

Sonifying Satellite Imagery: Exploring the Environmental Context of Architecture Faculties in Italy

Gabriele Stancato

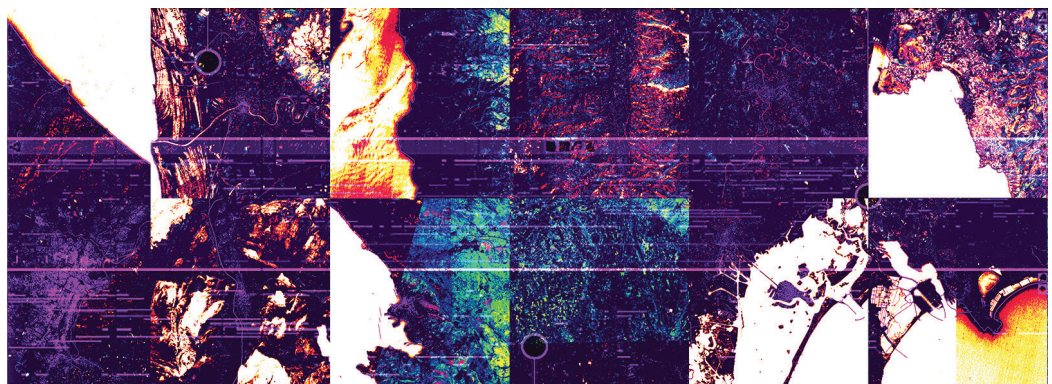
Abstract

This study presents a sonification method employing parameter mapping to analyze the environments surrounding 33 Italian architecture schools. Satellite imagery from *Sentinel-2* and *Landsat 8* was used to evaluate NDVI (vegetation), NDWI (water bodies), BU (built-up areas), and LST (land surface temperatures) within a bounding box of 20 km sides around each school. These indices were processed to create synthetic 'galaxy-like' images and translated into sound, where specific data streams were mapped to distinct musical instruments and timbres. The process incorporates differentiated BPMs to reflect territorial variations and uses the Lydian scale to harmonize note production. The sonification output is integrated into videos that dynamically combine acoustic spectra with satellite imagery. These visual-audio representations highlight the spatial heterogeneity of the landscapes while making complex data accessible to non-specialists and architecture students.

This approach transforms static satellite data into dynamic temporal features, illustrating progressive territorial differences of contexts surrounding architecture schools making them comparable at a more intuitive level.

Keywords

Sonification, satellite imagery, parameter mapping, data translation.



Composition of satellite
images merged as galaxies
pictures.

Introduction and background

Sonification is the process of translating data into non-speech audio to communicate information, often for educational or inclusion purposes [Dubus, Bresin 2013; Zanella *et al.* 2022, pp. 1241-1248]. A common approach for linking visualization and sound is parameter mapping, where visual elements such as position and color are translated into auditory components like pitch and timbre [Enge *et al.* 2024]. This mapping can either be literal and redundant, as in waveform representations, or complementary, combining visual information with sound production [Enge *et al.* 2024]. Sonification methods can be designed to support psychoacoustic interpretations, emulating positional object nature, or be more artificial, creating symbolic sounds to represent image qualities. These approaches can produce more realistic acoustic effects or more iconic music [Sanz *et al.* 2014, pp. 161-171]. The use of data sonification offers a means to overcome barriers related to science literacy, numeracy, and visualization, making complex data more accessible. This particularly benefits individuals with visual impairments, enabling them to engage with scientific data more inclusively and effectively [Sawe *et al.* 2020]. Astronomical communication has traditionally relied on visual representations, often excluding non-sighted audiences. In response, NASA initiated sonification projects aimed at engaging the public, including those with visual impairments, and several researchers have used NASA data for sonification experiments. For instance, in 2022, Arcand explored translating scientific data into various forms of communication, such as 3D printing, sonification, and visual descriptions, all aimed at effectively conveying complex astronomical information. Arcand emphasizes the importance of integrating verbal explanations with visual and tactile media to create a cohesive and accessible experience. It advocates for the application of these accessible content-creation methods across a broader range of scientific communication [Arcand 2022, pp. 53-56]. Further research by Arcand, involving feedback from 3,184 participants (including both blind or low-vision and sighted individuals), explored the sonification of data from three astronomical objects. The results demonstrated a significant improvement in data comprehension for visually impaired audiences, who are typically excluded from science fields where visual representation is the primary method of communication [Arcand *et al.* 2024]. The 'HARMONICES' project, led by Traver and Bergh in 2023, sonified planetary data in the solar system using a Markov model to create pitch collections based on various parameters [Traver, Bergh 2023]. This project allowed audiences to experience the sounds of orbiting planets through spatial audio techniques and interactive elements using a MIDI controller. Sonification also plays a valuable role in earth sciences, where it can simplify the interpretation of complex climate and geological data for laypeople [Russo *et al.* 2024, pp. 1-3]. Smith *et al.* presented the 'Accessible Oceans' project, which aims to create auditory displays to convey oceanographic data in informal learning environments like museums and aquariums [Smith *et al.* 2023, pp. 96-101]. These initiatives underscore the potential of sonification to bridge the gap between complex scientific data and public understanding, making the information more accessible and engaging. From a psychological perspective, the most significant sonification parameters include pitch, tempo, and volume [Beadling, Vickers 2023, pp. 15-22; Dubus, Bresin 2013]. These factors contribute to the coding of state transitions, identification of distinct elements, and the evolution of processes [Madhyastha, Reed 1995, pp. 45-56]. Furthermore, the selection of musical modes during data-to-sound conversion can evoke different emotional responses, with the major mode typically associated with happiness and the minor mode conveying sadness [Hallam *et al.* 2016].

Methods and materials

This study aims to convert satellite imagery data into sounds that intuitively communicate information about urban environments and climate conditions. The focus is not on creating artistic or aesthetically pleasing music, but rather on using the parameter mapping approach [Sarkar *et al.* 2012, pp. 86-90] to convey complex, static data in a dynamic and accessible manner. Satellite images are converted into a sequence of notes, and the resulting sound

is presented in a video, showing both the dynamic audio spectrum and the scanning of the image. This loop begins with the image and returns to it through sound. It is well-established that urban environments can exhibit anisotropy, with variations in feature continuities that influence cognitive processes and emotions [Stancato 2024a, pp. 2091-2101; Stancato 2024b]. These nuances play a significant role in shaping people's perceptions and experiences of urban space [Piga et al. 2021, pp. 2580-2586; Lou et al. 2024, pp. 691-706; Piga et al. 2023, pp. 849-858; Stancato, Piga 2024, pp. 443-458]. In many cases, for the studio's tasks, architectural students work on the urban layout of their town or region, developing a familiarity with the environments surrounding their schools. To illustrate various urban environments in which students usually live, work, or study, a set of images centred around architectural schools has been used in this work. The auditory description of these contexts can provide a synthetic interpretation of the territory, where elements of rhythm, the richness of notes played by different instruments, and the range of sounds illustrate the interplay between anthropogenic and natural components. The sonification process is elaborated through a sequence of automatic algorithms in three steps: i) satellite image download and elaboration, ii) parameter mapping to sound, and iii) video editing. The use of automation grants that the results are comparable and that the same sound effect always conveys the same meaning. The satellite base images for the sonification are obtained, using as centre the location of Italian architecture schools within a bounding box spanning 10 km per side from the centre (20 × 20 km). In some cases, one city can host multiple architecture schools (e.g., Rome); in that case, the location of the satellite picture is the centroid derived from representative schools' points. Another special case is the presence of detached architecture school centres belonging to the same institution (e.g., Milan). In this case, the rectorate building is the only one considered in the analysis; an exception is when the detached centre is the only one hosting an architecture programme (e.g., Syracuse for Catania).

Satellite image download and elaboration

For this study, *Sentinel-2* data were preferred for indices such as the NDVI (Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index), and NDBI (Normalized Difference Built-Up Index) due to its higher spatial resolution (10-20 m) and more frequent revisit time (5 days). These characteristics allow for a more detailed and timely analysis of vegetation, water bodies, and urban areas. Conversely, *Landsat 8* was

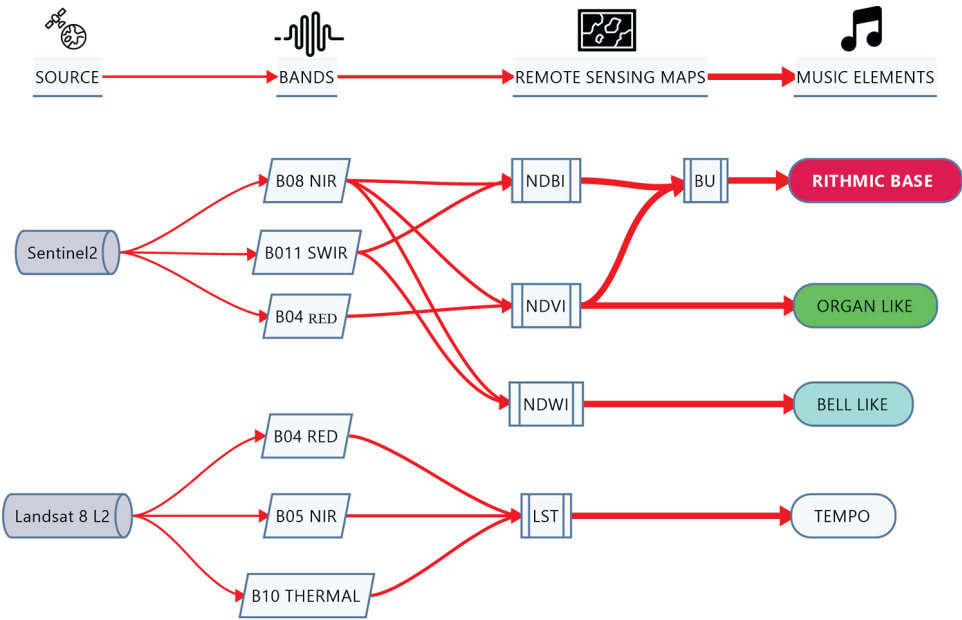
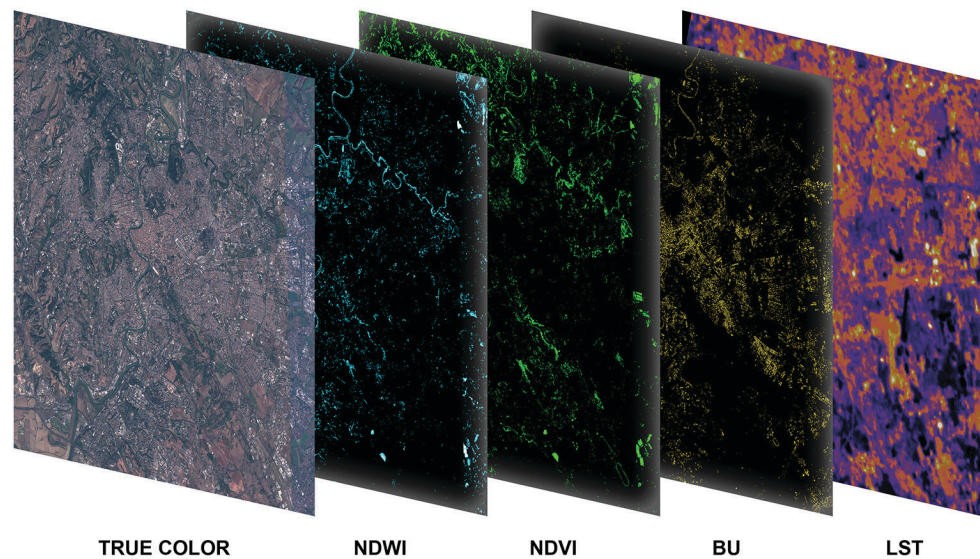


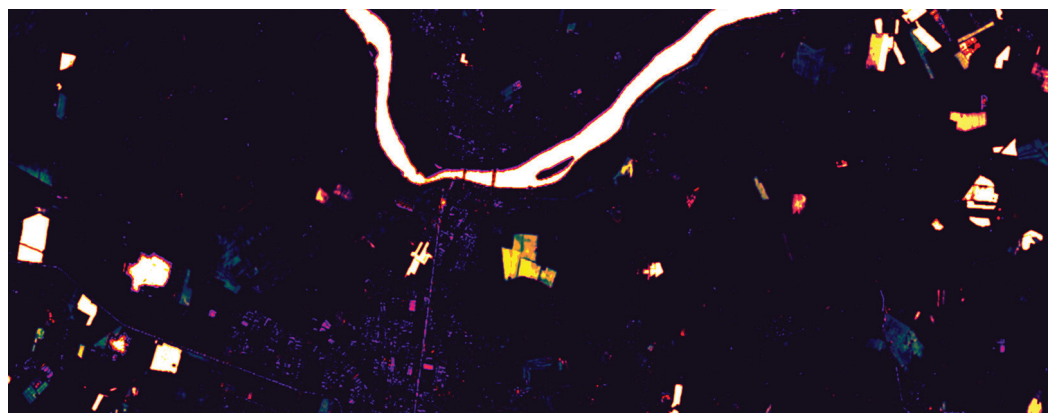
Fig. 1. Flowchart of the sonification process. Specific bands are downloaded from the satellite sources and fused to obtain maps of vegetation, water bodies, LST (land surface temperature), and build-up areas. Temperature. LST is used to determine the tempo, while the other maps determine the pitch values for instrumental tracks.

Fig. 2. Layers of satellite imagery. From left to right: true color, NDWI, NDVI, BU, LST. The first one is used for the video editing phase only; the other is used for the parameter mapping in the sonification process.



utilized for LST (Land Surface Temperature) calculations, as its TIRS (Thermal Infrared Bands) provide reliable and widely validated data for surface temperature estimation, which are not available in Sentinel-2 imagery (fig. 1). For consistency and optimal visibility, satellite scenes were retrieved via the Sentinel Hub using the least cloud coverage option, selecting the image with the lowest cloud cover within the summer 2023 period (from June 21 to September 23). BU (Build-up areas) are obtained by subtracting the NDVI values from the NDBI, obtaining a refined image without the influence of vegetation [Lu *et al.* 2014]. The images were filtered to obtain only the most intense values from the satellite surveys, except for LST. The reason for filtering is to limit the pixels involved in the sonification and thus obtain only the representative notes of notable features of the place in the sonification phase. Eventually, the NDVI, NDWI, and BU layers are merged to obtain a synthetic image inspired by NASA imagery used in famous sonification projects [NASA 2023]. The 'galaxyfication' code processes greyscale images, applying preprocessing operations such as Gaussian blurring to reduce noise and normalization to standardize pixel intensity ranges. Based on the satellite image, a suitable colour map is automatically assigned to enhance interpretability. The images are then converted to colour using this colour map with a mix of Viridis, Hot, and Inferno colour ramps. The satellite images are fused by adding their pixel values, followed by contrast stretching to enhance the dynamic range. Gamma correction is applied to further refine the image's appearance, adjusting brightness and contrast (figs. 2,3).

Fig. 3. Fusion of satellite images to obtain a galaxy-like representation of the features of the urban environment. This image is generated out of NDWI, NDVI, BU maps.



Parameter mapping to sound

There are several ways and directions in which the image can be read and converted to sound, consistent with the widely available literature [Dubus, Bresin 2013]. Scanning of the image proceeds from top to bottom; the geographic location from west to east was associated with the pitch value (notes) to mimic the reading of a score written from left to right. Considering the vast area in the satellite pictures, a differentiated BPM computation was obtained for the three horizontal slices of the LST image, calculating the average value of each and converting that value to a range of 66 BPM to 200 BPM. This differentiation described territorial differences and enriched musical dynamics (fig. 4). It is important to note that the sonification process of this work has no intention to be artistic or to create an actual musical composition. This process must balance the need for a consistent conversion of pixel matrix into notes and make the soundtracks as identifiable and harmonized as possible. The Lydian scale, a major scale mode, has been selected to characterize the music and harmonize the note production in the midi file. Research has shown that different modes, including Lydian, evoke varying emotional responses among listeners. For example, studies indicate that while the Lydian mode is perceived as less happy than the Ionian mode, it still holds a unique emotional character that can be used effectively in musical compositions [Temperley, Tan 2013, pp. 237-257]. This unique interval structure contributes to its distinct sound and emotional connotations in a piece of sweet and charming music [Straehley, Loebach 2014, pp. 21-34; Temperley, Tan 2013, pp. 237-257]. This scale is characterized by a raised fourth degree, distinguishing it from the Ionian (major) scale. The Lydian mode is often utilized in various musical contexts, including jazz and contemporary compositions, where it is preferred for its bright and ethereal quality. To make the different information streams recognizable, the NDWI, NDVI, and BU data are matched to different instruments and timbre. The BU index has been interpreted as a dotted chord bass to describe the environment as sound to give the rhythm and impression of the building process or mechanical movements. The NDVI has been translated as an organ, giving it a spiritual or holy character. The NDWI has been interpreted as a bell ring to recall the imagery of raindrops (fig. 5).

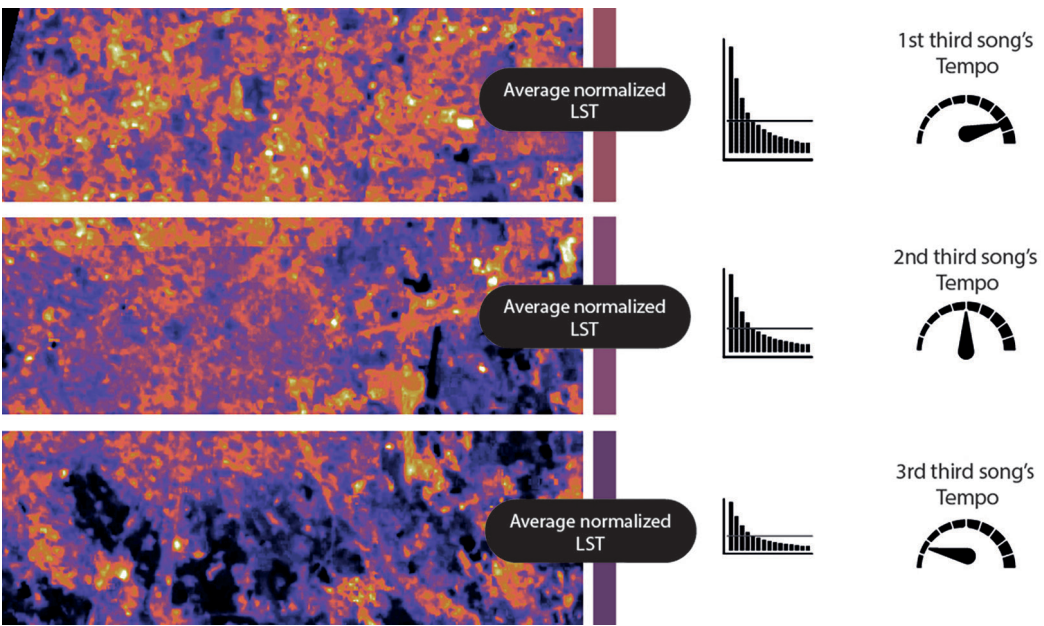


Fig. 4. Tempo computation. The LST image is sliced into three equal parts; for each, the average intensity is calculated and converted to BPM (Beats Per Minute) values.

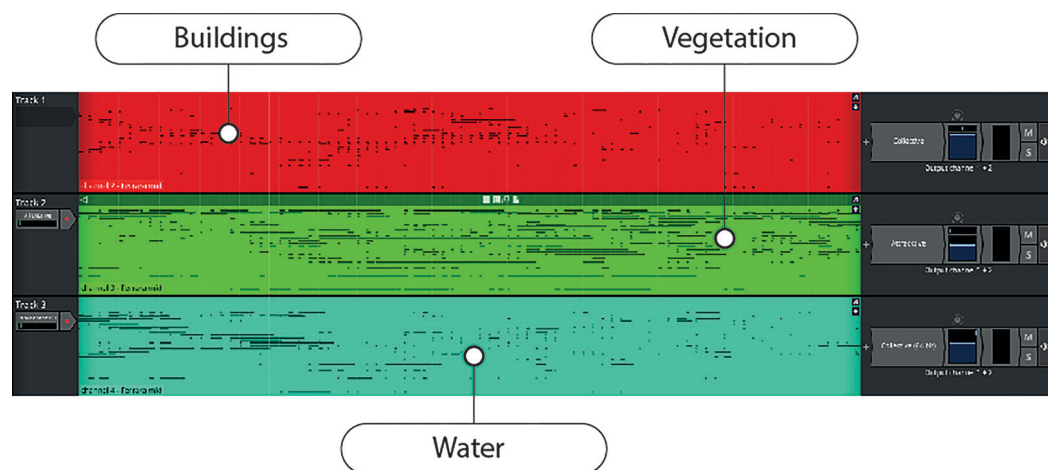
Video editing

Once that music is generated, it is needed to match the imagery with the sound to clarify the relationship between visual and acoustic information. This sort of back translation merges the two worlds of view and hearing, from image to sound and back to the image. Moreover, the translation of images into music implies the conversion from a static language to a dynamic one (picture to music); the use of a video provides a dynamic dimension to the visual information, aligning the two communication means. The setting of a video template helps keep the videos consistent and make them comparable. The foreground layer hosts the sound spectrum representation along a circular path; this tool dynamically shows how the different frequencies change in intensity while the music moves. The middle layer shows the true colour satellite image of the territorial portion involved in the sonification; whilst the music proceeds, the true colour picture vertically fades away, revealing the 'galaxy-like' composition of satellite images. This sequence of video elements guides the audience in identifying which part of the image is used in sound generation at each second (fig. 6).

Instruments

For image processing, an *ad hoc* Python 3.11 code has been developed by the author embedding the following libraries: *sentinelhub* 3.11.1 for remote sensing queries; *pillow* 11.1 for post-processing the images; *music21* 9.3.0, *miditool* 1.2.1, for midi file creation. Eventually, *Waveform Tracktion* 13.2 software was employed to associate musical instruments with midi tracks, and *Adobe After Effect* 25.1 was employed for video editing.

Fig. 5. The three soundtracks match the selected instruments. Black dashes indicate notes; longer lines show duration. The red track is a rhythmic base in stereo balance, green represents an organ-like instrument (left channel), and azure, a bell-like instrument for water (right channel), resembles a river's shape.



Results and insights

Following the abovementioned locations criteria, 33 GPS locations have been selected Italy, each place representing at least one architecture school (fig. 7): Ancona, Università Politecnica delle Marche; Aversa, Università degli Studi della Campania 'Luigi Vanvitelli'; Bari, Politecnico di Bari; Bologna, Università di Bologna; Brescia, Università degli Studi di Brescia; Cagliari, Università degli Studi di Cagliari; Camerino, Università di Camerino; Enna, Università degli Studi di Enna 'Kore'; Ferrara, Università degli Studi di Ferrara; Florence, Università degli Studi di Firenze; Fisciano, Università degli Studi di Salerno; Genoa, Università degli Studi di Genova; Gorizia, Università degli Studi di Trieste; L'Aquila, Università degli Studi dell'Aquila; Matera, Università degli studi della Basilicata; Milan, Politecnico di Milano; Naples, Università degli Studi 'Federico II'; Padua, Università di Padova; Palermo, Università degli Studi di Palermo; Parma, Università degli Studi di Parma; Pavia, Università di Pavia; Perugia, Università

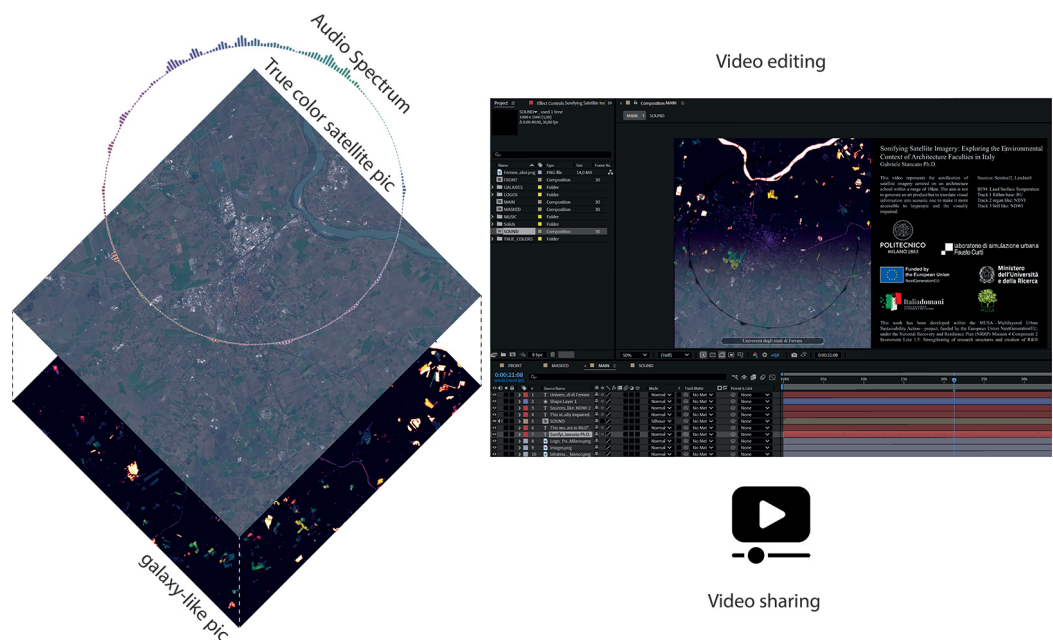


Fig. 6. Video editing phase. On the left side the true color image progressively fade out while the galaxy-like picture is revealed; the circular weave spectrum shows the sound dynamic of the related music.

degli Studi di Perugia; Pescara, Università degli Studi 'G. d'Annunzio' Chieti-Pescara; Pisa, Università di Pisa; Reggio Calabria, Università degli Studi 'Mediterranea' di Reggio Calabria; Rende, Università della Calabria; Rome, Sapienza Università di Roma, Università degli Studi Roma Tre, Tor Vergata; Sassari, Università degli Studi di Sassari; Syracuse, Università di Catania; Turin, Politecnico di Torino; Trento, Università di Trento; Udine, Università degli Studi di Udine; Venice, Università IUAV di Venezia.

A total of thirty-three 20 km x 20 km spatial tiles were generated to illustrate the heterogeneity of the landscapes surrounding various schools of architecture in Italy (fig. 8). These tiles cover a wide range of environments, including coastal, lacustrine, plain, and mountainous areas. Satellite data was extracted for each of the 33 locations to create representative images



Fig. 7. Map of the thirty-two Italian locations employed in this work, listed in alphabetical order.

that illustrate the quantitative and positional relationships between green areas, water bodies, and built structures (fig. 9). These spatial characteristics were further encoded into sounds that serve as both the soundtrack and the animation source for the thirty-three corresponding videos. The videos showing these dynamic interpretations of landscapes are available for viewing on YouTube™ at this URL: <http://bit.ly/40lzmXZ>.

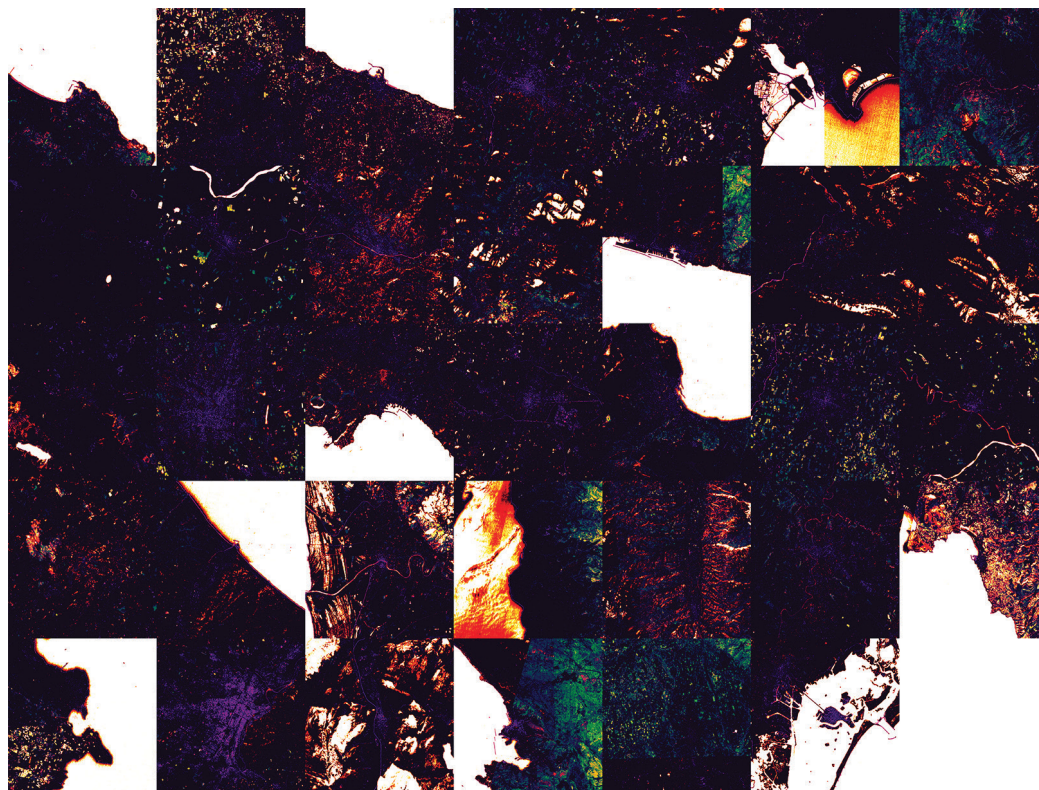


Fig. 8. True-color satellite image mosaic of the thirty-three locations; the images are sorted in location name alphabetical order from left to right and up to the bottom.

Discussion and conclusions

This study used a sonification method based on a parameter mapping process. Satellite images were obtained to describe the environments surrounding Italian architecture schools in terms of building density, vegetation presence, water bodies, and ground temperatures. The auditory translation of visual information transforms static elements into dynamic temporal features. This dynamism illustrates the progressive differences that can be identified within a given portion of the territory. To emphasize these dynamics, a series of videos were created in which sound animates an acoustic spectrum representation while gradually revealing the corresponding sonified satellite image. The choice of the Lydian scale in this study focused on the perception of territory is justified as a means to emphasize an optimistic or progressive narrative, reflecting the idea of interpreting satellite data as an opportunity to enhance our understanding of the land. Its distinctive augmented fourth creates a sense of openness and forward motion, aligning with the exploratory nature of the research. Furthermore, the Lydian scale's structure can be effectively adapted to various tonalities, making it well-suited for mapping specific numerical parameters without introducing harmonic conflicts, thus ensuring clarity and coherence in the sonification process. This approach allows for presenting complex information, such as satellite data, in a more intuitive and accessible form, making it suitable for non-specialists and early-year students. The environments presented in this work are the background of architecture students' education. Intuitively grasping the environmental features can offer them new perspectives and critical insights into the unique characteristics of the territory on both medium and large scales.

Fig. 9. False colour galaxy-like images mosaic of the thirty-three locations; the images are sorted in location name alphabetical order from left to right and up to the bottom.



Limitations

While a mathematical process determines the MIDI production, the choice of instruments remains subjective and may not align with common sense. Additionally, the number of locations selected for the analyses does not correspond to the total number of architecture faculty sites. This limitation was intentional, as the sample size was contained to optimize this process's exploratory phase and represent the major places where architecture is studied. Future research could benefit from a more comprehensive approach, extending the analysis to each specific location to provide a deeper and more nuanced understanding. This work can also serve as a foundation for collaboration with visually impaired associations to develop a more inclusive approach to geographic information. Including sound and communication design experts in the process could further enhance the inclusivity of urban space design, aligning with the objectives of UN Goal 11.

Acknowledgements

This research was developed through two EU-funded projects - NextGenerationEU (National Recovery and Resilience Plan), Mission 4 Education and Research - Component 2 From research to business; - MUSA - Multilayered Urban Sustainability Action - Spoke 1 'Urban Regeneration' - Investment Line 1.5.

Reference List

- Arcand, K. (2022). Chandra's accessible universe: From sight to sound & touch. *Revista Mexicana de Astronomía y Astrofísica Serie de Conferencias*, 54, pp. 53-56. <https://doi.org/10.22201/ia.14052059p.2022.54.11>.
- 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, Article 1288896. <https://doi.org/10.3389/fcomm.2024.1288896>.
- Beadling, A., Vickers, P. (2023). Listener perception of spatialized audio for embodied interaction in sonification. *Proceedings of the 28th*

- International Conference on Auditory Display (ICAD2023), pp. 15-22. <https://doi.org/10.21785/icad2023.3814>.
- 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>.
- Enge, K., Elmquist, E., Caiola, V., Rönnerberg, N., Rind, A., Iber, M., Aigner, W. (2024). Open your ears and take a look: A state-of-the-art report on the integration of sonification and visualization. In *Computer Graphics Forum*, 43(3), e15114. <https://doi.org/10.1111/cgf.15114>
- Hallam, S., Cross, I., & Thaut, M. (A. c. Di). (2014). *The Oxford handbook of music psychology* (Second edition). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198722946.001.0001>.
- Lou, S., Stancato, G., Piga, B. E. A. (2024). Assessing in-motion urban visual perception: Analyzing urban features, design qualities, and people's perception. In A. Giordano, M. Russo, R. Spallone (Eds.), *Advances in Representation: New AI- and XR-Driven Transdisciplinarity*, pp. 691-706. https://doi.org/10.1007/978-3-031-62963-1_42.
- Lu, L., Guo, H., Wang, C., Pesaresi, M., Ehrlich, D. (2014). Monitoring bi-decadal development of urban agglomeration with remote sensing images in the Jing-jin-Tang area, China. In *Journal of Applied Remote Sensing*, 8(1), 084592. <https://doi.org/10.1117/1.JRS.8.084592>.
- Madhyastha, T. M., Reed, D. A. (1995). Data sonification: Do you see what I hear? In *IEEE Software*, 12(2), pp. 45-56. <https://doi.org/10.1109/52.368264>.
- NASA. (2023). *Sonifications - NASA Science*. Retrieved January 4, 2025, from <https://science.nasa.gov/mission/hubble/multimedia/sonifications>.
- Piga, B. E. A., Stancato, G., Faccenda, G. (2023). Emotional and cognitive maps for urban design education: A human-centered design learning approach. In L.-Y. Cheng (Ed.), *ICGG 2022 - Proceedings of the 20th International Conference on Geometry and Graphics*. pp. 849-858. <https://doi.org/10.1007/978-3-031-13588-0>.
- Piga, B. E. A., Stancato, G., Rainisio, N., Boffi, M., Faccenda, G. (2021). Emotions and places. An investigation through virtual reality. In *CONNECTING - Drawing for weaving relationships. Languages Distances Technologies, 42nd International Conference of Representation Disciplines Teachers*. pp. 2580-2586. <https://doi.org/10.3280/oa-693.145>.
- Russo, M., Gemon, T. M., Santaguida, A., Hincks, T. K. (2024). Improving Earth science communication and accessibility with data sonification. In *Nature Reviews Earth & Environment*, 5(1), pp. 1-3. <https://doi.org/10.1038/s43017-023-00512-y>.
- Sanz, P., Garcia, R., Sobrino, A. (2014). Scenes and images into sounds: A taxonomy of image sonification methods for mobility applications. In *Journal of the Audio Engineering Society*, 62(3), pp. 161-171. <https://doi.org/10.17743/jaes.2014.0009>.
- Sarkar, R., Bakshi, S., Sa, P. K. (2012). Review on image sonification: A non-visual scene representation. In *2012 1st International Conference on Recent Advances in Information Technology (RAIT)*, pp. 86-90. <https://doi.org/10.1109/RAIT.2012.6194485>
- Sawe, N., Chafe, C., Treviño, J. (2020). Using data sonification to overcome science literacy, numeracy, and visualization barriers in science communication. In *Frontiers in Communication*, 5, Article 46. <https://doi.org/10.3389/fcomm.2020.00046>.
- Smith, L., Bower, A., Roberts, J., Bellona, J., Li, H. (2023). Expanding access to ocean science through inclusively designed data sonifications. *Oceanography*, 36(4), pp. 96-101. <https://doi.org/10.5670/oceanog.2024.124>.
- Stancato, G. (2024a). Quantifying city dynamics: Exploring the urban features representation of Milan's streets. In *Measure / Out of Measure - Devising, 45nd International Conference of Representation Disciplines Teachers* pp. 2091-2101. <https://doi.org/10.3280/oa-1180-c576>.
- Stancato, G. (2024b). The visual greenery field: Representing the urban green visual continuum with street view image analysis. *Sustainability*, 16(21), Article 9512. <https://doi.org/10.3390/su16219512>.
- Stancato, G., Piga, B. E. A. (2024). Image segmentation and emotional analysis of virtual and augmented reality urban scenes. In A. Giordano, M. Russo, R. Spallone (Eds.), *Beyond Digital Representation*, pp. 443-458. https://doi.org/10.1007/978-3-031-36155-5_28
- Straehley, I. C., Loebach, J. L. (2014). The influence of mode and musical experience on the attribution of emotions to melodic sequences. *Psychomusicology: Music, Mind, and Brain*, 24(1), pp. 21-34. <https://doi.org/10.1037/pmu0000032>.
- Temperley, D., Tan, D. (2013). Emotional connotations of diatonic modes. *Music Perception*, 30(3), pp. 237-257. <https://doi.org/10.1525/mp.2012.30.3.237>.
- Traver, P., Bergh, E. (2023). *Harmonices Solaris - Sonification of the Planets*. pp. 242-248. Retrieved from <https://hdl.handle.net/1853/72864>.
- Zanella, A., Harrison, C. M., Lenzi, S., Cooke, J., Damsma, P., and Fleming, S. W. (2022). Sonification and sound design for astronomy research, education and public engagement. In *Nature Astronomy*, 6(11), pp. 1241-1248. <https://doi.org/10.1038/s41550-022-01721-z>.

Author

Gabriele Stancato, Politecnico di Milano, gabriele.stancato@polimi.it

To cite this chapter: Gabriele Stancato. (2025). Sonifying Satellite Imagery: Exploring the Environmental Context of Architecture Faculties in Italy. In Carlevaris et al. (Eds.), *èkphrasis. Descrizioni nello spazio della rappresentazione/èkphrasis. Descriptions in the space of representation*. Proceedings of the 46th International Conference of Representation Disciplines Teachers. Milano: FrancoAngeli, pp. 3379-3388. DOI: 10.3280/oa-1430-c930.