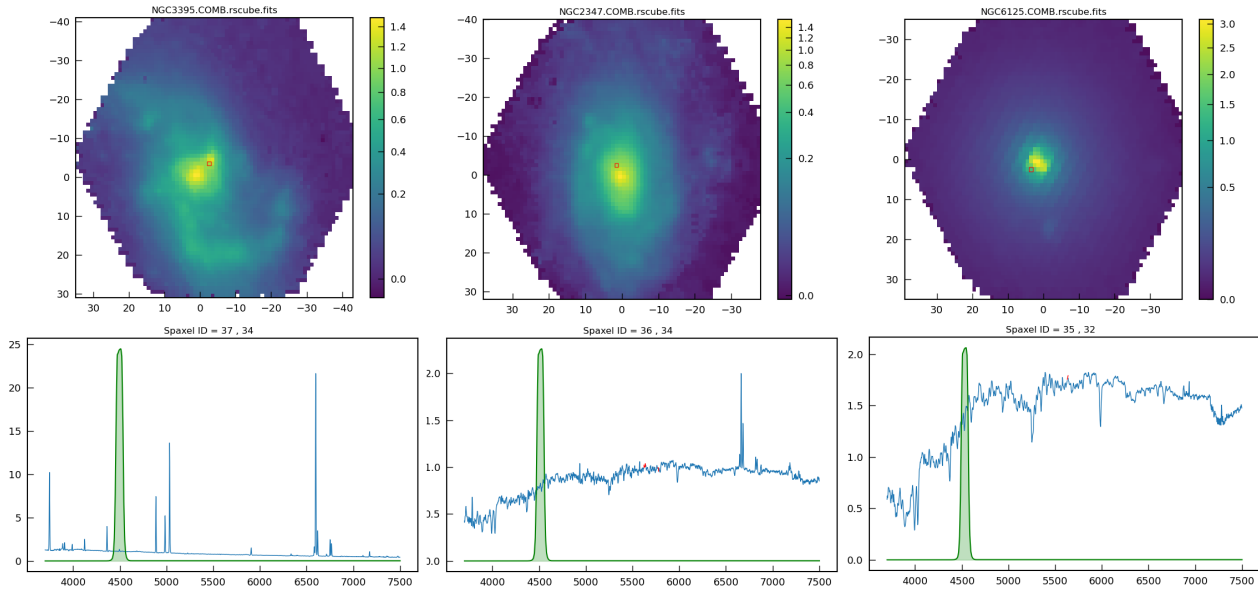


Figura 11.11: Ejemplos de edad/tipo de galaxia presentados en los vídeos de entrenamiento. Espectro de una región de formación estelar en la galaxia espiral (S0) NGC 3395 (izquierda), espectro cerca del centro de la galaxia espiral de edad intermedia (Sd) NGC 2347 (centro), y espectro de una región cerca del centro de la galaxia espiral retirada (Sb) NGC 6125 (derecha). Las figuras superiores muestran los mapas de continuo producidos con el filtro de banda estrecha indicado por las curvas verdes representadas en los espectros de las figuras inferiores. Los espectros corresponden a spaxels específicos, resaltados por cuadrados rojos en la imagen superior. Los ejes de las figuras superiores representan desplazamientos en segundos de arco relativos al centro de la galaxia. La barra de colores representa el flujo del cubo de datos convolucionado con el filtro, en unidades de 10^{-16} erg cm $^{-2}$ s $^{-1}$. Los espectros en las figuras inferiores se representan con longitudes de onda en Angstroms.

Esta sección exploró cómo la representación multimodal podría ayudar a diferenciar entre varios tipos de Edad/Galaxia. Los vídeos de entrenamiento presentaron tres ejemplos, ilustrados en la Figura 11.11. Estos ejemplos incluyen: el espectro (37,34) de una región de formación estelar de la galaxia espiral (S0) NGC 3395; el espectro (36,34) cerca del centro de la galaxia espiral de edad intermedia (Sd) NGC 2347; y el espectro (35,32) de una región cerca del centro de la galaxia espiral retirada (Sb) NGC 6125.

Sección 4. Preguntas combinadas. Finalmente, dos preguntas de opción múltiple analizaron el potencial de la aplicación para permitir la identificación de la posición del espectro representado (izquierda/derecha), su distancia al centro (cerca/lejos) y si correspondía a una región de formación estelar o a una galaxia retirada.

Pregunta cualitativa 1. Si has probado la aplicación en persona, califica la experiencia multimodal. Si no (sólo viste los vídeos de entrenamiento de este cuestionario), por favor, omite esta pregunta. Opciones: Muy mala, Mala, Aceptable, Buena o Muy buena.

Pregunta cualitativa 2. Califica la utilidad potencial de la propuesta multimodal para la exploración del estudio CALIFA. Opciones: Inútil, Dudosamente útil, Útil o Muy útil.

Pregunta cualitativa 3. Califica la estética de las sonificaciones. Opciones: Intolerable, Mala, Aceptable, Buena o Sonido agradable.

11.3.2 Resultados generales. Estudio III

El cuestionario fue completado por 67 participantes, incluyendo 31 astrónomos profesionales, dos de ellos identificados como BLV y 36 no astrónomos. Sus edades oscilaron entre menos de 21 (1) y más de 60 años (10), con la mayoría de los participantes entre 21 y 30 (18), y entre 41 y 50 (21). Los participantes residían principalmente de España (50 participantes) pero también de México, EE.UU., Japón, Alemania, China, Malta, Australia y Reino Unido. Se preguntó sobre sus preferencias musicales y lengua materna para explorar si el idioma podía influir en la capacidad de reconocer características sonoras. Aunque el estudio incluyó hablantes de once idiomas diferentes, el tamaño de la muestra obtenida fue demasiado pequeño y diverso como para extraer alguna conclusión. La misma limitación se encontró en el análisis de las preferencias musicales.

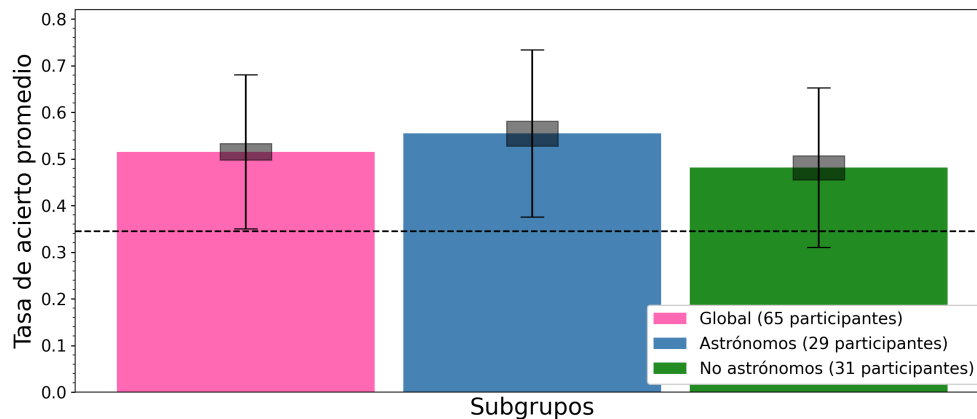


Figura 11.12: Tasas de acierto promedio para 65 participantes en las preguntas simples (izquierda, magenta, 0,515), para 29 astrónomos profesionales (centro, azul, 0,554) y para 31 no astrónomos (derecha, verde, 0,481), referidas al porcentaje promediado por pregunta de acierto aleatorio 0,34 (línea discontinua).

Siguiendo una metodología equivalente a la descrita en los estudios anteriores, las respuestas de los participantes BLV fueron separadas y los recuentos fueron analizados con el método del rango intercuartil (IQR) para detectar potenciales valores atípicos sin resultados destacables. También se realizó el test de normalidad Shapiro-Wilk sobre las distribuciones obtenidas, obteniendo normalidad en la distribución de la muestra completa y para los subgrupos formados por 29 astrónomos, 31 no astrónomos, 10 astrónomos músicos, 11 músicos y el subgrupo muestreado aleatoriamente con fines comparativos de 12 participantes sin experiencia.

Tal y como se muestra en la Figura 11.12, la tasa de acierto global promedio obtenida por 65 participantes fue de 0,5165, con un intervalo de confianza de Jeffreys de (0,497, 0,533), desviación estándar de 0,165 y un 68,3% de significación. El resultado fue comparado con la tasa promedio de acierto aleatorio (0.34) mediante la prueba t-test, confirmando que la diferencia del promedio observado con la probabilidad de acierto aleatorio resulta

estadísticamente significativa, $t(63) = 10,9$, $p < 0,001$, lo que indica que los participantes pudieron recuperar información de las sonificaciones gracias a los vídeos de entrenamiento.

En el análisis de acierto por participante de las doce preguntas simples, se obtuvo una media de 6,18, mediana 6,0 y valor más frecuente 6 (50 % de acierto), obtenido 15 veces (23,1 % de los participantes).

El subgrupo formado por astrónomos profesionales (29 participantes) obtuvo una tasa de acierto promedio de 0,554 con un intervalo de confianza de Jeffreys de (0,528, 0,581), desviación estándar de 0,179 y un 68,3 % de significación. Las tasas de acierto aleatorio por pregunta pueden consultarse en la Figura 11.15 (0,5 para las dos primeras de localización, 0,25 para las dos segundas, 0,33 para las de distancia y tipo de galaxia y 0,2 para las combinadas). Se utilizó el valor promedio (0,34) para comprobar mediante un t-test que el promedio observado respecto a la probabilidad de acierto aleatorio presentaba significación estadística, $t(28) = 10,50$, $p < 0,001$.

Por otra parte, el subgrupo de no astrónomos (submuestreado aleatoriamente de 36 a 31 participantes para permitir la comparación directa) obtuvo una tasa de acierto promedio de 0,481 con un intervalo de confianza de Jeffreys de (0,455, 0,507), desviación estándar de 0,171 y un 68,3 % de significación. La comparación de este resultado con la tasa promedio de acierto aleatorio de 0,34 también ofreció significación estadística, $t(30) = 9,53$, $p < 0,001$.

Respecto al análisis de acierto por participante, el grupo con experiencia obtuvo una media de 6,65, mediana 6,0 y valor más frecuente 6 (50 % de acierto), obtenido 8 veces (27,6 % de los participantes). El grupo sin experiencia obtuvo una media de 5,84, mediana 6 y valor más frecuente 5 (41,67 % de acierto), obtenido 6 veces (19,3 % de los participantes).

Para evaluar si existían diferencias significativas entre los promedios de los grupos de astrónomos y no astrónomos se utilizó la prueba ANOVA unidireccional, $F(1, 58) = 0,296$, $p = 0,59$. Este resultado confirmó que las diferencias observadas no son estadísticamente significativas. No obstante, los resultados indicativos parecen confirmar que todos los participantes pudieron comprender la información de las sonificaciones gracias a las pruebas personalizadas y a los vídeos de entrenamiento, incluso sin tener experiencia previa en Astronomía.

De acuerdo con los resultados obtenidos utilizando gráficos sonoros por Smith y Walker, 2005, el entrenamiento y el contexto proporcionados en el cuestionario pudieron mejorar el rendimiento de los participantes. Esto es particularmente notable en el análisis de las preguntas combinadas, en las que los participantes no astrónomos obtuvieron resultados 1,33 veces mejores que los astrónomos, tal y como puede apreciarse en la Figura 11.13. Este efecto podría estar relacionado con la decisión de permitir a los participantes repetir los vídeos de entrenamiento tantas veces como fuera necesario, algo que podría haber beneficiado a los participantes con conocimientos previos bajos o intermedios, tal y como reportan van Riesen et al., 2022.

11.3.3 Resultados cuantitativos por subgrupos. Estudio III

Para proporcionar un análisis más profundo de las respuestas registradas, los participantes se dividieron en cuatro subgrupos según su nivel de experiencia en Astronomía y Música. El

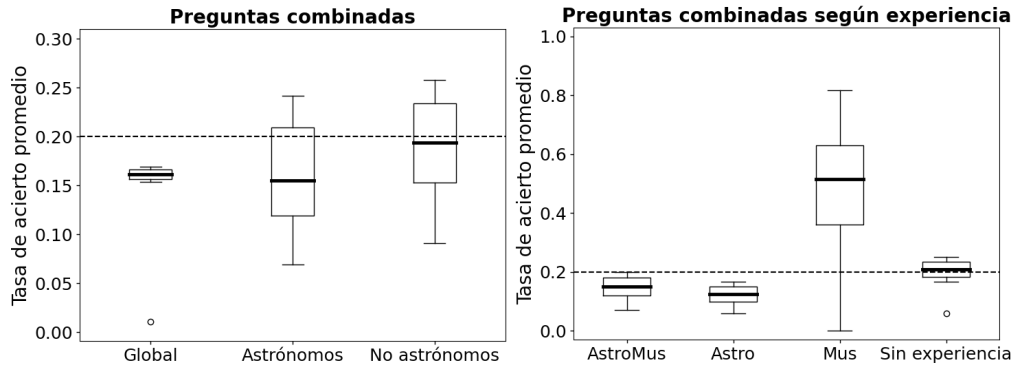


Figura 11.13: Tasas de acierto promedio en preguntas combinadas, resultados globales y por subgrupos.

subgrupo de Astrónomos músicos incluía astrónomos profesionales también identificados como músicos profesionales o aficionados. Un segundo grupo de Astrónomos no músicos se utilizó para analizar la influencia de la experiencia en análisis sonoro de posibles expertos en la tarea de análisis propuesta. El grupo de Músicos no astrónomos estaba formado por participantes identificados como músicos profesionales y aficionados. Finalmente, el grupo de participantes sin experiencia incluía al resto de los participantes.

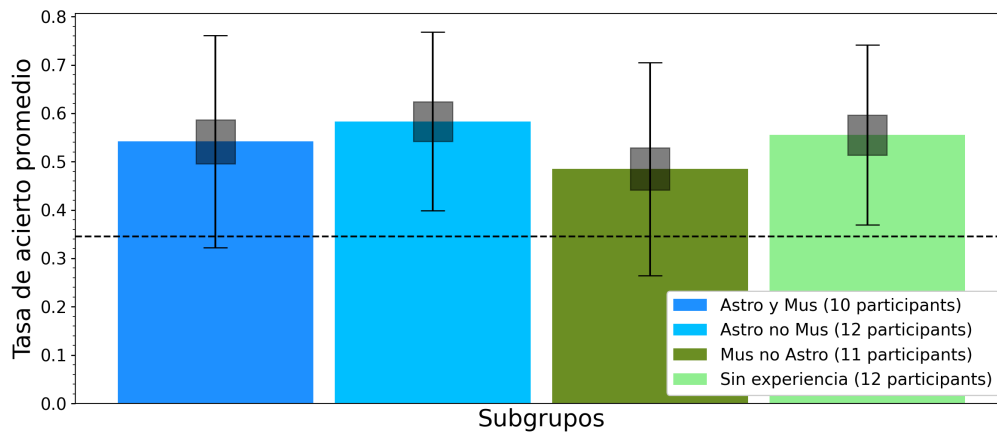


Figura 11.14: Tasas de acierto promedio en preguntas simples por campo de experiencia. De izquierda a derecha: Astrónomos músicos (0,54, 10 participantes), Astrónomos no músicos (0,58, 12 participantes), Músicos no astrónomos (0,48, 11 participantes) y submuestra de participantes sin experiencia (0,55, 12 participantes).

Tal y como muestra la Figura 11.14, el subgrupo de astrónomos no músicos obtuvo los mejores resultados en las secciones de preguntas simples, obteniendo una tasa de acierto promedio de 0,5873, con un intervalo de confianza de Jeffreys de (0,542, 0,624), desviación estándar de 0,185 y un 68,3 % de significación. Por su parte, los astrónomos músicos obtuvieron una tasa de acierto promedio de 0,5416, con un intervalo de confianza de Jeffreys de (0,496, 0,587), desviación estándar de 0,15, con un 68,3 % de significación. Los músicos obtuvieron una tasa de acierto promedio de 0,485, con un intervalo de confianza de Jeffreys de (0,441, 0,528), desviación estándar de 0,36 y el subgrupo de participantes sin experiencia obtuvo una tasa

de acierto promedio de 0,555, con un intervalo de confianza de Jeffreys de (0,514, 0,596), desviación estándar de 0,186, ambos con un 68,3% de significación. La comparación mediante t-test de los promedios observados con respecto a la tasa promedio de acierto aleatorio (0,34), resultó estadísticamente significativa en todos los subgrupos (astromus: $t(9) = 8,08$, $p < 0,001$, mus: $t(10) = 7,13$, $p < 0,001$, astro: $t(11) = 10,41$, $p < 0,001$ y sin experiencia: $t(11) = 9,83$, $p < 0,001$).

Un test ANOVA unidireccional, $F(1, 43) = 1,7$, $p = 0,18$ y un test Kruskal-Wallis, $H(3) = 3,034$, $p = 0,38$ (la distribución de astrónomos no es normal), confirmaron que las diferencias en los promedios observados no son estadísticamente significativas, por lo que los resultados de los subgrupos deben considerarse indicativos.

Las Figuras 11.15 y 11.16 muestra los porcentajes de acierto por pregunta y grupo de experiencia.

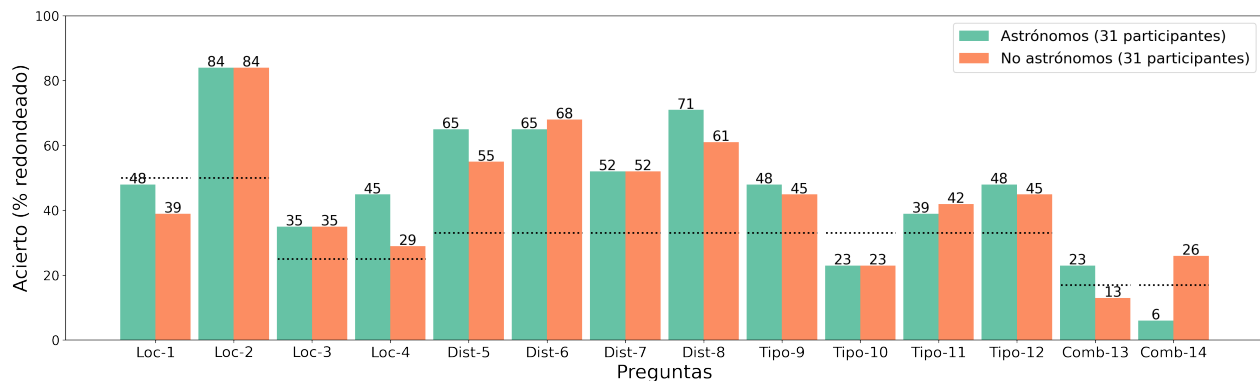


Figura 11.15: Porcentaje de acierto por pregunta y grupo de experiencia. Resultados para Astrónomos y No astrónomos. Preguntas de ubicación del sonido referenciadas como “Loc”, preguntas de distancia al centro de la galaxia referenciadas como “Dist”, preguntas de tipo de galaxia referenciadas como “Tipo” y preguntas combinadas (acierto = todas las opciones de opción múltiple correctas) referenciadas como “Comb”. Las líneas de puntos representan la tasa de referencia de elección aleatoria para cada pregunta.

Merece la pena mencionar que el grupo de participantes sin experiencia tuvo un rendimiento similar al de los astrónomos músicos, logrando una tasa de acierto promedio de 0,55 con una desviación estándar de 0,186. Este rendimiento fue 1,14 veces mejor que el de los Músicos, lo que sugiere que, una vez más, la atención y dedicación adicional requerida para aprender sobre campos desconocidos podría haber ayudado a los participantes sin experiencia en la tarea propuesta (van Riesen et al., 2022).

En cuanto a las preguntas combinadas, las tasas de acierto obtenidas fueron notablemente más bajas que las obtenidas en las preguntas simples para todos los grupos, con la excepción de los Músicos no astrónomos. Esto sugiere que la experiencia previa en el análisis de eventos sonoros fue más útil que el conocimiento previo de los datos para la tarea propuesta. Mediante un t-test se comprobó que la diferencia de los promedios obtenidos con la tasa de acierto aleatorio no resultó estadísticamente significativa.

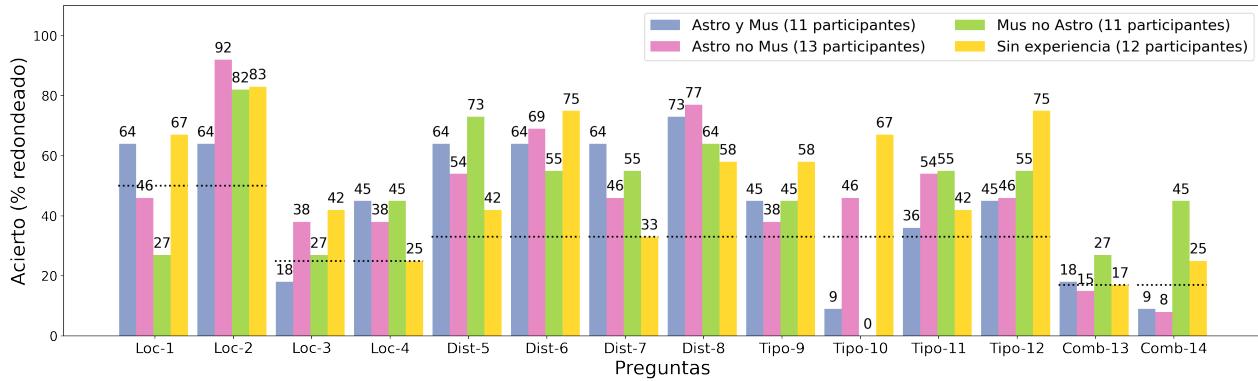


Figura 11.16: Porcentaje de acierto por pregunta y grupo de experiencia. Resultados para Astrónomos músicos, Astrónomos no músicos, Músicos no astrónomos y participantes sin experiencia. Preguntas de ubicación del sonido referenciadas como “Loc”, preguntas de distancia al centro de la galaxia referenciadas como “Dist”, preguntas de tipo de galaxia referenciadas como “Tipo” y preguntas combinadas (acierto = todas las opciones de opción múltiple correctas) referenciadas como “Comb”. Las líneas de puntos representan la tasa de acierto aleatorio para cada pregunta. Merece la pena destacar los resultados de elección aleatoria para los participantes sin experiencia en las preguntas Loc-4, Dist-7 y Comb-13.

No obstante, un análisis más profundo de las respuestas combinadas reveló cómo los participantes no lograron responder simultáneamente a los tres aspectos de las preguntas combinadas, pero tuvieron un buen rendimiento por sección. De la muestra completa, el 74,63% de los participantes ubicó correctamente la sonificación, el 46,27% marcó la distancia correcta al centro de la galaxia y el 36,57% interpretó con éxito la edad de la galaxia. Los respectivos porcentajes de acierto promedio de las preguntas simples por sección fueron del 53,2% para las preguntas de ubicación, del 63,30% para las preguntas de análisis de distancia y del 40,7% para las preguntas de tipo de galaxia.

Como puede verse en la Figura 11.17, las preguntas de Tipo/Edad tuvieron tasas de acierto relativas más bajas en todos los grupos, probablemente debido al nivel de abstracción involucrado en esta tarea. El resultado sugiere que interpretar el tipo o la edad de una galaxia a través del sonido en esta implementación de sonificación específica requiere más entrenamiento que las otras tareas, más intuitivas y alineadas con la experiencia previa de los participantes.

Vale la pena mencionar la excepción de los astrónomos no músicos, que obtuvieron los peores resultados en la interpretación de la distancia al centro de la galaxia. Este hecho tiene una correlación positiva con la falta de precisión en la percepción de la distancia, en comparación con la localización horizontal, tal y como se discute en Middlebrooks, 2015.

La Figura 11.18 ilustra las distribuciones de aciertos promedios para las preguntas simples y combinadas en función de la edad de los participantes. Para las preguntas simples se realizaron t-tests de comparación de las medias observadas con los valores de acierto aleatorio, obteniendo significación estadística en todos los casos salvo en el grupo de participantes menores de 21 años.

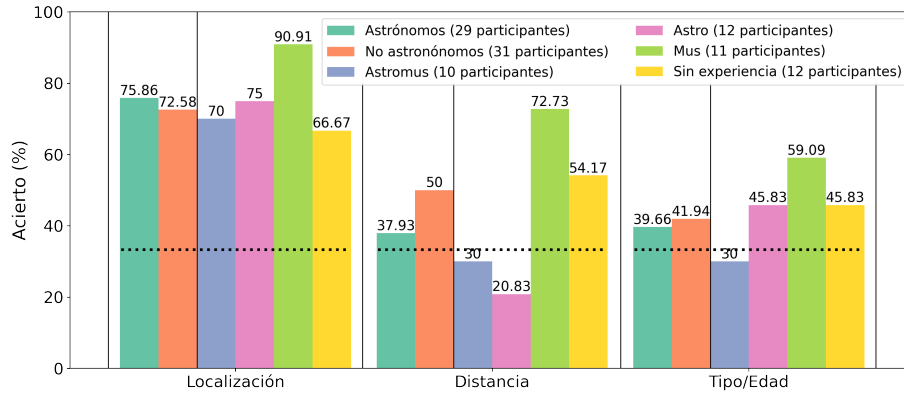


Figura 11.17: Tasa de acierto promedio de las dos preguntas combinadas por bloques (Ubicación, Distancia y Tipo/Edad). De izquierda a derecha, tasa de acierto para Astrónomos, No astrónomos y tasas de subgrupos para Astrónomos músicos, Astrónomos no músicos, Músicos no astrónomos y participantes sin experiencia. La línea de puntos indica la tasa promedio de acierto aleatorio (0,33), considerando tres posibles respuestas por bloque (correcta, incorrecta y no respondida), ya que 19 participantes no introdujeron ninguna respuesta en algunos bloques.

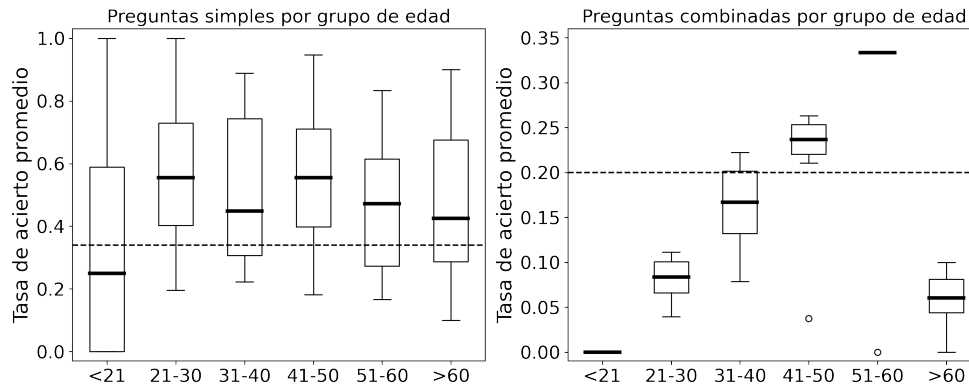


Figura 11.18: Tasas de acierto promedio en preguntas simples y combinadas por grupos de edad. La línea discontinua representa la tasa promedio de acierto aleatorio (valor promedio de las preguntas simples 0,34 y 0,2 en las combinadas). Preguntas simples (izquierda): Participantes <21 = 0,25, entre 21 y 30 = 0,55, entre 31 y 40 = 0,50, entre 41 y 50 = 0,56, entre 51 y 60 = 0,44 y >60 = 0,45. Preguntas combinadas (derecha): Participantes <21 = 0, entre 21 y 30 = 0,08, entre 31 y 40 = 0,17, entre 41 y 50 = 0,24, entre 51 y 60 = 0,33 y >60 = 0,05.

Se realizó un test de Kruskal-Wallis para comprobar que las diferencias observadas entre subgrupos eran estadísticamente significativas, $H(5) = 55,56$, $p < 0,001$.

En las preguntas combinadas por subgrupo de edad, las diferencias con la tasa de acierto aleatorio no resultaron estadísticamente significativas para ningún subgrupo. No obstante, puede destacarse que el subgrupo de 51 a 60 años (6 participantes) tuvo un resultados 1,56 veces mejores que los del subgrupo de 41 a 50 años (19 participantes) y 1,73 veces mejores que

los del subgrupo de 31 a 40 años (9 participantes), lo que podría sugerir que la experiencia y la atención al detalle jugaron un papel importante en la comprensión de información sonora compleja.

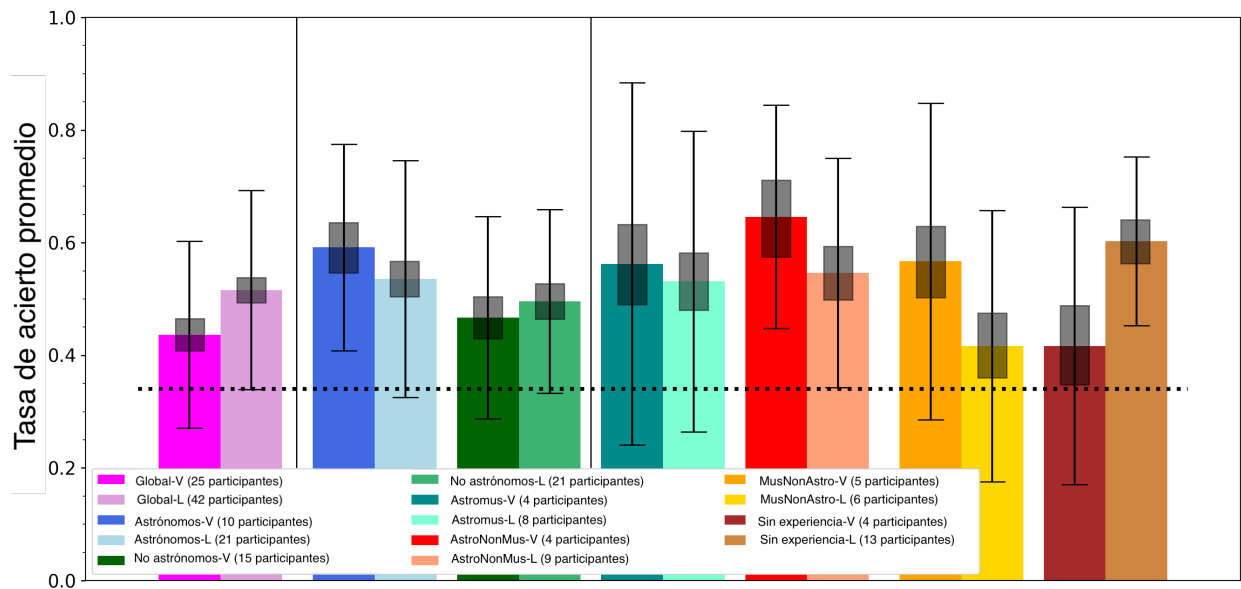


Figura 11.19: Tasa de acierto en preguntas simples para participantes entrenados sólo con vídeos (V) frente a los participantes que probaron la aplicación en vivo (L). De izquierda a derecha, resultados globales, astrónomos frente no astrónomos y subgrupos de experiencia. La línea de puntos muestra la tasa de elección aleatoria promedio (0,34) de preguntas con 2, 3 y 4 posibles respuestas.

La Figura 11.19 proporciona una comparación adicional de los resultados obtenidos por los 42 participantes que pudieron probar la aplicación en persona y los de los 25 participantes que sólo utilizaron los vídeos de entrenamiento. Las personas que pudieron probar *ViewCube* en vivo presentaron una tasa de acierto promedio de 0,51 frente al 0,43 de los participantes entrenados sólo con vídeo. Sin embargo, los mejores resultados fueron obtenidos por el subgrupo de astrónomos no músicos (4 participantes) utilizando sólo los vídeos de entrenamiento sin restricciones de repetición (0,65), seguido por el subgrupo de participantes sin experiencia (13 participantes) que probaron la aplicación en vivo (0,60). Estos resultados sugieren que ambos métodos de entrenamiento fueron útiles para las tareas propuestas.

En cuanto a los resultados de los astrónomos profesionales BLV (dos participantes), sus resultados en preguntas simples fueron similares, aunque ligeramente inferiores (4%) a los obtenidos por un grupo equivalente muestreado aleatoriamente de astrónomos profesionales no BLV, obteniendo una tasa promedio de acierto de 0,54, desviación estándar 0,396, intervalo de confianza de Jeffreys (0,44, 0,64) y 68,3% de significación.

Se realizó un t-test para comprobar que la diferencia del promedio observado con el valor de acierto aleatorio (0,34) era estadísticamente significativa, $p(1) = 3,99$, $p = 0,002$. El test de Kruskal-Wallis entre el subgrupo de astrónomos BLV y el aleatorio de dos astrónomos no BLV no permitió rechazar la hipótesis nula ($H(1) = 0,062$, $p = 0,8$), lo que habría indicado que los promedios observados son diferentes.

Estos resultados sugieren que la aplicación podría ayudar a acercar el análisis IFS a los astrónomos BLV de manera horizontal, aspecto que merece la pena ser abordado en futuros trabajos con una muestra de mayor tamaño. Respecto a las preguntas combinadas, ninguno de los dos astrónomos BLV respondió con éxito a la sección Tipo/Edad.

Todos los resultados pueden encontrarse resumidos en las Tablas 11.4 y 11.5.

Tabla 11.4: Evaluación cuantitativa. Resultados para secciones de preguntas simples y combinadas mostradas por grupo de experiencia y edad.

	Respuestas	Tasa de acierto	std	Comb.acierto	Comb.std
Completo	67	0,516	0,169	0,157	0,011
Global	65	0,516	0,165	0,161	0,010
Astro	29	0,554	0,179	0,155	0,122
NoAstro	31	0,481	0,171	0,193	0,091
AstroMus	10	0,542	0,219	0,150	0,070
AstNoMus	12	0,583	0,185	0,125	0,059
MusNoAst	11	0,485	0,220	0,364	0,128
Sin experiencia	12	0,555	0,186	0,208	0,059
<21	1	0,250	0,452	0,0	0,0
21-30	18	0,555	0,195	0,083	0,039
31-40	9	0,509	0,229	0,166	0,078
41-50	19	0,557	0,181	0,237	0,037
51-60	6	0,444	0,217	0,333	0,0
>60	10	0,450	0,247	0,05	0,070
BLV	2	0,542	0,396	0	0
No BLV	2	0,583	0,359	0	0

Tabla 11.5: Evaluación cualitativa. Porcentajes por grupo de experiencia. Los valores con asterisco corresponden a la muestra de participantes que probaron la aplicación en persona.

	Respuestas	Buena interactividad *	Útil	Buen sonido
Completo	42*/67	80,95	79,10	58,21
Astro	21*/31	76,19	74,19	51,61
NoAstro	17*/31	82,35	80,64	64,52
AstroMus	8*/11	75,0	63,64	45,45
AstNoMus	9*/13	66,66	79,92	53,85
MusNoAst	6*/11	83,33	81,82	63,64
Nothing	9*/12	55,55	83,33	58,33
BLV	1*/2	100,0	100,0	100,0

11.3.4 Resultados cualitativos. Estudio III

Para evaluar la interactividad, utilidad y estética de la propuesta, se incluyeron en el cuestionario las tres preguntas cualitativas descritas en la Sección 11.3.1. De los 67 participantes que

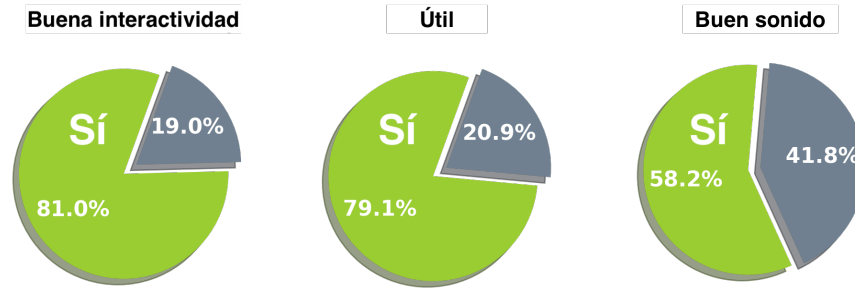


Figura 11.20: Evaluación cualitativa de *ViewCube*. Interactividad: resultados de los 42 participantes que probaron la aplicación en persona. Utilidad y estética: muestra completa, 67 participantes. El 81 % declaró que la aplicación tiene una “buena interactividad”, el 79,1 % la encontró “útil” y el 58,2 % con “buen sonido”.

completaron la encuesta, 42 probaron la aplicación en persona. Como se muestra en la Figura 11.20, el 81 % de estos participantes expresó la buena respuesta interactiva de la aplicación, el 79,1 % de la muestra completa de participantes (67) encontró la aplicación “Útil” o “Muy útil”, el 19,4 % “Dudosamente útil” y un participante (1,49%) la consideró “Inútil”. En cuanto a la estética sonora utilizada, el 58,2 % la calificó como “Buena” o de “Sonido agradable”, el 34,33 % como “Aceptable” y el 5,97 % con “Sonido Malo”. Además, 7 participantes expresaron explícitamente la dificultad encontrada para diferenciar las sonificaciones y 8 participantes expresaron explícitamente su entusiasmo por el proyecto.

Capítulo 12

Discusión

Los prototipos de sonificación y musificación presentados en esta tesis ofrecen un amplio conjunto de perspectivas y recursos de código abierto que podrían aplicarse a diferentes técnicas de estudio tales como la fotometría de tránsito, la espectroscopía de transmisión y la espectrofotometría, que se utilizan respectivamente en la búsqueda de tránsitos planetarios, en la caracterización de atmósferas de exoplanetas y en el estudio de la formación, evolución y clasificación de estrellas o galaxias. Además de una detallada descripción de las aplicaciones y metodologías de representación sonora desarrolladas, este trabajo proporciona tres evaluaciones cuantitativas y cualitativas realizadas con participantes especializados y no especializados en Astronomía, Astrofísica y Música. Estos estudios ofrecen tendencias y preferencias que podrían ser de utilidad como punto de partida en futuros desarrollos a gran escala.

12.1 Sobre la estética

El uso de ondas sinusoidales para la representación precisa de información científica ofrece resultados armónicos y enarmónicos que adquieren características musicales cuando se representan en forma de acorde o de manera secuencial. En las diferentes implementaciones desarrolladas en esta tesis, hemos utilizado simples complementos estéticos como la reverberación para enfatizar esta musicalidad, acercando las representaciones generadas al universo estético característico de la música electroacústica. Encontramos que mediante esta aproximación se puede reducir la fatiga en el oyente y mejorar la aceptación de las sonificaciones, tal y como avanzaron Barrass y Vickers, [2011](#), Risset, [2003](#), Kramer, [2000](#) y Roads, [1996](#).

Por otra parte, el protocolo MIDI puede ser de utilidad para facilitar la musificación y el mapeado de datos a instrumentos musicales virtuales, y para generar partituras de forma automática. Este enfoque puede ser de utilidad en aplicaciones de divulgación científica y para la difusión orientada al público en general, en conciertos y conferencias. Sin embargo, la cuantificación de frecuencia intrínseca en cada nota musical MIDI aumenta la incertidumbre de los resultados y hace que este tipo de conversión sea menos adecuado para aplicaciones científicas de precisión. Algunas excepciones podrían incluir el uso de representaciones sonoras preliminares para la exploración masiva de datos, o el uso de mapeados a diferentes instrumentos musicales para la detección mediante variación tímbrica de eventos casi coincidentes

en frecuencia, como en el caso de estudio descrito en el Anexo.6 del análisis de las líneas de absorción y emisión en espectros de galaxias. En cualquier caso, esta limitación puede resolverse mediante un mapeado doble, de nota y de la variación centesimal de pitch, tal y como se sugiere en la Sección 6.2.

Atendiendo a los resultados cualitativos del estudio I (Sección 11.1.3), las composiciones no supervisadas y las musicaciones de catálogos que hacen uso de instrumentos acústicos y formas musicales clásicas obtuvieron regularmente puntuaciones más altas, tal y como puede verse en la Figura 11.3. Esto sugiere que las representaciones sonoras abstractas podrían ser menos adecuadas que las musicales en aplicaciones divulgativas y en representaciones sonoras ampliamente aceptadas.

La creación de sonificaciones y musicaciones de estética inspirada en la música contemporánea occidental, como precedente con varias décadas de historia que también requirió un entrenamiento equivalente por parte de la audiencia, puede ayudar en la aceptación del análisis sonoro en el ámbito científico y reducir la posible fatiga asociada a largos tiempos de exposición. Como puede verse en la Figura 11.4, las calificaciones de las dos piezas musicales evaluadas por los 46 participantes del estudio I alcanzaron de media 7,71 y 6,5 respectivamente, lo que sugiere una buena aceptación de las propuestas.

Los criterios estéticos asociados a todos los aspectos creativos de las pruebas de concepto implementadas en esta tesis (elección de sintetizadores, instrumentación y postproducción), están orientados al contexto musical de creación de bandas sonoras para películas de ciencia ficción. Este género ofrece asociaciones entre los datos de carácter astronómico y los patrones perceptuales de la cultura audiovisual, que pueden considerarse equivalentes a la percepción ecológica (Walker y Kramer, 2004) de las sonificaciones de datos astronómicos, aportando significado y contexto para facilitar la conexión con la audiencia y la comprensión de la información presentada (Seiça et al., 2023). Encontramos que los resultados musicales de este trabajo de investigación (especialmente los de los Capítulos 6, 8 y 9) podrían formar parte de la banda sonora original de una película de ciencia ficción.

12.2 Sobre la sonificación de espectros de transmisión

Atendiendo a los resultados obtenidos para la tarea propuesta de detección mediante sonido de las líneas de absorción y emisión en galaxias LEGA-C (sin el apoyo de vídeos de entrenamiento), la incorporación de representaciones sonoras, presentada en el Anexo.6, podría proporcionar información útil complementaria a las representaciones gráficas de los espectros, ayudando en la detección de eventos ocultos e información superpuesta y haciendo este análisis más accesible para usuarios BLV. Los 44 participantes del estudio I (Sección 11.1.1) obtuvieron una tasa de acierto promedio (0,26) superior a la tasa de acierto aleatorio con significancia estadística [$t(43) = 4,78$, $p < 0,001$]. La diferencia del promedio de acierto de los participantes BLV con la tasa de acierto aleatorio no resultó estadísticamente significativa [$t(1) = 0,77$, $p = 0,46$], lo que no permite asegurar que sus resultados difieran de una respuesta al azar.

Los 15 participantes que se declararon con experiencia en Astronomía y Música tuvieron una tasa de acierto 3,3 veces superior a la de los 15 participantes que declararon no tener

experiencia, lo que confirmaría la influencia positiva del conocimiento previo del campo de estudio y del entrenamiento auditivo en la comprensión de las sonificaciones. No obstante, esta diferencia no alcanzó los niveles de confianza estadística convencionales [$F(1,28) = 0,241$, $p = 0,14$], por lo que el resultado debe considerarse indicativo a pesar de concordar con los reportes de Walker y Nees, 2011 y Tucker Brown et al., 2022.

Atendiendo a los resultados por subgrupos, tal y como muestra la Figura 11.1, los astrónomos músicos obtuvieron tasas de acierto promedio un 27,3% superiores a las de los astrónomos, un 33,3% superiores a las de los músicos y un 69,45% superiores a las de los participantes sin experiencia. No obstante, estas diferencias deben considerarse indicativas al no alcanzar significación estadística [$H(3) = 8,78$, $p = 0,83$].

Futuras investigaciones con una muestra de mayor tamaño podrían confirmar si la experiencia y el conocimiento de los participantes en ambos campos facilitan la comprensión de información astronómica mediante sonido, y si la Sonificación puede aportar un medio horizontal que equipare las posibilidades de análisis de participantes BLV y no BLV.

Acciones, a priori sencillas, tales como el recuento de eventos sonoros secuenciales, demostraron requerir entrenamiento y concentración a todos los participantes del estudio, que presentaron tasas de acierto relativamente bajas. Este resultado inesperado sugiere la necesidad de programas de educación y entrenamiento auditivo para mejorar las capacidades de escucha crítica y para permitir la exploración de librerías y catálogos estelares públicos a través del sonido, lo que concuerda con las apreciaciones de Neuhoff et al., 2002 y Sándor y Lane, 2003.

12.3 Sobre la sonificación simbólica

En cuanto a la sonificación simbólica propuesta en la Sección 6.2, los resultados obtenidos en la evaluación cuantitativa sugieren que el conocimiento musical básico de los participantes sirvió como punto de referencia para la comprensión de las sonificaciones, algo que concuerda con las conclusiones reportadas por Walker y Nees, 2011, Tucker Brown et al., 2022 y Fovino et al., 2024. El conjunto global de 38 participantes del estudio II (Sección 11.2.1) obtuvo una tasa de acierto promedio (0,57) estadísticamente significativa respecto a la tasa de acierto aleatorio [$t(37) = 7,77$, $p = 0,004$], lo que indica que los participantes pudieron recuperar información de las sonificaciones gracias a los vídeos de entrenamiento. Este resultado sugiere el potencial de este enfoque para su uso en tareas de clasificación espectral mediante sonido (incluso por participantes sin experiencia con suficiente entrenamiento), algo que puede ser de utilidad en el desarrollo de proyectos de ciencia ciudadana. La diferencia del promedio de acierto de los participantes BLV con la tasa de acierto aleatorio no resultó estadísticamente significativa [$t(1) = 1,25$, $p = 0,3$], lo que no permite asegurar que sus resultados difieran de una respuesta al azar.

Atendiendo a las diferencias entre participantes con y sin experiencia en Astronomía y Música, los 16 participantes que se declararon con experiencia obtuvieron una tasa de acierto promedio 1,62 veces superior a la obtenida por los 16 participantes que se declararon sin experiencia (utilizando vídeos de entrenamiento para las tareas de clasificación del tipo espectral estelar). Estas diferencias en las tasas de acierto promedio (respectivamente, 0,78 y 0,48) presentaron

significación estadística [$F(1,30) = 7,12$, $p = 0,036$], con un tamaño de efecto pequeño [$d = 0,118$], lo que confirma que aunque las diferencias en las tasas promedio obtenidas por los participantes con y sin experiencia presentan una diferencia real, su magnitud es baja y probablemente tenga una relevancia pequeña en términos prácticos.

Respecto a los resultados por subgrupos, como cabría esperar, los músicos obtuvieron las tasas de acierto más elevadas (0,94), seguidos de los astrónomos músicos (0,81), los astrónomos (0,58) y los participantes sin experiencia (0,5). Estas diferencias en las tasas de acierto promedio resultaron estadísticamente significativas [$H(3) = 7,91$, $p = 0,048$]. El análisis en profundidad demostró que la tasa promedio del grupo de músicos es significativamente diferente de la del grupo de astrónomos [$p = 0,017$] y de la del grupo de participantes sin experiencia [$p = 0,031$]. Respecto al tamaño del efecto, en ambos casos se obtuvo un efecto medio, con $d = 0,52$ en la comparación con los astrónomos y $d = 0,49$ en la comparación con los participantes sin experiencia.

Los resultados notablemente superiores obtenidos por los músicos, confirman que el entrenamiento en percepción auditiva aumenta la comprensión de las sonificaciones, lo que concuerda con los reportes de Walker y Nees, 2005, 2011. Del mismo modo, los resultados aceptables obtenidos por el grupo de participantes sin experiencia sugieren que incluso careciendo de conocimientos previos en los campos relacionados, y a pesar de la dificultad para diferenciar sonidos declarada por algunos participantes en el apartado de observaciones del cuestionario, los vídeos de autoentrenamiento facilitaron la comprensión de la propuesta, algo que puede ser de utilidad en proyectos de divulgación y ciencia ciudadana.

La diferencia observada entre los grupos de músicos y astrónomos confirma que el entrenamiento en el análisis de fuentes sonoras fue más útil que el conocimiento previo de los datos en la tarea propuesta. En este sentido, cabe mencionar cómo varios de los astrónomos profesionales que participaron en el estudio comentaron que el mapeado simbólico utilizado en la Sección 6.2 les había resultado ligeramente desorientador debido a su expectativa de una asociación ordenada de la altura de nota musical (pitch) con la temperatura, según los tipos espectrales (de O a M o viceversa. Ver Figura 6.2). Este comentario subrayó la importancia de la definición de los usuarios finales en el diseño de sonificación y confirmó cómo las aplicaciones orientadas al análisis científico deben respetar la lógica asociada a cada variable representada. El mapeado descrito en la Sección 6.2, se realizó con la intención de proporcionar un método basado en el conocimiento común musical básico de la escala de Do mayor y en la coincidencia simbólica casi directa de las letras, bajo la hipótesis de que esta conversión permitiría a cualquier persona identificar los espectros, con independencia de sus conocimientos previos de Astronomía. El objetivo final consistía en analizar si el conocimiento musical previo podría ayudar en las tareas de clasificación y, al mismo tiempo, proporcionar una metodología con la que comparar el método de sonificación mediante deep learning. No obstante, el comentario de los astrónomos profesionales demuestra cómo la subjetividad en la toma de decisiones relacionadas con el diseño de sonificación determina en gran medida la adopción y aceptación de las propuestas.

12.4 Sobre la sonificación basada en deep learning

Aunque el espacio latente de los autoencoders a menudo es difícil de interpretar, los estudios presentados en este trabajo, que utilizan los espectros de galaxias de CALIFA y las librerías estelares STELIB y MILES, muestran respectivamente el alto potencial de la conversión de vectores latentes en frecuencia para permitir la clasificación sonora de espectros estelares y la diferenciación de zonas de formación estelar y galaxias viejas. Tal y como se muestra en la Sección 7.2 y en el Anexo .15, la metodología basada en autoencoders captura eficazmente las características generales de los datos espectrales, permitiendo la reducción de cada espectro a una representación sonora inteligible de un solo evento.

El estudio presentado en la Sección 11.2.3 demuestra además que los participantes entrenados en el análisis sonoro pueden tener éxito en las tareas de clasificación espectral mediante sonido utilizando esta estrategia de sonificación, e ilustra el potencial de la sonificación basada en deep learning para la clasificación sonora de tipos espectrales.

La conversión de los vectores latentes generados mediante esta metodología en acordes de ondas sinusoidales genera sonificaciones que minimizan la subjetividad y arbitrariedad en el diseño de sonificación, maximizando la objetividad de la representación y haciendo de la sonificación basada en deep learning un método formal de alto potencial para el análisis científico.

El conjunto global de 38 participantes del estudio II (Sección 11.2.1) obtuvo una tasa de acierto promedio (0,41) estadísticamente significativa respecto a la tasa de acierto aleatorio [$t(37) = 23,79$, $p < 0,001$], lo que indica que los participantes pudieron recuperar información de las sonificaciones gracias a los vídeos de entrenamiento. La diferencia del promedio de acierto de los participantes BLV con la tasa de acierto aleatorio no resultó estadísticamente significativa [$t(1) = 1,25$, $p = 0,3$], lo que no permite asegurar que sus resultados difieran de una respuesta al azar.

Atendiendo a las diferencias entre participantes con y sin experiencia en Astronomía y Música, los 16 participantes que se declararon con experiencia obtuvieron una tasa de acierto promedio 1,24 veces superior a la obtenida por los 16 participantes que se declararon sin experiencia (utilizando vídeos de entrenamiento para las tareas de clasificación de tipo espectral estelar). Estas diferencias en las tasas de acierto promedio (respectivamente, 0,48 y 0,39) no presentaron significación estadística [$F(1,30) = 2$, $p = 0,2$] por lo que los resultados deben considerarse indicativos.

Respecto a los resultados por subgrupos, de nuevo los músicos obtuvieron las tasas de acierto más elevadas (0,69), seguidos de los astrónomos músicos (0,5), los participantes sin experiencia (0,44) y finalmente los astrónomos (0,08), que obtuvieron resultados muy bajos. Estas diferencias en las tasas de acierto promedio resultaron estadísticamente significativas [$H(3) = 10,15$, $p = 0,017$]. El análisis en profundidad demostró que la tasa promedio del grupo de músicos es significativamente diferente de la del grupo de astrónomos [$p = 0,012$], con un efecto elevado [$d = 0,81$]. Este resultado apunta nuevamente a la necesidad de programas de entrenamiento específico para la consolidación del análisis sonoro en los procesos científicos.

Los resultados cualitativos del estudio II (11.10), obtenidos a partir de las respuestas de

40 participantes, parecen revelar una buena aceptación de las metodologías de sonificación propuestas. Un 90 % de los participantes consideró que los enfoques de sonificación simbólica y mediante deep learning son útiles para introducir la clasificación espectral a un público no experimentado en actividades de divulgación. Un 72,5 % de los participantes declaró interés en escuchar más sonificaciones después de completar la encuesta.

12.5 Sobre el análisis comparativo de metodologías de sonificación

Mediante el establecimiento de un marco de referencia, representado por la propuesta de sonificación simbólica (considerada a priori como potencialmente sencilla debido a que se basa en el reconocimiento de las notas musicales de la escala de Do mayor), hemos podido extraer conclusiones comparativas sobre la metodología experimental de sonificación basada en deep learning que hemos desarrollado en esta investigación. La comparación de ambos enfoques, presentada en la Sección 11.2.4, constituye así mismo una propuesta metodológica para la evaluación práctica del potencial de futuras estrategias de sonificación.

Tal y como muestra la Figura 11.9, la comparativa permite apreciar una reducción de las tasas de acierto promedio utilizando la sonificación mediante deep learning, lo que confirma que este enfoque requiere más entrenamiento que el enfoque simbólico. Las diferencias observadas resultaron estadísticamente significativas [$t(74) = 2,2278$, $p = 0,041$] y relevantes en términos prácticos [$d = 1,05$].

Atendiendo a los subgrupos, también puede verse cómo las personas con experiencia en Música clasificaron de manera más eficaz los espectros estelares con independencia de la metodología utilizada, sugiriendo que, con suficiente entrenamiento, la sonificación mediante deep learning no introduciría dificultades de interpretación adicionales.

Los conocimientos musicales básicos de los participantes resultaron de utilidad en las preguntas de mapeado simbólico, incluso para los grupos con menos experiencia. Por otra parte, la abstracción de las sonificaciones generadas mediante deep learning redujo la diferencia entre los aciertos de los participantes con y sin experiencia ya que las representaciones resultaron completamente nuevas para todos los participantes.

12.6 Sobre la espacialización de las sonificaciones

La espacialización binaural y ambisónica de los sonidos representados a lo largo de los capítulos de esta tesis proporciona una dimensión adicional inmersiva al conjunto de datos, que podría implementarse en la sonificación de archivos astronómicos masivos de acceso abierto. Tal y como mostramos en los Capítulos 6, 7 y 10, la representación sonora espacializada a partir de las coordenadas reales de los objetos tiene un papel importante en la generación de experiencias inmersivas, facilita la diferenciación entre sonificaciones y puede proporcionar una experiencia de mayor interés para la audiencia (Carlile, 2011), completando la información contenida en los mapas de cobertura de los archivos astronómicos.

El desarrollo de sonificaciones ambisónicas y binaurales puede ser de utilidad tanto en actividades de divulgación científica con público, como en análisis individuales utilizando auriculares. Futuras mejoras de las implementaciones binaurales propuestas podrían incluir la captura de HRTF personalizadas, ya que se ha demostrado que este aspecto puede tener un impacto significativo en la precisión de las espacializaciones (Jones y Fenton, 2023). Por otra parte, dado que las representaciones ambisónicas permiten comunicar resultados a audiencias más amplias, utilizarlas en exposiciones e instalaciones educativas o artísticas puede ayudar a fomentar la divulgación científica y las disciplinas STEM. Futuros desarrollos de espacialización ambisónica podrían explorar implementaciones de orden superior, mejorando también la precisión espacial de las representaciones.

12.7 Sobre las composiciones musicales generadas a partir de datos astronómicos

Asumiendo que “para que una sonificación represente información de manera significativa, la información debe ser parte de la experiencia de la representación” (Worrall, 2019), el sistema de composición autónomo descrito en el Capítulo 8 ha demostrado ser una herramienta capaz de representar información estelar a través del sonido, que podría utilizarse de forma eficaz en la creación de bandas sonoras originales para películas y contenido audiovisual. Como puede apreciarse en las composiciones presentadas y publicadas a lo largo del desarrollo de este trabajo (ver Sección 8.5 y Anexo .18), el sistema tiene el potencial de generar piezas completamente originales a partir de los datos de entrada de diversos archivos astronómicos.

Las composiciones musicales incluidas en esta tesis partieron del contexto relacionado con los intereses musicales de Johannes Kepler (Capítulo 8), explorando las posibilidades de la Música Renacentista en el entrenamiento de redes neuronales para realizar composiciones autónomas relacionadas con uno de los pioneros del binomio Música-Astronomía, en lo que podría interpretarse como una revisión empírica y automatizada de las teorías clásicas sobre *Armonías de las Esferas*.

Al mismo tiempo, estas creaciones musicales han servido para proponer nuevas líneas de desarrollo en la representación multimodal a través del sonido. A modo de ejemplo, las características altamente deterministas de la música renacentista podrían ayudar en el análisis y entendimiento del comportamiento de algoritmos de aprendizaje automático entrenados con esta música, algo que también podría resultar de utilidad para la monitorización sonora del comportamiento de redes neuronales.

12.8 Sobre la utilización de instrumentos musicales en la exploración de catálogos

El uso de instrumentos musicales digitales en la exploración de datos astronómicos se propone en esta investigación como un enfoque interdisciplinar que combina ciencia y arte para proporcionar una herramienta innovadora que fomenta la creatividad, la accesibilidad y la inclusión dentro del marco de análisis de datos astronómicos. Se espera que este enfoque

experimental sea de utilidad para futuros desarrollos que podrían incluir cualquier tipo de controlador MIDI y/o OSC para explorar las librerías y bases de datos del Observatorio Virtual, aumentando el interés y la participación del público en la investigación científica.

12.9 Sobre las herramientas multimodales interactivas

En cuanto a la aplicación descrita en el Capítulo 10, *ViewCube* ha demostrado su capacidad para transmitir información concreta de las galaxias de CALIFA. Los resultados de las pruebas realizadas por participantes con experiencia y sin experiencia en el campo de la Astronomía y la Música sugieren que el modelo interactivo multimodal propuesto puede mejorar el proceso de análisis de cubos de datos y hacer que la espectroscopía 3D sea más accesible para los investigadores BLV. Este enfoque sirve como una herramienta de sonificación inicial diseñada para proporcionar una visión general de las propiedades espectrales, lo que puede ayudar en el análisis de los estudios espectroscópicos masivos actuales y futuros.

La evaluación cuantitativa de la aplicación, realizada en el estudio III (Sección 11.3), analizó su potencial para la estimación mediante sonido de: 1) la posición de un spaxel seleccionado en el paisaje sonoro virtual generado por un cubo de datos (izquierda, derecha, frontal o trasera); 2) la distancia del spaxel seleccionado al centro de la galaxia representada (cerca del centro, distancia intermedia o lejos del centro); 3) el tipo/edad del espectro del spaxel seleccionado (región de formación estelar, edad intermedia o galaxia vieja); y 4) las tres características combinadas.

El conjunto global de 65 participantes obtuvo una tasa de acierto promedio (0,51) estadísticamente significativa respecto a la tasa de acierto aleatorio [$t(64) = 10,9$, $p < 0,001$]. Los participantes BLV también obtuvieron una tasa de acierto promedio (0,54) estadísticamente significativa respecto a la tasa de acierto aleatorio [$t(1) = 3,25$, $p = 0,007$].

Estos resultados indican que todos los participantes (con y sin experiencia en Música y Astronomía, BLV y no BLV) pudieron recuperar información de las sonificaciones gracias a las pruebas personalizadas y a los vídeos de entrenamiento, sugiriendo el potencial de la aplicación para proporcionar un entorno horizontal de análisis para personas BLV y no BLV.

Los astrónomos profesionales obtuvieron resultados 1,15 veces superiores a los resultados de los participantes no astrónomos, sugiriendo una vez más la importancia del contexto y los conocimientos previos en la interpretación de la sonificaciones. No obstante, estas diferencias en las tasas de acierto promedio (respectivamente, 0,55 y 0,48) no presentaron significación estadística [$F(1, 58) = 0,296$, $p = 0,59$], por lo que deben considerarse indicativas.

Respecto a los resultados por subgrupos, los astrónomos no músicos obtuvieron las tasas de acierto más elevadas (0,58), seguidos de los astrónomos músicos (0,54), los músicos (0,48) y los participantes sin experiencia (0,55). Estas diferencias en las tasas de acierto promedio tampoco resultaron estadísticamente significativas [$H(3) = 3,03$, $p = 0,38$], por lo que también deben considerarse indicativas y analizarse con una muestra de mayor tamaño en futuros estudios. No obstante, los resultados sugieren que la atención adicional requerida para aprender aspectos de campos desconocidos podría haber ayudado a los participantes sin experiencia con las tareas propuestas, tal y como reportan van Riesen et al., 2022.

En cuanto a la evaluación cualitativa de la interactividad de la aplicación, obtenida a partir de las respuestas de los 42 participantes (incluidos 21 astrónomos profesionales) que pudieron evaluarla en persona, se puede concluir que es buena o muy buena (80,95 %). Respecto a su utilidad y calidad sonora, un 79,1 % de la muestra completa (67 participantes, incluidos 31 astrónomos profesionales) la encontraron “Útil” o “Muy útil” y un 58,21 % calificaron su sonido como “Sonido bueno” o “Sonido agradable”.

Capítulo 13

Conclusión

Esta tesis tiene como objetivo principal demostrar el valor añadido de las sinergias multimodales que se producen entre sonido e imagen en la exploración de conjuntos de datos astronómicos. Para ello, desarrolla dos líneas de investigación paralelas. La primera, centrada en la sonificación aplicada al análisis científico. La segunda, dedicada a la exploración de las posibilidades de musificación y composición musical autónoma basada en datos. Ambos pilares se basan en el uso de deep learning para proponer soluciones de sonificación y musificación aplicadas a casos de estudio reales extraídos de las principales misiones internacionales de observación astronómica. Como resultado, este trabajo proporciona un amplio conjunto de implementaciones (ver Anexo.17), que pueden utilizarse para inspirar futuros diseños aplicables a diversas líneas de investigación y desarrollo.

Por un lado, en el ámbito científico, los prototipos y pruebas de concepto presentadas en cada capítulo pueden servir para permitir a investigadores con y sin problemas de visión analizar sus casos de estudio de una manera inclusiva y horizontal, para mejorar la accesibilidad de archivos y catálogos astronómicos y para complementar las representaciones gráficas utilizadas en comunicaciones científicas, en el desarrollo de actividades educativas y de divulgación o en actividades para fomentar las disciplinas STEM.

Por otra parte, en el ámbito creativo, las herramientas desarrolladas pueden tener aplicación en diseño sonoro, en composición de bandas sonoras originales o incluso en interpretaciones y presentaciones musicales en vivo. Desde el punto de vista de la tecnología musical, este trabajo analiza las posibilidades que ofrecen los sistemas binaurales y ambisónicos y desarrolla sintetizadores de audio que permiten generar texturas sonoras a partir de espectros estelares. Además, el capítulo dedicado al diseño de un sistema de composición autónomo permite la generación de composiciones musicales originales a partir de conjuntos de datos astronómicos.

El proceso de investigación se ha apoyado en cuatro líneas metodológicas (análisis bibliográfico multidisciplinar, implementación práctica de los diseños y estudio cuantitativo y cualitativo de las propuestas), las cuales han permitido cumplir con los objetivos iniciales del proyecto, centrados en la exploración de archivos de las agencias espaciales internacionales a través del sonido y en la promoción del uso de información astronómica para aplicaciones creativas y de composición musical.

A través de ejemplos que ilustran las relaciones entre Música y Astronomía a lo largo de la historia, hemos descrito el estado del arte en materia de análisis de datos astronómicos, sonificación, música algorítmica e inteligencia artificial aplicada a la creación musical. Partiendo de estos precedentes, hemos identificado oportunidades interdisciplinares de investigación que son descritas y resueltas mediante implementaciones y pruebas de concepto. Entre estas oportunidades, destacan dos proyectos de colaboración que hemos realizado respectivamente con el Space Telescope Science Institute (gracias al soporte del Sonification Lab de la Georgia Tech) y con el Instituto de Astrofísica de Andalucía, para la implementación de soluciones de sonificación en aplicaciones de análisis espectral. Finalmente, hemos realizado tres estudios cuantitativos y cualitativos basados en cuestionarios online y pruebas presenciales que nos han permitido analizar la funcionalidad de las pruebas de concepto desarrolladas y registrar las preferencias de un total de 153 participantes con y sin experiencia en Astronomía, Astrofísica y Música.

Centrada en la construcción de conexiones sólidas entre la Tecnología Musical y la Astronomía o Astrofísica, esta investigación comenzó con la formulación de tres preguntas principales destinadas a evaluar las capacidades del deep learning aplicado a la sonificación y a la generación de música a partir de catálogos y bases de datos astronómicos.

En cuanto a la primera pregunta (**Q1**), que cuestionaba si la conversión de datos astronómicos en sonido podría ser útil para el análisis científico en Astronomía y Astrofísica, el trabajo ha demostrado el potencial de las representaciones sonoras para complementar las representaciones gráficas de información científica con los proyectos de colaboración descritos en la Sección 3.5 y en el Capítulo 10. Ambas colaboraciones proporcionan dos ejemplos prácticos de prototipos de sonificación integrados en entornos profesionales de análisis científico en el campo de la Astronomía y la Astrofísica.

Por una parte, la colaboración con el Space Telescope Science Institute (STScI) para el diseño e implementación de un módulo de sonificación de espectros dentro de *Astronify*, se centró en la determinación de un modelo capaz de proporcionar sonificaciones preliminares que puedan ser útiles en la exploración rápida de archivos astronómicos. La utilidad de *Astronify* para el descubrimiento científico y la mejora de la accesibilidad en el análisis de curvas de luz ha sido evaluada por Harrison et al., 2022 y demostrada en el proyecto de ciencia ciudadana *Whispers from the Universe* (Sulis, 2024), en estudios dedicados al análisis de la influencia de la relación señal ruido en la detección de eventos (Tucker Brown et al., 2022) y en la evaluación de las diferencias entre astrónomos y no astrónomos en la tarea de detección de tránsitos en series temporales (Fovino et al., 2024). La ampliación de estos estudios, basados en curvas de luz, al campo de análisis espectral utilizando el módulo desarrollado durante esta colaboración, constituye una línea de trabajo futuro que tiene el potencial de acercar este tipo de análisis a personas dedicadas a la investigación, estudiantes y entusiastas de la Astronomía con y sin problemas de visión.

Esta colaboración, motivó además el estudio I (Sección 11.1.1), dedicado a la identificación de la estrategia de sonificación más eficaz para revelar el número de líneas de emisión y absorción en espectros de galaxias. Los recuentos cuantitativos de acierto por pregunta para los 46 participantes del estudio permitieron apuntar a la representación secuencial mediante osciladores sinusoidales como la representación más efectiva para la tarea propuesta. Este resultado

confirmó la idoneidad del diseño de sonificación del módulo espectral de *Astronify*. El estudio desveló diferencias apreciables pero no estadísticamente significativas entre las tasas promedio de acierto de los subgrupos de astrónomos músicos, músicos, astrónomos y participantes sin experiencia. Todas ellas con diferencias estadísticamente significativas respecto a la tasa de acierto aleatoria, lo que demuestra que los participantes extrajeron información útil de las sonificaciones. Los mejores resultados fueron obtenidos por los astrónomos músicos, aunque la dificultad encontrada por los participantes para contabilizar a través del sonido las líneas espectrales superpuestas en las representaciones gráficas, hizo que todos los participantes obtuvieran resultados relativamente bajos. Este hecho pone de manifiesto, en concordancia con las conclusiones reportadas por Walker y Nees, 2011, Tucker Brown et al., 2022 y Harrison et al., 2022, la importancia del entrenamiento auditivo para conseguir una integración efectiva de la sonificación en este tipo de análisis altamente especializados.

Por otro lado, la colaboración con el Instituto de Astrofísica de Andalucía (IAA) para la integración del módulo *SoniCube* en *ViewCube*, permitió la creación de una aplicación de IFS multimodal interactiva que proporciona un entorno inmersivo para la exploración de cubos de datos. La evaluación cuantitativa de esta aplicación, descrita en la Sección 11.3 y realizada por 42 participantes que pudieron probarla en persona y 25 que utilizaron vídeos de entrenamiento, demostró que todos los participantes (con y sin experiencia en Astronomía y Música, BLV y no BLV) fueron capaces de recuperar información de las sonificaciones, presentando tasas de acierto promedio con diferencias estadísticamente significativas respecto a la tasa de acierto aleatorio. Aunque también se observaron diferencias apreciables en las tasas de acierto promedio de los subgrupos de astrónomos músicos, músicos, astrónomos, participantes sin experiencia, participantes BLV y no BLV, dichas diferencias no alcanzaron los niveles convencionales de confianza estadística y deberán ser examinadas con una muestra de mayor tamaño en futuros estudios. No obstante, los resultados indicativos obtenidos sugieren el potencial de la aplicación para proporcionar un entorno horizontal de análisis para personas dedicadas a la investigación, con y sin problemas de visión. Los astrónomos no músicos obtuvieron los mejores resultados, confirmando la importancia del conocimiento previo del campo de estudio y del contexto para la comprensión eficaz de las sonificaciones, especialmente en aplicaciones altamente específicas. Los participantes pudieron diferenciar con éxito mediante sonido diversas características morfológicas y evolutivas en el estudio de galaxias de CALIFA, tales como la ubicación de los puntos seleccionados dentro de una galaxia, la distancia al centro de la galaxia de los spaxels seleccionados y la edad o el tipo de galaxia, permitiendo la diferenciación sonora entre regiones de formación estelar, galaxias intermedias y galaxias viejas. En cuanto a la evaluación cualitativa (Sección 11.3.4), los resultados sugieren que la interactividad, utilidad y estética sonora de la aplicación son buenas, lo que potencialmente puede contribuir a mejorar la experiencia de análisis de cubos de datos.

La segunda pregunta de investigación (**Q2**) cuestionaba si es posible extraer información esencial de catálogos estelares utilizando inteligencia artificial, que pueda entenderse mediante sonificación basada en criterios no arbitrarios.

Para contestar esta pregunta, este trabajo propone y evalúa el uso de deep learning en la sonificación de datos astronómicos (Capítulo 7). El principal valor de esta metodología radica

en la reducción de la arbitrariedad asociada a las decisiones de diseño que requieren otros tipos de metodologías de sonificación, especialmente cuando los resultados se mapean directamente a conjuntos de osciladores sinusoidales. Dado que el proceso desarrollado por redes neuronales puede considerarse en última instancia como una optimización o minimización del error, las decisiones involucradas en el proceso de sonificación pueden justificarse formalmente, tal y como se muestra en los Anexos .8, .7 y .8.3, proporcionando una “objetividad” a las representaciones sonoras que hace que la sonificación basada en deep learning sea especialmente apropiada para el análisis científico.

El estudio II (Capítulo 11.2.1, Sección 11.2.3), realizado con 40 participantes con y sin experiencia en Astronomía y Música, ofrece resultados prometedores que sugieren que los participantes de la encuesta en línea realizada pudieron recuperar información relevante de las sonificaciones basadas en deep learning con la ayuda de vídeos de entrenamiento. Todos los grupos de participantes, con la excepción del grupo BLV, obtuvieron tasas de acierto promedio estadísticamente significativas respecto a la tasa de acierto aleatorio. Los músicos obtuvieron los mejores resultados y se observaron diferencias estadísticamente significativas en las tasas promedio de músicos y astrónomos. Este resultado apunta nuevamente a la necesidad de programas de entrenamiento específico para la consolidación del análisis sonoro en los procesos científicos.

Con la intención de establecer una metodología de referencia que permitiera realizar comparaciones, también evaluamos la conversión directa simbólica de los tipos de espectros estelares en notas musicales (Sección 11.2.2). Esta sonificación simbólica se consideró como la metodología que probablemente sería más efectiva para la clasificación de los tipos espectrales por usuarios con y sin experiencia en Astronomía y Música, ya que se basa en los conocimientos musicales básicos de los usuarios (escala de Do mayor). Los resultados obtenidos en su evaluación confirmaron la hipótesis de partida. Todos los grupos de participantes, con la excepción de los BLV, obtuvieron tasas de acierto promedio estadísticamente significativas respecto a la tasa de acierto aleatorio. También se observaron diferencias estadísticamente significativas entre los grupos con y sin experiencia. Destacaron los resultados de los subgrupos de músicos y astrónomos músicos, aunque las diferencias con el resto de grupos no resultaron estadísticamente significativas, por lo que los resultados deben considerarse indicativos y estudiarse en futuros trabajos con una muestra de mayor tamaño. No obstante, la tendencia observada sugiere que el entrenamiento en percepción auditiva aumenta la comprensión de las sonificaciones, tal y como apuntan Walker y Nees, 2005, 2011. Del mismo modo, los resultados aceptables obtenidos por el grupo de participantes sin experiencia sugieren que los vídeos de autoentrenamiento facilitaron la comprensión de la propuesta, algo que puede ser de utilidad en proyectos de divulgación y ciencia ciudadana. Cabe destacar, no obstante, que la arbitrariedad intrínseca en el mapeado simbólico directo resultó desorientadora para el grupo de astrónomos profesionales, que esperaban una ordenación secuencial asociada a la temperatura efectiva. Este hecho pone de manifiesto la necesidad de respetar la lógica asociada a cada variable representada en aplicaciones destinadas al análisis científico.

Finalmente, se realizó un análisis comparativo de las dos metodologías de sonificación, simbólica y basada en deep learning (Sección 11.2.4). La comparación de las tasas de acierto promedio obtenidas con ambos métodos de sonificación resultó estadísticamente significativa y

relevante en términos prácticos. Los resultados obtenidos mediante deep learning fueron más bajos en todos los grupos de participantes, confirmando que esta metodología requiere más entrenamiento que el enfoque simbólico para la tarea de clasificación sonora propuesta. También se observó una reducción en la diferencia entre los promedios de acierto de los participantes con y sin experiencia, lo que sugiere que el conocimiento musical básico ayudó en la clasificación de espectros mediante sonificación simbólica, mientras que las representaciones basadas en deep learning resultaron completamente nuevas para todos los participantes. El grupo de músicos, “entrenados” en el análisis de programas sonoros, fue capaz de clasificar eficazmente los espectros estelares utilizando indistintamente mapeados simbólicos y sonificaciones basadas en deep learning, lo que sugiere que, con suficiente entrenamiento, la sonificación basada en deep learning no introduciría dificultades de interpretación adicionales en comparación con una representación menos abstracta. Adicionalmente, la evaluación cualitativa registrada en la Sección 11.2.5 mostró una buena aceptación de las propuestas, con altos porcentajes de participantes encontrándolas de utilidad para tareas de clasificación sonora y manifestando interés en escuchar más sonificaciones después de completar el cuestionario.

La tercera pregunta de investigación (**Q3**), que cuestionaba si las técnicas actuales de deep learning basadas en criterios objetivables podrían generar música a partir de datos astronómicos para ofrecer una revisión empírica de la idea clásica de la *Música de las Esferas*, fue respondida en el Capítulo 8 con el diseño e implementación de un sistema de composición autónomo basado en deep learning. Una vez más, el formalismo subyacente en los modelos de autoencoders y redes neuronales recurrentes, proporciona una “objetividad” a las sonificaciones y musificaciones propuestas en este trabajo que permiten su utilización en el ámbito científico con un grado mínimo de arbitrariedad.

La pregunta Q3 también inspiró la creación de composiciones musicales a partir de lo que podría haber sido la inspiración musical de Johannes Kepler durante el proceso de escritura de su obra *Harmonices Mundi*. En la sección cualitativa del estudio I (Sección 11.1.3), se incluyeron varias preguntas orientadas a la evaluación cualitativa de alguna de estas composiciones. Las calificaciones obtenidas muestran una buena aceptación de los resultados musicales generados con el sistema y proporcionan información sobre las preferencias de la audiencia, algo que podría ser de utilidad para guiar la estética de futuros diseños de sonificación y musificación. Tal y como se muestra en la Sección 8.5 y en los Anexos .18 y .22, el sistema permitió generar piezas musicales originales a partir de las obras seleccionadas de Orlando di Lasso, Giovanni Pierluigi da Palestrina, Josquin Des Prez, Jean Fèry Rebel, J. S. Bach, Erik Satie, W.A. Mozart, Frédéric Chopin y Miles Davis.

El uso de instrumentos musicales digitales para la exploración de catálogos astronómicos también fue propuesto en el Capítulo 9 como método de análisis “musicalmente ordenado”. Este enfoque ha contribuido a demostrar que los datos astronómicos pueden utilizarse en la creación de sonidos originales que puedan emplearse en interpretaciones y presentaciones públicas en vivo. La propuesta, basada en deep learning pre-procesado “a tiempo real”, constituye un puente entre la interpretación musical y la exploración de datos astronómicos, con un elevado potencial para su uso en espectáculos y demostraciones audiovisuales.

A modo de conclusión final, podemos afirmar que las estrategias de sonificación y musificación presentadas en esta tesis proporcionan un amplio ecosistema de herramientas que pueden

utilizarse en aplicaciones de análisis científico, divulgación científica y actividades que fomenten las disciplinas STEM, así como en aplicaciones de composición musical, diseño de sonido y arte sonoro. No obstante, resulta necesario mencionar la necesidad de invertir esfuerzos y recursos en el desarrollo de futuros estudios que permitan alcanzar la representatividad estadística que confirme las tendencias presentadas y que puedan facilitar el avance hacia una mayor aceptación de las representaciones sonoras en aplicaciones de análisis científico.

En su conjunto, esta investigación puede considerarse un punto de partida para futuros desarrollos hacia la creación de un observatorio virtual sonoro, que podría permitir la exploración multimodal de librerías y catálogos astronómicos públicos, mejorando su accesibilidad y favoreciendo la inclusión en el campo de la ciencia espacial.

Referencias

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). TensorFlow: a system for Large-Scale machine learning [<https://dl.acm.org/doi/10.5555/3026877.3026899>]. *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 265-283.
- Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., et al. (2023). Musiclm: Generating music from text. *arXiv preprint arXiv:2301.11325*. <https://doi.org/10.48550/arXiv.2301.11325>
- Albini, G. (2018). Combinatorics, probability and choice in music composition: Towards an aesthetics of composing systems for non-musicians. *Proceedings of Bridges 2018: Mathematics, Art, Music, Architecture, Education, Culture*, 395-398.
- Alexander, R. L., O'Modhrain, S., Roberts, D. A., Gilbert, J. A., & Zurbuchen, T. H. (2014). The bird's ear view of space physics: Audification as a tool for the spectral analysis of time series data. *Journal of Geophysical Research: Space Physics*, 119(7), 5259-5271. <https://doi.org/10.1002/2014ja020025>
- Allington-Smith, J. (2006). Basic principles of integral field spectroscopy. *New Astronomy Reviews*, 50(4-5), 244-251. <https://doi.org/10.1016/j.newar.2006.02.024>
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Van Eesn, B. C., Awwal, A. A. S., & Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*. <https://doi.org/10.48550/arXiv.1803.01164>
- Anderton, C. (1987). The MIDI protocol. *Audio Engineering Society Conference: 5th International Conference: Music and Digital Technology*.
- Arcand, K., Russo, M., Santaguida, A., & Malec, C. (2020). A Universe of Sound [<https://chandra.harvard.edu/sound/>].
- Arcand, K. K., Schonhut-Stasik, J. S., Kane, S. G., Sturdevant, G., Russo, M., Watzke, M., Hsu, B., & Smith, L. F. (2024). A Universe of Sound: processing NASA data into sonifications to explore participant response. *Frontiers in Communication*, 9, 1288896. <https://doi.org/10.3389/fcomm.2024.1288896>
- Arviset, C., Gaudet, S., & IVOA Technical Coordination Group. (2012). The IVOA Architecture. *European Planetary Science Congress 2012*, Artículo EPSC2012-626, EPSC2012-626.
- Arviset, C., & Gaudet, S. (2010). *IVOA Architecture Version 1.0* (inf. téc.). IVOA Note 23 November 2010. arXiv 2011. <https://doi.org/10.5479/ads/bib/2010ivoa.rept.1123a>
- Astropy. (2011-2022a). Astropy [<https://docs.astropy.org/en/stable/index.html>].

- Astropy. (2011-2022b). Astropy Astrostatistics Tools [<https://docs.astropy.org/en/stable/stats/index.html>].
- AURA. (2016). Astro Data Lab [<https://datalab.noirlab.edu/docs/manual/Appendices/Notebooks/Notebooks.html>].
- Auvergne, M., Bodin, P., Boissard, L., Buey, J.-T., Chaintreuil, S., Epstein, G., Jouret, M., Lam-Trong, T., Levacher, P., Magnan, A., et al. (2009). The CoRoT satellite in flight: description and performance. *Astronomy & Astrophysics*, 506(1), 411-424. <https://doi.org/10.1051/0004-6361/200810860>
- Babcock, J., & Bali, R. (2021). *Generative AI with Python and TensorFlow 2: Harness the power of generative models to create images, text, and music*. Packt Publishing, Limited.
- Bahdanau, D. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*. <https://doi.org/10.48550/arXiv.1409.0473>
- Baldi, P. (2012). Autoencoders, unsupervised learning, and deep architectures [<https://dl.acm.org/doi/10.5555/3045796.3045801>]. *Proceedings of ICML workshop on unsupervised and transfer learning*, 37-49.
- Ball, P. (2009). *The music instinct: how music works and why we can't do without it*. Vintage Books.
- Ballora, M. (2014). Sonification, Science and Popular Music: In search of the 'wow'. *Organised Sound*, 19(1), 30-40. <https://doi.org/10.1017/s1355771813000381>
- Barahona-Ríos, A. (2023). *Deep Learning for the Synthesis of Sound Effects* [Tesis doctoral, University of York].
- Barchet, A. V., Rimmele, J. M., & Pelofi, C. (2024). TenseMusic: An automatic prediction model for musical tension. *Plos one*, 19(1), e0296385. <https://doi.org/10.31234/osf.io/xck3w>
- Bardelli, S., Ferretti, C., Ludovico, L. A., Presti, G., & Rinaldi, M. (2021). A Sonification of the zCOSMOS Galaxy Dataset. *International Conference on Human-Computer Interaction*, 171-188. https://doi.org/10.1007/978-3-030-77411-0_12
- Barrass, S. (2012). The aesthetic turn in sonification towards a social and cultural medium. *AI & society*, 27, 177-181. <https://doi.org/10.1007/s00146-011-0335-5>
- Barrass, S., & Vickers, P. (2011). Sonification Design and Aesthetics. En T. Hermann, A. Hunt & J. G. Neuhoff (Eds.), *The Sonification Handbook* (pp. 145-171). Logos Publishing House.
- Barrett, G. D. (2021). Deep (Space) Listening: Posthuman Moonbounce in Pauline Oliveros's Echoes from the Moon. *Discourse*, 43(3), 321-350. <https://doi.org/10.13110/discourse.43.3.0321>
- Bartlett, B., Bartlett, J., & Bartlett, J. (2013). *On Location Recording Techniques*. Routledge. <https://doi.org/10.4324/9780080513041>
- Bartlett, B., & Bartlett, J. (2016). *Practical Recording Techniques: The step-by-step approach to professional audio recording*. Routledge. <https://doi.org/10.4324/9781315696331>
- Bastien, F. A. (2014). *Empirically Interrelating Stellar Chromospheric Activity, Photometric Variability and Radial Velocity Variations to Enhance Planet Discovery*. Vanderbilt University.
- Baudin, F., Baglin, A., Orcesi, J.-L., Nguyen-Kim, K., Solano, E., Ochseneim, F., Committee, C. S., et al. (2006). Everything You Always Wanted to Know about CoRoT Data.

- "Proceedings of The CoRoT Mission Pre-Launch Status-Stellar Seismology and Planet Finding"(ESA SP-1306). Editors: M. Fridlund, A. Baglin, J. Lochard and L. Conroy. ISBN 92-9092-465-9., p. 145", 1306, 145.*
- Baum, D. (2006). Emomusic-Classifying music according to emotion. *Proceedings of the 7th Workshop on Data Analysis (WDA2006)*.
- Bayo, A., Rodrigo, C., y Navascués, D. B., Solano, E., Gutiérrez, R., Morales-Calderón, M., & Allard, F. (2008). VOSA: virtual observatory SED analyzer-An application to the Collinder 69 open cluster. *Astronomy & Astrophysics*, 492(1), 277-287. <https://doi.org/10.1051/0004-6361:200810395>
- Belanger, O. (2016). Pyo, the python dsp toolbox. *Proceedings of the 24th ACM international conference on Multimedia*, 1214-1217. <https://doi.org/10.1145/2964284.2973804>
- Bernhardt, M., Cowell, C., & Oxford, W. (2007, abril). Sonification of hyperspectral image data. En S. S. Shen & P. E. Lewis (Eds.), *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIII* (p. 65650D, Vol. 6565). <https://doi.org/10.1117/12.719348>
- Biles, J., et al. (1994). GenJam: A genetic algorithm for generating jazz solos. *ICMC*, 94, 131-137.
- Biles, J. A., & Eign, W. G. (1995). GenJam Populi: Training an IGA via audience-mediated performance. *San Francisco, USA*, 347-348.
- Bly, S. (1982). *Sound and computer information presentation* (inf. téc.). Lawrence Livermore National Lab., CA (USA); California Univ., Davis (USA). <https://doi.org/10.2172/5221536>
- Boch, T., & Fernique, P. (2014). Aladin Lite: Embed your Sky in the browser. *Astronomical data analysis software and systems XXIII*, 485, 277.
- Boehringer, J., Pietruszewski, M., Bowers, J., Hogg, B., Newbold, J., Sharma, G., Shaw, T., & Vickers, P. (2022). Listener-centred Sonification Practice as Transdisciplinary Experimental Artistic Engagement. <https://doi.org/10.7203/PUV-OA-532-0>
- Bonet, N., Kirke, A., & Miranda, E. R. (2016). Sonification of dark matter: Challenges and opportunities.
- Bonet Filella, N. (2019). *Data Sonification in Creative Practice* [Tesis doctoral, University of Plymouth]. <https://doi.org/10.24382/937>
- Bonnarel, F., Fernique, P., Bienaymé, O., Egret, D., Genova, F., Louys, M., Ochsenbein, F., Wenger, M., & Bartlett, J. G. (2000). The ALADIN interactive sky atlas-a reference tool for identification of astronomical sources. *Astronomy and Astrophysics Supplement Series*, 143(1), 33-40. <https://doi.org/10.1051/aas:2000331>
- Borsos, Z., Marinier, R., Vincent, D., Kharitonov, E., Pietquin, O., Sharifi, M., Roblek, D., Teboul, O., Grangier, D., Tagliasacchi, M., et al. (2023). Audiolm: a language modeling approach to audio generation. *IEEE/ACM transactions on audio, speech, and language processing*, 31, 2523-2533. <https://doi.org/10.1109/TASLP.2023.3288409>
- Boulanger, R. (2000). *The Csound book: perspectives in software synthesis, sound design, signal processing, and programming*. MIT press.
- Bourlard, H., & Kamp, Y. (1988). Auto-association by multilayer perceptrons and singular value decomposition. *Biological cybernetics*, 59(4), 291-294. <https://doi.org/10.1007/bf00332918>

- Brasseur, C. E., Fleming, S. W., & Kotler, J. (2020). Astronify: An Open-Source Python package to sonify one-dimensional astronomical data [<https://astronify.readthedocs.io/en/latest/>].
- Briot, J.-P., Hadjeres, G., & Pachet, F.-D. (2020). *Deep learning techniques for music generation* (Vol. 1). Springer. <https://doi.org/10.1007/978-3-319-70163-9>
- Bristow, P., Kerber, F., & Rosa, M. R. (2008). Advanced calibration techniques for astronomical spectrographs. *The Messenger*, 131, p2.
- Brown, C. P., & Duda, R. O. (1997). An efficient HRTF model for 3-D sound. *Proceedings of 1997 Workshop on Applications of Signal Processing to Audio and Acoustics*, 4-pp. <https://doi.org/10.1109/aspaa.1997.625596>
- Brown, D. (2015). Exploring Skyscape in Stellarium. *Journal of Skyscape Archaeology*, 1(1). <https://doi.org/10.1558/jsa.v1i1.26956>
- Bufano, F., Riggi, S., Sciacca, E., & Schilliro, F. (2022). Machine Learning for Astrophysics. *International Conference 30 May*. <https://doi.org/10.1007/978-3-031-34167-0>
- Caballero, J., Arias, A., Machuca, J., & Morente, S. (2016). Music and astronomy. II. united-soundsofcosmos. *arXiv preprint arXiv:1611.07401*.
- Caballero, J. A. (2024). Music and astronomy. VI. Harmonia Mundi 2024.
- Caballero, J. A., Sánchez, S. G., & Caballero, I. (2008). Music and Astronomy. *arXiv preprint arXiv:0810.2032*.
- Caballero, J. A. (2012). Multiverso: Rock'n'Astronomy. *Rock'n'Astronomy*, 6.
- CAHA. (2025). Observatorio Astronómico de Calar Alto [<https://www.caha.es/>].
- Caltech. (2020). Educational resources. GROWTH AStromy School [<https://www.growth.caltech.edu/growth-astro-school-2020-resources.html>].
- Cameron, A. C., & Windmeijer, F. A. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of econometrics*, 77(2), 329-342. [https://doi.org/10.1016/S0304-4076\(96\)01818-0](https://doi.org/10.1016/S0304-4076(96)01818-0)
- Candey, R. M., Schertenleib, A. M., & Diaz Merced, W. (2006). Xsonify sonification tool for space physics.
- Cantrell, S. J., Walker, B. N., & Moseng, Ø. (2021). Highcharts Sonification Studio: an online, open-source, extensible, and accessible data sonification tool. *Proceedings of the International Conference on Auditory Display*. <https://doi.org/10.21785/icad2021.005>
- Carlile, S. (2011). Psychoacoustics. En T. Hermann, A. Hunt & J. G. Neuhoff (Eds.), *The Sonification Handbook* (pp. 41-61). Logos Publishing House.
- Carty, B. (2008). hrtfmove2 opcode [<http://www.csounds.com/manual/html/hrtfmove2.html>].
- Carver, N. (2023). WorlWide Telescope [<https://www.worldwidetelescope.org/>].
- Casado, J., & García, B. (2024). A multimodal approach to data analysis in astronomy: SONOUNO applications in photometry and spectroscopy. *RAS Techniques and Instruments*, 3(1), 625-635. <https://doi.org/https://doi.org/10.1093/rasti/rzae042>
- Casado, J., Diaz-Merced, W., & García, B. (2024). Analysis of astronomical data through sonification: reaching more inclusion for visual disable scientists. *arXiv preprint arXiv:2402.00611*. <https://doi.org/10.48550/arXiv.2402.00611>
- CDS. (1983). Strasbourg astronomical Data Center / Centre de données astronomiques de Strasbourg [<https://cds.unistra.fr/>].
- CDS. (2000a). Aladin [<http://aladin.cds.unistra.fr/aladin.gml>].
- CDS. (2000b). Simbad [<https://simbad.cds.unistra.fr/simbad/>].

- CDS. (2000c). VizieR [<https://vizier.cds.unistra.fr/index.gml>].
- Cenarro, A., Cardiel, N., Gorgas, J., Peletier, R., Vazdekis, A., & Prada, F. (2001). Empirical calibration of the near-infrared Ca II triplet—I. The stellar library and index definition. *Monthly Notices of the Royal Astronomical Society*, *326*(3), 959-980. <https://doi.org/10.1046/j.1365-8711.2001.04688.x>
- CfA. (2025). For Educators and Students. Harvard & Smithsonian Center for Astrophysics [<https://pweb.cfa.harvard.edu/resources/educators-and-students>].
- Charbonneau, J., Novak, C., Gaspar, R., & Ule, H. (2012). A-weighting the equal loudness contours. *J. Acoust. Soc. Am*, *131*, 3502-3508. <https://doi.org/10.1121/1.4709236>
- Chéreau, F., Wolf, A., Zotti, G., Hoffmann, S., Kabatsayev, R., Boonplod, W., Chéreau, G., & Bernardi, M. (2025). Stellarium (v25.1). <https://doi.org/10.5281/zenodo.15071984>
- CHIAPPETTI, L., Currie, M. J., Allen, S., Dobrzycki, A., Pence, W. D., Rots, A., Shaw, R., & Thompson, W. D. (2018). Definition of the Flexible Image Transport System (FITS) The FITS Standard Version 4.0: updated 2016 July 22 by the IAUFWG Original document publication date: 2016 July 22 Language-edited document publication date: 2018 August 13.
- Childs, E. (2002). Achorripsis: A sonification of probability distributions. *Proceedings of the 2002 International Conference on Auditory Display*, 2-5.
- Civit, M., Civit-Masot, J., Cuadrado, F., & Escalona, M. J. (2022). A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends. *Expert Systems with Applications*, *209*, 118190. <https://doi.org/10.1016/j.eswa.2022.118190>
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge. <https://doi.org/10.4324/9780203771587>
- Cointepas, M., Almenara, J., Bonfils, X., Bouchy, F., Astudillo-Defru, N., Murgas, F., Otegi, J., Wyttenbach, A., Anderson, D., Artigau, É., et al. (2021). TOI-269 b: an eccentric sub-Neptune transiting a M2 dwarf revisited with ExTrA. *Astronomy & Astrophysics*, *650*, A145. <https://doi.org/10.1051/0004-6361/202140328>
- Collins, N., McLean, A., Rohrhuber, J., & Ward, A. (2003). Live coding in laptop performance. *Organised sound*, *8*(3), 321-330. <https://doi.org/10.1017/s135577180300030x>
- Conway, J. H. (1978). Loopy games. En *Annals of Discrete Mathematics* (pp. 55-74, Vol. 3). Elsevier. [https://doi.org/10.1016/s0167-5060\(08\)70498-9](https://doi.org/10.1016/s0167-5060(08)70498-9)
- Cooke, J., Díaz-Merced, W., Foran, G., Hannam, J., & Garcia, B. (2017). Exploring data sonification to enable, enhance, and accelerate the analysis of big, noisy, and multi-dimensional data: workshop 9. *Proceedings of the International Astronomical Union*, *14* (S339), 251-256. <https://doi.org/10.1017/s1743921318002703>
- Cope, D. (1989). Experiments in musical intelligence (EMI): Non-linear linguistic-based composition. *Journal of New Music Research*, *18*(1-2), 117-139. <https://doi.org/10.1080/09298218908570541>
- Cope, D. (1991). *Computers and musical style* (Vol. 6). Citeseer.
- Copeland, J., & Long, J. (2016). Restoring the first recording of computer music.
- Copet, J., Kreuk, F., Gat, I., Remez, T., Kant, D., Synnaeve, G., Adi, Y., & Défossez, A. (2023). Simple and controllable music generation. *Advances in Neural Information Processing Systems*, *36*, 47704-47720. <https://doi.org/10.48550/arXiv.2306.05284>

- CSound. (2003). CSound signal generation opcodes [<https://csound.com/docs/manual/PartOpcodesOverview.html>].
- Ctcsound. (2022). Ctcsound library [<https://pypi.org/project/ctcsound/>].
- Cuthbert, M. S., & Ariza, C. (2010). music21: A toolkit for computer-aided musicology and symbolic music data.
- Dalba, P. A., Kane, S. R., Dragomir, D., Villanueva, S., Collins, K. A., Jacobs, T. L., LaCourse, D. M., Gagliano, R., Kristiansen, M. H., Omohundro, M., et al. (2022). The TESS-Keck Survey. VIII. Confirmation of a Transiting Giant Planet on an Eccentric 261 Day Orbit with the Automated Planet Finder Telescope. *The Astronomical Journal*, 163(2), 61. <https://doi.org/10.3847/1538-3881/ac415b>
- Damm, S., Forster, D., Velychko, D., Dai, Z., Fischer, A., & Lücke, J. (2023). The ELBO of variational autoencoders converges to a sum of entropies. *International Conference on Artificial Intelligence and Statistics*, 3931-3960. <https://doi.org/10.48550/arXiv.2010.14860>
- D'Arezzo, G. (1996). *Micrologus, traduction et commentaires de Marie-Noël Colette et Jean-Christophe Jolivet. Paris, Cité de la Musique.*
- de Amorim, A. L., García-Benito, R., Cid Fernandes, R., Cortijo-Ferrero, C., González Delgado, R. M., Lacerda, E. A. D., López Fernández, R., Pérez, E., & Vale Asari, N. (2017). The PyCASSO database: spatially resolved stellar population properties for CALIFA galaxies. *mnras*, 471(3), 3727-3752. <https://doi.org/10.1093/mnras/stx1805>
- de Campo, A. (2009). *Science by ear: an interdisciplinary approach to sonifying scientific data* [Tesis doctoral, Institute of Electronic Music, Acoustics - IEM, University of Music y Dramatic Arts, Graz].
- De La Vega, G., Dominguez, L. M. E., Casado, J., & García, B. (2022). SonoUno web: an innovative user centred web interface. *International Conference on Human-Computer Interaction*, 628-633. https://doi.org/10.1007/978-3-031-19679-9_79
- De Mantaras, R. L., & Arcos, J. L. (2002). AI and music: From composition to expressive performance. *AI magazine*, 23(3), 43-43. <https://doi.org/10.1609/aimag.v23i3.1656>
- Dean, R. T. (2018). *The Oxford handbook of algorithmic music*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190226992.001.0001>
- Dekking, F. M. (2005). *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media. <https://doi.org/10.1007/1-84628-168-7>
- Deleuil, M., Meunier, J., Moutou, C., Surace, C., Deeg, H., Barbieri, M., Debosscher, J., Almenara, J., Agneray, F., Granet, Y., et al. (2009). Exo-Dat: an information system in support of the CoRoT/exoplanet science. *The Astronomical Journal*, 138(2), 649. <https://doi.org/10.1088/0004-6256/138/2/649>
- Deng, Q., Yang, Q., Yuan, R., Huang, Y., Wang, Y., Liu, X., Tian, Z., Pan, J., Zhang, G., Lin, H., et al. (2024). Composerx: Multi-agent symbolic music composition with llms. *arXiv preprint arXiv:2404.18081*. <https://doi.org/10.48550/arXiv.2404.18081>
- Dhariwal, P., Jun, H., Payne, C., Kim, J. W., Radford, A., & Sutskever, I. (2020). Jukebox: A generative model for music. *arXiv preprint arXiv:2005.00341*. <https://doi.org/10.48550/arXiv.2005.00341>
- Diaz-Merced, W. L., Candey, R. M., Brickhouse, N., Schneps, M., Mannone, J. C., Brewster, S., & Kolenberg, K. (2011). Sonification of astronomical data. *Proceedings of the*

- International Astronomical Union*, 7(S285), 133-136. <https://doi.org/10.1017/s1743921312000440>
- Djorgovski, S. G., & Williams, R. (2005). Virtual observatory: From concept to implementation. *arXiv preprint astro-ph/0504006*. <https://doi.org/10.48550/arXiv.astro-ph/0504006>
- Doersch, C. (2016). Tutorial on variational autoencoders. *arXiv preprint arXiv:1606.05908*. <https://doi.org/10.48550/arXiv.1606.05908>
- Donahue, C., McAuley, J., & Puckette, M. (2018). Adversarial audio synthesis. *arXiv preprint arXiv:1802.04208*. <https://doi.org/10.48550/arXiv.1802.04208>
- Doornbusch, P. (2004). Computer sound synthesis in 1951: The music of CSIRAC. *Computer Music Journal*, 28(1), 10-25. <https://doi.org/10.1162/014892604322970616>
- Dowler, P., Evans, J., Arviset, C., Gaudet, S., Group, T. C., et al. (2021). IVOA Architecture Version 2.0. *IVOA Endorsed Note 01 November 2021*, 1101. <https://doi.org/10.5479/ads/bib/2021ivoa.spec.1101d>
- Downie, J. S. (2003). Music information retrieval. *Annual review of information science and technology*, 37(1), 295-340. <https://doi.org/10.1002/aris.1440370108>
- Dubnov, S., Assayag, G., & Gokul, V. (2022). Creative improvised interaction with generative musical systems. *2022 IEEE 5th International Conference on Multimedia Information Processing and Retrieval (MIPR)*, 121-126. <https://doi.org/10.1109/MIPR54900.2022.00028>
- Dubreuil, A. (2020). *Hands-on music generation with magenta: Explore the role of deep learning in music generation and assisted music composition*. Packt Publishing Ltd. <https://doi.org/10.0000/9781838825768-001>
- Dubus, G., & Bresin, R. (2013). A systematic review of mapping strategies for the sonification of physical quantities. *PloS one*, 8(12), e82491. <https://doi.org/10.1371/journal.pone.0082491>
- Ducharme, M. J. J. (2023). *CMB (Cosmic Microwave Background)* [Tesis doctoral, University of British Columbia].
- Dunn, D. (1992). A history of electronic music pioneers. *Eigenwelt der apparatewelt: Pioneers of electronic art*.
- Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American statistical association*, 56(293), 52-64. <https://doi.org/10.1080/01621459.1961.10482090>
- Enge, K., Elmquist, E., Caiola, V., Rönnerberg, N., Rind, A., Iber, M., Lenzi, S., Lan, F., Höldrich, R., & Aigner, W. (2024). Open Your Ears and Take a Look: A State-of-the-Art Report on the Integration of Sonification and Visualization. *Computer Graphics Forum*, 43(3), e15114. <https://doi.org/10.1111/cgf.15114>
- Engel, J., Agrawal, K. K., Chen, S., Gulrajani, I., Donahue, C., & Roberts, A. (2019). Gansynth: Adversarial neural audio synthesis. *arXiv preprint arXiv:1902.08710*. <https://doi.org/10.48550/arXiv.1902.08710>
- ESA. (2025). Gaia. ESA's billion star surveyor [https://www.esa.int/Science_Exploration/Space_Science/Gaia].
- ESO. (2010). Educational Material for ALMA [https://www.eso.org/public/products/education/edu_0068/].
- ESO. (2018). Large Early Galaxy Census (LEGA-C) Public Spectroscopic Survey [<http://www.eso.org/sci/publications/announcements/sciann17120.html>].

- ESO. (2025). Atacama Large Milimeter/submillimeter Array [<https://www.eso.org/public/teles-instr/alma/>].
- Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., Ricciardelli, E., Cardiel, N., Cenarro, A., Gorgas, J., & Peletier, R. (2011). An updated MILES stellar library and stellar population models. *Astronomy & Astrophysics*, 532, A95. <https://doi.org/10.1051/0004-6361/201116842>
- Feigelson, E. D., De Souza, R. S., Ishida, E. E., & Babu, G. J. (2021). Twenty-first-century statistical and computational challenges in astrophysics. *Annual Review of Statistics and Its Application*, 8(1), 493-517. <https://doi.org/10.1146/annurev-statistics-042720-112045>
- Fiebrink, R., & Cook, P. R. (2010). The Wekinator: a system for real-time, interactive machine learning in music. *Proceedings of The Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)(Utrecht)*, 3, 2-1.
- Fletcher, H., & Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *Bell System Technical Journal*, 12(4), 377-430. <https://doi.org/10.1002/j.1538-7305.1933.tb00403.x>
- Flowers, J. H. (2005). Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions.
- Foran, G., Cooke, J., & Hannam, J. (2022). The power of listening to your data: Opening doors and enhancing discovery using sonification. *Revista Mexicana de Astronomía y Astrofísica Serie de Conferencias (RMxAC)*, 54, 1-8. <https://doi.org/10.22201/ia.14052059p.2022.54.01>
- Forte, A. (1973). *The Structure of Atonal Music*. Yale University Press. <https://doi.org/10.12987/9780300156720>
- Fovino, L. G. N., Zanella, A., & Grassi, M. (2024). Evaluation of the Effectiveness of Sonification for Time-series Data Exploration. *The Astronomical Journal*, 167(4), 150. <https://doi.org/10.3847/1538-3881/ad2943>
- Fraknoi, A. (2008). Music Inspired by Astronomy. *Preparing for the International Year of Astronomy 2009*.
- Fraknoi, A. (2012). Music Inspired by Astronomy: A Resource Guide Organized by Topic. *Astronomy Education Review*, 11(1). <https://doi.org/10.3847/aer2012043>
- Franinovic, K., & Serafin, S. (2013). *Sonic interaction design*. Mit Press. <https://doi.org/10.7551/mitpress/8555.001.0001>
- Freudling, W., Romaniello, M., Bramich, D., Ballester, P., Forchi, V., García-Dabló, C., Moehler, S., & Neeser, M. (2013). Automated data reduction workflows for astronomy-The ESO Reflex environment. *Astronomy & Astrophysics*, 559, A96. <https://doi.org/10.1051/0004-6361/201322494>
- Gao, X., Chen, D. K., Gou, Z., Ma, L., Liu, R., Zhao, D., & Ham, J. (2024). AI-Driven Music Generation and Emotion Conversion. *Affective and Pleasurable Design*, 123(123). <https://doi.org/10.54941/ahfe1004679>
- Garcia, B., Diaz-Merced, W., Casado, J., & Cancio, A. (2019). Evolving from xSonify: a new digital platform for sonorization. *EPJ Web of Conferences*, 200, 01013. <https://doi.org/10.1051/epjconf/201920001013>
- García Martín, R., et al. (2009). La teoría de la armonía de las esferas en el libro quinto de *Harmonices Mundi* de Johannes Kepler.

- García Riber, A., & Serradilla, F. (2024a). Deep learning sonification synthesizer for the CaT stellar spectra library [aes2.org/publications/elibrary-page/?id=22537]. *AES Convention 156, Madrid, Paper Number 191*, 1-8.
- García Riber, A. (2019). Sonifigrapher. Sonified light curve synthesizer. *Proc. 25th International Conference on Auditory Display (ICAD 2019)*, 62-66. <https://doi.org/10.21785/icad2019.016>
- García Riber, A. (2022). FITS2OSC-A Sonification Pipeline for Light Curve Interactive Auditory Exploration. *Astronomical Society of the Pacific Conference Series*, 532, 121. <https://doi.org/2022ASPC..532..121G>
- García Riber, A., & Serradilla, F. (2022). Sonification of TESS data validation timeseries files. <https://doi.org/10.21785/icad2022.024>
- García Riber, A., García-Benito, R., & Serradilla, F. (2024). Interactive multimodal integral field spectroscopy. *RAS Techniques and Instruments*, 3(1), 748-758. <https://doi.org/10.1093/rasti/rzae049>
- García Riber, A., & Serradilla, F. (2023a). AI-rmonies of the Spheres. *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*, 132-147. https://doi.org/10.1007/978-3-031-29956-8_9
- García Riber, A., & Serradilla, F. (2023b). Spectral AI-rmonies of Classic Forms [<https://hdl.handle.net/1853/72890>]. *Proc. 28th International Conference on Auditory Display (ICAD 2023)*, 191-192.
- García Riber, A., & Serradilla, F. (2024b). CoRoT light curve Ambisonics/Binaural Sonification. *The 29th International Conference on Auditory Display (ICAD 2024)*. <https://doi.org/10.21785/icad2024.010>
- García Riber, A., & Serradilla, F. (2024c). Toward an Auditory Virtual Observatory. *Journal of the Audio Engineering Society*, 72(5), 341-351. <https://doi.org/10.17743/jaes.2022.0146>
- García-Benito, R. (2019). Cosmonic [<http://rgb.iaa.es/cosmonic/>].
- García-Benito, R. (2023). Astronomy & Astrophysics in ICAD History. *arXiv preprint arXiv:2311.12101*.
- García-Benito, R., Zibetti, S., Sánchez, S., Husemann, B., De Amorim, A., Castillo-Morales, A., Fernandes, R. C., Ellis, S., Falcón-Barroso, J., Galbany, L., et al. (2015). Califa, the calar alto legacy integral field area survey-iii. second public data release. *Astronomy & Astrophysics*, 576, A135. <https://doi.org/10.1051/0004-6361/201425080>
- Gardner, W. G. (1998). Reverberation algorithms. En *Applications of digital signal processing to audio and acoustics* (pp. 85-131). Springer. https://doi.org/10.1007/0-306-47042-x_3
- Gargaud, M., & Amils, R. (2011). *Encyclopedia of astrobiology* (Vol. 1). Springer Science & Business Media.
- Giacalone, S., Dressing, C. D., Hedges, C., Kostov, V. B., Collins, K. A., Jensen, E. L., Yahalom, D. A., Bieryla, A., Ciardi, D. R., Howell, S. B., et al. (2022). Validation of 13 Hot and Potentially Terrestrial TESS Planets. *The Astronomical Journal*, 163(2), 99. <https://doi.org/10.3847/1538-3881/ac4334>
- GLAS-Education. (2020). Sonification World Chat [<https://geosonnet.org/>].
- Gong, J., Zhao, W., Wang, S., Xu, S., & Guo, J. (2025). ACE-Step: A Step Towards Music Generation Foundation Model [<https://github.com/ace-step/ACE-Step>].

- Good, M. D. (2012). MusicXML. *Structuring Music through Markup Language: Designs and Architectures: Designs and Architectures*, 187. <https://doi.org/10.4018/978-1-4666-2497-9.ch009>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Gouk, P. (1986). Newton and music: From the microcosm to the macrocosm. *International studies in the philosophy of science*, 1(1), 36-59. <https://doi.org/10.1080/02698598608573281>
- Graves, A., Mohamed, A.-r., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. *2013 IEEE international conference on acoustics, speech and signal processing*, 6645-6649. <https://doi.org/10.1109/icassp.2013.6638947>
- Gray, R. O., & Corbally, C. J. (2009). *Stellar spectral classification*. Princeton university press. <https://doi.org/10.1515/9781400833368>
- Greisen, E. W., & Harten, R. H. (1981). An extension of FITS for groups of small arrays of data. *Astronomy and Astrophysics Supplement, Vol. 44, P. 371, 1981, 44*, 371.
- Gresham-Lancaster, S. (2012). Relationships of sonification to music and sound art. *AI & society*, 27, 207-212. <https://doi.org/10.1007/s00146-011-0337-3>
- Grond, F., & Berger, J. (2011). Parameter Mapping Sonification. En T. Hermann, A. Hunt & J. G. Neuhoff (Eds.), *The Sonification Handbook* (pp. 363-397). Logos Publishing House.
- Grosbøl, P. (1991). The FITS data format. En *Databases & On-line Data in Astronomy* (pp. 253-257). Springer. https://doi.org/10.1007/978-94-011-3250-3_25
- Grunblatt, S. K., Saunders, N., Sun, M., Chontos, A., Soares-Furtado, M., Eisner, N., Pereira, F., Komacek, T., Huber, D., Collins, K., et al. (2022). TESS giants transiting giants. II. The hottest Jupiters orbiting evolved stars. *The Astronomical Journal*, 163(3), 120. <https://doi.org/10.3847/1538-3881/ac4972>
- Guillochon, J., Parrent, J., Kelley, L. Z., & Margutti, R. (2017). An open catalog for supernova data. *The Astrophysical Journal*, 835(1), 64. <https://doi.org/10.3847/1538-4357/835/1/64>
- Hadjeres, G., Pachet, F., & Nielsen, F. (2017). Deepbach: a steerable model for bach chorales generation. *International conference on machine learning*, 1362-1371.
- Hall, J. C. (2008). Stellar chromospheric activity. *Living Reviews in Solar Physics*, 5(1), 1-53. <https://doi.org/10.12942/lrsp-2008-2>
- Hall, K. W., Bradley, A. J., Hinrichs, U., Huron, S., Wood, J., Collins, C., & Carpendale, S. (2019). Design by immersion: A transdisciplinary approach to problem-driven visualizations. *IEEE transactions on visualization and computer graphics*, 26(1), 109-118. <https://doi.org/10.1109/tvcg.2019.2934790>
- Hanisch, R. J., Berriman, G. B., Lazio, T., Bunn, S. E., Evans, J., McGlynn, T., & Plante, R. (2015). The virtual astronomical observatory: re-engineering access to astronomical data. *Astronomy and Computing*, 11, 190-209. <https://doi.org/10.1016/j.ascom.2015.03.007>
- Hanno Rein, e. a. (2020). Open Exoplanet Catalog [<http://www.openexoplanetcatalogue.com>].
- Hansen, B., Burchett, J. N., & Forbes, A. G. (2020). Quasar Spectroscopy Sound: Analyzing Intergalactic and Circumgalactic Media via Data Sonification. *Journal of the Audio Engineering Society*, 68(11), 865-875. <https://doi.org/10.17743/jaes.2020.0011>

- Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., et al. (2020). Array programming with NumPy. *Nature*, 585(7825), 357-362. <https://doi.org/10.1038/s41586-020-2649-2>
- Harrison, C., Zanella, A., Bonne, N., Meredith, K., & Misdariis, N. (2022). Audible universe. *Nature Astronomy*, 6(1), 22-23. <https://doi.org/10.1038/s41550-021-01582-y>
- Harvard-Smithsonian. (2015). Pulsating Variable Stars and the Hertzsprung-Russell Diagram [https://chandra.harvard.edu/edu/formal/variable_stars/bg_info.html].
- Harvey, J. (2000). Spectralism. *Contemporary music review*, 19(3), 11-14. <https://doi.org/10.1080/07494460000640331>
- Haseeb, M. T., Hammoudeh, A., & Xia, G. (2024). Gpt-4 driven cinematic music generation through text processing. *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6995-6999.
- Hawthorne, C., Stasyuk, A., Roberts, A., Simon, I., Huang, C.-Z. A., Dieleman, S., Elsen, E., Engel, J., & Eck, D. (2018). Enabling factorized piano music modeling and generation with the MAESTRO dataset. *arXiv preprint arXiv:1810.12247*. <https://doi.org/10.48550/arXiv.1810.12247>
- Hedges, S. A. (1978). Dice music in the eighteenth century. *Music & Letters*, 59(2), 180-187. <https://doi.org/10.1093/ml/59.2.180>
- Hermann, T. (2008). *Taxonomy and definitions for sonification and auditory display* (F. I. Proceedings of the 14th International Conference on Auditory Display (ICAD 2008). Susini P Warusfel O (Eds); Paris, Ed.).
- Hernandez-Olivan, C., & Beltran, J. R. (2022). Music composition with deep learning: A review. *Advances in speech and music technology: computational aspects and applications*, 25-50. https://doi.org/10.1007/978-3-031-18444-4_2
- Herremans, D., Chuan, C.-H., & Chew, E. (2017). A functional taxonomy of music generation systems. *ACM Computing Surveys (CSUR)*, 50(5), 1-30. <https://doi.org/10.1145/3108242>
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507. <https://doi.org/10.1126/science.1127647>
- Hochreiter, S. (1997). Long Short-term Memory. *Neural Computation MIT-Press*. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Holland, S. (2013). Artificial intelligence in music education: A critical review. *Readings in Music and Artificial Intelligence*, 239-274.
- Holloway, L. M., Goncu, C., Ilsar, A., Butler, M., & Marriott, K. (2022). Infosonics: Accessible infographics for people who are blind using sonification and voice. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1-13. <https://doi.org/10.1145/3491102.3517465>
- Huang, C.-Z. A., Vaswani, A., Uszkoreit, J., Shazeer, N., Hawthorne, C., Dai, A. M., Hoffman, M. D., & Eck, D. (2018). Music Transformer: Generating Music with Long-Term Structure. *arXiv preprint arXiv:1809.04281*. <https://doi.org/10.48550/arXiv.1809.04281>
- Huang, Q., Park, D. S., Wang, T., Denk, T. I., Ly, A., Chen, N., Zhang, Z., Zhang, Z., Yu, J., Frank, C., et al. (2023). Noise2music: Text-conditioned music generation with diffusion models. *arXiv preprint arXiv:2302.03917*.

- Huber, D. M., & Runstein, R. (2013). *Modern recording techniques*. Routledge. <https://doi.org/10.4324/9781003260530>
- Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), 90-95. <https://doi.org/10.1109/MCSE.2007.55>
- Huppenkothen, D., Pampin, J., Davenport, J. R., & Wenlock, J. (2023). The sonified hertzprung-russell diagram. *arXiv preprint arXiv:2401.00488*. <https://doi.org/10.21785/icad2023.6263>
- Husarik, S. (1983). John Cage and LeJaren Hiller: HPSCHD, 1969. *American Music*, 1-21. <https://doi.org/10.2307/3051496>
- Hussein, K., Tilevich, E., Bukvic, I. I., & Kim, S. (2009). Sonification design guidelines to enhance program comprehension. *2009 IEEE 17th International Conference on Program Comprehension*, 120-129. <https://doi.org/10.1109/icpc.2009.5090035>
- Hutchings, P. E., & McCormack, J. (2019). Adaptive music composition for games. *IEEE Transactions on Games*, 12(3), 270-280. <https://doi.org/10.1109/TG.2019.2921979>
- IAA. (2025). Revista IAA: Información y Actualidad Astronómica [<https://divulgacion.iaa.es/>].
- IAC. (2025). Outreach projects and activities [<https://www.iac.es/en/outreach/projects-and-activities>].
- IAS. (2016). The COROT archive [http://idoc-corot.ias.u-psud.fr/sitools/client-user/COROT_N2_PUBLIC_DATA/project-index.html].
- IAU. (2018). FITS standard [https://fits.gsfc.nasa.gov/fits_standard.html].
- IAU. (2020). IAU office of Astronomy for Education [<https://www.astro4edu.org/>].
- ICCUB. (2025). Outreach portal. Institut de Ciències del Cosmos [<https://icc.ub.edu/outreach>].
- IEC. (2013). Standard 61672-1: 2013, Electroacoustics—Sound Level Meters—Part 1: Specifications.
- IPython. (2014). Jupyter notebook [<https://jupyter.org/>].
- Ivezić, Ž., Connolly, A. J., VanderPlas, J. T., & Gray, A. (2019). *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data*. Princeton University Press.
- Jansky, K. G. (1933). Electrical disturbances apparently of extraterrestrial origin. *Proceedings of the Institute of Radio Engineers*, 21(10), 1387-1398. <https://doi.org/10.1109/jrproc.1933.227458>
- Jaschek, C. (1989). *Data in astronomy*. Cambridge University Press.
- Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 186(1007), 453-461. <https://doi.org/10.1098/rspa.1946.0056>
- Ji, S., Wu, S., Wang, Z., Li, S., & Zhang, K. (2025). A Comprehensive Survey on Generative AI for Video-to-Music Generation. *arXiv preprint arXiv:2502.12489*. <https://doi.org/10.48550/arXiv.2502.12489>
- Ji, S., Yang, X., & Luo, J. (2023). A survey on deep learning for symbolic music generation: Representations, algorithms, evaluations, and challenges. *ACM Computing Surveys*, 56(1), 1-39. <https://doi.org/10.1145/3597493>
- Jing, Y., Yang, Y., Feng, Z., Ye, J., Yu, Y., & Song, M. (2019). Neural style transfer: A review. *IEEE transactions on visualization and computer graphics*, 26(11), 3365-3385. <https://doi.org/10.1109/TVCG.2019.2921336>

- Jones, C., & Fenton, S. (2023). Use of Immersive Audio as an Assistive Technology for the Visually Impaired—A Systematic Review. *Audio Engineering Society Conference: AES 2023 International Conference on Spatial and Immersive Audio*.
- Kaae, J. (2017). Theoretical approaches to composing dynamic music for video games. En *From pac-man to pop music* (pp. 75-91). Routledge.
- Kahrs, M., & Brandenburg, K. (1998). *Applications of digital signal processing to audio and acoustics*. Springer Science & Business Media.
- Kates, J. M. (2005). Principles of digital dynamic-range compression. *Trends in amplification*, 9(2), 45-76. <https://doi.org/10.1177/108471380500900202>
- Kellermann, K. I., & Bouton, E. N. (2023). *Star Noise: Discovering the Radio Universe*. Cambridge University Press. <https://doi.org/10.1017/9781009023443>
- Kemmer, J., Dreizler, S., Kossakowski, D., Stock, S., Quirrenbach, A., Caballero, J., Amado, P., Collins, K., Espinoza, N., Herrero, E., et al. (2022). Discovery and mass measurement of the hot, transiting, Earth-sized planet, GJ 3929 b. *Astronomy & Astrophysics*, 659, A17. <https://doi.org/10.1051/0004-6361/202142653>
- Kennicutt Jr, R. C. (1992a). The integrated spectra of nearby galaxies-General properties and emission-line spectra. *The Astrophysical Journal*, 388, 310-327. <https://doi.org/10.1086/171154>
- Kennicutt Jr, R. C. (1992b). A spectrophotometric atlas of galaxies. *The Astrophysical Journal Supplement Series*, 79, 255-284. <https://doi.org/10.1086/191653>
- Kepler, J. (1969). *Harmonices mundi libri V*.
- Khostovan, A. A., Malhotra, S., Rhoads, J. E., Harish, S., Jiang, C., Wang, J., Wold, I., Zheng, Z.-Y., Barrientos, L. F., Coughlin, A., et al. (2021). Correlations between H α equivalent width and galaxy properties at $z=0.47$: Physical or selection-driven? *Monthly Notices of the Royal Astronomical Society*, 503(4), 5115-5133. <https://doi.org/10.1093/mnras/stab778>
- Kim, S., Lee, S.-g., Song, J., Kim, J., & Yoon, S. (2018). FloWaveNet: A generative flow for raw audio. *arXiv preprint arXiv:1811.02155*. <https://doi.org/10.48550/arXiv.1811.02155>
- Kingma, D. P. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*. <https://doi.org/10.48550/arXiv.1312.6114>
- Kingma, D. P., Welling, M., et al. (2019). An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12(4), 307-392. <https://doi.org/10.1561/22000000056>
- Kleczkowski, P. (1989). Group additive synthesis. *Computer Music Journal*, 13(1), 12-20. <https://doi.org/10.2307/3679851>
- Kong, Z., Ping, W., Huang, J., Zhao, K., & Catanzaro, B. (2020). Diffwave: A versatile diffusion model for audio synthesis. *arXiv preprint arXiv:2009.09761*. <https://doi.org/10.48550/arXiv.2009.09761>
- Kostov, V. B., Powell, B. P., Orosz, J. A., Welsh, W. F., Cochran, W., Collins, K. A., Endl, M., Hellier, C., Latham, D. W., MacQueen, P., et al. (2021). TIC 172900988: A Transiting Circumbinary Planet Detected in One Sector of TESS Data. *The Astronomical Journal*, 162(6), 234. <https://doi.org/10.3847/1538-3881/ac223a>
- Kovács, G., Zucker, S., & Mazeh, T. (2002). A box-fitting algorithm in the search for periodic transits. *Astronomy & Astrophysics*, 391(1), 369-377. <https://doi.org/10.1051/0004-6361:20020802>

- Kramer, G. (2000). *Auditory display: sonification, audification and auditory interfaces*. Addison-Wesley Longman Publishing Co., Inc.
- Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260), 583-621. <https://doi.org/10.2307/2280779>
- Kumar, K., Kumar, R., De Boissiere, T., Gestin, L., Teoh, W. Z., Sotelo, J., De Brebisson, A., Bengio, Y., & Courville, A. C. (2019). Melgan: Generative adversarial networks for conditional waveform synthesis. *Advances in neural information processing systems*, 32.
- Laidlow, R. (2024). Generative AI and music composition. En *Music, Technology, Innovation* (pp. 253-273). Routledge.
- Lazzarini, V., & Carty, B. (2008). New csound opcodes for binaural processing. *Proc. 6th Linux Audio Conference, Köln*.
- Lazzarini, V., Yi, S., Heintz, J., Brandtsegg, Ø., McCurdy, I., et al. (2016). *Csound: a sound and music computing system*. Springer.
- Le Borgne, J.-F., Bruzual, G., Pelló, R., Lançon, A., Rocca-Volmerange, B., Sanahuja, B., Schaerer, D., Soubiran, C., & Vílchez-Gómez, R. (2003). STELIB: A library of stellar spectra at $R \sim 2000$. *Astronomy & Astrophysics*, 402(2), 433-442. <https://doi.org/10.1051/0004-6361:20030243>
- Le Cun, Y. (1986). Learning process in an asymmetric threshold network. En *Disordered systems and biological organization* (pp. 233-240). Springer. https://doi.org/10.1007/978-3-642-82657-3_24
- Lenzi, S., Ciucarelli, P., & Geere, D. (2024, noviembre). Sonification Awards [<https://www.sonificationawards.org/>].
- Lenzi, S., & Ciucarelli, P. (2024). Designing tools for designers: The data Sonification Canvas. <https://doi.org/10.21606/drs.2024.730>
- Li, P.-p., & Wang, B. (2024). Artificial intelligence in music education. *International Journal of Human-Computer Interaction*, 40(16), 4183-4192. <https://doi.org/10.1080/10447318.2023.2209984>
- LIGO Laboratory, Gravitational Wave Open Science Center [<https://gwosc.org/audio/>]. (2021).
- Lillo-Box, J., Morales-Calderón, M., Barrado, D., Balsalobre-Ruza, O., Castro-González, A., Mendigutía, I., Huélamo, N., Montesinos, B., & Vioque, M. (2024). The AstraLux-TESS high spatial resolution imaging survey-Search for stellar companions of 215 planet candidates from TESS. *Astronomy & Astrophysics*, 686, A232. <https://doi.org/10.1051/0004-6361/202449687>
- Lohani, B., Gautam, C. K., Kushwaha, P. K., & Gupta, A. (2024). Deep Learning Approaches for Enhanced Audio Quality Through Noise Reduction. *2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE)*, 447-453. <https://doi.org/10.1109/ic3se62002.2024.10593073>
- Loy, G. (1985). Musicians make a standard: The MIDI phenomenon. *Computer Music Journal*, 9(4), 8-26. <https://doi.org/10.2307/3679619>
- Lu, Y.-C., & Cooke, M. (2010). Binaural estimation of sound source distance via the direct-to-reverberant energy ratio for static and moving sources. *IEEE Transactions on Audio*,

- Speech, and Language Processing*, 18(7), 1793-1805. <https://doi.org/10.1109/tasl.2010.2050718>
- Luo, S., Yan, C., Hu, C., & Zhao, H. (2023). Diff-foley: Synchronized video-to-audio synthesis with latent diffusion models. *Advances in Neural Information Processing Systems*, 36, 48855-48876.
- Lynch, J. F. (2017). Acoustics and astronomy. *Acoustics Today*, 13(4), 27-34.
- Ma, Y., Øland, A., Ragni, A., Del Sette, B. M., Saitis, C., Donahue, C., Lin, C., Plachouras, C., Benetos, E., Shatri, E., et al. (2024). Foundation models for music: A survey. *arXiv preprint arXiv:2408.14340*. <https://doi.org/10.48550/arXiv.2408.14340>
- Ma, Z., Reiss, J. D., & Black, D. A. (2013). Implementation of an intelligent equalization tool using Yule-Walker for music mixing and mastering. *Audio Engineering Society Convention 134*.
- Malham, D. G., & Myatt, A. (1995). 3-D sound spatialization using ambisonic techniques. *Computer music journal*, 19(4), 58-70. <https://doi.org/10.2307/3680991>
- Malikova., E., Pilyugin., V., Adzhiev., V., Pasko., G., & Pasko., A. Multisensory Analytics: Case of Visual-auditory Analysis of Scalar Fields. En: *En Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - IVAPP, (VISIGRAPP 2017)*. INSTICC. SciTePress, 2017, 322-329. <https://doi.org/10.5220/0006255003220329>
- Marra, F., & Ferreira, L. N. (2024). Long-Form Text-to-Music Generation with Adaptive Prompts: A Case of Study in Tabletop Role-Playing Games Soundtracks. *arXiv preprint arXiv:2411.03948*. <https://doi.org/10.48550/arXiv.2411.03948>
- Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Jia, Y., Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, . . . Xiaoqiang Zheng. (2015). TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems [<https://dl.acm.org/doi/10.5555/3026877.3026899>].
- Mas-Buitrago, P., González-Marcos, A., Solano, E., Passegger, V., Cortés-Contreras, M., Ordieres-Meré, J., Bello-García, A., Caballero, J., Schweitzer, A., Tabernero, H., et al. (2024). Using autoencoders and deep transfer learning to determine the stellar parameters of 286 CARMENES M dwarfs. *Astronomy & Astrophysics*, 687, A205. <https://doi.org/10.1051/0004-6361/202449865>
- Mathews, M. V. (1963). The Digital Computer as a Musical Instrument: A computer can be programmed to play “instrumental” music, to aid the composer, or to compose unaided. *Science*, 142(3592), 553-557. <https://doi.org/10.1126/science.142.3592.553>
- McCallum, M. C. (2019). Unsupervised learning of deep features for music segmentation. *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 346-350. <https://doi.org/10.1109/icassp.2019.8683407>
- McCartney, J. (1996). Supercollider Software [<https://supercollider.github.io/>].
- McCartney, J. (2002). Rethinking the computer music language: Super collider. *Computer Music Journal*, 26(4), 61-68. <https://doi.org/10.1162/014892602320991383>
- McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015). librosa: Audio and music signal analysis in python. *Proceedings of the 14th Python in Science conference (SCIPY)*, 18-24. <https://doi.org/10.25080/majora-7b98e3ed-003>

- Melchior, P., Hahn, C., & Liang, Y. (2022). Autoencoding Galaxy Spectra. *ICML Workshop Machine Learning for Astrophysics*. <https://doi.org/10.3847/1538-3881/ace0ff>
- Merced, W. D., & Gastrow, M. (2018). Astronomy and inclusive development: access to astronomy for people with disabilities. *Proceedings of the International Astronomical Union*, 14(A30), 596-597. <https://doi.org/10.1017/s1743921319005593>
- Metallinou, F.-A. (2022). Sonification and the Art of Sound in Interdisciplinary Space Science Communication. *44th COSPAR Scientific Assembly. Held 16-24 July, 44*, 1745.
- Middlebrooks, J. C. (2015). Sound localization. *Handbook of clinical neurology*, 129, 99-116. <https://doi.org/10.1016/B978-0-444-62630-1.00006-8>
- Middleton, J., Hakulinen, J., Tiitinen, K., Hella, J., Keskinen, T., Huuskonen, P., Linna, J., Turunen, M., Ziat, M., & Raisamo, R. (2018). Sonification with musical characteristics: a path guided by user engagement. *Proceedings of the International Conference on Auditory Display, Michigan Technological University*. <https://doi.org/10.21785/icad2018.006>
- Miranda, E. R. (2021). *Handbook of artificial intelligence for music*. Springer. <https://doi.org/10.1007/978-3-030-72116-9>
- Misdariis, N., Özcan, E., Grassi, M., Pauletto, S., Barrass, S., Bresin, R., & Susini, P. (2022). Sound experts' perspectives on astronomy sonification projects. *Nature Astronomy*, 6(11), 1249-1255. <https://doi.org/10.1038/s41550-022-01821-w>
- Mittal, G., Engel, J., Hawthorne, C., & Simon, I. (2021). Symbolic music generation with diffusion models. *arXiv preprint arXiv:2103.16091*. <https://doi.org/10.48550/arXiv.2103.16091>
- Molinaro, M., Allen, M., Genova, F., Schaaff, A., Neves, M. C., Demleitner, M., Bertocco, S., Morris, D., Bonnarel, F., Voutsinas, S., et al. (2020). The Virtual Observatory ecosystem facing the European Open Science Cloud. *Observatory Operations: Strategies, Processes, and Systems VIII*, 11449, 400-413. <https://doi.org/10.1117/12.2562322>
- Møller, H. (1992). Fundamentals of binaural technology. *Applied acoustics*, 36(3-4), 171-218. [https://doi.org/10.1016/0003-682X\(92\)90046-U](https://doi.org/10.1016/0003-682X(92)90046-U)
- Montalto, M., Malavolta, L., Gregorio, J., Mantovan, G., Desidera, S., Piotto, G., Nascimbeni, V., Granata, V., Manthopoulou, E., & Claudi, R. (2022). TIC 257060897b: An inflated, low-density, hot-Jupiter transiting a rapidly evolving subgiant star. *Monthly Notices of the Royal Astronomical Society*, 509(2), 2908-2919. <https://doi.org/10.1093/mnras/stab2923>
- Montmerle, T., & Fauque, D. (2022). *Astronomers as diplomats*. Springer. <https://doi.org/10.1007/978-3-030-98625-4>
- Moore, D. S., McCabe, G. P., & Craig, B. A. (2009). *Introduction to the Practice of Statistics* (Vol. 4). WH Freeman New York.
- Moorer, J. A. (1977). Signal processing aspects of computer music: A survey. *Proceedings of the IEEE*, 65(8), 1108-1137. <https://doi.org/10.1109/PROC.1977.10660>
- Morello, G., Parviainen, H., Murgas, F., Pallé, E., Oshagh, M., Fukui, A., Hirano, T., Ishikawa, H., Mori, M., Narita, N., et al. (2023). TOI-1442 b and TOI-2445 b: Two potentially rocky ultra-short period planets around M dwarfs. *Astronomy & Astrophysics*, 673, A32. <https://doi.org/10.1051/0004-6361/202243592>
- Morgan, W. W., Keenan, P. C., & Kellman, E. (1943). An atlas of stellar spectra. *University of Chicago Press*. <https://doi.org/10.1038/152147a0>

- Moysis, L., Iliadis, L. A., Sotiroidis, S. P., Boursianis, A. D., Papadopoulou, M. S., Kokkinidis, K.-I. D., Volos, C., Sarigiannidis, P., Nikolaidis, S., & Goudos, S. K. (2023). Music deep learning: deep learning methods for music signal processing—a review of the state-of-the-art. *Ieee Access*, *11*, 17031-17052. <https://doi.org/10.1109/access.2023.3244620>
- Mullally, S. E., & Fleming, S. W. (2018). Beginner: Read and Plot A TESS Data Validation Timeseries File [https://github.com/spacetelescope/notebooks/blob/master/notebooks/MAST/TESS/beginner_how_to_use_dvt/beginner_how_to_use_dvt.ipynb].
- NASA. (2011a). Exoplanet archive [<https://exoplanetarchive.ipac.caltech.edu/overview/KOI-3506>].
- NASA. (2011b). Kepler Data Characteristics Handbook [https://archive.stsci.edu/kepler/manuals/Data_Characteristics_Handbook_20110201.pdf].
- NASA. (2015, enero). TESS Documentation [<https://heasarc.gsfc.nasa.gov/docs/tess/documentation.html>].
- NASA. (2016). Kepler Archive Manual [https://archive.stsci.edu/kepler/manuals/archive_manual.pdf].
- NASA. (2020a). Kepler and K2 Science Center [<https://keplergo.github.io/KeplerScienceWebsite/>].
- NASA. (2020b). Nasa Exoplanet Catalog [<https://exoplanetarchive.ipac.caltech.edu/>].
- NASA. (2023). STEM Resources [<https://science.gsfc.nasa.gov/astrophysics/eduresources>].
- NASA. (2025a). Hubble Space Telescope [<https://science.nasa.gov/mission/hubble/>].
- NASA. (2025b). James Webb Space Telescope [<https://science.nasa.gov/mission/webb/>].
- NASA. (2025c). Nancy Grace Roman Space Telescope [<https://roman.gsfc.nasa.gov/>].
- NASA & the Harvard-Smithsonian Center for Astrophysics. (2019). Data Sonification: Sounds from Around the Milky Way [<https://chandra.si.edu/photo/2020/sonify/>].
- Neuhoff, J. G., Knight, R., & Wayand, J. (2002). Pitch change, sonification, and musical expertise: Which way is up. *Proceedings of the 2002 International Conference on Auditory Display*, 1-6.
- Neukom, M. (2014). Ambisonics User Defined Opcodes for Csound. *ICMC*.
- Ng, A., et al. (2011). Sparse autoencoder. *CS294A Lecture notes*, *72*(2011), 1-19.
- Noé, J. V. G. (2022). La música eléctrica y el electrocompositor de Juan García Castillejo (1933-1944). *Enclaves. Revista de Literatura, Música y Artes Escénicas*, (2), 97-112. <https://doi.org/10.12795/enclaves.2022.i02.07>
- Noel-Storr, J., & Willebrands, M. (2022). Accessibility in astronomy for the visually impaired. *Nature Astronomy*, *6*(11), 1216-1218. <https://doi.org/10.1038/s41550-022-01691-2>
- NSF. (2023). Vera Rubin Observatory [<https://rubinobservatory.org/>].
- NSF. (2025). SDSS-V: Pioneering panoptic spectroscopy [<https://www.sdss.org/>].
- Numpy. (2008). Composite trapezoidal rule integration [<https://numpy.org/doc/1.25/reference/generated/numpy.trapz.html>].
- Odaibo, S. (2019). Tutorial: Deriving the standard variational autoencoder (vae) loss function. *arXiv preprint arXiv:1907.08956*. <https://doi.org/10.48550/arXiv.1907.08956>
- Oord, A. v. d., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., & Kavukcuoglu, K. (2016). Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*. <https://doi.org/10.48550/arXiv.1609.03499>
- Ortega-Alonso, D., Pérez-Montero, E., Gibaja, J. F., García-Benito, R., Mozota, M., & Remolins, G. (2024). Narrativas transmedia y activismo, ¿ingredientes secretos para una divulgación científica inclusiva? *Arte, Individuo y Sociedad [2024];[36 (3)]: 687-699*.

- Osborn, H. P., Bonfanti, A., Gandolfi, D., Hedges, C., Leleu, A., Fortier, A., Futyan, D., Gutermann, P., Maxted, P. F., Borsato, L., et al. (2022). Uncovering the true periods of the young sub-Neptunes orbiting TOI-2076. *Astronomy & Astrophysics*, *664*, A156. <https://doi.org/https://doi.org/10.1051/0004-6361/202243065>
- Pabón, G. C. (2009). *Numerus-Proportio en el De Música De San Agustín:(Libros I y VI): la tradición pitagórico-platónica* [Tesis doctoral, Universidad de Salamanca]. <https://doi.org/10.14201/gredos.76465>
- Pacholczyk, J. (1996). Music and Astronomy in the Muslim World. *Leonardo*, *29*(2), 145-150. <https://doi.org/10.2307/1576351>
- Papers from the 1977 International Computer Music Conference. (1978). *Computer Music Journal*, *2*(3), 43-43.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, *107*(10), 4758-4763. <https://doi.org/10.1073/pnas.0909074107>
- Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annu. Rev. Psychol.*, *56*, 89-114. <https://doi.org/10.1146/annurev.psych.56.091103.070225>
- Pérez-Montero, E., Gómez-Caro, E. G., Sánchez, Y., Molina, A. O.-G., de Lacalle, S. L., & Tamayo, A. (2017). Astroaccesible: Bringing the study of the Universe to the visually impaired. *Highlights of the Spanish Astrophysics IX. Proceedings of the Scientific Meeting of the Spanish Astronomical Society*, 742-747.
- Pesic, P. (2005). Earthly Music and Cosmic Harmony: Johannes Kepler's Interest in Practical Music, Especially Orlando di Lasso. *Journal of Seventeenth-Century Music*, *11*(1).
- Pesic, P. (2006). Isaac Newton and the mystery of the major sixth: a transcription of his manuscript 'Of Musick' with commentary. *Interdisciplinary Science Reviews*, *31*(4), 291-306. <https://doi.org/10.1179/030801806x143268>
- Pesnell, W. D., Ingram-Johnson, K., & Addison, K. (2021). Using Hilbert curves to organize, sample, and sonify solar data. *American Journal of Physics*, *89*(10), 943-943. <https://doi.org/10.1119/10.0005403>
- PHOENIX [<http://astro.vaporia.com/start/phoenixcode.html>]. (1990).
- Pons, J., Slizovskaia, O., Gong, R., Gómez, E., & Serra, X. (2017). Timbre analysis of music audio signals with convolutional neural networks. *2017 25th European Signal Processing Conference (EUSIPCO)*, 2744-2748. <https://doi.org/10.23919/eusipco.2017.8081710>
- Portillo, S. K., Parejko, J. K., Vergara, J. R., & Connolly, A. J. (2020). Dimensionality reduction of SDSS spectra with variational autoencoders. *The Astronomical Journal*, *160*(1), 45. <https://doi.org/10.3847/1538-3881/ab9644>
- Price-Whelan, A. M., Sipőcz, B., Günther, H., Lim, P., Crawford, S., Conseil, S., Shupe, D., Craig, M., Dencheva, N., Ginsburg, A., et al. (2018). The astropy project: building an open-science project and status of the v2. 0 core package. *The Astronomical Journal*, *156*(3), 123. <https://doi.org/10.3847/1538-3881/aabc4f>
- Prša, A., Batalha, N., Slawson, R. W., Doyle, L. R., Welsh, W. F., Orosz, J. A., Seager, S., Rucker, M., Mjaseth, K., Engle, S. G., et al. (2011). Kepler eclipsing binary stars. I. Catalog and principal characterization of 1879 eclipsing binaries in the first data release. *The Astronomical Journal*, *141*(3), 83. <https://doi.org/10.1088/0004-6256/141/3/83>

- Prugniel, P., Soubiran, C., Koleva, M., & Borgne, D. L. (2007). New release of the ELODIE library: Version 3.1. *arXiv preprint astro-ph/0703658*. <https://doi.org/10.48550/arXiv.astro-ph/0703658>
- Puckette, M. (1985). Max-MSP Software [<https://cycling74.com/products/max>].
- Puckette, M. (1991). Combining event and signal processing in the MAX graphical programming environment. *Computer music journal*, 15(3), 68-77. <https://doi.org/10.2307/3680767>
- Puckette, M., et al. (1996). Pure Data: another integrated computer music environment. *Proceedings of the second intercollege computer music concerts*, 37-41.
- Purwins, H., Li, B., Virtanen, T., Schlüter, J., Chang, S.-Y., & Sainath, T. (2019). Deep learning for audio signal processing. *IEEE Journal of Selected Topics in Signal Processing*, 13(2), 206-219. <https://doi.org/10.1109/jstsp.2019.2908700>
- Pygame [<https://www.pygame.org/docs/ref/midi.html>]. (2000).
- Python-OSC [<https://pypi.org/project/python-osc/>]. (2013).
- Quatieri, T., & McAulay, R. (1986). Speech transformations based on a sinusoidal representation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(6), 1449-1464. <https://doi.org/10.1109/TASSP.1986.1164985>
- Quinn, P. J. (2003). Virtual Observatories: the Future of Astronomical Information. *Library and Information Services in Astronomy IV (LISA IV)*, 55.
- Quinn, P. J., & Górski, K. M. (2004). *Toward an International Virtual Observatory: Proceedings of the ESO/ESA/NASA/NSF Conference Held at Garching, Germany, 10-14 June 2002*. Springer Science & Business Media. <https://doi.org/10.1007/b82981>
- Quinton, M., McGregor, I., & Benyon, D. (2020). Sonification of an exoplanetary atmosphere. *Proceedings of the 15th International Audio Mostly Conference*, 191-198. <https://doi.org/10.1145/3411109.3411117>
- Quinton, M., McGregor, I., & Benyon, D. (2021). Sonification of Planetary Orbits in Asteroid Belts. En *Audio Mostly 2021* (pp. 72-80). <https://doi.org/10.1145/3478384.3478390>
- Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., Ahmad, J., Ali, M. E., & Azam, S. (2024). A review on large language models: Architectures, applications, taxonomies, open issues and challenges. *IEEE access*, 12, 26839-26874. <https://doi.org/10.1109/ACCESS.2024.3365742>
- Rezende, D. J., Mohamed, S., & Wierstra, D. (2014). Stochastic backpropagation and variational inference in deep latent gaussian models. *International conference on machine learning*, 2, 2.
- Risset, J.-C. (2003). The perception of musical sound. *Computer music: Why*, 1-12.
- Rixon, G., Linde, T., Auden, E., & Walton, N. (2004). Authentication and Authorization Architecture for AstroGrid and the Virtual Observatory. *Toward an International Virtual Observatory: Proceedings of the ESO/ESA/NASA/NSF Conference Held at Garching, Germany, 10-14 June 2002*, 75-79. https://doi.org/10.1007/10857598_11
- Rizzuto, A. C., Newton, E. R., Mann, A. W., Tofflemire, B. M., Vanderburg, A., Kraus, A. L., Wood, M. L., Quinn, S. N., Zhou, G., Thao, P. C., et al. (2020). TESS hunt for young and maturing exoplanets (THYME). II. A 17 Myr old transiting hot Jupiter in the Sco-cen association. *The Astronomical Journal*, 160(1), 33. <https://doi.org/10.3847/1538-3881/ab94b7>

- Roads, C. (1978). A Report on the 1978 International Computer Music Conference. *Computer Music Journal*, 2(4), 21-26.
- Roads, C. (1996). *The computer music tutorial*. MIT press.
- Roberts, A., Engel, J., Raffel, C., Hawthorne, C., & Eck, D. (2018). A hierarchical latent vector model for learning long-term structure in music. *International conference on machine learning*, 4364-4373.
- Robinson, D. W., & Dadson, R. S. (1956). A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics*, 7(5), 166. <https://doi.org/10.1088/0508-3443/7/5/302>
- Robitaille, T., & Bressert, E. (2012, agosto). APLpy: Astronomical Plotting Library in Python.
- Robitaille, T. (2019, febrero). APLpy v2.0: The Astronomical Plotting Library in Python. <https://doi.org/10.5281/zenodo.2567476>
- Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., et al. (2013). Astropy: A community Python package for astronomy. *Astronomy & Astrophysics*, 558, A33. <https://doi.org/10.1051/0004-6361/201322068>
- Roginska, A., & Geluso, P. (2017). *Immersive Sound*. Focal Press. <https://doi.org/10.4324/9781315707525>
- Ronan, D. M. (2019). *Intelligent Subgrouping of Multitrack Audio* [Tesis doctoral, Queen Mary University of London].
- Rönnerberg, N., & Jimmy, J. (2016). Interactive sonification for visual dense data displays. *ISon 2016, 5th Interactive Sonification Workshop, CITEC, Bielefeld University, Germany, December 16, 2016*, 63-67.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536. <https://doi.org/10.1038/323533a0>
- Saariaho, K. (2005). Asteroid 4179: Toutatis [<https://saariaho.org/works/asteroid-4179-toutatis>].
- Sánchez, S. F., García-Benito, R., Zibetti, S., Walcher, C., Husemann, B., Mendoza, M., Galbany, L., Falcón-Barroso, J., Mast, D., Aceituno, J., et al. (2016). CALIFA, the Calar Alto Legacy Integral Field Area survey-IV. Third public data release. *Astronomy & Astrophysics*, 594, A36. <https://doi.org/10.1051/0004-6361/201628661>
- Sánchez, S., Kennicutt, R., De Paz, A. G., Van de Ven, G., Vílchez, J., Wisotzki, L., Walcher, C., Mast, D., Aguerri, J. A. L., Albiol-Pérez, S., et al. (2012). Califa, the calar alto legacy integral field area survey-i. survey presentation. *Astronomy & Astrophysics*, 538, A8. <https://doi.org/10.1051/0004-6361/201117353>
- Sánchez-Blázquez, P., Peletier, R., Jiménez-Vicente, J., Cardiel, N., Cenarro, A. J., Falcon-Barroso, J., Gorgas, J., Selam, S., & Vazdekis, A. (2006). Medium-resolution Isaac Newton Telescope library of empirical spectra. *Monthly Notices of the Royal Astronomical Society*, 371(2), 703-718. <https://doi.org/10.1111/j.1365-2966.2006.10699.x>
- Sándor, A., & Lane, D. M. (2003). Sonification of absolute values with single and multiple dimensions.
- Saunders, N., Grunblatt, S. K., Huber, D., Collins, K. A., Jensen, E. L., Vanderburg, A., Brahm, R., Jordán, A., Espinoza, N., Henning, T., et al. (2022). TESS Giants Transiting Giants. I.: A Noninflated Hot Jupiter Orbiting a Massive Subgiant. *The Astronomical Journal*, 163(2), 53. <https://doi.org/10.3847/1538-3881/ac38a1>

- Schaeffer, P. (2012). *In search of a concrete music* (Vol. 15). Univ of California Press.
- Schanche, N., Pozuelos, F., Günther, M., Wells, R. D., Burgasser, A., Chinchilla, P., Delrez, L., Ducrot, E., Garcia, L., Chew, Y. G. M., et al. (2022). TOI-2257 b: A highly eccentric long-period sub-Neptune transiting a nearby M dwarf. *Astronomy & Astrophysics*, 657, A45. <https://doi.org/10.1051/0004-6361/202142280>
- Schedl, M. (2019). Deep learning in music recommendation systems. *Frontiers in Applied Mathematics and Statistics*, 5, 457883. <https://doi.org/10.3389/fams.2019.00044>
- Schlecker, M., Kossakowski, D., Brahm, R., Espinoza, N., Henning, T., Carone, K., Trifonov, L., Molaverdikhani, T., Mollière, P., Hobson, M. J., et al. (2020). A highly eccentric warm jupiter orbiting TIC 237913194. *The Astronomical Journal*, 160(6), 275. <https://doi.org/10.3847/1538-3881/abbe03>
- Schubert, E. (2009). The fundamental function of music. *Musicae Scientiae*, 13(2_suppl), 63-81. <https://doi.org/10.1177/1029864909013002051>
- Schweer, W. (2002). MuseScore Software [<https://musescore.org/es>].
- Scikit-learn [<https://scikit-learn.org/stable/>]. (2007).
- SciPy [<https://scipy.org/>]. (2001).
- Seager, S., & Mallen-Ornelas, G. (2003). A Unique Solution of Planet and Star Parameters from an Extrasolar Planet Transit Light Curve. *The Astrophysical Journal*, 585(2), 1038-1055. <https://doi.org/10.1086/346105>
- Seiça, M., Roque, L., Martins, P., & Cardoso, F. A. (2023). An interdisciplinary journey towards an aesthetics of sonification experience. *Journal on Multimodal User Interfaces*, 17(4), 263-284. <https://doi.org/10.1007/s12193-023-00416-7>
- Serra, X. (2013). Musical sound modeling with sinusoids plus noise. En *Musical signal processing* (pp. 91-122). Routledge.
- Shah, N., & Ganatra, A. (2022). Comparative Study of Autoencoders-Its Types and Application. *2022 6th International Conference on Electronics, Communication and Aerospace Technology*, 175-180. <https://doi.org/10.1109/iceca55336.2022.10009387>
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611. <https://doi.org/10.2307/2333709>
- Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306. <https://doi.org/10.1016/j.physd.2019.132306>
- Shih, Y.-J., Wu, S.-L., Zalkow, F., Müller, M., & Yang, Y.-H. (2022). Theme transformer: Symbolic music generation with theme-conditioned transformer. *IEEE Transactions on Multimedia*, 25, 3495-3508. <https://doi.org/10.1109/TMM.2022.3161851>
- Shiloah, A. (1978). *The epistle on music of the Ikhwan al-Safa:(Bagdad, 10th century)* (Vol. 3). Tel Aviv University.
- Shirlaw, M. (1951). The music and tone-systems of ancient Greece. *Music & Letters*, 131-139. <https://doi.org/10.1093/ml/xxxii.2.131>
- Slawson, R. W., Prša, A., Welsh, W. F., Orosz, J. A., Rucker, M., Batalha, N., Doyle, L. R., Engle, S. G., Conroy, K., Coughlin, J., et al. (2011). Kepler eclipsing binary stars. II. 2165 eclipsing binaries in the second data release. *The Astronomical Journal*, 142(5), 160. <https://doi.org/10.1088/0004-6256/142/5/160>

- Smirnov, V. (1999). Music theory and the harmony method in J. Kepler's work the harmony of the universe. *Astronomical and Astrophysical Transactions*, 18(3), 521-532. <https://doi.org/10.1080/10556799908203010>
- Smith, D. R., & Walker, B. N. (2005). Effects of auditory context cues and training on performance of a point estimation sonification task. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 19(8), 1065-1087. <https://doi.org/10.1002/acp.1146>
- Snook, K., Barri, T., Bolles, M., Ericson, P., Fravel, C., Goßmann, J., Green-Mateu, S. E., Luck, A., Schedel, M., & Thomas, R. (2020). Concordia: A musical XR instrument for playing the solar system. *Journal of New Music Research*, 49(1), 88-103. <https://doi.org/10.1080/09298215.2020.1714666>
- Socia, Q. J., Welsh, W. F., Orosz, J. A., Cochran, W. D., Endl, M., Quarles, B., Short, D. R., Torres, G., Windmiller, G., & Yenawine, M. (2020a). Kepler-1661 b: A Neptune-sized Kepler Transiting Circumbinary Planet around a Grazing Eclipsing Binary. *The Astronomical Journal*, 159(3), 94. <https://doi.org/10.3847/1538-3881/ab665b>
- Socia, Q. J., Welsh, W. F., Orosz, J. A., Cochran, W. D., Endl, M., Quarles, B., Short, D. R., Torres, G., Windmiller, G., & Yenawine, M. (2020b). Kepler-1661 b: a neptune-sized kepler transiting circumbinary planet around a grazing eclipsing binary. *The Astronomical Journal*, 159(3), 94. <https://doi.org/10.3847/1538-3881/ab665b>
- Solomos, M. (2005). Cellular automata in Xenakis's music. Theory and Practice. *International Symposium Iannis Xenakis (Athens, May 2005)*, 11-p.
- St George, J., Kim, S. Y., & Bischof, H.-P. (2018). Sonification of simulated black hole merger data. *Proceedings of the International Conference on Modeling, Simulation and Visualization Methods (MSV)*, 3-9.
- Starks, G. (2018). Soniverse [<https://soniverse.space/2018/02/15/listen-to-the-radio-cube-of-the-antennae-galaxies>].
- Stephenson, B. (1994). *The music of the heavens: Kepler's harmonic astronomy*. Princeton University Press. <https://doi.org/10.1515/9781400863822>
- STScI. (1981, septiembre). STScI [<https://www.stsci.edu/>].
- STScI. (2008). Mikulsky Archive for Space Telescopes [<https://archive.stsci.edu/>].
- STScI. (2013). KOI search results [<https://archive.stsci.edu/kepler/koi/search.php>].
- STScI. (2020). MOS Spectroscopy of Extragalactic Field [https://github.com/spacetelescope/jdat_notebooks].
- STScI, M. (2017). Kepler Mission [<https://archive.stsci.edu/missions-and-data/kepler>].
- Student. (1908). The probable error of a mean. *Biometrika*, 1-25. <https://doi.org/10.2307/2331554>
- Sulis, V. (2024). Sonification of Astronomical Data. Available at SSRN 4891458. <https://doi.org/10.2139/ssrn.4891458>
- Sullivan, B. (1989). Interpretive Models of Guido of Arezzo's Micrologus. *Comitatus: A Journal of Medieval and Renaissance Studies*, 20(1).
- Survey, S. D. S. (2003). SkyServer [<http://skyserver.sdss.org/dr1/en/proj/basic/spectraltypes/lines.asp>].
- Suzuki, Y., & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *The Journal of the Acoustical Society of America*, 116(2), 918-933. <https://doi.org/10.1121/1.1763601>

- SVO. (2003). Stelib stellar library [<http://svocats.cab.inta-csic.es/stelib/index.php>].
- SVO. (2008). Vosa [<http://svo2.cab.inta-csic.es/theory/vosa/>].
- SVO. (2019). Stellar Spectral Libraries [<https://svo2.cab.inta-csic.es/theory/libraries/>].
- Szalay, A. S., Gray, J., Kunszt, P., Thakar, A., & Slutz, D. (2001). Large databases in astronomy. *Mining the Sky: Proceedings of the MPA/ESO/MPE Workshop Held at Garching, Germany, July 31-August 4, 2000*, 99-116. https://doi.org/10.1007/10849171_7
- Tasca, L., Le Fèvre, O., Hathi, N., Schaerer, D., Ilbert, O., Zamorani, G., Lemaux, B., Cassata, P., Garilli, B., Le Brun, V., et al. (2015). The evolving star formation rate: M* relation and sSFR since z= 5 from the VUDS spectroscopic survey. *Astronomy & Astrophysics*, 581, A54. <https://doi.org/10.1051/0004-6361/201425379>
- Taylor, M. (2005). Topcat [<http://www.star.bristol.ac.uk/~mbt/topcat/>].
- Taylor, M. B. (2005). TOPCAT & STIL: Starlink table/VOTable processing software. *Astronomical data analysis software and systems XIV*, 347, 29.
- Temperley, D. (2004). *The cognition of basic musical structures*. MIT press.
- Thompson, S. E., Coughlin, J. L., Hoffman, K., Mullally, F., Christiansen, J. L., Burke, C. J., Bryson, S., Batalha, N., Haas, M. R., Catanzarite, J., et al. (2018). Planetary candidates observed by Kepler. VIII. A fully automated catalog with measured completeness and reliability based on data release 25. *The Astrophysical Journal Supplement Series*, 235(2), 38. <https://doi.org/10.3847/1538-4365/aab4f9>
- Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychological review*, 107(4), 885. <https://doi.org/10.1037//0033-295x.107.4.885>
- Tody, D. (1986). The IRAF data reduction and analysis system. *Instrumentation in astronomy VI*, 627, 733-748. <https://doi.org/10.1117/12.968154>
- Tody, D. (1993). IRAF in the Nineties. *Proceedings of the Astronomical Data Analysis Software and Systems II, Boston, MA, USA*, 2-4.
- Torres, J. (2020). *Python deep learning: Introducción práctica con Keras y TensorFlow 2*. Alpha Editorial.
- Trayford, J. W., & Harrison, C. M. (2023). Introducing STRAUSS: A flexible sonification Python package. *arXiv preprint arXiv:2311.16847*. <https://doi.org/10.21785/icad2023.1978>
- Trayford, J. W., Harrison, C., Hinz, R., Kavanagh Blatt, M., Dougherty, S., & Girdhar, A. (2023). Inspecting spectra with sound: proof-of-concept and extension to datacubes. *RAS Techniques and Instruments*, 2(1), 387-392. <https://doi.org/10.1093/rasti/rzad021>
- Trifonov, T., Brahm, R., Espinoza, N., Henning, T., Jordán, A., Nesvorný, D., Dawson, R. I., Lissauer, J. J., Lee, M. H., Kossakowski, D., et al. (2021). A pair of warm giant planets near the 2: 1 mean motion resonance around the K-dwarf star TOI-2202. *The Astronomical Journal*, 162(6), 283. <https://doi.org/10.3847/1538-3881/ac1bbe>
- Tucker Brown, J., Harrison, C., Zanella, A., & Trayford, J. (2022). Evaluating the efficacy of sonification for signal detection in univariate, evenly sampled light curves using astronify. *Monthly Notices of the Royal Astronomical Society*, 516(4), 5674-5683. <https://doi.org/10.1093/mnras/stac2590>
- Turing, A. M. (2009). *Computing machinery and intelligence*. Springer. https://doi.org/10.1007/978-1-4020-6710-5_3

- Tutchton, R. M., Wood, M. A., Still, M. D., Howell, S. B., Cannizzo, J. K., & Smale, A. P. (2012). Sonification of Kepler field SU UMA cataclysmic variable stars V344 Lyr and V1504 Cyg. *Journal of the Southeastern Association for Research in Astronomy*, 6(GSFC-E-DAA-TN6380).
- Twicken, J. D., Caldwell, D. A., Jenkins, J. M., Vanderspek, R. K., Tenenbaum, P., Smith, J. C., Wohler, B., Rose, M., Ting, E. B., Vanderspek, R., et al. (2020). TESS Science Data Products Description Document: EXP-TESS-ARC-ICD-0014 Rev F.
- UNOOSA. (2022). Sonification: a tool for research, outreach and inclusion in space sciences [<https://www.unoosa.org/oosa/en/ourwork/space4personswithdisabilites/index.html>].
- Valdes, F., Gupta, R., Rose, J. A., Singh, H. P., & Bell, D. J. (2004). The Indo-US library of Coudé feed stellar spectra. *The Astrophysical Journal Supplement Series*, 152(2), 251. <https://doi.org/10.1086/386343>
- Valle, A., & Korol, V. (2022). For LISA. A piano-based sonification project of gravitational waves. *arXiv preprint arXiv:2202.04621*. <https://doi.org/10.48550/arXiv.2202.04621>
- Van Der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: a structure for efficient numerical computation. *Computing in science & engineering*, 13(2), 22-30. <https://doi.org/10.1109/mcse.2011.37>
- Van der Wel, A., Bezanson, R., D'Eugenio, F., Straatman, C., Franx, M., van Houdt, J., Maseda, M. V., Gallazzi, A., Wu, P.-F., Pacifici, C., et al. (2021). The Large Early Galaxy Astrophysics Census (LEGA-C) Data Release 3: 3000 High-quality Spectra of K s-selected Galaxies at $z > 0.6$. *The Astrophysical Journal Supplement Series*, 256(2), 44. <https://doi.org/10.3847/1538-4365/ac1356>
- Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. *Artificial Intelligence Review*, 53(8), 5929-5955.
- VanderPlas, J. T. (2018). Understanding the lomb–scargle periodogram. *The Astrophysical Journal Supplement Series*, 236(1), 16. <https://doi.org/10.3847/1538-4365/aab766>
- van Riesen, S. A., Gijlers, H., Anjewierden, A. A., & de Jong, T. (2022). The influence of prior knowledge on the effectiveness of guided experiment design. *Interactive Learning Environments*, 30(1), 17-33. <https://doi.org/10.1080/10494820.2019.1631193>
- Vasireddy, A., Kumar, T. N., Chebrolu, M., Nair, K., & Prasad, V. B. (2024). AudioPalette: Generating Soundtracks through Feature Extraction from Images. *2024 International Conference on Inventive Computation Technologies (ICICT)*, 714-719. <https://doi.org/10.1109/ICICT60155.2024.10544569>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30. <https://doi.org/https://doi.org/10.48550/arXiv.1706.03762>
- Vercoe, B., et al. (1986). Csound. *The CSound Manual Version*, 3.
- Vercoe, B., Ffitch, J., Piche, J., Nix, P., Boulanger, R., Ekman, R., Boothe, D., Conder, K., Yi, S., Gogins, M., et al. (2007). The canonical Csound reference manual [<https://www.csounds.com/manual/html/title.html>].
- Vickers. (2013). Ways of Listening and Modes of Being: Electroacoustic Auditory Display. *arXiv, abs/1311.5880*. <https://doi.org/10.48550/arXiv.1311.5880>

- Vickers, P. (2006). Lemma 4: Haptic input+ auditory display= musical instrument? *International Workshop on Haptic and Audio Interaction Design*, 56-67. https://doi.org/10.1007/11821731_6
- Vickers, P., & Hogg, B. (2006). Sonification abstraite/sonification concrete: An 'aesthetic perspective space' for classifying auditory displays in the ars musica domain. *Proceedings of the International Conference on Auditory Display, London, UK*.
- Villaescusa-Navarro, F., Anglés-Alcázar, D., Genel, S., Spergel, D. N., Somerville, R. S., Dave, R., Pillepich, A., Hernquist, L., Nelson, D., Torrey, P., et al. (2021). The camels project: Cosmology and astrophysics with machine-learning simulations. *The Astrophysical Journal*, 915(1), 71. <https://doi.org/10.3847/1538-4357/abf7ba>
- Walcher, C., Wisotzki, L., Bekeraité, S., Husemann, B., Iglesias-Páramo, J., Backsmann, N., Ballesteros, J. B., Catalán-Torrecilla, C., Cortijo, C., Del Olmo, A., et al. (2014). CALIFA: a diameter-selected sample for an integral field spectroscopy galaxy survey. *Astronomy & Astrophysics*, 569, A1. <https://doi.org/10.1051/0004-6361/201424198>
- Walker, B. N. (2023). Pasado, Presente y Promesas de la Sonificación. *ARBOR Ciencia, Pensamiento y Cultura*, 199(810), e728. <https://doi.org/10.3989/arbor.2023.810008>
- Walker, B. N., & Cothran, J. T. (2003). Sonification Sandbox: A graphical toolkit for auditory graphs. *Proceedings of the International Conference on Auditory Display, Boston, MA, USA*.
- Walker, B. N., & Kramer, G. (2004). Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. En *Ecological psychoacoustics* (pp. 149-174). Brill. https://doi.org/10.1163/9780080477442_009
- Walker, B. N., & Kramer, G. (2005). Mappings and metaphors in auditory displays: An experimental assessment. *ACM Transactions on Applied Perception (TAP)*, 2(4), 407-412. <https://doi.org/10.1145/1101530.1101534>
- Walker, B. N., & Lane, D. M. (2001). Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners. *Proceedings of the International Conference on Auditory Display*, 90-94.
- Walker, B. N., & Mauney, L. M. (2010). Universal design of auditory graphs: A comparison of sonification mappings for visually impaired and sighted listeners. *ACM Transactions on Accessible Computing (TACCESS)*, 2(3), 1-16. <https://doi.org/10.1145/1714458.1714459>
- Walker, B. N., & Nees, M. A. (2005). Brief training for performance of a point estimation sonification task. *Proceedings of the International Conference on Auditory Display*, 276-279.
- Walker, B. N., & Nees, M. A. (2011). Theory of sonification. *The sonification handbook*, 1, 9-39.
- Walsh, R., McCurdy, I., & Boyle, G. (2008). Cabbage-Audio software [<https://cabbageaudio.com/>].
- Wang, W. (2006). Astronomy and Music. *ESSAI*, 4(1), 40.
- Wang, Z., Tang, C.-K., & Tai, Y.-W. (2024). Audio-agent: Leveraging llms for audio generation, editing and composition. *arXiv preprint arXiv:2410.03335*. <https://doi.org/10.48550/arXiv.2410.03335>
- Wei, S., Wei, M., Wang, H., Zhao, Y., & Kou, G. (2024). Melody is all you need for music generation. *arXiv preprint arXiv:2409.20196*.

- Weinberg, G., Bretan, M., Hoffman, G., & Driscoll, S. (2020). *Robotic musicianship: embodied artificial creativity and mechatronic musical expression* (Vol. 8). Springer Nature. <https://doi.org/10.1007/978-3-030-38930-7>
- Weinberg, G., Bretan, M., Hoffman, G., & Driscoll, S. (2008). Robotic musicianship. *RO-MAN, Springer*. <https://doi.org/10.1109/roman.2008.4600629>
- Wells, R., Rackham, B., Schanche, N., Petrucci, R., Chew, Y. G. M., Demory, B.-O., Burgasser, A., Burn, R., Pozuelos, F., Günther, M., et al. (2021). A large sub-Neptune transiting the thick-disk M4 V TOI-2406. *Astronomy & Astrophysics*, 653, A97. <https://doi.org/10.1051/0004-6361/202141277>
- Wenger, M., Ochsenbein, F., Egret, D., Dubois, P., Bonnarel, F., Borde, S., Genova, F., Jasniewicz, G., Laloë, S., Lesteven, S., et al. (2000). The SIMBAD astronomical database-The CDS reference database for astronomical objects. *Astronomy and Astrophysics Supplement Series*, 143(1), 9-22. <https://doi.org/10.1051/aas:2000332>
- West, R., Johnson, V., Yeh, I. C., Thomas, Z., & Tarlton, M. (2018). Experiencing a Slice of the Sky: Immersive Rendering and Sonification of Antarctic Astronomy Data. *Electronic Imaging*, 30, 1-10. <https://doi.org/10.2352/issn.2470-1173.2018.03.ervr-449>
- Wittrock, J. M., Dreizler, S., Reefe, M. A., Morris, B. M., Plavchan, P. P., Lowrance, P. J., Demory, B.-O., Ingalls, J. G., Gilbert, E. A., Barclay, T., et al. (2022). Transit Timing Variations for AU Microscopii b & c. *arXiv preprint arXiv:2202.05813*. <https://doi.org/10.48550/arXiv.2202.05813>
- Wolfram, S. (1983). Statistical mechanics of cellular automata. *Reviews of modern physics*, 55(3), 601. <https://doi.org/10.1201/9780429494093-1>
- Woodward, D., Stevens, E., & Linstead, E. (2019). Generating Transit Light Curves with Variational Autoencoders. *2019 IEEE International Conference on Space Mission Challenges for Information Technology (SMC-IT)*, 24-32. <https://doi.org/10.1109/SMC-IT.2019.00008>
- Worrall, D. (2019). *Sonification Design*. Cham: Springer. <https://doi.org/10.1007/978-3-030-01497-1>
- Wright, M. (2002). OSC specification [http://opensoundcontrol.org/spec-1_0.html].
- Wu, S., Fei, H., Qu, L., Ji, W., & Chua, T.-S. (2024). Next-gpt: Any-to-any multimodal llm. *Forty-first International Conference on Machine Learning*. <https://doi.org/10.48550/arXiv.2309.05519>
- Xenakis, I. (1992). *Formalized music: thought and mathematics in composition*. Pendragon Press. <https://doi.org/10.2307/896037>
- Xiang, Y., Gu, S., & Cao, D. (2022). Investigation of stellar magnetic activity using variational autoencoder based on low-resolution spectroscopic survey. *Monthly Notices of the Royal Astronomical Society*, 514(4), 4781-4793. <https://doi.org/10.1093/mnras/stac1693>
- Xie, Z., He, Q., Zhu, Y., He, Q., & Li, M. (2025). FilmComposer: LLM-Driven Music Production for Silent Film Clips. *arXiv preprint arXiv:2503.08147*. <https://doi.org/10.48550/arXiv.2503.08147>
- Xu, W., McAuley, J., Berg-Kirkpatrick, T., Dubnov, S., & Dong, H.-W. (2024). Generating symbolic music from natural language prompts using an llm-enhanced dataset. *arXiv preprint arXiv:2410.02084*. <https://doi.org/10.48550/arXiv.2410.02084>

- Yang, T., & Li, X. (2015). An autoencoder of stellar spectra and its application in automatically estimating atmospheric parameters. *Monthly Notices of the Royal Astronomical Society*, 452(1), 158-168. <https://doi.org/10.1093/mnras/stv1210>
- Zanella, A., Harrison, C., Lenzi, S., Cooke, J., Damsma, P., & Fleming, S. (2022). Sonification and Sound Design for Astronomy Research, Education and Public Engagement. *arXiv preprint arXiv:2206.13536*. <https://doi.org/10.1038/s41550-022-01721-z>
- Zanella, A., Noel-Storr, J., Misdariis, N., Diaz-Merced, W., Scaletti, C., Gionfrida, L., Roginska, A., Sawe, N., Chafe, C., & Trevino, J. (2022). Sounds of the stars: how scientists are listening in on space. *NATURE*, 611(7935), 204-204. <https://doi.org/10.1038/d41586-022-03597-5>
- Zhang, D. (2023). The Aesthetic Exploration and Ideological Implication of Mozart's Jupiter Symphony. *Advances in Education, Humanities and Social Science Research*, 7(1), 469-469. <https://doi.org/10.56028/aehtsr.7.1.469.2023>
- Zhang, L., & Fuentes, M. (2025). Sonique: Video background music generation using unpaired audio-visual data. *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1-5. <https://doi.org/10.1109/ICASSP49660.2025.10889053>
- Zhang, Y., Guo, W., Pan, C., Zhu, Z., Li, R., Lu, J., Huang, R., Zhang, R., Hong, Z., Jiang, Z., et al. (2025). Versatile Framework for Song Generation with Prompt-based Control. *arXiv preprint arXiv:2504.19062*. <https://doi.org/10.48550/arXiv.2504.19062>
- Zhao, Y., Yang, M., Lin, Y., Zhang, X., Shi, F., Wang, Z., Ding, J., & Ning, H. (2025). AI-Enabled Text-to-Music Generation: A Comprehensive Review of Methods, Frameworks, and Future Directions. *Electronics*, 14(6). <https://doi.org/10.3390/electronics14061197>
- Zheng, W., Zhao, P., Huang, K., & Chen, G. (2021). Understanding the property of long term memory for the LSTM with attention mechanism. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 2708-2717. <https://doi.org/10.1145/3459637.3482399>
- Zicarelli, D. (1987). M and jam factory. *Computer Music Journal*, 11(4), 13-29. <https://doi.org/10.2307/3680237>
- Zooniverse. (2025). People-powered research [<https://www.zooniverse.org/>].
- Zotter, F., & Frank, M. (2019). *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*. Springer Nature. <https://doi.org/10.1007/978-3-030-17207-7>
- Zotti, G., & Neubauer, W. (2024). Advanced Virtual Archaeoastronomy. *Cosmovisiones*, 5. <https://doi.org/10.24215/26840162e018>
- Zwicker, E., & Fastl, H. (2013). *Psychoacoustics: Facts and models* (Vol. 22). Springer Science & Business Media.

Anexos

.1 Formato de archivos FITS

El Sistema de Transporte de Imágenes Flexible (FITS), es el formato de archivo estándar para conjuntos de datos astronómicos. Inicialmente diseñado para transportar imágenes astronómicas entre instituciones de investigación en cintas magnéticas, su estructura de datos está completamente especificada en varios artículos publicados desde 1981 (el primer artículo) hasta 2016 (el estándar actual) (CHIAPPETTI et al., 2018; Greisen y Harten, 1981; Grosbøl, 1991).

Este formato, que no depende de ningún software específico, ofrece la posibilidad de definir reglas para calcular las coordenadas mundiales que corresponden a las coordenadas de los píxeles, y no tiene un límite fijo en su tamaño total. Esto permite que diferentes misiones almacenen sus datos usando este formato con una estructura personalizada. Por ejemplo, los archivos FITS de curvas de luz de K2 y Kepler, usados en el Capítulo 5, contienen un encabezado primario y una unidad de datos (HDU) con metadatos almacenados en el encabezado, y dos extensiones. La primera extensión incluye las marcas de tiempo, los flujos SAP y los flujos PDCSAP representados en la curva de luz. La segunda extensión consiste en una imagen formada por los píxeles recolectados del objetivo. Algunos conjuntos de datos, como los archivos DVT de TESS usados en el Capítulo 4, proporcionan un HDU primario con metadatos y varias extensiones para cada objeto observado, así como una extensión adicional con estadísticas sobre la búsqueda. Otros archivos, como el estudio CALIFA descrito en el Capítulo 10, usan el formato FITS para almacenar cubos de datos que contienen miles de espectros de una sola galaxia.

.2 El protocolo OSC

El protocolo Open Sound Control (OSC) fue desarrollado originalmente en 2002 por el Centro de Nuevas Tecnologías de Música y Audio (CNMAT) de la Universidad de California en Berkeley para la comunicación entre computadoras, sintetizadores de sonido y otros dispositivos multimedia (Wright, 2002). Basado en tecnología de redes y con una filosofía de servidor-cliente, se utiliza comúnmente en proyectos basados en computadora para mapear datos no musicales a sonido, controlar la expresión musical basada en gestos, soportar LAN en red y presentaciones musicales compartidas multiusuario o incluso incrustar otros protocolos, como el MIDI, para mejorar la compatibilidad. Los paquetes de datos OSC consisten en un

solo mensaje o un paquete (un grupo de mensajes con una marca de tiempo), que pueden ser representados por datagramas en los protocolos de red UDP y TCP.

.3 Tránsitos planetarios

La técnica de detección de tránsitos fotométricos se basa en la observación de las variaciones del flujo de luz de objetos celestes. Como se ilustra en la Figura 1, cuando un planeta pasa frente a una estrella, produce un eclipse parcial, y la consiguiente pérdida de flujo percibido se puede medir en términos de tiempo y profundidad. La duración del eclipse total tiene una correspondencia directa con la disminución de energía observada en la curva de luz. Según la Ecuación (1), la profundidad del tránsito es proporcional a las relaciones de superficie del planeta y la estrella, R_p/R_* .

$$\frac{R_p}{R_*} = \sqrt{\Delta F} \tag{1}$$

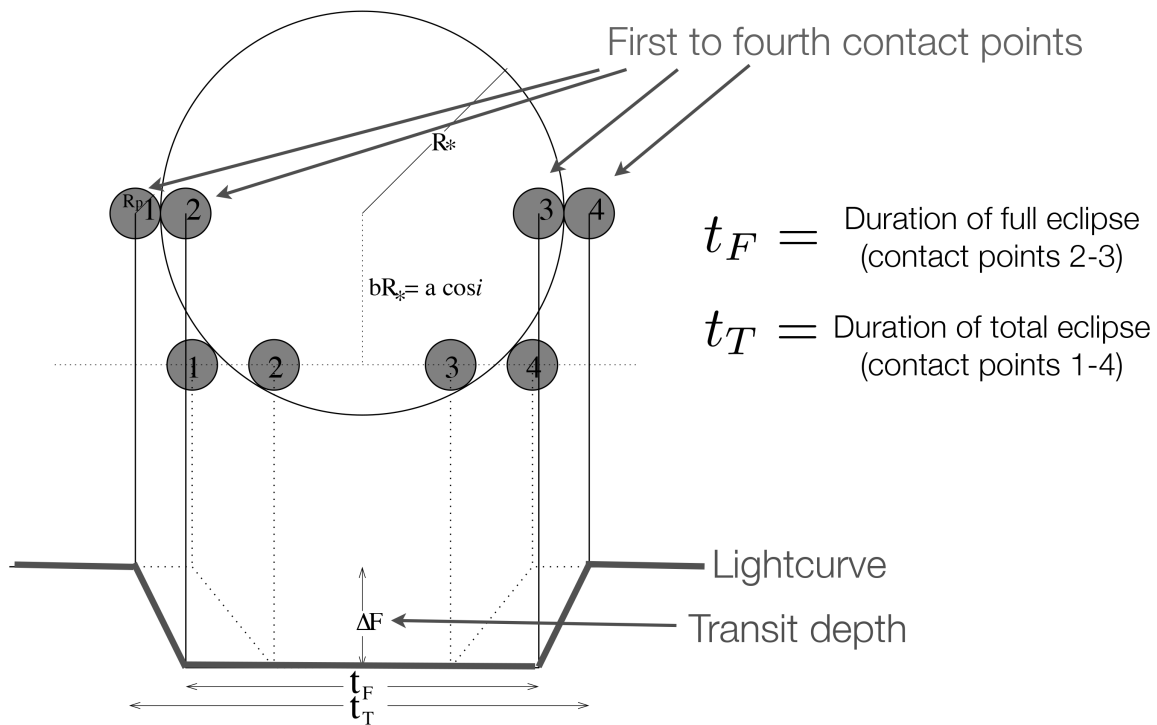


Figura 1: Caracterización del tránsito planetario (Seager y Mallen-Ornelas, 2003).

La Ecuación (2) proporciona la inclinación orbital del planeta, donde a es el radio orbital (asumiendo órbita circular, 3ª ley de Kepler), y b es el parámetro de impacto que representa el aumento de la duración total del eclipse para órbitas cercanas al ecuador de la estrella. También se puede calcular a partir de la curva de luz utilizando la Ecuación (3).

$$i = \cos^{-1} b \frac{R_*}{a} \tag{2}$$

$$b = \sqrt{\frac{(1 - \sqrt{\Delta F})^2 - \left(\frac{t_F}{t_T}\right)^2(1 + \sqrt{\Delta F})^2}{1 - \sqrt{\frac{t_F}{t_T}}}} \quad (3)$$

El periodo orbital de cada planeta se puede calcular a partir de la curva de luz, midiendo el intervalo de tiempo entre tránsitos. Se necesitan al menos tres tránsitos para obtener resultados concluyentes.

.4 Clasificación de espectros estelares

Las líneas de absorción y emisión dibujadas en un espectro representan las longitudes de onda en las que los elementos químicos absorben o emiten luz. La Tabla 1 ofrece una vista esquemática de algunas líneas espectrales comunes y sus longitudes de onda correspondientes. Su análisis no sólo se utiliza en el estudio de la caracterización y clasificación de estrellas, sino que también se aplica a espectros completos de galaxias como una forma de obtener “un poderoso diagnóstico de su contenido estelar y propiedades evolutivas” (Kennicutt Jr, 1992a, 1992b).

Tabla 1: Longitudes de onda electromagnéticas para líneas espectrales comunes (Survey, 2003).

Líneas espectrales	Longitudes de onda (Å)
$H_\alpha, H_\beta, H_\gamma$	6600, 4800, 4350
H and K (Calcio ionizado)	3800-4000
Óxido de Titanio	Multiples líneas en rangos: 4900-5200, 5400-5700 6200-6300, 6700-6900
Banda G	4250
Sodio	5800
Helio (neutro)	4200
Helio (ionizado)	4400

Los espectros estelares se pueden clasificar según sus características de absorción y temperatura efectiva utilizando la clasificación espectral de Harvard. Este sistema utiliza la secuencia O, B, A, F, G, K, M que va desde las estrellas más calientes (más de 28.000K) hasta las más frías (menos de 3.500K). Dentro de cada tipo espectral, 10 subclases adicionales numeradas del 0 al 9 permiten la clasificación de los espectros en función de la intensidad de sus líneas de absorción (Gray y Corbally, 2009). Se requiere una clase de luminosidad adicional para diferenciar estrellas con la misma temperatura pero con diferente luminosidad. La clasificación de Morgan-Keenan (Morgan et al., 1943) utiliza números romanos para diferenciar entre estrellas supergigantes extremadamente luminosas (clase *I*) y estrellas enanas de la secuencia principal (clase *V*). La Figura 2 proporciona un ejemplo para cada tipo espectral, incluyendo bandas de longitud de onda.

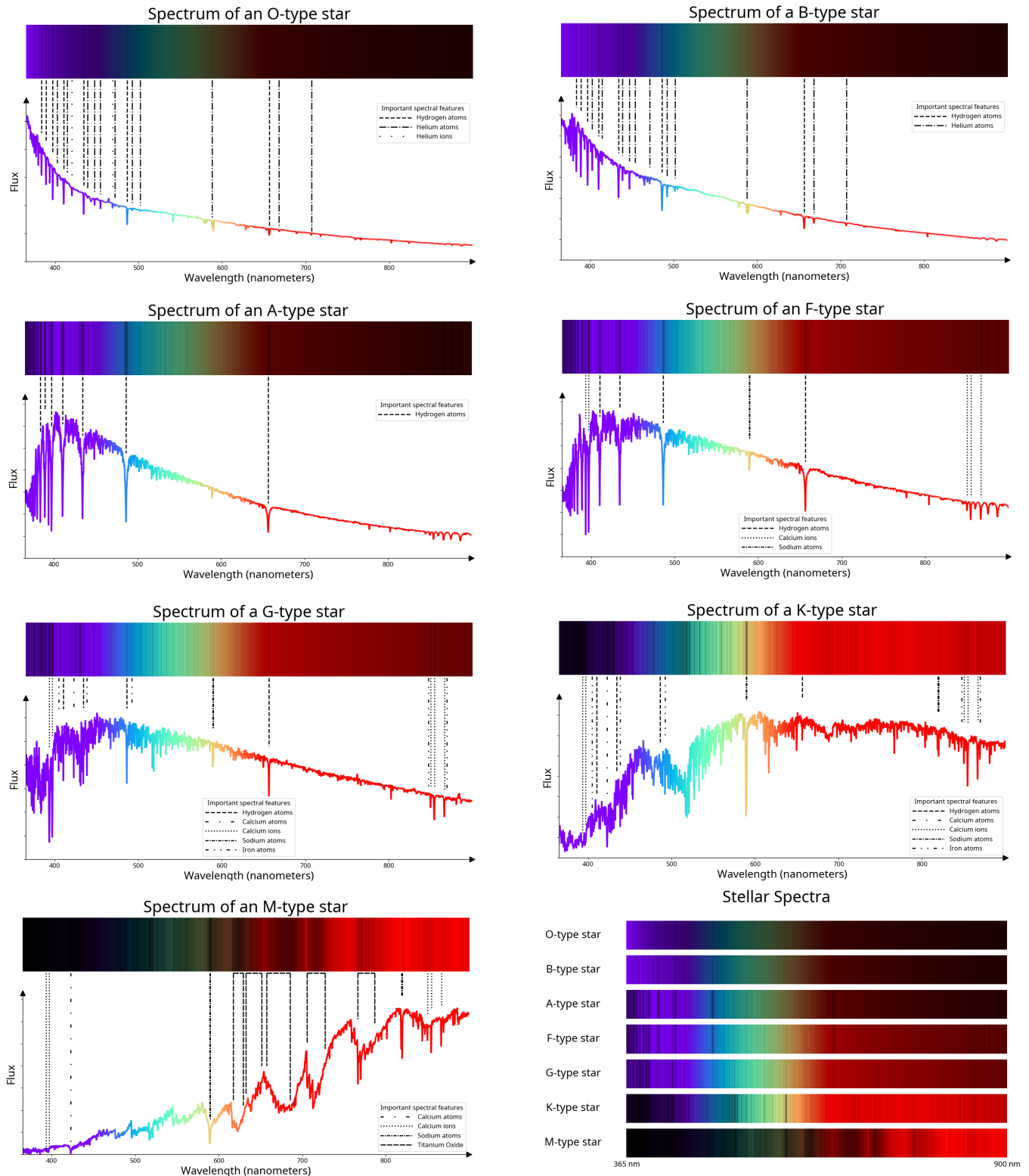


Figura 2: Clasificación de espectros estelares (IAU, 2020).

La clase espectral *O* fue definida originalmente por la presencia de líneas de absorción y, a veces, de emisión de He II en longitudes de onda azul-violeta. El estudio de las estrellas *B* condujo al descubrimiento y mapeado de la estructura espiral de la Vía Láctea. Esta clase fue definida originalmente para aquellas estrellas que mostraban líneas de He I en ausencia de He II en el azul-violeta. Las estrellas *A* se caracterizan por la desaparición de las líneas de He I y, a menudo, la presencia de peculiaridades químicas. Una estrella de tipo *F* presenta una atmósfera en la que ocurren cambios físicos importantes. El sol es un ejemplo de estrella de tipo *G*, marcadas por su abundancia de características espectrales. Los tipos *G* y *K* son las estrellas con más probabilidades de tener planetas habitables a su alrededor, mientras que las estrellas de tipo *M* son las más numerosas de todas las clases (Gray y Corbally, 2009).

La Figura 3 muestra dos espectros de una estrella A2V y una M2V que alcanzan su punto máximo respectivamente en las bandas UV e infrarroja. El gráfico proporciona una comparación de los filtros Kepler, CoRoT y Johnson B,V,R,I. Se puede encontrar información adicional sobre los perfiles de los filtros en:

<http://svo2.cab.inta-csic.es/theory/fps/>

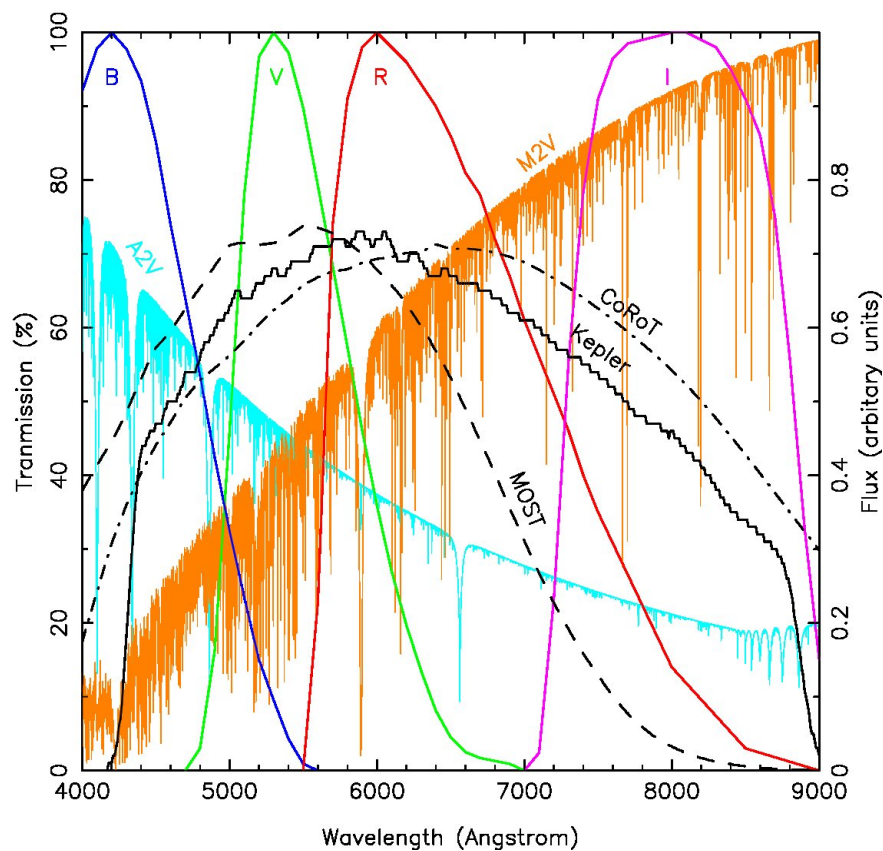


Figura 3: Comparación de las curvas de respuesta de Kepler, CoRoT y Johnson (Jason Rowe, extraído de 2009IAUS..253..121R) (NASA, 2020a). Funciones de transmisión escaladas para alcanzar el 100% de transmisión. Espectro de la estrella A2V alcanzando su punto máximo en la banda UV (cian), y espectro de la estrella M2V alcanzando su punto máximo en el infrarrojo (naranja).

.5 Temperatura efectiva, diagrama HR, metalicidad y gravedad superficial

La temperatura efectiva (T_{eff}) de una estrella es la temperatura de un cuerpo negro del mismo tamaño que la estrella y que irradiaría la misma cantidad total de energía electromagnética emitida por la estrella. Es uno de los parámetros utilizados para caracterizar las estrellas, ya que está directamente relacionado con su color (cuanto mayor es la temperatura, más azul es la luz emitida por la estrella). Se utiliza junto con la luminosidad estelar para construir el diagrama de clasificación HR clásico. A modo de ejemplo, la temperatura efectiva del Sol es de 5.780 K, para una estrella de tipo O es de 40.000 K y para una estrella de tipo M de 3.000 K (Gargaud y Amils, 2011).

El diagrama de Hertzsprung-Russell (HR) representa el brillo intrínseco o magnitud absoluta (M_v) de las estrellas frente a su temperatura superficial. En el diagrama HR, las estrellas representadas aparecen restringidas a unas pocas regiones bien definidas. Dentro de estas regiones, las estrellas comparten características comunes. Además, a medida que las características físicas de una estrella cambian con el tiempo, el diagrama HR también se utiliza para trazar la evolución estelar. La Figura 4 muestra ambas aplicaciones de los diagramas HR.

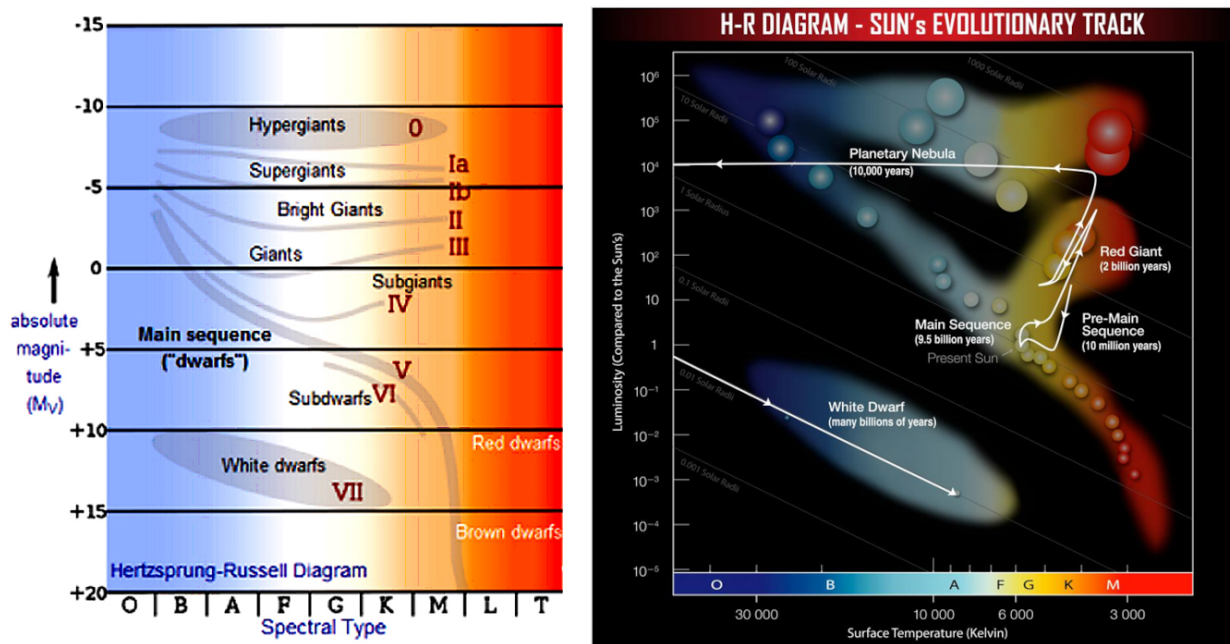


Figura 4: Dos ejemplos de diagramas HR utilizados respectivamente para la clasificación y evolución estelar (Harvard-Smithsonian, 2015). Clasificaciones espectrales de Harvard y Morgan-Keenan sobre un diagrama HR (izquierda), y la trayectoria evolutiva del sol (derecha).

La metalicidad (Z), especifica la cantidad relativa de elementos más pesados que el helio (llamados *metales* por los astrónomos) en una estrella, galaxia o alguna parte del medio interestelar. La metalicidad del sol es $Z_{\odot} = 1,5\%$ en masa. El rango de metalicidades de las estrellas en la Vía Láctea va desde $4,5 Z_{\odot}$ hasta $0,0001 Z_{\odot}$ (Gargaud y Amils, 2011).

La gravedad superficial ($\log g$) se define como la aceleración de una masa en caída libre en la superficie de un cuerpo masivo bajo la fuerza gravitacional producida por el cuerpo. La gravedad superficial en la Tierra es de $9,81 \text{ ms}^{-2}$. La gravedad superficial (g) de un cuerpo esférico se deriva de la expresión de Newton, Ecuación 4, donde G es la constante gravitacional ($6,672 \times 10^{-11} \text{ Nm}^2/\text{kg}^2$), y M y R son la masa y el radio del cuerpo (Gargaud y Amils, 2011).

$$g = \frac{GM}{R^2} \quad (4)$$

.6 Espectroscopía de transmisión sonificada

Las líneas de absorción y emisión dibujadas en un espectro representan las longitudes de onda en las que los elementos químicos absorben o emiten luz. Su análisis se utiliza en la caracterización y clasificación de estrellas y galaxias. Aplicando una conversión directa de cubos de datos a sonido (técnica de sonificación que recibe el nombre de audificación), se han logrado resultados prometedores en la sonificación de espectros de galaxias (Trayford et al., 2023).

En un enfoque completamente diferente, basado en la representación de datos pre-procesados, se implementó el prototipo *AbsEmis2Sound* para la sonificación de espectros del Large Early Galaxy Census (LEGA-C) (ESO, 2018). Esta herramienta proporciona una sonificación simple de las líneas de absorción y emisión. Basado en una filosofía de detección de eventos, el prototipo puede revelar la existencia de líneas ocultas adicionales superpuestas en representaciones gráficas como la mostrada en la Figura 5. Esta estrategia podría ser útil en el análisis de datos científicos y podría proporcionar una metodología que haga más accesible este análisis para los investigadores ciegos o con poca visión (BLV).

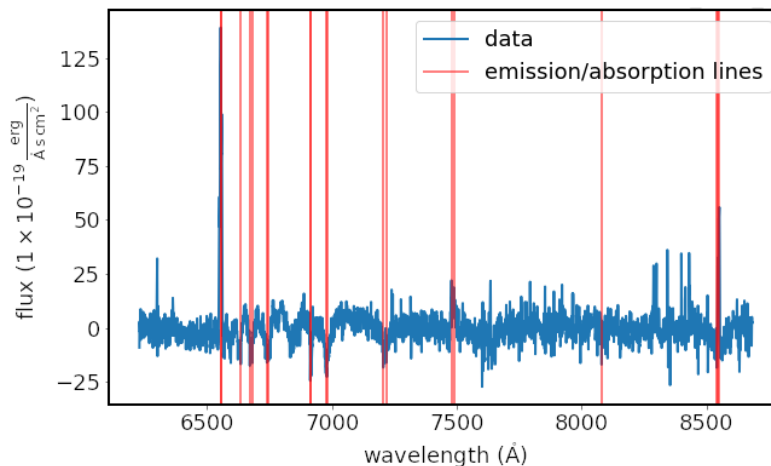


Figura 5: Ejemplo de espectro LEGA-C pre-procesado. Espectro restado del continuo y líneas de emisión marcadas mediante el módulo `find lines threshold` para la galaxia COSMOS M01 con coordenadas RA: 150.220555, DEC: 2.41104.

En cuanto al mapeado de parámetros, se utilizó una estrategia directa de conversión de la longitud de onda en frecuencia, con el fin de hacer el modelo lo más intuitivo posible. El

análisis espectroscópico propuesto incluye el suavizado, ajuste y substracción del continuo, como parte del proceso de identificación de líneas, que se basa en el ecosistema de herramientas de análisis de datos post-pipeline del STScI, 2020. El prototipo desarrollado *ad hoc* para este análisis está disponible en:

<https://github.com/AuditoryVO/AbsEmis2Sound>

El primer vídeo proporcionado a continuación muestra cómo el diseño permite a los usuarios finales conmutar entre la representación de líneas de emisión y absorción pre-calculadas, para analizar su número, amplitud y frecuencia. En el segundo vídeo, se presentan cuatro enfoques diferentes de sonificación. La primera representación genera un acorde simultáneo en el que cada línea de emisión o absorción está representada por un instrumento musical diferente. La segunda proporciona una representación similar utilizando ondas sinusoidales. El efecto de “batimiento” producido cuando dos o más ondas sinusoidales están cerca en frecuencia ayuda en la detección de líneas de absorción o emisión adicionales, pero para percibir con precisión su número se necesita otro tipo de enfoque. Una solución es presentar las líneas secuencialmente. El tercer ejemplo muestra este tipo de sonificación, que utiliza notas de piano, y en el que la altura de cada nota corresponde directamente a la longitud de onda de cada línea de absorción o emisión. Este enfoque permite contar los eventos, pero aquellas líneas que están próximas en frecuencia aparecen representadas con la misma nota debido a la cuantificación MIDI. El ejemplo final realiza una representación secuencial con ondas sinusoidales, lo que mejora la diferenciación de eventos de frecuencia cercana y la apreciación de su amplitud. Los eventos de sonido también se han panoramizado de izquierda a derecha según su posición en el gráfico para facilitar su identificación.

<https://vimeo.com/771366086>

<https://vimeo.com/793542919>

.7 Mapeado de parámetros formalizado

El mapeado de parámetros es una metodología de sonificación que establece una correspondencia bidireccional entre los parámetros de datos y los parámetros de sonido. Este método se utiliza ampliamente en el análisis, exploración y comunicación de datos, ya que permite una percepción intuitiva de patrones y cambios de tendencia, la retroalimentación en tiempo real necesaria para aplicaciones interactivas y la representación multidimensional.

A modo de ejemplo, el mapeado de parámetros más sencillo realizado en este trabajo de investigación puede encontrarse en la Sección 3.3, donde el flujo de luz se hace corresponder a la frecuencia fundamental de una senoide a lo largo del eje x , que se interpreta como línea de tiempo. El proceso se puede formalizar como se expresa en la Ecuación 5:

$$S(x) = \mathcal{M}(\text{flux}(x)) = A_o \sin(\text{flux}(x)x) \quad (5)$$

donde S es el sonido generado para cada valor de flujo en el instante x (en curvas de luz/series de tiempo) o longitud de onda (en espectros), y \mathcal{M} es la función de mapeado que convierte

los valores de flujo en ondas sinusoidales, con amplitud constante $A_o = 1$ para la señal útil y $A_o = 0$ para la ausencia de datos.

En situaciones más complejas, la función de mapeado (\mathcal{M}) se puede personalizar para enfatizar aspectos específicos de los datos, resaltando características interesantes o dirigiendo la atención del oyente a objetivos específicos.

.8 Descripción formal de un autoencoder

Esta sección proporciona una descripción formal de los autoencoders utilizados en esta tesis para la reducción de espectros estelares a vectores de baja dimensión sonificados.

.8.1 Sparse autoencoders

Los autoencoders son redes neuronales entrenadas para intentar reconstruir una copia imperfecta de su entrada. La red consta de una función de codificación $h = f(x)$, que extrae información significativa de la entrada (restringiendo h a tener una dimensión menor que x), y una función de decodificación $r = g(h)$, que genera la reconstrucción a partir del espacio latente codificado h .

El proceso de aprendizaje consiste en minimizar una función de pérdida (como el MSE), que penaliza a $g(f(x))$ por ser diferente de x (Goodfellow et al., 2016). La función de pérdida \mathcal{L} se puede expresar como se muestra en la Ecuación 6.

$$\mathcal{L}(x, g(f(x))) \tag{6}$$

Un sparse autoencoder agrega una penalización de dispersión $\Omega(h)$ al error de reconstrucción en el criterio de entrenamiento, tal y como muestra en la Ecuación 7. Este término de regularización responde a características estadísticas únicas del conjunto de datos, que pueden ser útiles en tareas de clasificación.

$$\mathcal{L}(x, g(f(x))) + \Omega(h) \tag{7}$$

.8.2 Autoencoders variacionales

Los autoencoders variacionales combinan redes neuronales con métodos bayesianos variacionales para proporcionar una distribución latente cercana a la distribución previa, que generalmente es una distribución normal estándar como la mostrada en la Ecuación 8.

El codificador mapea los datos de entrada $x \in \mathbb{R}^j$ a un espacio latente $z \in \mathbb{R}^1$. La salida del codificador es el espacio latente descrito por la Ecuación 9, que representa la distribución posterior aproximada ($q_\phi(z|x)$) con media μ y varianza σ^2 . θ y ϕ son, respectivamente, los parámetros optimizados en el proceso de entrenamiento (Kingma, 2013).

$$p_\theta(z) = \mathcal{N}(z; 0, I). \tag{8}$$

$$z \sim q_\phi(z|x) = \mathcal{N}(z; \mu_\phi(x), \sigma_\phi^2(x)) \quad (9)$$

El decodificador reconstruye los datos de entrada a partir de la variable latente z . Su salida es la probabilidad de verosimilitud ($p_\theta(x|z)$) expresada en la Ecuación 10.

$$p_\theta(x|z) = \mathcal{N}(x; \mu_\theta(z), \sigma_\theta^2(z)) \quad (10)$$

El proceso de entrenamiento de un VAE consiste en encontrar los parámetros óptimos de la red que maximizan la función objetivo, denominada Evidence Lower Bound Optimization (ELBO) (Damm et al., 2023) y mostrada en la Ecuación 11. El primer término de la ecuación representa las pérdidas de reconstrucción (\mathbb{E}), el segundo término es la divergencia de Kullback-Leibler (\mathbb{D}_{KL}) entre la distribución posterior aproximada y la distribución previa $p_\theta(z)$.

$$\mathcal{L}(\theta, \phi; x) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - \mathbb{D}_{KL}(q_\phi(z|x) || p_\theta(z)) \quad (11)$$

Tal y como se describe en la Sección .8.3, el término KL se puede expresar con la Ecuación 12:

$$\mathbb{D}_{KL}(q_\phi(z|x) || p_\theta(z)) = \int q_\phi(z|x) \log \frac{q_\phi(z|x)}{p_\theta(z)} dz \quad (12)$$

Finalmente, para el caso gaussiano, la ELBO se puede expresar según muestra la Ecuación 13.

$$\mathcal{L}(\theta, \phi; x) \simeq \frac{1}{L} \sum_{l=1}^L \log p_\theta(x|z) + \frac{1}{2} \sum_{j=1}^J (1 + \log((\sigma_j)^2) - (\mu_j)^2 - (\sigma_j)^2) \quad (13)$$

.8.3 Divergencia de Kullback-Leibler

La divergencia de Kullback-Leibler se utiliza comúnmente para medir la similitud entre dos distribuciones de probabilidad (Odaibo, 2019). Asumiendo que el contenido de información de eventos altamente probables es bajo, la información está inversamente relacionada con la probabilidad de un evento, y se puede expresar como se muestra en la Ecuación 14, donde I representa el contenido de información del evento x con respecto a diferentes probabilidades $p(x)$ y $q(x)$.

$$I_p(x) = -\log p(x), \text{ y } I_q(x) = -\log q(x) \quad (14)$$

La diferencia de información entre $q(x)$ y $p(x)$ es:

$$\Delta I = I_p(x) - I_q(x) = -\log p(x) + \log q(x) = \log\left(\frac{q(x)}{p(x)}\right) \quad (15)$$

Y la divergencia de Kullback-Leibler (\mathbb{D}_{KL}) es la expectativa (\mathbb{E}) de esta diferencia (no simétrica):

$$\mathbb{D}_{\text{KL}}(q(x)||p(x)) := \mathbb{E}_{\sim q}[\Delta I] = \int (\Delta I)q(x)dx = \int q(x)\log\left(\frac{q(x)}{p(x)}\right)dx \quad (16)$$

$$\mathbb{D}_{\text{KL}}(p(x)||q(x)) := \mathbb{E}_{\sim p}[\Delta I] = \int (\Delta I)p(x)dx = \int p(x)\log\left(\frac{p(x)}{q(x)}\right)dx \quad (17)$$

En el caso de los autoencoders variacionales, durante el proceso de entrenamiento, la divergencia de Kullback-Leibler de las probabilidades en los espacios original y latente se minimiza mediante el algoritmo del descenso de gradiente, asegurando que la distribución del espacio latente aprendido sea cercana a la distribución previa.

.9 Descripción formal de una red LSTM

Las redes LSTM (Hochreiter, 1997) fueron desarrolladas para resolver los problemas de desvanecimiento y explosión de gradiente que aparecen en redes neuronales con retro-propagación tales como las Redes Neuronales Recurrentes (RNN) tradicionales. Este problema es crítico cuando el número de pasos de propagación es alto (redes profundas), ya que los gradientes de los pesos anteriores se vuelven exponencialmente más pequeños que los gradientes de los pesos posteriores. Esta diferencia produce inestabilidad en el proceso de entrenamiento. Las redes LSTM lo resuelven introduciendo un “Carrusel de Error Constante (CEC) o auto-bucle, que mantiene la señal de error dentro de la celda de cada unidad” (Van Houdt et al., 2020). La red acumula información de dependencias a largo plazo en datos secuenciales y, una vez que se utiliza la información, olvida el estado anterior en un ciclo recurrente “puertado”.

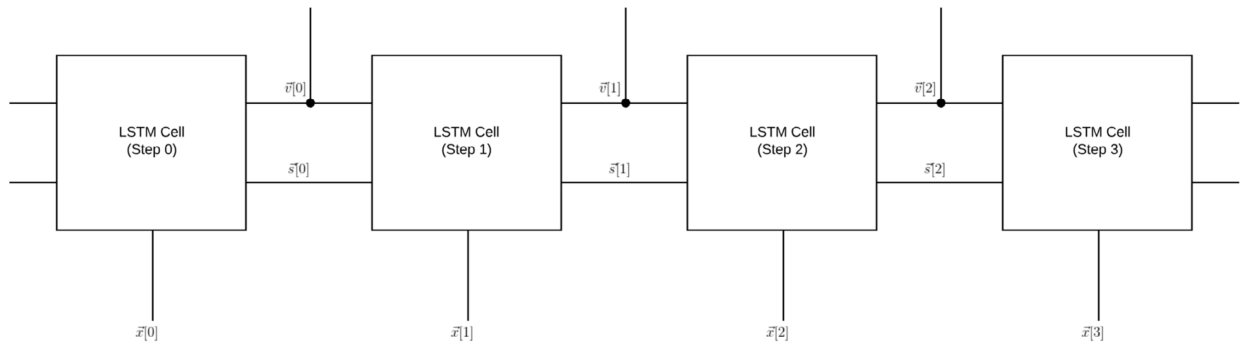


Figura 6: Secuencia de cuatro pasos de una celda de red LSTM (Sherstinsky, 2020), con \vec{x} representando el vector de entrada a la celda, \vec{s} representando la señal de estado de la celda, y \vec{v} representando la señal de valor observable de la celda para propósitos externos.

Las redes LSTM constan de celdas (RNN en sí mismas), puertas de entrada, puertas de salida y puertas de borrado de memoria. La siguiente descripción formal proporciona las ecuaciones de propagación que definen completamente la celda LSTM simple mostrada en las Figuras 6 y 7 (Sherstinsky, 2020).

Tal y como se describe en Goodfellow et al., 2016, el componente más importante de una red LSTM es la unidad de estado $s_i^{(t)}$, expresada en la Ecuación 18. En esta ecuación, $x^{(t)}$

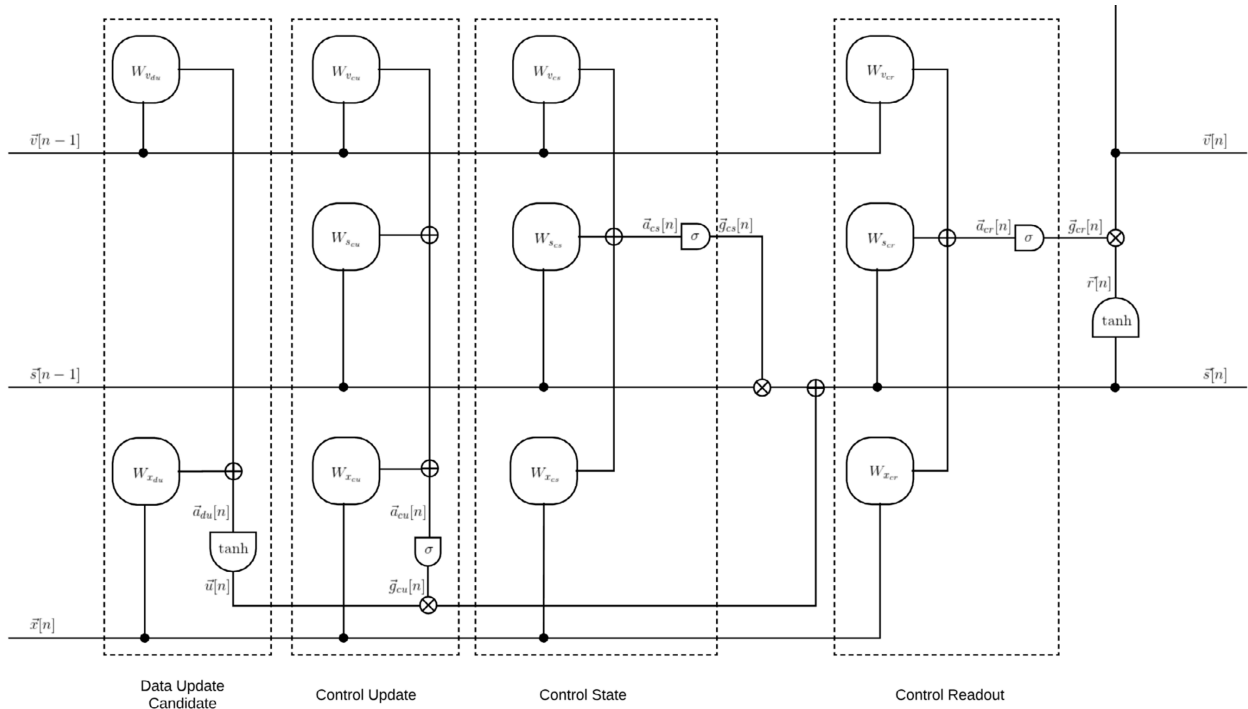


Figura 7: Celda de red LSTM simple con las etapas del sistema delineadas por rectángulos discontinuos (Sherstinsky, 2020), donde \vec{x} es el vector de entrada, \vec{s} es el estado de la celda, y \vec{v} es el valor observable de la celda para propósitos externos. $\vec{a}_{cu}[n]$ es el nodo de acumulación, con parámetros $W_{x_{cu}}$, $W_{s_{cu}}$, $W_{v_{cu}}$, y \vec{b}_{cu} . $\vec{a}_{cs}[n]$ es el nodo de acumulación de la puerta que controla la fracción de la señal de estado de la celda ($\vec{s}[n-1]$), con parámetros $W_{x_{cs}}$, $W_{s_{cs}}$, $W_{v_{cs}}$, y \vec{b}_{cs} . $\vec{a}_{cr}[n]$ es el nodo de acumulación de la puerta que controla la fracción de la señal estimada de lectura ($\vec{r}[n]$) utilizada para liberar la señal observable de la celda, con parámetros $W_{x_{cr}}$, $W_{s_{cr}}$, $W_{v_{cr}}$, y \vec{b}_{cr} . $\vec{a}_{du}[n]$ es el nodo de acumulación para la función de transformación de datos, que actualiza la señal estimada ($\vec{u}[n]$), con parámetros $W_{x_{du}}$, $W_{v_{du}}$, y \vec{b}_{du} .

representa el vector de entrada actual, $h^{(t)}$ es el vector de capa oculta que contiene las salidas de todas las celdas LSTM, y b^f , U^f y W^f son, respectivamente, los sesgos, pesos de entrada y pesos recurrentes para las puertas de borrado de memoria. El peso del auto-bucle está controlado por $f_i^{(t)}$, la unidad de puerta de borrado de la Ecuación 19.

$$s_i^{(t)} = f_i^{(t)} s_i^{(t-1)} + g_i^{(t)} \sigma \left(b_i + \sum_j U_{i,j} x_j^{(t)} + \sum_j W_{i,j} h_j^{(t-1)} \right) \quad (18)$$

$$f_i^{(t)} = \sigma \left(b_i + \sum_j U_{i,j} x_j^{(t)} + \sum_j W_{i,j} h_j^{(t-1)} \right) \quad (19)$$

La unidad de puerta de entrada externa se calcula con una unidad sigmoide, de manera similar a la puerta de borrado de memoria pero con sus propios parámetros, a través de la expresión de la Ecuación 20.

$$g_i^{(t)} = \sigma \left(b_i^g + \sum_j U_{i,j}^g x_j^{(t)} + \sum_j W_{i,j}^g h_j^{(t-1)} \right) \quad (20)$$

Finalmente, la salida de la celda LSTM, $h_i^{(t)}$, está controlada por la puerta de salida $q_i^{(t)}$, que usa una unidad sigmoide tal y como describen las Ecuaciones 21 y 22.

$$h_i^{(t)} = \tanh \left(s_i^{(t)} \right) q_i^{(t)} \quad (21)$$

$$q_i^{(t)} = \sigma \left(b_i^o + \sum_j U_{i,j}^o x_j^{(t)} + \sum_j W_{i,j}^o h_j^{(t-1)} \right) \quad (22)$$

.10 Mecanismo de atención en redes LSTM

La inclusión de un mecanismo de atención en las redes LSTM mejora su rendimiento de pronóstico y permite a las redes capturar información temporal de observaciones históricas. Este tipo de mecanismos reducen el error de pronóstico capturando potencialmente la influencia dinámica de las series exógenas más relevantes en la serie objetivo (Zheng et al., 2021).

Usando el estado oculto histórico $h^{(t-1)}$ y el estado de la celda de memoria $s^{(t-1)}$ en el codificador, un mecanismo de atención estándar (e_i) se puede describir como se muestra en la Ecuación 23, donde x_i representa los datos de entrada, y v_e , W_e , U_e y b_e representan los parámetros entrenables.

$$e_i^{(t)} = v_e \tanh \left(W_e [h_i^{(t-1)}; s_i^{(t-1)}] + U_e x_i^{(t)} + b_e \right) \quad (23)$$

La puntuación de atención de la Ecuación 24 se utiliza para medir la importancia de las series exógenas de entrada para la serie objetivo, obteniendo finalmente los estados ocultos $\tilde{h}^{(t)}$, tal y como se expresa en la Ecuación 25.

$$\alpha_i^{(t)} = \text{Softmax}(e_i^{(t)}) \quad (24)$$

$$\tilde{h}^{(t)} = \sum_j \alpha_j^{(t)} h_j^{(t)} \quad (25)$$

.11 Descripción formal de un sintetizador

La siguiente formalización describe el sintetizador incluido en *SoniCube*, que fue presentado en el Capítulo 10, como ejemplo representativo de los sintetizadores implementados en este trabajo.

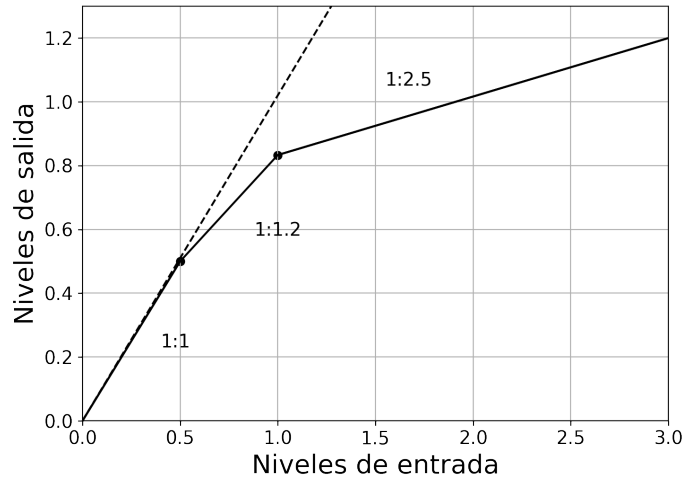


Figura 8: Curva de nivel de transferencia utilizada para el control de valores de flujo salientes que pueden causar daños auditivos al representar cubos de datos inexplorados.

La ecuación del sintetizador aditivo de seis osciladores implementado en *SoniCube* se puede expresar para cada sonificación como:

$$S(t) = \sum_{i=0}^6 A_i F r \sin(2\pi f_i t + \phi_i) \quad (26)$$

donde A_i es el coeficiente de ponderación A para cada frecuencia, F es la mediana del flujo absoluto del espectro de galaxia representado, r es la relación o pendiente del limitador/compresor de rango dinámico utilizado para controlar los valores de flujo salientes, f_i son las frecuencias fundamentales obtenidas a través del proceso de reducción de dimensión del autoencoder (6 dimensiones), y ϕ_i , las fases relativas de los osciladores (en nuestro caso, todas establecidas en cero).

Atendiendo a la sonoridad de cada sonificación, los coeficientes de ponderación A se calculan para cada una de las seis frecuencias o formantes como se describe en la Ecuación 27 (IEC, 2013).

$$A_i = 20 \log R_A(f_i) - 20 \log R_A(1000) \approx 20 \log R_A(f_i) + 2 \quad (27)$$

con $R_A(f_i)$ calculado a partir de la Ecuación 28:

$$R_A(f) = \frac{12194^2 f^4}{(f^2 + 20, 6^2) \sqrt{(f^2 + 1007, 7^2)(f^2 + 737, 9^2)(f^2 + 12194^2)}} \quad (28)$$

La Figura 8 muestra la curva de nivel de transferencia del limitador/compresor de rango dinámico de dos etapas incluido en *SoniCube* para prevenir daños auditivos. La pendiente de cada etapa está representada por r en la Ecuación 26.

El sintetizador incorpora un procesador de reverberación estéreo basado en ocho líneas de retardo (Kahrs y Brandenburg, 1998) que se suma al flujo de señal para proporcionar información sobre la distancia del spaxel seleccionado respecto al centro de la galaxia. La señal final enviada al codificador binaural se puede expresar como:

$$X_{L,R}(t) = S(t) + d \sum_{j=0}^8 g^{j-1} S(t - j\tau) \quad (29)$$

donde $S(t)$ es la salida del sintetizador aditivo, d es la distancia del spaxel seleccionado por el usuario al punto de referencia (centro de la galaxia), g es la ganancia del bucle de retroalimentación (en nuestro caso de estudio se establece en 0,9 para proporcionar un efecto de reverberación largo), y τ es el tiempo de retardo fijo aplicado a cada línea.

.12 Codificación binaural

El oído humano puede localizar fuentes de sonido en un espacio 3D a través del análisis de las diferencias de nivel interaural (ILD) y los retardos de tiempo interaural (ITD). Las alteraciones producidas en un sonido al viajar desde la fuente al oyente se pueden caracterizar utilizando Head Related Transfer Functions (HRTF) (C. P. Brown y Duda, 1997). Los sistemas binaurales se basan en la convolución de la HRTF de ambos oídos con las fuentes de sonido, lo que permite la espacialización de ubicaciones estáticas y dinámicas (Lazzarini y Carty, 2008).

La relación entre la fuente de sonido y la señal que llega a los oídos del oyente se puede expresar en términos de los ángulos de azimut (AZ) y elevación (EL), la distancia a la fuente (d) y la frecuencia angular (w), como se muestra en la Ecuación 30.

$$Y_{L,R}(AZ, EL, d, w) = H_{L,R}(AZ, EL, d, w) * X_{L,R}(w) \quad (30)$$

donde $Y_{L,R}$ son los espectros de audio de las señales acústicas en los oídos del oyente, $H_{L,R}$ es la HRTF y $X_{L,R}$ son los espectros de audio de la fuente de sonido.

Los prototipos binaurales desarrollados en esta tesis utilizan el opcode binaural *hrtfmove* *Csound* (Carty, 2008) para representar objetos reales del cielo en un paisaje sonoro 2D inmersivo correlacionado con una visualización de pantalla plana. El azimut se calcula a partir de coordenadas reales (x,y) y la elevación se establece en cero para descartar la tercera dimensión. La distancia no se utiliza para mantener los niveles de las sonificaciones dependientes del flujo representado, reduciendo la expresión general a la Ecuación 31.

$$Y_{L,R}(AZ, 0, w) = H_{L,R}(AZ, 0, w) * X_{L,R}(w) \quad (31)$$

.13 Codificación ambisónica

Los sistemas de sonido ambisónico permiten generar campos sonoros envolventes a partir de la representación de fuentes sonoras virtuales en un entorno 3D. En este formato, la fuente de

sonido se codifica y almacena en un número variable de canales llamados formato-B, que no está relacionado con el número de altavoces o canales utilizados en la proyección final. La codificación ambisónica de primer orden utiliza un formato-B de 4 canales, la de segundo orden utiliza 9 canales y la de tercer orden 16. Las codificaciones de orden superior conducen, en general, a una mejor resolución espacial (Zotter y Frank, 2019). La representación inmersiva final requiere un proceso de decodificación para convertir el formato-B en las señales de audio finales que alimentan la configuración de altavoces elegida. La codificación ambisónica de primer orden requiere un mínimo de 4 altavoces para una representación correcta, la de segundo orden 6 o 9 altavoces si se requiere tridimensionalidad, y la de tercer orden 16 también en representaciones 3D (Roginska y Geluso, 2017).

Este trabajo de investigación implementa codificación ambisónica de primer orden (FOA) en algunos prototipos, requiriendo la codificación en un formato-B con tres dimensiones básicas ($a_0, a_{1,1}, a_{1,2}$), y su decodificación en ocho señales de audio finales que alimentan la configuración de ocho altavoces propuesta. Para implementar esta configuración en *CSound* se ha desarrollado un opcode definido por el usuario (UDO) (Neukom, 2014). Las expresiones que rigen la codificación se muestran en las Ecuaciones 32, donde φ representa el ángulo de las coordenadas polares de cada punto representado, y $a_0, a_{1,1}$ y $a_{1,2}$ las señales codificadas en formato-B. La ascensión recta (RA) de las coordenadas reales de cada objeto celeste se interpretan como φ en la implementación final.

$$a_0 = Source(mono) \quad (32a)$$

$$a_{1,1} = X = Source \cdot \cos\varphi \quad (32b)$$

$$a_{1,2} = Y = Source \cdot \sin\varphi \quad (32c)$$

Este mapeado directo de la ascensión recta a las coordenadas de azimut se realizó para proporcionar representaciones 2D correlacionadas con la proyección de las imágenes en pantalla plana ($RA = AZ = \varphi$). Pruebas preliminares revelaron cómo la conversión precisa de la ascensión recta y la declinación a azimut y elevación produce sonificaciones no correlacionadas con la esfera celeste mostrada en la representación gráfica y, además, requieren la introducción de las coordenadas de referencia de ubicación del usuario, produciendo potencialmente resultados variables.

Las ecuaciones de decodificación para ocho altavoces colocados simétricamente en un círculo y espaciados $\phi_n = 45 \cdot (n - 1)$ grados, se pueden expresar como se muestra en la Ecuación 33, con a_n representando la señal de salida que alimenta el altavoz “n”.

$$a_n = 2/3(a_0/2 + a_{1,1} \cdot \cos\phi_n + a_{1,2} \cdot \sin\phi_n) \quad (33)$$

.14 Coeficiente de determinación

El coeficiente de determinación R^2 (Cameron y Windmeijer, 1997), se utiliza para medir cuantitativamente la capacidad de un modelo de predicción para replicar los datos originales.

Al igual que la varianza, no proporciona resultados comparables entre diferentes conjuntos de datos. Su valor máximo es 1 y puede ser negativo si el modelo predice resultados arbitrarios. Para un conjunto de datos de entrada no constantes, si el modelo predice constantemente la media (ignorando las variaciones en los datos originales), el R^2 será cero.

El R^2 puede definirse según la Ecuación 34 (“Scikit-learn”, 2007):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (34)$$

Donde $\hat{y}_i \in \mathbb{R}^n$ representa el valor de la predicción de la muestra i , $y_i \in \mathbb{R}^n$ el valor de la muestra original, $\bar{y}_i = 1/n \sum_{i=1}^n y_i$ y $\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (\epsilon_i)^2$.

.15 Análisis de correlación

A lo largo del desarrollo de esta investigación, se han probado diversos tipos de autoencoders y configuraciones de redes neuronales. El coeficiente de determinación R^2 se ha utilizado en todas las pruebas para evaluar el proceso de codificación utilizando diferentes modelos y casos de estudio. Tal y como muestra el Anexo .14, este coeficiente proporciona un método de cuantificación de la correlación entre los espectros originales y los espectros decodificados recuperados del espacio latente, con 1 indicando que la salida del modelo es una copia perfecta de los datos de entrada. El cálculo del coeficiente se ha implementado con la librería de Python *Scikit-learn* (“Scikit-learn”, 2007).

En este anexo se presentan algunos resultados obtenidos con las librerías de espectros estelares STELIB y MILES, y el estudio de galaxias CALIFA. En todos los casos, se ha encontrado que los resultados finales dependen en gran medida de las peculiaridades y dimensiones de cada conjunto de datos, lo que hace necesario desarrollar modelos *ad hoc* para cada librería, así como pruebas intensivas para su optimización.

Usando el caso de estudio de los 256 espectros de la librería STELIB, podemos formalizar el proceso de codificación del VAE de seis dimensiones que reduce a seis valores los 6700 puntos de cada espectro. El modelo intenta aprender un espacio latente $Z \in \mathbb{R}^6$ a partir de un conjunto de muestras de entrada $x_m \subseteq X \in \mathbb{R}^{6700}$. El objetivo del VAE es intentar muestrear valores de Z que probablemente hayan producido X , y calcular la probabilidad $P(X)$ sólo a partir de esos valores. Esto significa que necesitamos una nueva función $Q(Z|X)$ que pueda tomar un valor de X y darnos una distribución sobre los valores de Z que probablemente produzcan X (probabilidad posterior). Con suerte, el espacio de valores Z que son probables bajo Q será mucho más pequeño que el espacio de todos los Z que son probables bajo la probabilidad previa $P(Z)$ (Doersch, 2016).

La Figura 9 muestra el rango de coeficientes de determinación R^2 obtenidos en la codificación de la librería STELIB (0,622 - 0,999), utilizando el VAE de seis dimensiones. Se necesitó una combinación de una pequeña tasa de aprendizaje ($1e-5$) y un alto factor de aumento del corpus (repetición x15) para optimizar la respuesta del VAE y maximizar la precisión del modelo, alcanzando finalmente un $R^2 = 0,9773$ para el conjunto de datos completo.

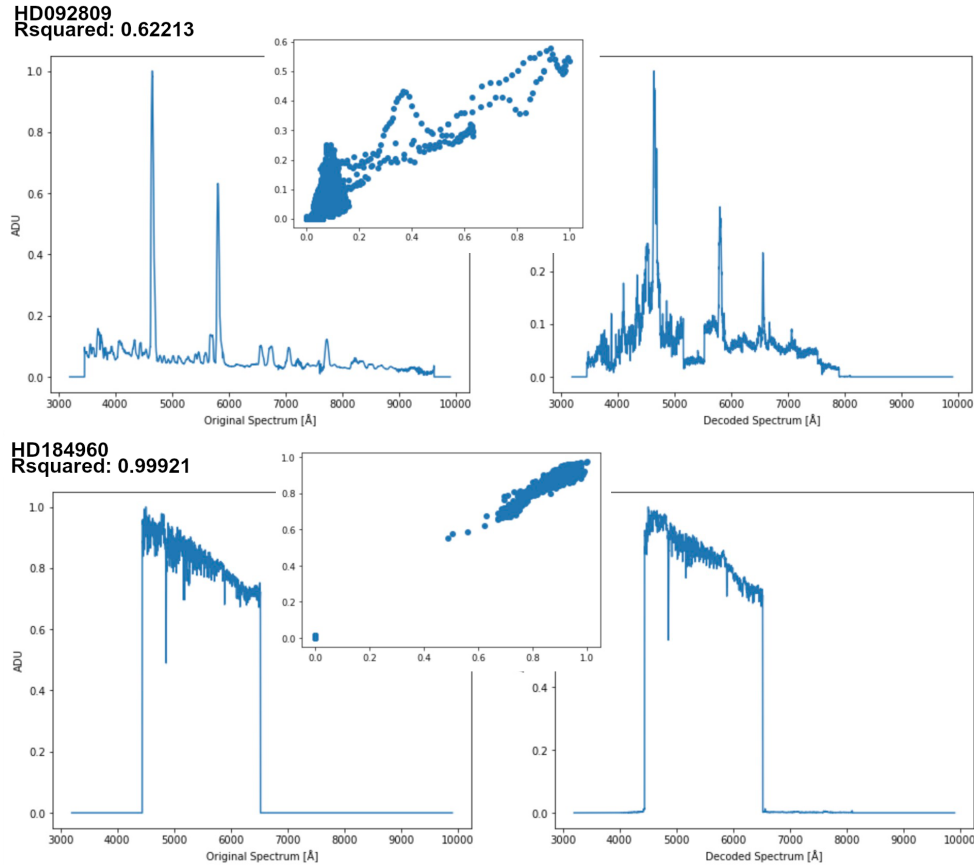


Figura 9: Resultados individuales de R^2 mostrando la peor codificación (arriba) y la mejor (abajo) para la librería STELIB, correspondiendo respectivamente a las estrellas HD 092809 y HD 184960. Vale la pena observar cómo incluso para $R^2 = 0,62213$, el modelo representa las principales características espectrales del espectro de entrada alrededor de 4700 y 5900 Å. Sin embargo, genera una línea H α falsa alrededor de 6600 Å.

Por otra parte, la Figura 10 proporciona una comparación entre la distribución de coeficientes de determinación R^2 para cada espectro de la librería MILES, codificada con un sparse autoencoder y con un VAE, para los conjuntos de datos original y aumentado.

Entrenando el VAE de seis dimensiones en la librería MILES original, que contiene 985 espectros, se obtuvo un R^2 total de 0,834 con el 45,79% de los espectros presentando un R^2 superior a 0,9. Al aumentar el conjunto de datos repitiendo cada espectro 20 veces, se obtuvo un R^2 total de 0,976 para el conjunto de datos completo, con el 96,92% de los espectros presentando un R^2 superior a 0,9. R^2 mínimo = 0,3283, R^2 máximo = 0,9981, mediana = 0,9784, desviación estándar = 0,038.

Aplicando el mismo análisis a un sparse autoencoder de seis dimensiones con el mismo conjunto de datos, se obtuvo un R^2 total de 0,96 con el 94,11% de los espectros presentando un R^2 superior a 0,9 para el conjunto de datos original, y un R^2 total de 0,985, con el 97,868% de los espectros presentando un R^2 superior a 0,9 para el conjunto de datos aumentado. R^2 mínimo = 0,4648, R^2 máximo = 0,9983, mediana = 0,9743, desviación estándar = 0,064.

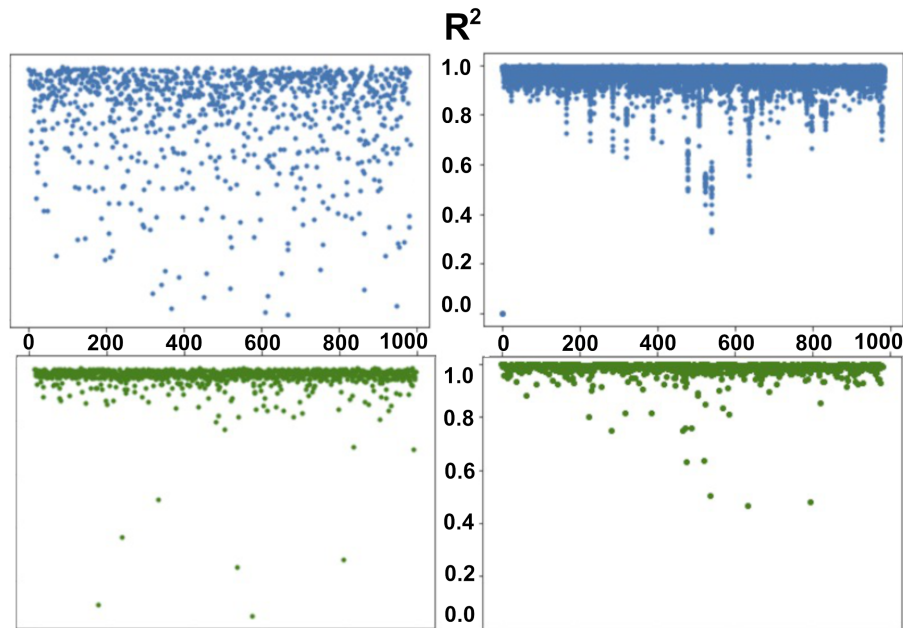


Figura 10: Distribución de los resultados de R^2 para la librería MILES utilizando un VAE (azul/arriba) y un sparse autoencoder (verde/abajo), entrenados en el conjunto de datos original (izquierda) y en el conjunto de datos aumentado x20 (derecha). R^2 para cada espectro analizado.

Cabe destacar cómo el sparse autoencoder proporcionó mejores resultados para el conjunto de datos original. El tiempo de cálculo también fue 10 veces menor para el sparse autoencoder (8 s/epoch frente a 83 s/epoch para el conjunto de datos aumentado).

Además de este tipo de análisis, para el caso de estudio de las galaxias de CALIFA, se realizaron mediciones de correlación adicionales. La Figura 11 proporciona dos matrices de correlación entre los valores latentes obtenidos en el proceso de codificación y los anchos equivalentes (EW) de $H\alpha$ para las galaxias NGC3395 y NGC2347.

El EW de $H\alpha$ es un indicador observacional utilizado en la caracterización de regiones de formación estelar. El estudio de estas regiones es crucial para comprender cómo se forman y evolucionan las galaxias. El EW de $H\alpha$ se define como la relación entre el flujo de $H\alpha$ (que dibuja la actividad de formación estelar instantánea, SF) y la densidad de flujo continuo (indicador de la masa estelar) (Khostovan et al., 2021). Funciona como un proxy equivalente a la Tasa de Formación Estelar específica ($sSFR = SFR/M_*$) (Tasca et al., 2015), inversamente proporcional a la masa estelar (M_*).

En cuanto a los resultados obtenidos, aunque se deben realizar pruebas adicionales para obtener datos estadísticamente representativos, las galaxias con EW de $H\alpha$ fuerte o intermedio se pueden diferenciar mediante sonido utilizando este enfoque de sonificación, que presenta una correlación directa con los valores latentes L3, L1 y L4. La Figura 12 muestra las distribuciones de EW de $H\alpha$ por spaxel, vector latente por spaxel y EW de $H\alpha$ vs vector latente por spaxel para NGC3395 y NGC2347.

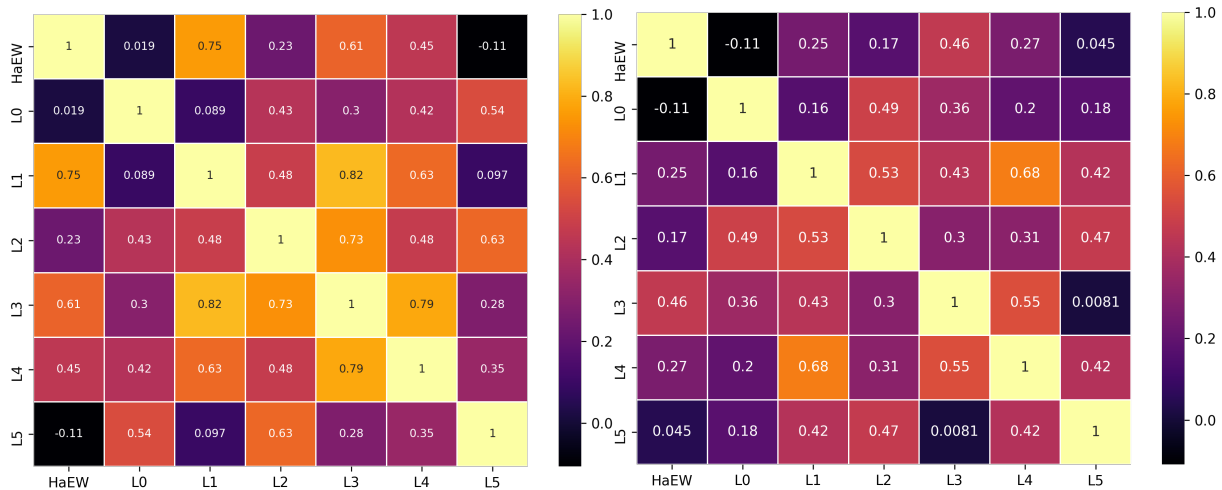


Figura 11: Correlación entre los seis valores latentes generados por el autoencoder y los anchos equivalentes (EW) de H α de NGC3395 (izquierda, EW alto) y NGC2347 (derecha, EW intermedio). Ambos ejemplos muestran la tendencia de correlación general más alta compartida entre los valores latentes L3, L1 y L4. CALIFA survey.

.16 Intervalos de confianza de Jeffreys

El intervalo de confianza de Jeffreys se usa para espacios muestrales de reducido tamaño, especialmente en análisis de proporciones binomiales. El enfoque utiliza el teorema de Bayes con una distribución a priori (sin información previa) expresada por la Ecuación 35:

$$f(\epsilon) = \pi^{-1} \epsilon^{-1/2} (1 - \epsilon)^{-1/2} \tag{35}$$

El cálculo de los intervalos se ha realizado con la función `binom_conf_interval` de *astropy* (https://docs.astropy.org/en/stable/api/astropy.stats.binom_conf_interval.html).

.17 Listado de aplicaciones desarrolladas

La Tabla 2 muestra la relación de aplicaciones que hemos desarrollado para proporcionar las pruebas de concepto descritas en los capítulos de este trabajo.

Tabla 2: Listado de aplicaciones desarrolladas

Nombre	Descripción	GitHub
FITS2OSC	Conversión de variaciones de flujo en mensajes OSC	https://github.com/AuditoryVO/AutoFITS2Sound
DVT Explorer	Sonificación multivariable de curvas de luz	https://github.com/AuditoryVO/DVT-Explorer
LC2Score	Conversión de periodogramas en notas musicales	https://github.com/AuditoryVO/LC2Score
LC2Music	Sintetizador de flauta a partir de periodogramas	https://github.com/AuditoryVO/LC2Music
CoRoT Explorer	Sonificación simbólica de curvas de luz	https://github.com/AuditoryVO/CoRoT_Explorer
Autoencoder2Sound	Sonificación mediante deep learning (notas y acordes)	https://github.com/AuditoryVO/Autoencoder2Sound
VAE2Sound	Sonificación mediante deep learning (VAE)	https://github.com/AuditoryVO/VAE2Sound
AI-rmonizer	Sistema de composición basado en deep learning	https://github.com/AuditoryVO/AI-rmonizer
CaT Synth	Sintetizador de audio basado en deep learning	https://github.com/AuditoryVO/CaT_Sonification_Synthesizer
SoniCube	Sonificación de cubos de datos mediante deep learning	https://github.com/rgbIAA/ViewCube-Evaluation/tree/main/2024

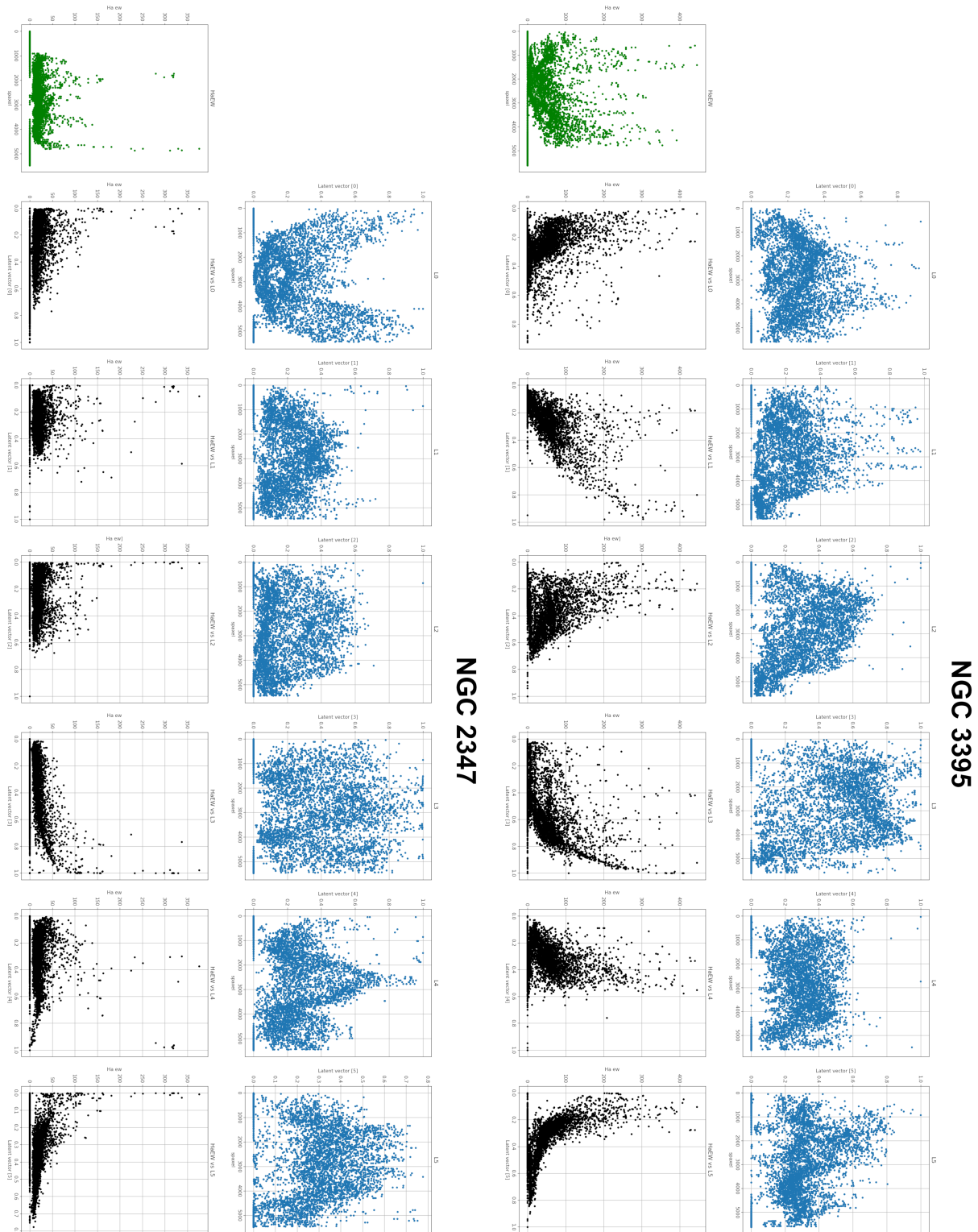


Figura 12: EW de H α por spaxel (verde), vector latente por spaxel (azul) y EW de H α vs vector latente (negro) para NGC3395 y NGC2347.

.18 Composiciones musicales generadas

Las siguientes piezas fueron generadas con el sistema de composición desarrollado en el Capítulo 8 y estrenadas durante el transcurso de esta tesis.

Spectral AI-rmonies of Classic Forms. Pieza para piano y electrónica generada e interpretada en directo por el autor para el concierto de sonificación de la International Conference for Auditory Display, ICAD 2023, en Linköping, Suecia. La pieza fue activada por *Le Chaos*, la introducción instrumental de la simfonie nouvelle para orquesta *Lès èlèmens*, de Jean Fèry Rebel. Se puede encontrar una descripción completa del trabajo en García Riber y Serradilla, 2023b.

El siguiente vídeo proporciona la interpretación en vivo de la pieza en la ESMUC, Barcelona, junto con otra composición para flauta y electrónica llamada *Spectral suite No. 1*, que fue generada utilizando las notas de activación de la *Suite para violonchelo No. 1* de J. S. Bach para la semana de música y tecnología TECSMUC 2023, y estrenada en directo por Helena Bas Cañas.

<https://www.youtube.com/watch?v=ugq-MZQIMBc&t=4713s>

Galaxy Spectra Suites. Pieza electrónica, en formato fijo audiovisual, activada por las Suites completas para violonchelo de J. S. Bach. La composición, formada por seis movimientos, se estrenará en el concierto de sonificación de la Audio Mostly & ICAD Joint Conference, AM.ICAD 2025, en Coimbra, Portugal. El trabajo explora los espectros de galaxias del estudio CALIFA (Calar Alto Legacy Integral Field Area survey), convertidos en acordes de seis notas, utilizando la estructura musical subyacente creada a partir de las seis obras maestras barrocas.

<https://vimeo.com/902748176>

Spectral Gymnopédie. Composición finalista en el AI Song Contest 2023. La pieza fue generada mediante la entrada de activación de la obra *Gymnopédie No. 1* de Erik Satie y la conversión de los espectros estelares de la librería MILES en acordes de cuatro notas. El proceso de creación incluye una edición en bloque mínima, realizada por los autores a partir de la salida generativa para repetir el leitmotiv musical de la pieza final. En el renderizado se utilizaron instrumentos acústicos muestreados para dar a la pista de audio final un carácter de cuarteto de cámara.

<https://www.aisongcontest.com/participants-2023/auditory-virtual-observatory>

Miles away from Miles y *Miles away from Jupiter*. Ejemplos adicionales generados respectivamente a partir de *All Blues* de Miles Davis y de la *Sinfonía No. 41 en Do mayor (K.551)*, *Júpiter* de W.A. Mozart. Las piezas se presentan sin ningún proceso de edición para mostrar el potencial de la propuesta en la composición de bandas sonoras originales para películas y contenido audiovisual. Los autores realizaron el proceso de diseño sonoro para dar a las pistas de audio final su respectivo carácter orquestal y de banda sonora original de película de ciencia ficción.

<https://soundcloud.com/user-659343643/miles-away-from-miles-raw-output>

<https://soundcloud.com/user-659343643/miles-from-jupiter-raw-output>

.19 Material Complementario

Todo el material complementario presentado en esta tesis (imágenes, audio y vídeo), se encuentra alojado en:

<https://doi.org/10.5281/zenodo.15514645>

.20 Música relacionada con la astronomía

La siguiente lista presenta las piezas musicales relacionadas con la astronomía que han estado orbitando alrededor del proceso de escritura de esta tesis. Se pueden encontrar referencias adicionales de música relacionada con la astronomía en W. Wang, 2006, J. A. Caballero et al., 2008, Fraknoi, 2008, J. A. Caballero, 2012, Fraknoi, 2012 y J. Caballero et al., 2016.

W.A. Mozart, *Sinfonía No. 41 en Do mayor (K.551)* (1788).

Rued Langgaard, *Sfærernes Musik* (1910).

Enrique Granados, *Cant de les estrelles* (1911).

Alexander Scriabin, *Vers la flamme* (1914).

Gustav Holst, *The Planets, Op.32* (1917).

Hector Villa-Lobos, *As Três Marias* (1939).

Paul Hindemith, *Die Harmonie der Welt* (1957).

Willie Ruff and John Rodgers, *The Harmony of the World: A Realization for the Ear of Johannes Kepler's Astronomical Data from Harmonices Mundi* (2003).

Tarik O'Regan's, *A Celestial Map of the Sky* (2014).

Tim Watts, *Kepler's Trial* (2017).

Kaija Sariaho, *Asteroid 4179: Toutatis* (2014) y *Solar* (2018).

Sophie Katsner, *Where Parallel Lines Converge* (2020).

.21 Código fuente

Las siguientes muestras de código fuente se proporcionan como ejemplos representativos de los módulos de *Python* que hemos implementado durante el desarrollo de esta tesis. El primero, `autoencoding.py`, se utilizó para automatizar la codificación de los cubos de datos de CALIFA con un VAE, tal y como se describe en el Capítulo 10. El segundo, `SoniCube.py`, se incluyó en la aplicación *ViewCube* para representar mediante sonido los espectros de los cubos de datos de CALIFA.

```
Autoencoding.py
```

```
#!/usr/bin/env python
```

```
# coding: utf-8
# Adrián García Riber 2023

import astropy
from astropy.io import fits
from astroquery.simbad import Simbad
import numpy as np
from numpy import save
import matplotlib.pyplot as plt
import os
from pathlib import Path
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers, losses
from keras.layers import Dense, Input
from keras.models import Model
from sklearn.metrics import r2_score

path = 'Cubes'
num = 0
for path, subdirs, files in os.walk(path):
    for name in files:
        num += 1

import hashlib
def md5(fname):
    hash_md5 = hashlib.md5()
    with open(fname, "rb") as f:
        for chunk in iter(lambda: f.read(4096), b''):
            hash_md5.update(chunk)
    return hash_md5.hexdigest()

#Importing the model
autoencoder = tf.keras.models.load_model("Autoencoder.tf")

for name in files:
    fits_file = os.path.join(path, name)
    hdulist = fits.open(fits_file)
    checksum = md5(fits_file)# For the dictionary

    data = hdulist[0].data
    dimensions = data.shape # for the dictionary
    spectrum_dim = data.shape[0]
    dim1 = data.shape[0]

    #Dimensions of the dataset and autoencoder
    num = data.shape[1]*data.shape[2]

    latent_dim = 6 #For the dictionary
    intermediate_dim = int(dim1/latent_dim)

    custom_set = np.zeros((num, spectrum_dim))
    custom_set = np.transpose(np.reshape(data, (spectrum_dim, num), order= 'F'))
```

```

#Normalization MAX-MIN and null spectrum search
label_set = []
index = []
zeros_counter = 0
idx = 0

custom_set[np.isnan(custom_set)] = 0

for spectrum in custom_set:
    max_flux = np.nanmax(spectrum)
    min_flux = np.nanmin(spectrum)

    if np.nanmax(spectrum) != 0:
        # MAX-MIN Normalizing spectra
        custom_set[idx] = np.reshape((spectrum - min_flux) /
            (max_flux - min_flux), spectrum_dim)
    else:
        index.append(idx)
        zeros_counter += 1
        idx += 1
        label_set.append(idx)

new_set = np.delete(custom_set,index,0) #Eliminating null spectra
new_label = np.delete(label_set,index,0)
augmentation = 50
augmented_set = np.repeat(new_set, augmentation, axis=0)
augmented_label = np.repeat(new_label, augmentation, axis=0)
# Splitting and randomizing the training and test sets
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(augmented_set,augmented_label,
test_size=0.2,random_state=123) #for reproducibility
hdulist.close()

#Autoencoder
input_shape = (spectrum_dim,)
autoencoder.compile(optimizer='adam', loss=losses.MeanSquaredError())
callback = tf.keras.callbacks.EarlyStopping(monitor='loss', patience=20)
history = autoencoder.fit(x_train, x_train,
                        epochs=100,
                        shuffle=True, callbacks=[callback],
                        validation_data=(x_test, x_test))
encoded_imgs = autoencoder.encoder(x_test).numpy()
decoded_imgs = autoencoder.decoder(encoded_imgs).numpy()

#R2 for the test set
r2_test = r2_score(x_test, decoded_imgs, multioutput='variance_weighted')
encoded_imgs2 = autoencoder.encoder(x_train).numpy()
decoded_imgs2 = autoencoder.decoder(encoded_imgs2).numpy()
#R2 for the training set
r2_train = r2_score(x_train, decoded_imgs2, multioutput='variance_weighted')
parent_folder = 'Autoencoded/'

#Saving
save_path = os.path.join(parent_folder, os.path.splitext(name)[0])

```

```
os.makedirs(save_path)
    #Learning rate
plt.title(os.path.splitext(name)[0] + ' - Learning rates -'+ R2= {}'.format(r2_test))
plt.plot(history.history['loss'], label='loss')
plt.plot(history.history['val_loss'], label='validation_loss')
plt.legend()
plt.savefig(save_path + '/' + os.path.splitext(name)[0] + '_learning_rate.png')
plt.close()

#We just need one model for all the cubes
#autoencoder.save(save_path + '/' + name + '_Autoencoder.tf', save_format='tf')
weights = autoencoder.get_weights()
autoencoder.save_weights(save_path + '/' + os.path.splitext(name)[0] + '_Weights', weights)

info = np.dtype([('file', 'U50'), ('checksum', 'U50'), ('encoder_dim', 'i4'),
                ('cube_dimensions', 'i4', (1, 3)), ('r2_train', 'U50'), ('r2_test', 'U50')])
Reference = np.zeros(1, dtype=info)
Reference[0][0] = os.path.splitext(name)[0]
Reference[0][1] = checksum
Reference[0][2] = latent_dim
Reference[0][3] = dimensions
Reference[0][4] = r2_train
Reference[0][5] = r2_test

np.save(save_path + '/' + os.path.splitext(name)[0] + '_Reference', Reference)
print(Reference[0])
```

```
#-----
```

```
SoniCube.py
```

```
#!/usr/bin/env python
# coding: utf-8
# Adrián García Riber 2023

from astropy.io import fits
import numpy as np
import hashlib
import tensorflow as tf
from tensorflow import keras
from pythonosc import udp_client
import ctcsound
import librosa
import image
import math
import pyautogui #To be eliminated when including ViewCube's (y,x) capture method
import time

class SoniCube():
    '''Califa survey Sonification'''

    def _init_(self):
        '''CSound and Autoencoder initialization with file checksum test'''
```

```

print("IMPORTANT ADVICE: Some spectra could produce high sound pressure levels.
Perform a first scan of the cube to adapt the volume of your headphones.")
self.auto_path = "sound/data/"
try:
    reference = np.load(self.auto_path + file + "/" + file + "_Reference.npy")
    self.lat_dim = reference[0][2]
    self.checksum = sonicube.md5(file)

    if self.checksum == reference[0][1]:
        left = "sound/binaural/hrtf-48000-left.dat"
        right = "sound/binaural/hrtf-48000-right.dat"
        self.port = 9970      # OSC port
        self.cs = ctcsound.Csound()      #Opens CSound
        self.cs.compileCsd("sound/csound/SoniCube.csd") #SoniCube_Pre.csd
        self.cs.start()

        self.cs.setStringChannel("hrtf_L", left)
        self.cs.setStringChannel("hrtf_R", right)

        #Autoencoder initialization
        #Importing the model
        self.autoencoder = tf.keras.models.load_model(self.auto_path + file +
        "/" + file + "_Autoencoder.tf")
        #Importing the weights
        self.weights = self.autoencoder.load_weights(self.auto_path +file +
        "/" + file + "_Weights")

    else:
        print("This file doesn't match the encoded cube: ", file)
except:
    print("No encoded file found. This cube is not processed for sonification yet.")

def md5(self, fname):
    '''Calculates the checksum of a file'''

    hash_md5 = hashlib.md5()
    with open(fname, "rb") as f:
        for chunk in iter(lambda: f.read(4096), b''):
            hash_md5.update(chunk)
    return hash_md5.hexdigest()

def preprocessing (self, file, **kwargs):
    ''' Preprocesses and extracts variables from the cube'''

    hdulist = fits.open(file)      #Opening data cube

    data = hdulist[0].data
    #extracting the image dimensions to mark sonification limits
    self.spectrum_dim, self.y_tot, self.x_tot = data.shape
    #number of spectra
    self.num = self.y_tot*self.x_tot

```

```

hdulist.close()                                #closing fits file

custom_set = np.transpose(np.reshape(data, (self.spectrum_dim, self.num), order= 'F'))
custom_set[np.isnan(custom_set)] = 0            #removing nan values

#Calculating median absolut fluxes for audio normalization
median_set = np.nanmedian(custom_set, axis=1)
self.max_median = np.nanmax(median_set)
self.min_median = np.nanmin(median_set)

(self.y_ref,self.x_ref) = image.cube_median(file) #Central pixel, from Rubén's image.py

return custom_set

def autoencoding (self, spectrum, **kwargs):
    '''Calculates the latent space of a single spectrum'''

    #We need to adapt the dimensions to the autoencoder input
    redim_spectrum = np.zeros((1, self.spectrum_dim))
    #We need to normalize and kill nans before autoencoding
    redim_spectrum[0] = spectrum
    redim_spectrum[np.isnan(redim_spectrum)] = 0
    max_flux = np.nanmax(redim_spectrum)
    min_flux = np.nanmin(redim_spectrum)

    if np.nanmax(redim_spectrum) != 0:          # MAX-MIN Normalizing
        normalized_spectrum = (redim_spectrum - min_flux) / (max_flux - min_flux)
        #Encoding
        encoded_spectrum = sonicube.autoencoder.encoder(normalized_spectrum).numpy()
    else:
        normalized_spectrum = (redim_spectrum - min_flux) / 1e-16
        #Encoding
        encoded_spectrum = sonicube.autoencoder.encoder(normalized_spectrum).numpy()

    lat = encoded_spectrum*10000              #Scaling latent values to audible frequencies

    return (lat[0])

def coord_converter (self, y, x, **kwargs):
    ''' Calculates azimuth and distance factor (normalized between 0 and 2)
    for the selected pixel (y,x)'''

    #Calculates azimuth
    azimuth = (90 - math.degrees(math.atan2((y-self.y_ref), (x-self.x_ref)))) % 360
    #Calculates x,y distance to x_ref, y_ref
    radial_distance = math.sqrt((y-self.y_ref)**2 + (x-self.x_ref)**2)

    #Calculates max_distance to normalize
    coordinates = [(0, 0), (0, self.y_tot), (self.x_tot, 0), (self.x_tot, self.y_tot)]
    max_distance = 0
    for k, j in coordinates:

```

```

        distance = math.sqrt((k - self.y_ref)**2 + (j - self.x_ref)**2)
        max_distance = max(max_distance, distance)
#Calculates a distance factor between 0 and 2 to control Reverb(distance emulator)
norm_distance = (radial_distance/max_distance)*2

return azimuth, norm_distance

#returns the A-weighted amplification factor to compensate loudness
def ponderating (self, frequencies, **kwargs):
    '''Applies the A-weighted levels curve to the input array of frequencies to return
        their normalized amplification factors'''

    amp_factors = []
    for freq in frequencies:
        if freq != 0:
            weight = librosa.A_weighting(freq)
            #MAX-MIN normalization for 24KHz bandwidth
            norm_weight = (weight-(-80))/((1.27134)-(-80))
            if (norm_weight) != 0:
                amp_factor = round(1/norm_weight, 2)
            else:
                amp_factor = 0
        else:
            amp_factor = 0
        amp_factors.append(amp_factor)
    return amp_factors

#Calculates the index of the spectrum (figure)from the (y,x) coordinates
def get_index (self, y, x, **kwargs):
    if (x <= self.x_tot) and (y <= self.y_tot):
        figure = (self.x_tot*y)+x
    else:
        #Control value for positions out of the image.
        #Not necessary if the spectrum matches the dimensions
        figure = 6000
    return figure

def sonifying (self, y, x, spectrum, **kwargs):
    '''Sonifies the spectrum selected with the cursor'''
    try:
        #Sonification

        amps_label = "amp0"      #Generating labels for the OSC messages
        freqs_label = "lat0"
        for dimension in range(1, self.lat_dim):
            amps_label = amps_label + ("/amp%i"%dimension)
            freqs_label = freqs_label + ("/lat%i"%dimension)

        if (0 <= x < self.x_tot) and (0 <= y < self.y_tot):

            #Calculating coordinates and radial distance
            print('pixel coordinates (y,x): ', (y,x))
            (azimuth, dist) = sonicube.coord_converter(y, x)
            azimuth = round(azimuth,2)

```

```
dist = round(dist,2)

#Calculates the index of the spectrum (figure)from the (y,x) coordinates
#figure = (self.x_tot*y)+x

lat = np.round(sonicube.autoencoding(spectrum))

#Creates the list of frequencies to send
lat_vect = []
for dimension in range(self.lat_dim):
    lat_vect.append(round(lat[dimension]))

#calculating median of absolut flux for current spectrum
median_flux = round(np.nanmedian(spectrum, axis=0), 5)

if (median_flux != 0):
    # MAX-MIN Normalizing fluxes between 1&10
    median_norm = 1 + ((median_flux - self.min_median) /
        (self.max_median - self.min_median))*9
else:
    median_norm = 1
log_norm_flux = round(np.log(median_norm), 5)

if log_norm_flux >= 1: #Dynamic range compression in two stages
    log_norm_flux = round(log_norm_flux/2.5, 5)
if 0.5 <= log_norm_flux < 1:
    log_norm_flux = round(log_norm_flux/1.2, 5)

#Ponderating latent vector amplitudes (A-weighted)
amp_weights = sonicube.ponderating(lat_vect)

coords = [azimuth, dist]

# Sending OSC messages to CSound
client_amps = udp_client.SimpleUDPClient("127.0.0.1", self.port)
client_amps.send_message(amps_label, amp_weights)
print("Amp. factors =", amp_weights)

client_freqs = udp_client.SimpleUDPClient("127.0.0.1", self.port+1)
client_freqs.send_message(freqs_label, lat_vect)
print("Latent vector =", lat_vect)

client_coords = udp_client.SimpleUDPClient("127.0.0.1", self.port+2)
client_coords.send_message("az/dist", coords)
print("Azimuth =", azimuth, "Distance =", dist)

client_coords = udp_client.SimpleUDPClient("127.0.0.1", self.port+3)
client_coords.send_message("flux", float(log_norm_flux))
print("Median absoulft flux=", median_flux)
print("Amplitude factor=", log_norm_flux)

else:
    print("Selected pixel out of image bounds")
    client_freqs = udp_client.SimpleUDPClient("127.0.0.1", self.port+1)
```

```

        client_freqs.send_message(freqs_label, (0,0,0,0,0,0))
        client_coords = udp_client.SimpleUDPClient("127.0.0.1", self.port+2)
        client_coords.send_message("az/dist", (0,0))
        client_coords = udp_client.SimpleUDPClient("127.0.0.1", self.port+3)
        client_coords.send_message("flux", 0)

#         SoniCube.pt.stop()         #Stops the synthesizer thread
#         SoniCube.cs.reset()        #Stops CSound

    except KeyboardInterrupt:
        print("Sonification error")

##-----TESTING-----
file = 'NGC5406.COMB.rscube.fits'          #Loading the cube and initializing

sonicube = SoniCube()
sonicube._init_()

absolut_set = sonicube.preprocessing(file)

#Starts the synthesizer thread
sonicube.pt = ctcsound.CsoundPerformanceThread(sonicube.cs.csound())
sonicube.pt.play()

                                                #Launching the test loop
try:
    y_pre = 0
    x_pre = 0

    while True:
        #TO DO: Replace by ViewCube's y,x coordinate capture
        x,y = pyautogui.position()
        if (x != x_pre) or (y != y_pre):
            spectrum_index = sonicube.get_index(y,x)
            #Limit to avoid crashing when exceeding absolut_set dimensions(5544)
            if spectrum_index < 5544:
                sonicube.sonifying (y, x, absolut_set[spectrum_index])
            else:
                #Sends silence message. Uses any spectrum (absolut[0])
                sonicube.sonifying (y, x, absolut_set[0])
            y_pre, x_pre = y, x

        else:
            y_pre, x_pre = y, x
except KeyboardInterrupt:
    print("End")

```

.22 Partitura musical

La partitura de la Figura 13 se incluye como ejemplo representativo de las partituras musicales (incluyendo la referencia de las fuentes estelares) que se obtienen a la salida del sistema de

composición descrito en el Capítulo 8. La pieza, denominada *INT* en referencia al Telescopio Isaac Newton utilizado en las observaciones de los objetos de la librería MILES, fue generada a partir de las notas de activación del *Preludio en do sostenido menor, Op. 45* (1841) de Frédéric Chopin. Esta obra se considera una de las menos convencionales de Chopin debido a su desviación de la estructura típica del preludio y a la incorporación de disonancia y cromatismo. El siguiente notebook permite la regeneración y reproducción “en tiempo real” de la pieza a partir de pesos, acordes y duraciones pre-calculadas.

<https://github.com/AuditoryVO/INT>

Alternativamente, el GitHub proporcionado en la Sección 8.4 puede utilizarse para obtener los pesos, los acordes espectrales y las duraciones, reproduciendo el proceso generativo completo desde cero. Es necesario tener en cuenta que en ambos casos los resultados finales dependen del proceso de codificación, cuyas probabilidades proporcionan valores ligeramente diferentes en cada entrenamiento y, por tanto, las partituras generadas diferirán de la de la Figura 13.

INT

A.G. Riber and F. Serradilla

♩ = 90

HD084937 BD+19_5116 HD150012 HD237846 HD146624

4

HD036395 HD209459 HD003546 M71_AH-A2 M71_AH-A2

9

M71_AH-A2 HD212454 HD250792 HD250792 HD250792

Figura 13: *INT*. Composición polifónica no supervisada, generada a partir de espectros estelares.

2

14

HD088609 HD009408 HD048329 M71_AH-I HD009408 HD212454

This block contains the musical notation for measures 14 through 17. It consists of four staves of music. The notation includes various note values, rests, and accidentals. Below the staves, there are six alphanumeric identifiers: HD088609, HD009408, HD048329, M71_AH-I, HD009408, and HD212454.

18

HD088609 HD009408 HD218502 HD125184 HD136726

This block contains the musical notation for measures 18 through 20. It consists of four staves of music. The notation includes various note values, rests, and accidentals. Below the staves, there are five alphanumeric identifiers: HD088609, HD009408, HD218502, HD125184, and HD136726.

21

BD+37_1665 HD084937 HD150453 HD105740 HD120136 HD120136

This block contains the musical notation for measures 21 through 23. It consists of four staves of music. The notation includes various note values, rests, and accidentals. Below the staves, there are seven alphanumeric identifiers: BD+37_1665, HD084937, HD150453, HD105740, HD120136, and HD120136.

25

HD122956 HD153286 HD163990 HD003795

Detailed description: This block contains the first system of musical notation, measures 25 through 28. It consists of four staves of music. The notation includes treble clefs, a key signature of one sharp (F#), and a time signature of 4/4. The music features a variety of note values including quarter, eighth, and sixteenth notes, often grouped with slurs and beams. Measure numbers 25, 26, 27, and 28 are indicated above the first staff. Below the staves, four star names are listed: HD122956, HD153286, HD163990, and HD003795.

29

HD094028 HD044691 HD105740 HD161227 HD004656 NGC6121-2303

Detailed description: This block contains the second system of musical notation, measures 29 through 32. It consists of four staves of music. The notation includes treble clefs, a key signature of one sharp (F#), and a time signature of 4/4. The music features a variety of note values including quarter, eighth, and sixteenth notes, often grouped with slurs and beams. Measure numbers 29, 30, 31, and 32 are indicated above the first staff. Below the staves, five star names are listed: HD094028, HD044691, HD105740, HD161227, and HD004656, followed by NGC6121-2303.

33

NGC5272_III-28 HD009408 HD089744 HD209459 HD126660

Detailed description: This block contains the third system of musical notation, measures 33 through 36. It consists of four staves of music. The notation includes treble clefs, a key signature of one sharp (F#), and a time signature of 4/4. The music features a variety of note values including quarter, eighth, and sixteenth notes, often grouped with slurs and beams. Measure numbers 33, 34, 35, and 36 are indicated above the first staff. Below the staves, five star names are listed: NGC5272_III-28, HD009408, HD089744, HD209459, and HD126660.

4
37

HD116544 HD108177 HD219116 HD166620 HD219116

This system contains four staves of music for measures 37 through 41. The notation includes various note values, rests, and slurs. The first staff begins with a treble clef and a key signature of one sharp (F#). The notes are: 37: G4 (quarter), A4 (quarter), B4 (quarter), C5 (quarter); 38: D5 (quarter), E5 (quarter), F#5 (quarter), G5 (quarter); 39: A5 (quarter), B5 (quarter), C6 (quarter), D6 (quarter); 40: E6 (quarter), F#6 (quarter), G6 (quarter), A6 (quarter); 41: B6 (quarter), C7 (quarter), D7 (quarter), E7 (quarter).

42

HD168720 BD+18_2890 HD023261 HD176301 BD+37_1665 HD083632

This system contains four staves of music for measures 42 through 47. The notation includes various note values, rests, and slurs. The first staff begins with a treble clef and a key signature of one sharp (F#). The notes are: 42: G4 (quarter), A4 (quarter), B4 (quarter), C5 (quarter); 43: D5 (quarter), E5 (quarter), F#5 (quarter), G5 (quarter); 44: A5 (quarter), B5 (quarter), C6 (quarter), D6 (quarter); 45: E6 (quarter), F#6 (quarter), G6 (quarter), A6 (quarter); 46: B6 (quarter), C7 (quarter), D7 (quarter), E7 (quarter); 47: F#7 (quarter), G7 (quarter), A7 (quarter), B7 (quarter).

48

HD221345 HD223385 HD237846 HD237846 HD006805 HD038545

This system contains four staves of music for measures 48 through 53. The notation includes various note values, rests, and slurs. The first staff begins with a treble clef and a key signature of one sharp (F#). The notes are: 48: G4 (quarter), A4 (quarter), B4 (quarter), C5 (quarter); 49: D5 (quarter), E5 (quarter), F#5 (quarter), G5 (quarter); 50: A5 (quarter), B5 (quarter), C6 (quarter), D6 (quarter); 51: E6 (quarter), F#6 (quarter), G6 (quarter), A6 (quarter); 52: B6 (quarter), C7 (quarter), D7 (quarter), E7 (quarter); 53: F#7 (quarter), G7 (quarter), A7 (quarter), B7 (quarter).

54

HD076813M71_021 HD001326 HD212454 HD250792 HD250792

58

HD250792 HD088609 HD009408 HD009408 M5_II-51

62

HD009408 HD202109 HD088609 HD212454 HD088609 HD009408

6

66

Musical score for measures 66-69, consisting of four staves. The notation includes various note values, rests, and slurs. The key signature has one sharp (F#).

HD009408 HD126327 HD122956 HD048433 M67_F-108

70

Musical score for measures 70-73, consisting of four staves. The notation includes various note values, rests, and slurs. The key signature has one sharp (F#).

HD164136 HD003546 HD128959 HD090508 HD111631

74

Musical score for measures 74-77, consisting of four staves. The notation includes various note values, rests, and slurs. The key signature has one sharp (F#).

BD+37_1665 HD219978 HD186408 HD072324

78

HD009408 HD237846 HD001326 HD237846 HD237846

Detailed description: This block contains the musical notation for measures 78 through 81. It consists of four staves of music. The notation includes various note values, rests, and slurs. Below the staves, five specific notes are identified with their corresponding HD numbers: HD009408, HD237846, HD001326, HD237846, and HD237846.

82

HD212454 NGC5272_III-28 HD221170 HD171496 HD001326

Detailed description: This block contains the musical notation for measures 82 through 86. It consists of four staves of music. The notation includes various note values, rests, and slurs. Below the staves, five specific notes are identified with their corresponding HD numbers: HD212454, NGC5272_III-28, HD221170, HD171496, and HD001326.

87

HD212454 HD209975 HD009408 HD072324

Detailed description: This block contains the musical notation for measures 87 through 90. It consists of four staves of music. The notation includes various note values, rests, and slurs. Below the staves, four specific notes are identified with their corresponding HD numbers: HD212454, HD209975, HD009408, and HD072324.

8
91

HD009408 HD009408 HD009408 BD+15_1305 HD009408

94

HD114642 HD166620 HD081817 HD157089 HD139641

98

HD076813HD209975 HD056274HD237846 HD030743 HD036395

102

HD134169 BD+36_2165 BD+80_0245 HD187691 HD157089 HD209975

Detailed description: This block contains the musical notation for measures 102 through 106. It consists of four staves of music. The notation includes various note values, rests, and dynamic markings. Below the staves, there are six alphanumeric labels: HD134169, BD+36_2165, BD+80_0245, HD187691, HD157089, and HD209975.

107

HD163990 HD163990 HD121258 HD165195 NGC288_77

Detailed description: This block contains the musical notation for measures 107 through 111. It consists of four staves of music. The notation includes various note values, rests, and dynamic markings. Below the staves, there are five alphanumeric labels: HD163990, HD163990, HD121258, HD165195, and NGC288_77.

112

M71_AH-A2 HD009408 HD126141 HD009408 HD149661

Detailed description: This block contains the musical notation for measures 112 through 116. It consists of four staves of music. The notation includes various note values, rests, and dynamic markings. Below the staves, there are five alphanumeric labels: M71_AH-A2, HD009408, HD126141, HD009408, and HD149661.

10

115



HD102224 HD221345

HD001326

HD064332

HD223047

119



HD088609

G-156-31

HD221345

HD174912

HD121258

124



HD121258

M71_AH-A9

HD037394 HD173819

BD+09_3223

128

HD223524 HD209975 HD163990 HD168608 HD176301

132

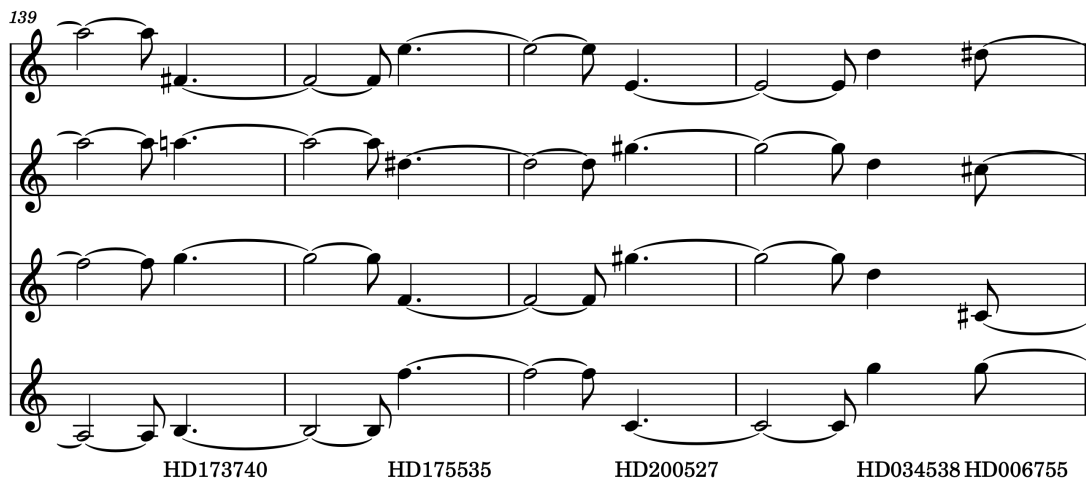
HD079765 HD009138 HD036395 HD031767

136

HD119288 HD003546 HD119288 HD251611 HD205021 HD187879

12

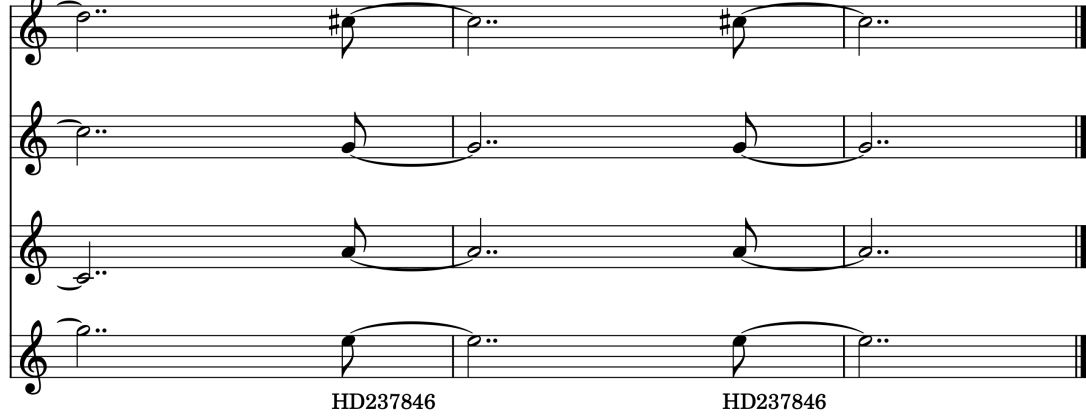
139



HD173740 HD175535 HD200527 HD034538 HD006755

This block contains a musical score for measures 139 through 142. It consists of four staves of music. The notation includes various note values, rests, and accidentals, with some notes beamed together. Below the staves, there are five alphanumeric codes: HD173740, HD175535, HD200527, HD034538, and HD006755.

143



HD237846 HD237846

This block contains a musical score for measures 143 and 144. It consists of four staves of music. The notation features longer note values and rests, with some notes beamed together. Below the staves, there are two alphanumeric codes: HD237846 and HD237846.