Highlights of Spanish Astrophysics XII, Proceedings of the XVI Scientific Meeting of the Spanish Astronomical Society held on July 15 - 19, 2024, in Granada, Spain. M. Manteiga, F. González Galindo, A. Labiano Ortega, M. Martínez González, N. Rea, M. Romero Gómez, A. Ulla Miguel, G. Yepes, C. Rodríguez López, A. Gómez García and C. Dafonte (eds.), 2025

Sonification strategies for astronomical data analysis and outreach

García Riber, A.[1], and Serradilla, F.[2]

[1] DMCT, ETSIST, Universidad Politécnica de Madrid (UPM), 28031, Spain
[2] Departamento de Sistemas Informáticos, ETSISI, UPM, 28031, Spain

Abstract

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 graphic representations, enhancing the accessibility of the archives, offering alternative display strategies to specific case studies, and bringing scientific concepts to potential audiences through multimodal and transmedia approaches. This article summarizes a set of sonification and musification strategies implemented for the auditory exploration of the MILES and STELIB stellar spectra libraries, the TESS light curve database, and the Lega-C and CALIFA galaxy surveys. Mainly based on automation, parameter mapping, and deep learning algorithms, these resources show alternative possibilities for the representation of astronomical data through sound and music, ranging from symbolic representations of spectral types, through binaural applications based on autoencoders, to complete data-driven music composition systems based on neural networks and attention mechanisms. The work aims to contribute to the discussion exploring the implications and potential goals of future developments toward a proposal of an auditory virtual observatory, which 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.

1 Introduction

The current development of massive astronomical archives and virtual observatory technology represents a great opportunity to consolidate the use of sound in the analysis and communication of scientific information. Sonification is a prolific tool for studying the universe, with potential benefits for the representation of astronomical data and concepts [1], which has been successfully used to study stellar and planetary structure [2], to communicate information to blind and low-vision (BLV) individuals [3], and to foster public engagement showcasing several decades of astronomical data [4]. The rationale for using auditory displays is based on the ability of the human auditory system to recognize temporal changes and patterns, its fast temporal response, and the ability to monitor and process multiple auditory datasets in parallel [5]. The design of a sonification tool requires the definition of a use case, a relationship with the phenomenon, and the creation of a listening experience [6], involving both, technical and perceptual challenges [7]. Our research aims at contributing with the exploration of the potential of deep learning to provide unsupervised auditory representations that could be used on scientific data analysis, inclusive complementary exploration methodologies, and public outreach.

2 Sonification strategies

This section provides a brief summary of a collection of sonification and musification strategies[1] developed for the analysis and representation of astronomical data. Each implementation is applied to a specific case study, allowing the exploration of different galaxy and stellar libraries through sound. All subsections include references to source publications. These resources are expected to be useful in the search for a proposal of an auditory virtual observatory.

2.1 AI-rmonies of the Spheres. MILES library

This musical approach[2] converts the almost 1000 stellar spectra of the MILES library (Instituto de Astrofísica de Canarias, IAC) into a data base of "stellar chords" using a variational autoencoder architecture. As shown in Fig.1, these chords are cross-matched with the musical chords generated by a Long Short Term Memory (LSTM) with attention neural network, trained with 200 hours of virtuous piano pieces from the MAESTRO dataset [8]. The work uses the piece that could have inspired Johannes Kepler during the writing of his treatise *Harmonices Mundi* to trigger the generation of the final musical composition [9].

2.2 Stellar spectra sound synthesis. CaT library

This proposal introduces the use of sound synthesizers for the exploration of stellar spectra using the Ca II triplet (CaT) stellar spectra library from the Spanish Virtual Observatory (SVO). The system reacts "in real time" to each note of the piano keyboard, providing a deep learning-based harmonic accompaniment. The algorithm uses an autoencoder to convert the stellar spectra of real sky objects data[3] into musical chords, and searches for the "stellar chords" matching the note of the piano key pressed by the user. The proposal can be used in sound design and musical composition for the creation of complex timbres and textures, as well as in educational and outreach activities to bring Astronomy concepts closer to all audiences using the engagement potential of sound and music [10].

[1] All the applications presented in this article are available at:
`https://github.com/AuditoryVO` and `https://github.com/rgbIAA/viewcube`

[2] *AI-rmonies of the spheres* video: `https://vimeo.com/user82659899`

[3] CaT synthesizer video: `https://vimeo.com/913841717`

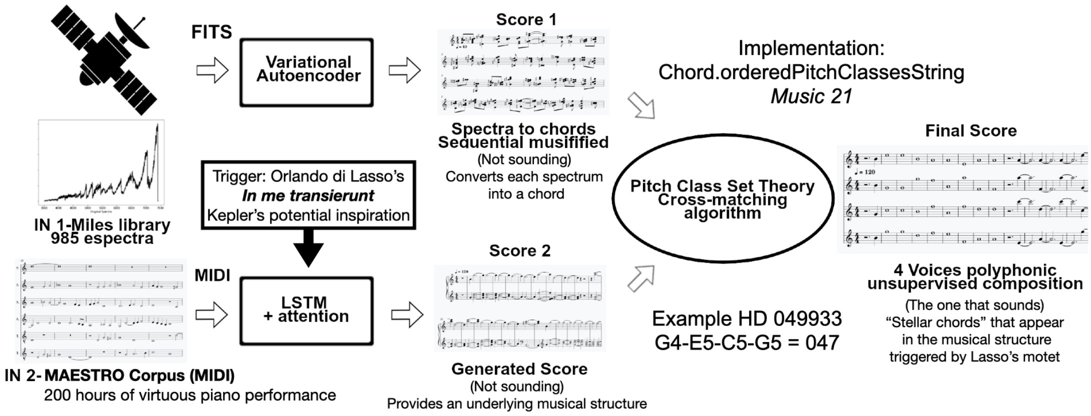

Figure 1: Structure of the astronomical composition system based on deep learning.

2.3 TESS DVT file exploration with VO tools

Focused on 132 examples of publicly available light curves observed by the Transiting Exoplanet Survey Satelite (TESS), this proposal displays sequential soundscapes[4] generated from the Data Validation Time-series (DVT) files of the TESS Object of Interest (TOI) catalog. The soundscapes include the sonification of the effective temperature, metallicity, surface gravity, and stellar radius of targeted host stars, as well as the orbital period, transit duration, and depth of planet transit candidates. The sonifications can be interactively controlled in terms of complexity and sound balance, and the application allows the analysis of the objects via *Aladin Lite* and *Vizier* queries [11].

2.4 Ambisonic and binaural sonification. STELIB and CoRoT libraries

Aimed at analyzing the potential of spatialized sonifications, we implemented a deep learning prototype with binaural encoding [12] to provide an immersive representation of the STELIB library[5]. Additionally, we developed an application[6] for the sonification of light curve catalogs in a user-selectable ambisonic or binaural configuration using the Convection, Rotation and planetary Transits (CoRoT) archive. These sonifications are based on a symbolic spectral-type to musical-note mapping strategy. This approach provides musical aesthetics and auditory spectral differentiation capability. It also includes an auditory representation of the "best fit" period found in each light curve, represented by the low frequency oscillator (LFO) of a tremolo. The user interface shows the footprint of 474 objects from the CoRoT N2 legacy data release and a sequential representation of its light curves. Both applications were based on the actual coordinates of the objects, and they could be useful in outreach activities and archival accessibility improvement [13].

[4]TESS DVT explorer video: `https://vimeo.com/723815913`
[5]STELIB binaural representation video: `https://vimeo.com/878943483`
[6]CoRoT binaural representation video: `https://vimeo.com/921824459`

2.5 Multimodal IFS. CALIFA survey

Exploring the potential of multimodal Integral Field Spectroscopy (IFS), and focused on the case study of the galaxies from the Calar Alto Integral Field Spectroscopy Area (CALIFA) survey, we implemented a deep learning sonification module into *ViewCube*[7], an IFS analysis application developed by Rubén García Benito (IAA-CSIC). The final multimodal application provides interactive visualizations and binaural sonifications of data cube spectral information. It is designed to incorporate future sonification approaches for analyzing different aspects of the spectra. *ViewCube* [14] aims to complement scientific graphic analysis and improve the accessibility of data cube analysis processes.

3 Evaluations

We conducted two online surveys to evaluate some of the proposals. The first questionnaire[8] was supported by training videos. It presented 8 quantitative questions focused on auditory classification tasks using the CoRoT light curves, and 2 qualitative questions asking the participants about the usefulness of the proposal and their interest in listening to more sonifications after completing the survey. Fig.2 shows the results obtained from 40 participants.

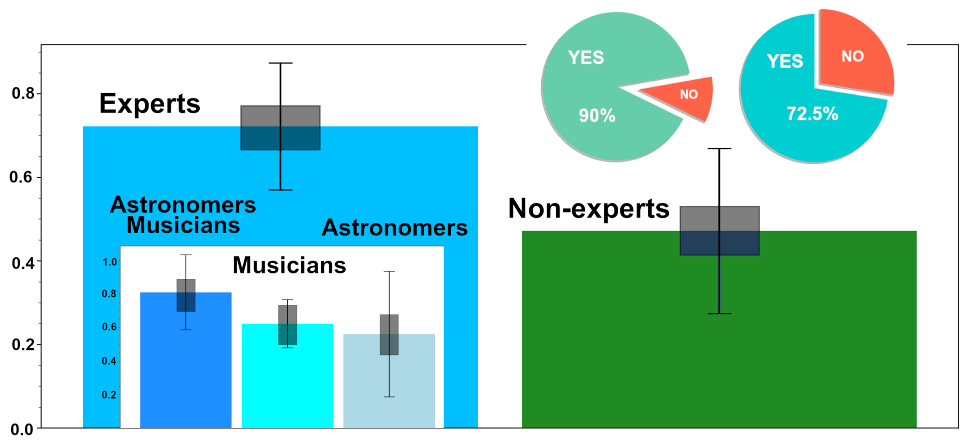

Figure 2: Quantitative results for spectral classification tasks using the CoRoT light curves. Average success rate by group of expertise, 18 experts vs 18 non-experts using training videos. Global qualitative feedback (top-right). Usefulness of the proposal (left), and interest of the participants in listening to more sonifications after completing the test (right).

The second questionnaire[9] was supported only by text descriptions. It presented 8 quantitative questions aimed at detecting the number of absorption lines in galaxy spectra, and 8 qualitative questions in which participants could rate the sonification proposals. Fig.3 shows the results from 46 participants [15].

[7] *ViewCube* video: `https://vimeo.com/1005208084`
[8] Light curve classification questionnaire: `https://forms.office.com/e/veisXfB6fw`
[9] Absorption line detection questionnaire: `https://forms.office.com/e/iCFyXfwsdT`

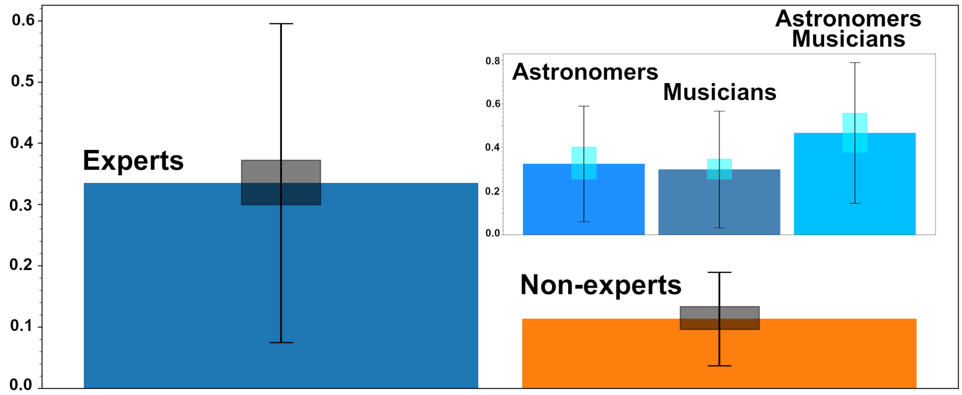

Figure 3: Quantitative results for absorption line recognition tasks using Large Early Galaxy Astrophysics Census (LEGA-C) spectra. Average success rate by group of expertise, 17 experts vs 17 non-experts without training videos.

4 Conclusions

Incorporating auditory representations into scientific problems can provide useful information for the detection of hidden events and overlapped information. Sonification can also be used as a complement to visualization for classification tasks, which can make astronomical analysis more accessible for blind and low vision (BLV) researchers.

The results presented in this summary show that simple actions, such as counting sound events, require training and concentration even for experienced participants. Education and training programs seem crucial for improving listening skills and thus enabling the exploration of public stellar libraries and catalogs through sound.

Regarding quantitative feedback, our results are consistent with the conclusions reported by Walker and Nees (2011) [16], Tucker(2022) [17], and Fovino(2024) [18]. Astronomers musicians reported the best results, what seems to confirm the importance of prior knowledge and training for a complete understanding of the sonifications. Participants self-declared experienced in Astronomy and Music performed around 1.53 times better than self-declared non-experienced (0.72 vs 0.47 success rate) using training videos for galaxy spectral type classification tasks. Nevertheless, all participants could retrieve information from the sonifications with the support of self-training videos. In galaxy absorption line detection tasks without the support of training videos, participants self-declared experienced in Astronomy and Music performed 2,94 times better than self-declared non-experienced.

Attending qualitative feedback, the acceptance of our proposal can be considered good, with 90% of the participants finding it useful and 72.5% declaring interest in listening to more sonifications after completing the survey. Abstract sound representations obtained lower rates than recognizable music forms, suggesting the use of musical sounds for outreach applications and widely accepted auditory representations.

Acknowledgments

Special thanks to Rubén García-Benito and Enrique Pérez Montero for your close collaboration and guidance through the *ViewCube* project. This contribution was presented at the Symposium on Sonification 'Exploring New Horizons: Transforming Astronomy Through Sonification and Multidisciplinary Innovation,' organized by Rubén García-Benito (Chair, IAA-CSIC), Sandra Pauletto (Co-chair, KTH), Enrique Pérez-Montero (IAA-CSIC), and Anita Zanella (INAF). The research includes data from the STELIB, MILES and CaT library services developed by the Spanish Virtual Observatory in the framework of the IAU Commission G5 Working Group: Spectral Stellar Libraries. It also uses data collected with the TESS mission, obtained from the MAST data archive at the Space Telescope Science Institute (STScI). Funding for the TESS mission is provided by the NASA Explorer Program and by the NASA Science Mission Directorate. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555. The CoRoT space mission, launched on 2006 December 27, was developed and operated by the CNES, with participation of the Science Programs of ESA, ESA's RSSD, Austria, Belgium, Brazil, Germany and Spain. The study also uses data provided by the Calar Alto Legacy Integral Field Area (CALIFA) survey (`https://califa.caha.es/`). Based on observations collected at the Centro Astronómico Hispano Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck-Institut fűr Astronomie and the Instituto de Astrofísica de Andalucía (CSIC).

References

[1] Zanella, A. et al. 2022, Nature Astronomy, 6(11), 1241-1248.

[2] Lynch, J.F. 2017. Acoustics Today, 13(4), 27-34.

[3] Merced, W.D., & Gastrow, M. 2018. Proceedings of the IAU, 14(A30), 596-597.

[4] Arcand, K.K. et al. 2024. Frontiers in Communication, 9, 1288896.

[5] Walker, B.N. 2023. ARBOR Ciencia, Pensamiento y Cultura, 199(810), e728.

[6] Lenzi, S., & Ciuccarelli, P. 2024, DRS.

[7] Worrall, D. 2019, Springer.

[8] Hawthorne, C. et al. 2018, arXiv preprint.

[9] García Riber, A., & Serradilla, F. 2023, Springer, 132-147.

[10] García Riber, A., & Serradilla, F. 2024, AES Proceedings.

[11] García Riber, A., & Serradilla, F. 2022, ICAD Proceedings.

[12] García Riber, A., & Serradilla, F. 2023, ADASS Proceedings.

[13] García Riber, A., & Serradilla, F. 2024, ICAD Proceedings.

[14] García Riber, A., García-Benito, R., & Serradilla, F. 2024, RAS TI, 3(1), 748–758.

[15] García Riber, A., & Serradilla, F. 2024, AES Journal, 72(5), 341-351.

[16] Walker, B.N., & Nees, M.A. 2011. Logos Publishing House, 1, 9-39.

[17] Tucker Brown, J. et al. 2022, Royal Astronomical Society, 516(4), 5674-5683.

[18] Fovino, L.G.N. et al. 2024, The Astronomical Journal, 167(4), 150.