Sonicolour: Exploring Colour Control of Sound Synthesis with Interactive Machine Learning

Tug F. O'Flaherty
Centre for Digital Music
Queen Mary University of London
London, UK
t.f.oflaherty@se24.qmul.ac.uk

Luigi Marino
Centre for Digital Music
Queen Mary University of London
London, UK
l.marino@qmul.ac.uk

Charalampos Saitis
Centre for Digital Music
Queen Mary University of London
London, UK
c.saitis@qmul.ac.uk

Anna Xambó Sedó
Centre for Digital Music
Queen Mary University of London
London, UK
a.xambosedo@qmul.ac.uk

Abstract

This paper explores crossmodal mappings of colour to sound. The instrument presented analyses the colour of physical objects via a colour light-to-frequency sensor and maps the corresponding red, green, and blue data values to parameters of a synthesiser. Interactive machine learning is used to facilitate the discovery of new relationships between sound and colour. The role of interactive machine learning is to find unexpected relationships between the visual features of the objects and the sound synthesis. The performance is evaluated by its ability to provide the user with a playful interaction between the visual and tactile exploration of coloured objects, and the generation of synthetic sounds. We conclude by outlining the potential of this approach for musical interaction design and music performance.

Keywords

Supervised learning, crossmodal mapping, additive synthesis, synaesthesia

1 Introduction

Generating audio through colour-to-sound mapping is a developing area in which many devices have been produced, including visual impairment aids and art installations [10]. Previous research commonly uses pitch height as a basis for colour and sound associations [33]. In contrast, whilst many sound-to-colour mapping methods have been proposed, most involve audio analysis of spectral features, such as the spectral centroid [32]. Studies have shown an association between soft timbres and blue, green, or light greyscales, and harsh timbres with red, yellow, or black [1]. Nonetheless, fewer research exists on creating performable instruments capable of facilitating the desired colour-to-sound associations of instrumentalists and composers.

Sound-to-colour mappings may be presented through a relatively common cognitive phenomenon known as crossmodal correspondences, where associations between music, and notably timbres, are often made with specific colours [16]. A rare medical form of such colour and sound mappings, with little research and creative potential, is known as audiovisual synaesthesia, for which mappings can be consistent and memorable to the listener [19]. Audiovisual synaesthetes may precisely and idiosyncratically associate particular sounds with distinct shapes, colours, brightness, and spatial location [14, 17].

There exists a long history of music and colour associations, although commonly related to tonality, perceived by famous composers including Messiaen and Rimsky-Korsakov [16, 27]. Such literature suggests crossmodal correspondences explored through an instrument, associating performers' perceptions of visual colour and sound colour (i.e., timbre), would offer performative benefits, for instance, facilitating expressivity in musical interaction through action-perception loops, subtlety and even creative ambiguity [8]. Inspired by NIME infra-instrument aesthetics [6], this paper seeks to explore the performability and creative freedom afforded by mapping colour and brightness to sound, through the creation of a custom instrument, utilising supervised machine learning and additive synthesis. A focus is placed on answering the following research question: *what can colour-to-sound mapping offer regarding unexplored benefits to performability and creative affordances?*

Unlike common mappings between pitch height and colour, the *Sonicolour* instrument presents colour mappings using spectral features and frequency manipulations to produce varied timbres. This increases sonic variation and facilitates investigation into alternative mappings for creative expression from colour, alongside the instrument's pitch control, retaining pitch playability. Such potential findings may benefit sonic expression, particularly in assisting visual or neurodivergent learners with exploring and creating music. Neurodivergent listeners are shown to often require learning through actions, and some individuals also diagnosed with synaesthesia attempt to remember through associations between senses [31].

Additive synthesis permits a greater control of the harmonics used to create the timbre [11], and supervised learning techniques are ideal for generating output from small known datasets [13], including the colour-to-sound mappings. However, a significant challenge lies in generating accurate representations of timbre colour, as current understanding of timbre perception is limited [36], and non-standardised. This instrument builds upon Adeli and colleagues' [1] identified colour-to-sound mappings, with the first author generating their own timbres using their perceived colour associations.

NIME '25, June 24–27, 2025, Canberra, Australia

2 Related Work

2.1 Crossmodal Associations Between Sound and Colour

Much is known about crossmodal correspondences [7], although previous studies of colour-to-sound mapping provide inconclusive results [30]. Some studies show a primary association between natural mappings of colours and pitch ranges, known as pitch registers, although colour-to-timbre mapping often yields no consistent association between individuals [27]. It is suggested that timbre quantification provides challenges, due to its variation over pitch and loudness, when played on acoustic instruments [27].

In constrast, Reymore and Lindsey discovered timbre-to-colour mappings do show some consistency, varying most significantly across registers and by brightness, with some effects of saturation and warm-cool differences [23]. Siddiq and colleagues also noted similar differences across registers; although most instruments sounded yellow, brown, or orange to non-synaesthetes, synaesthetes offered more timbre-to-colour variation [27].

Spence and Di Stefano suggest that colour-to-sound mappings are disputed, claiming that emotion is responsible for colour and pitch associations, due to insufficient evidence or agreement on distinct colour-to-sound mappings [30]. They argue that brightness-to-pitch correspondence, and potentially hue-to-timbre associations, are evidenced, but other colour-to-sound mappings are represented by emotion. Moreover, they note that crossmodal correspondences offer significance to studying and understanding music, and leveraging emotional connotations with colour can benefit musical expression.

Although the existence and methods of sound-to-colour associations are inconclusive, these studies evidence a pitch-to-brightness correspondence. This interplay led to the consideration of unified pitch and brightness controls for the *Sonicolour* instrument, to allow a performer to manipulate the pitch and brightness independently, offering greater creative freedom, with certain timbral brightness accuracy over the pitch range.

2.2 Pitch and Timbral Brightness

Previous studies have concluded that a clear interplay exists between pitch and timbral brightness [3, 12, 21, 25]. Evidence suggests that when a timbre is brighter, the pitch is interpreted as higher, and vice versa [25], also applying to pitch changes. Similarly, studies have shown that the pitch affects the brightness correlated with the spectral centroid [18].

It is generally acknowledged that the spectral centroid controls the perceived timbre brightness [29]. Audio with a low spectral centroid, having more concentrated low-frequency energy, sounds dull, and conversely, a high spectral centroid appears bright. The application of the spectral centroid-timbral brightness interplay with the pitch is a minor focus of this work. This seeks to identify how brightness and pitch may be controlled for increased performability and variation of colour-to-sound mappings within an instrument.

Colour-to-sound mapping exploration offers wide-ranging benefits. Using colour to generate music can evoke significant emotional responses from listeners, particularly in conveying emotion in films [15], offering possibilities in generating impactful music. Moreover, such mappings allow synaesthetic-inspired composers to fully convey their emotions and ideas through music and visuals, to non-synaesthetic listeners [2, 4]. In turn, by affecting musical expectations, through knowledge of colour-to sound-mappings, composers may develop more engaging and emotive pieces using audiovisual crossmodal correspondences [23].

2.3 Colour to Sound Mapping in Musical Interaction

Work has been conducted in the field of colour-to-sound mapping for musical interaction for many years. Castel's 1735 colour organ mapped colours to pitch class, Rimington's 1893 colour organ attempted to map physical properties of colour to sound, and Pridmore's 1992 sound-to-light transducer mapped tone and hue [30]. The 1915 "clavier lumière" (colour-light keyboard) was a musical instrument invented by Alexander Scriabin for his piece Prometheus: The Poem of Fire to visually accompany the music [22].

More recently, two-dimensional images have been converted into music, partially based on their hue values [20]. Images were analysed and processed using Max/MSP and Jitter,[1] before being converted to Hue, Saturation, and Lightness (HSL) for sound generation [20]. Colour scanning was used to iterate over the pixels, to relate the musical timbres to the image content, aligning with the notion that timbre requires colour and form [28]. Moreover, colours have a weaker relationship to timbre than shapes, as proven by most orchestral instruments reportedly sounding blue, red or green, by participants [1].

3D image HSL pixels have been translated into music as an auditory aid for visually impaired users to represent images: hue mapped to timbre, saturation to pitch, and luminosity to instruments, according to timbral brightness (high-frequency content) [5].

Related work found in interactive installations is beyond the scope of this paper, such as Golan Levin's projects that foster dynamic relationships between visual inputs (including colour)[2] and sound or Zach Lieberman's multimedia installations where visual inputs like colour play a role in shaping the audio output.[3]

This paper further develops the idea of crafting sound based on colour, although instead using machine learning, offering the possibility of creating a wider range of timbres than with traditional parameter mapping approaches. Furthermore, trainable machine learning models facilitate customisation of the instrument's colour-to-timbre mappings, enabling performers to generate new sounds tailored to their creative intent, crossmodal correspondences, or synaesthetic experiences. This aligns with the literature on supporting open mappings between colour and sound. It also extends the performability of colour-to-sound mapping by offering dynamic colour, pitch and brightness control, focussing on accurate timbre manipulation of one synthesised sound, without the constraints of colour-to-sound mapping through iterating over predefined coloured pixels of an existing image.

3 Prototype Development

This section outlines the developmental process followed to produce the instrument. The physical instrument design and construction is examined, before detailing the creation of an additive synthesiser using Max. Moreover, explanation on training a supervised machine learning model to design the colour-to-timbre

[1] https://cycling74.com
[2] https://www.flong.com
[3] http://zach.li

mappings is provided. The instrument's firmware and sound generation utilises many existing libraries, explained on the companion website, which also provides all instrument designs and code.[4]

3.1 Physical Instrument

The crafted instrument uses a wooden case of 315mm x 265mm x 90mm dimensions, designed to be relatively compact. The box was designed using a template from Makercase,[5] with custom cutouts for all sensors, and simple but clear engraving, to identify the sensors' sonic functions. To facilitate a low entry barrier for use, as shown in Figure 1, all sensors were labelled according to their sonic variation properties. For instance "brightness" controls the spectral centroid, such that non-technical users would suitably understand its controls.

For durability and long-lasting use, the instrument was crafted from 6.5 mm plywood, offering strength, with acrylic inlays protecting sensor components. The instrument was sanded and polished with beeswax, a natural, renewable, and biodegradable resource that is considered an eco-friendly choice. A custom-designed colour wheel was 3D-printed and attached to the instrument. Using sustainable materials was a priority; wood, reclaimed, or scrap materials were used where possible.

Sensors were used to elicit the performer's input and control the sonic output. To test the research question, a TCS3200 colour sensor[6] was mounted to the instrument's inside case. This reads the colours placed above it, either directly on the colour wheel or through optional filters, to later determine the timbral colour.

A Bela Trill Square[7] facilitates control of the brightness and pitch simultaneously, through its X-axis and Y-axis readings, respectively. This sensor was selected for its multi-input control, to test the second research focus.

Moreover, to offer the performer additional sonic controls, thus facilitating performance, a 10 kΩ linear potentiometer varies the oscillator depth of the additive synthesiser's voices. An MPU-6050 accelerometer[8] provides control over panning, volume, and amount of feedback.

To facilitate effective user feedback, thus improving usability, an LED was added, outputting a visual representation of the perceived sonic brightness to the performer. A liquid-crystal display (LCD), displays the instrument's, including pitch and oscillator wave information, allowing the performer to relate the produced audio to visual forms.

All sensors were connected to an Arduino UNO R4 WiFi,[9] running custom firmware, to interface with the sensors. The code receives the sensors' current status and transmits the values over Wi-Fi using Open Sound Control (OSC) [34], received on the connected computer running the sound generation program. Bidirectional communication permits the Arduino to receive OSC messages from the sound generation program, to update the LED brightness and LCD status, as the generated audio properties change.

[4]https://sites.google.com/view/sonicolour-sc-1
[5]https://en.makercase.com/#/basicbox
[6]https://ams-osram.com/products/sensor-solutions/ambient-light-color-spectral-proximity-sensors/ams-tcs3200-color-sensor
[7]https://learn.bela.io/products/trill/about-trill/
[8]https://invensense.tdk.com/products/motion-tracking/6-axis/mpu-6050
[9]https://docs.arduino.cc/hardware/uno-r4-wifi/

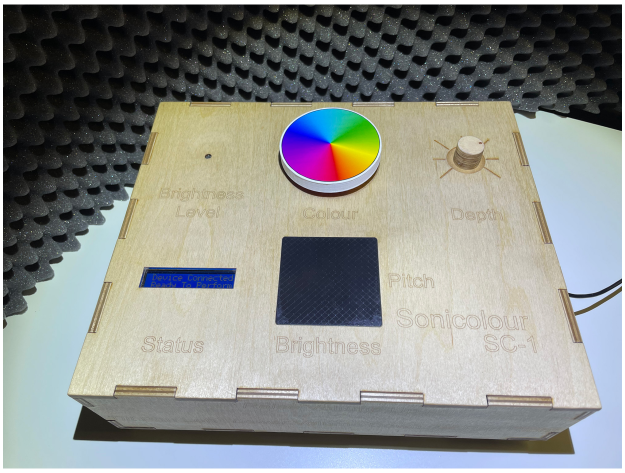

Figure 1: Designed and produced physical instrument controller

3.2 Additive Synthesis

The generated sound utilises additive synthesis techniques, with output generated by 25 voices, permitting exact control over the audio produced. Additive synthesis offers accurate control over the instrument's spectral properties and output frequencies [26], facilitating effective timbral manipulation to create precise colour-to-sound mappings.

Sound was synthesised using Max 8, due to its high performance and efficient tools to prototype the specificities of the desired audio synthesis (Figure 2). When the host computer is connected to the instrument's Wi-Fi network, the Max Patch receives the sensor data in User Datagram Protocol (UDP) packets, via OSC. The data is extracted according to the message route, rescaled if necessary, and transmitted to the appropriate synthesiser control.

Received colour sensor messages are re-transmitted to the supervised machine learning model, with its output controlling the additive synthesiser voices' ratios and depths, and the generated sound's Attack, Decay, Sustain, and Release (ADSR) envelope. The fundamental frequency is controlled by the Y-axis reading from the Bela Trill Square sensor, with the X-axis determining the multiplication factor of each voice's amplitude, thus manipulating the spectral centroid. This adjusts the perceived timbral brightness, providing flexibility to create "dull" to "bright" timbres with respect to the currently-selected timbral colour.

The linear potentiometer manipulates voice depth, through varying the frequency of each voice's low-frequency oscillator (LFO) modulation wave. The accelerometer's X-axis, the instrument's height, controls the sound output master volume, whilst the Y-axis, for left and right movement, determines the audio's left and right panning respectively. A delay and feedback loop, adding controllable amplitude levels of combined delay and feedback, is activated using the Z-axis movements, for forward and backward tilt. This offers simple sonic manipulation, enhancing creative flexibility, and introducing learners with little musical background to audio effects processing.

Jitter is used to generate graphics, providing visual accompaniment following the output received from the colour sensor, and the brightness from the Bela Trill Square. This permits the audio with accompanied colours and brightness to be recorded, for later review, or evaluated by an audience in real-time, to reflect on our research question.

3.3 Interactive Machine Learning

Colour-to-sound mappings were generated using a supervised machine learning approach. Wekinator [13] was used due to its simple interface, and its efficiency to interactively generate a machine learning model. The Wekinator project was set to receive OSC messages from the colour sensor, and a supervised neural network model was generated for each ratio and LFO frequency of the 25 additive synthesis voices, offering a powerful method to generate the customised mappings. Neural network models were also generated for the attack, decay, sustain, and release components of the synthesiser's ADSR envelope.

Each neural network model was a small multilayer perceptron with three inputs corresponding to three RGB colour values (red, green, blue), a single hidden layer containing three neurons (nodes) with sigmoid activation functions, and a single output neuron with a linear activation function.

Training data output values were selected through using the Wekinator `randomize` button to select initial parameter values, until a timbre emerged that the first author believed broadly matched their perception of a particular colour. Manual changes were made to each synthesiser voice's LFO, ratio, and ADSR envelope times, to create suitable consonant and dissonant harmonics.

Colour mappings were based on common findings [1], with the primary colours of red offering an aggressive, harsh timbre, yellow appearing bright and shrill, and blue offering a calm and less dissonant timbre. The secondary colours of orange produced a moderately aggressive and dissonant timbre, green a moderately calm and neutral timbre, and purple providing a dark and dull timbre, in the opinions of the first author.

Upon crafting a suitable timbre, the associated colour was placed above the colour wheel, and 50 colour sensor reading input values were recorded for each synthesiser voice's ratio, LFO frequency, and ADSR envelope machine learning model in Wekinator. This ensured mappings remained clear and accurate, even when slight variations of the received colour occurred, due to colour sensor inaccuracies, through capturing all minimal sensor reading variance around a static colour. The process was repeated for all primary and secondary colours, to equally cover the full colour spectrum. After repeating this process for all primary and secondary colours of the colour wheel, to produce varied and accurate colour-to-sound mapping across the colour wheel, the model was trained, to learn the colour-to-value output. Previously-unseen minor hues are interpolated from nearby colour sensor values.

Formal validation was performed on the trained model, via creating a train-validation-test model. The train dataset comprised all colour wheel primary and secondary colours, the validation set included all tertiary colours, and differing major hues formed the test set. The model was formally evaluated, to validate its efficacy on unseen colours.

Supervised machine learning was effective in this use case. As both, the input—received from the colour sensor connected to Arduino—and output values—based on the colour-to-timbre trained Wekinator mappings detailed above—were known, mappings could be crafted. Moreover, a machine learning approach offered the benefit that unseen colours are automatically mapped to corresponding timbres. Consequently, colours in-between those trained on the colour wheel, or unseen colours, produced by using additional filters above the colour sensor, need not be manually mapped to produce a sensible output.

4 Composition

To address the research question, examining the interplay between pitch and spectral centroid, and validating the concept of colour-to-sound mapping, a short musical piece titled "The Colour Space Zoom" was produced. The composition utilises 25 additive synthesis voices, initially exploring the pitch-spectral centroid interaction using the colour green. This is followed by repeating the same musical exploration through the colours of turquoise, blue, purple, pink, red, orange, and yellow. Such process allowed the author to evaluate both the accuracy and musical affordance of the colour-to-timbre mappings.

A performance of the composition may be viewed at https://youtu.be/OA2S8dAP-I8.

A first-person investigation was subsequently produced, determining the perceived accuracy of the colour-to-timbre mappings, and allowing identification of any interplay between pitch, spectral centroid and the associated timbre colour. Following the performance, the piece was reviewed, and notes were recorded on the instrument's relation to the research questions.

5 Reflection

Upon reviewing the completed instrument performance, analysis was conducted on the audio with respect to the initial research question: *What can colour-to-sound mapping offer regarding unexplored benefits to performability and creative affordances?*. This section discusses the reflections relating to colour-to-sound mapping, the interplay between pitch and spectral centroid in determining perceived brightness, and any other notable observations.

5.1 Colour to Sound Mapping Analysis

It was observed that, in the opinions of the first author, colour-to-sound mapping can be customised to personal views and needs. During review of the performance, colours green and blue evoked associations with peaceful timbres, using consonant harmonics and sustained ADSR envelopes. As colours were changed to purple, pink, and red, so too did the timbre's aggression, resulting in chaotic, dissonant harmonics, and shorter ADSR envelopes. Moreover, orange and yellow timbres naturally contained increased brightness, through utilising more voices with higher harmonics.

These observations appear to support the theory of the existence of colour-to-sound mappings, through crossmodal correspondences. As is presented, one form of successful mapping involves tranquil timbres with blue, aggressive with red, and bright with yellow, although it must be noted that other colour mappings do exist, and may be equally valid. Previous research suggests that blue and green are associated with soft timbres, and red and yellow with harsh timbres [1], which supports the observations in this study.

It must be noted that no other colour-to-sound mappings were tested during this study, thus comparisons between mapping accuracies are not possible but with potential for future work. Moreover, a mapping approach utilising previous colour psychology research could offer significant developments in accurate colour-to-sound mappings.

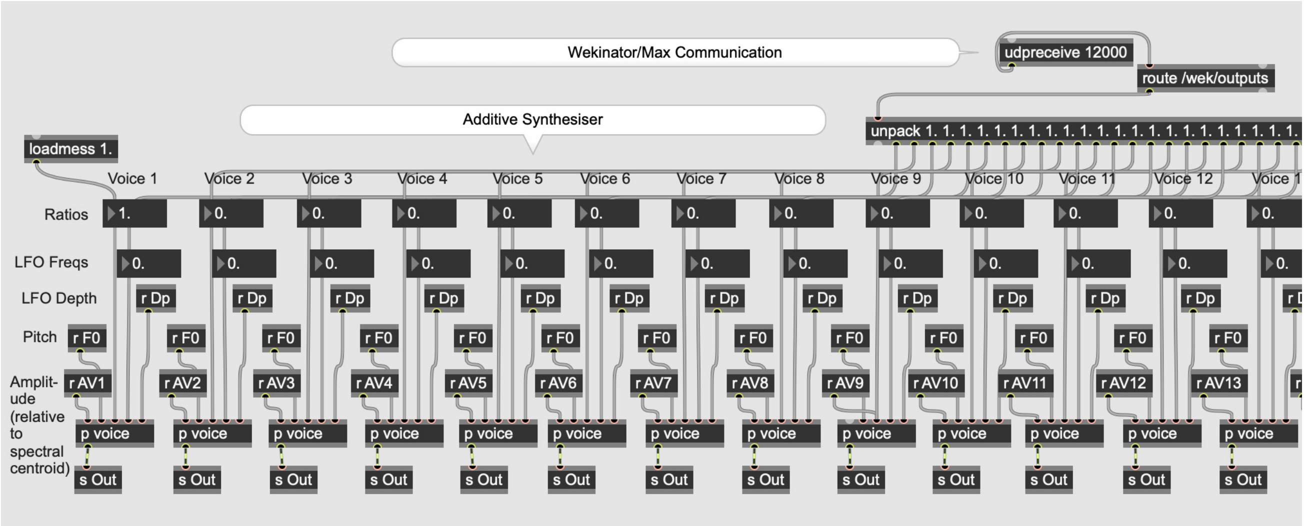

Figure 2: Custom-made Max 8 additive synthesiser patch for sound generation

5.2 Brightness via Pitch and Spectral Centroid

An association between pitch and brightness was observed. As the pitch (here fundamental frequency) is lowered, so too is the perceived brightness of the timbre, likely due to the highest energy, thus most audible, harmonics in turn having a lowered frequency. Similarly, if the pitch is increased, the timbral brightness also appears to increase, suggesting brightness has a dependence on pitch.

Similarly, spectral centroid has a direct connection to brightness. Through controlling the spectral centroid, it was noted that a lower spectral centroid, thus more energy in the lower harmonics of the audio closer to the fundamental frequency, sounds notably more dull, as less energy is present in the higher harmonics. Similarly, with a higher spectral centroid, the timbre appears brighter, with more energy focussed in the higher harmonics.

A connection between pitch and spectral centroid was also observed upon reviewing the performance. It was noted that a low pitch with a high spectral centroid results in a moderately high timbral brightness, as the spectral centroid lowers some of the effects of the dull timbre associated with low pitches. Moreover, a high pitch with low spectral centroid results in a moderately low timbral brightness, with both factors averaging to produce a combined result. This performance shows promise that the instrument can be used to create more consistently bright timbres across a wider range of pitch, through varying the spectral centroid relative to the pitch, however further research is required to develop a precise and musically useful model.

5.3 Additional Observations

During the performance, an interplay between timbre colour and brightness was identified, with some colours, including yellow, appearing brighter than others, such as blue, independent of the instrument's brightness control. Such observation suggests that timbral brightness and colour cannot truly be discretely managed, and such connection must be considered when attempting to accurately control timbral brightness alongside timbral colour.

Notably, the developed instrument does have some limitations, including its delayed response to note input, causing it to appear less responsive, thus more difficult to fully hear the ADSR envelope used in the timbre colour. Furthermore, the colour sensor used is not fully accurate, therefore the timbre colour may sometimes be slightly different to the selected colour, hindering the performability of the instrument. Due to the colour-to-sound mappings being generated by the supervised machine learning algorithm, there is the potential for inaccuracies in the mappings, in this case hindering the accurate control of the spectral brightness, due to the technique used in its manipulation.

Overall, the *Sonicolour* instrument has offered a suitable testbed to explore some of the benefits that colour-to-sound mapping can bring to performability and creative affordances. Affordances here are considered more like open processes tailored to each musician [24]. Although the instrument is not paradigm-shifting of interactive music performance, it offers an original perspective that can contribute incrementally towards a new way of making music. The instrument allows for an intuitive way of training and exploring customised colour-to-sound mappings that any user could investigate. We envision the instrument as especially suitable for neurodivergent, synaesthetic-inspired and visual-minded artists and learners. Future work can confirm if the instrument is an easy-to-use tool for beginners in electronic music based on an audiovisual interaction metaphor of colouring timbre textures.

6 Conclusions and Future Work

This paper explored the existence of crossmodal correspondences, in relation to colour-to-timbre mapping. It examined the creation of an instrument, using machine learning and additive synthesis, to investigate colour-to-sound mapping, and the interplay between pitch and spectral centroid. Evidence was found to support the existence of crossmodal correspondences through colour-to-sound mapping, alongside an interplay between pitch, spectral centroid, and perceived timbre brightness.

This study offers an initial exploration into a novel approach to performative colour-to-sound mapping, although there are some limitations that warrant further work. The creation and evaluation of the colour-to-sound mappings utilised a limited first-person retrospective, thus more systematic user studies would be necessary to assess performability and creative affordances for non-musicians, musicians, and audiences. It is also of special interest to continue NIME practice-based research using a first-person approach applying related creative methodologies [9, 35].

Moreover, the instrument's lack of precise pitch height/class control, and MIDI or external software support, make the current instrument difficult to use as a compositional or structured performance tool. Future studies must be conducted to validate that results remain consistent when the instrument is performed alongside additional instrumentation, and identify more varied timbre mappings associated with colours. Nonetheless, this initial work shows promise in the potential to perform colour-to-sound mapping using machine learning approaches.

Ethical Standards

The authors have no known conflicts of interest. The reflections utilised first-person studies, with no external participation beyond the first author. The instrument was developed by the first author with a sustainable approach, using reclaimed, recycled and natural materials where possible.

Acknowledgments

Gratitude is expressed to Ms Geetha Bommireddy, for assistance in the manufacture of the instrument. We thank anonymous reviewers for insightful comments and suggestions that have helped improve the presentation of this work.

References

[1] Mohammad Adeli, Jean Rouat, and Stéphane Molotchnikoff. 2014. Audiovisual Correspondence Between Musical Timbre and Visual Shapes. *Frontiers in Human Neuroscience* 8, 352 (2014). https://doi.org/10.3389/fnhum.2014.00352

[2] Matthew Joseph Adiletta and Oliver Thomas. 2020. An Artistic Visualization of Music Modeling a Synesthetic Experience. (2020). https://doi.org/10.48550/arXiv.2012.08034 arXiv:2012.08034

[3] Emily J. Allen and Andrew J. Oxenham. 2014. Symmetric Interactions and Interference Between Pitch and Timbre. *The Journal of the Acoustical Society of America* 135, 3 (03 2014), 1371–1379. https://doi.org/10.1121/1.4863269

[4] Angeliki Antoniou, Berardina De-Carolis, George Raptis, Cristina Gena, Tsvi Kuflik, Alan Dix, Antonio Origlia, and George Lepouras. 2020. AVI2CH 2020: Workshop on Advanced Visual Interfaces and Interactions in Cultural Heritage. In *Proceedings of the International Conference on Advanced Visual Interfaces (AVI '20)*. ACM, 1–2. https://doi.org/10.1145/3399715.3400869

[5] Guido Bologna, Deville Benoît, Pun Thierry, and Vinckenbosch Michel. 2007. Transforming 3D Coloured Pixels into Musical Instrument Notes for Vision Substitution Applications. *EURASIP Journal on Image and Video Processing* 2007 (01 2007). https://doi.org/10.1155/2007/76204

[6] John Bowers and Phil Archer. 2005. Not Hyper, Not Meta, Not Cyber but Infra-Instruments. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Vancouver, BC, Canada, 5–10. https://doi.org/10.5281/zenodo.1176713

[7] Riccardo Brunetti, Allegra Indraccolo, Claudia Del Gatto, Charles Spence, and Valerio Santangelo. 2017. Are Crossmodal Correspondences Relative or Absolute? Sequential Effects on Speeded Classification. *Attention, Perception, Psychophysics* 80 (11 2017). https://doi.org/10.3758/s13414-017-1445-z

[8] Miguel Bruns, Stijn Ossevoort, and Marianne Graves Petersen. 2021. Expressivity in Interaction: A Framework for Design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[9] Doga Cavdir. 2024. Creative Practice as an Evaluation Method: A Case Study with a Movement-based Musical Instrument. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, S M Astrid Bin and Courtney N. Reed (Eds.). Utrecht, Netherlands, 456–464. https://doi.org/10.5281/zenodo.13904909

[10] Costanza Cenerini, Luca Vollero, Giorgio Pennazza, Marco Santonico, Nicola Di Stefano, and Flavio Keller. 2023. Investigating Colour-Sound Mapping in Children and Adults: A Pilot Study. 129–135. https://doi.org/10.21785/icad2023.6710

[11] Peter Comerford. 1993. Simulating an Organ with Additive Synthesis. *Computer Music Journal* 17, 2 (1993), 55–65. https://doi.org/10.2307/3680869

[12] Laurent Demany and Catherine Semal. 1993. Pitch versus Brightness of Timbre: Detecting Combined Shifts in Fundamental and Formant Frequency. *Music Perception* 11, 1 (10 1993), 1–13. https://doi.org/10.2307/40285596 arXiv:https://online.ucpress.edu/mp/article-pdf/11/1/1/547842/40285596.pdf

[13] Rebecca Fiebrink, Dan Trueman, and Perry R. Cook. 2009. A Meta-Instrument for Interactive, On-the-Fly Machine Learning. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Pittsburgh, PA, United States, 280–285. https://doi.org/10.5281/zenodo.1177513

[14] Yanjie Li. 2020. Psychological Mechanism of Audio-Visual Synaesthesia in the Teaching of Instrumental Performance. *Revista Argentina de Clínica Psicológica* 29, 1 (2020), 1182. https://doi.org/10.24205/03276716.2020.169

[15] PerMagnus Lindborg and Anders K. Friberg. 2015. Colour Association with Music Is Mediated by Emotion: Evidence from an Experiment Using a CIE Lab Interface and Interviews. *PLOS ONE* 10, 12 (12 2015), 1–26. https://doi.org/10.1371/journal.pone.0144013

[16] Jingyu Liu, Anni Zhao, Shuang Wang, Yiyang Li, and Hui Ren. 2021. Research on the Correlation Between the Timbre Attributes of Musical Sound and Visual Color. *IEEE Access* 9 (2021), 97855–97877. https://doi.org/10.1109/ACCESS.2021.3095197

[17] David Reby Louise Fernay and Jamie Ward. 2012. Visualized Voices: A Case Study of Audio-Visual Synesthesia. *Neurocase* 18, 1 (2012), 50–56. https://doi.org/10.1080/13554794.2010.547863

[18] Jeremy Marozeau and Alain de Cheveigné. 2007. The Effect of Fundamental Frequency on the Brightness Dimension of Timbre. *The Journal of the Acoustical Society of America* 121, 1 (2007), 383–387. https://doi.org/10.1121/1.2384910

[19] Konstantina Orlandatou, E. Cambouropoulos, C. Tsougras, P. Mavromatis, and K. Pastiadis. 2012. The Role of Pitch and Timbre in the Synaesthetic Experience. In *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music*. Thessaloniki, Greece, 751–758.

[20] Dave Payling, Stella Mills, and Tim Howle. 2007. Hue Music – Creating Timbral Soundscapes From Coloured Pictures. In *Proceedings of the 13th International Conference on Auditory Display*. Montréal, Canada.

[21] Marco Pitteri, Mauro Marchetti, Konstantinos Priftis, and Massimo Grassi. 2017. Naturally Together: Pitch-Height and Brightness as Coupled Factors for Eliciting the SMARC Effect in Non-Musicians. *Psychological Research* 81 (01 2017). https://doi.org/10.1007/s00426-015-0713-6

[22] Harry Chapin Plummer. 1915. Color Music—A New Art Created With the Aid of Science. *Scientific American* 112, 15 (1915), 343–351.

[23] Lindsey Reymore and Delwin T. Lindsey. 2025. Color and Tone Color: Audiovisual Crossmodal Correspondences with Musical Instrument Timbre. *Frontiers in Psychology* 15 (2025). https://doi.org/10.3389/fpsyg.2024.1520131

[24] Matthew Rodger, Paul Stapleton, Maarten van Walstijn, Miguel Ortiz, and Laurel S Pardue. 2020. What Makes a Good Musical Instrument? A Matter of Processes, Ecologies and Specificities. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Romain Michon and Franziska Schroeder (Eds.). Birmingham City University, Birmingham, UK, 405–410. https://doi.org/10.5281/zenodo.4813438

[25] Charalampos Saitis and Zachary Wallmark. 2024. Timbral Brightness Perception Investigated Through Multimodal Interference. *Attention, Perception, Psychophysics* 86 (2024), 1835–1845. https://doi.org/10.3758/s13414-024-02934-2

[26] Lawrence H. Sasaki and Kenneth C. Smith. 1980. A Simple Data Reduction Scheme for Additive Synthesis. *Computer Music Journal* 4, 1 (1980), 22–24.

[27] Saleh Siddiq, Isabella Czedik-Eysenberg, Jörg Jewanski, Charalampos Saitis, Sascha Kruchten, Rustem Sakhabiev, Michael Oehler, and Christoph Reuter. 2024. Colored Timbres. Do Crossmodal Correspondences Between Musical Instrument Sounds and Visual Colors Rather Depend on Pitch Instead of Timbre? *Musik- Tanz- Kunsttherapie* 34 (01 2024), 182–201. https://doi.org/10.2440/005

[28] Mattias Sköld. 2022. The Visual Representation of Timbre. *Organised Sound* 27, 3 (2022), 387–400. https://doi.org/10.1017/S1355771822000541

[29] Sean Soraghan, Felix Faire, Alain Renaud, and Ben Supper. 2018. A New Timbre Visualization Technique Based on Semantic Descriptors. *Computer Music Journal* 42, 1 (2018), 23–36. https://doi.org/10.1162/comj_a_00449

[30] Charles Spence and Nicola Di Stefano. 2022. Coloured Hearing, Colour Music, Colour Organs, and the Search for Perceptually Meaningful Correspondences Between Colour and Sound. *i-Perception* 13, 3 (2022). https://doi.org/10.1177/20416695221092802

[31] B. Sundararajan, M. Sundararajan, K. Haigh, and S. Vimalraj. 2023. Every Teacher Should Have a Student Like Chris: Embracing Neurodivergence, UDL, and Educating a Variety of Learners. In *EDULEARN23 Proceedings* (Palma, Spain) *(15th International Conference on Education and New Learning Technologies)*. IATED, 7678–7687. https://doi.org/10.21125/edulearn.2023.1989

[32] Lindsay Vickery. 2018. Some Approaches to Representing Sound with Colour and Shape. In *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'18*, Sandeep Bhagwati and Jean Bresson (Eds.). Concordia University, Montreal, Canada, 165–173.

[33] Jamie Ward, Brett Huckstep, and Elias Tsakanikos. 2006. Sound-Colour Synaesthesia: to What Extent Does it Use Cross-Modal Mechanisms Common to Us All? *Cortex* 42, 2 (2006), 264–280. https://doi.org/10.1016/S0010-9452(08)70352-6

[34] Matthew Wright, Adrian Freed, and Ali Momeni. 2003. OpenSound Control: State of the Art 2003. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (22-24 May, 2003). Montreal, Canada, 153–159. https://doi.org/10.5281/zenodo.1176575

[35] Eevee Zayas-Garin, Charlotte Nordmoen, and Andrew McPherson. 2023. Transmitting Digital Lutherie Knowledge: The Rashomon Effect for DMI Designers. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Miguel Ortiz and Adnan Marquez-Borbon (Eds.). Mexico City, Mexico, 350–357. https://doi.org/10.5281/zenodo.11189206

[36] Hong Zhang, Jie Lin, and Shengxuan Chen. 2024. Timbre Perception, Representation, and its Neuroscientific Exploration: A Comprehensive Review. arXiv:2405.13661 https://arxiv.org/abs/2405.13661