Image Sonification Application to Art and Performance

Michele Mengucci[1], Francisco Medeiros[1], and Miguel Amaral[2]

[1] LabIO – Laboratório de Interação e Oralidade, Lisboa, Portugal
labio.pt@gmail.com, mengucci@gmail.com, excalibas@gmail.com
[2] IN+ Center for Innovation, Technology and Policy Research,
Instituto Superior Técnico – Universidade de Lisboa, Portugal
miguel.amaral@tecnico.ulisboa.pt

Abstract.

An application for image sonification originally developed to aid blind users associating sounds to colors and light, is able to convey a sound-image synesthesia effect to common people. Since a digital image is a composition of pixels (elementary picture elements), it is converted into a composition of their elementary sounds. A straightforward mapping of pixels position and color to sound parameters makes it robust and able to detect presence of particular shapes. Its use in artistic field provides means for painting sonification, installations in art exhibitions, and live performance. Its audio output is customizable by the user. The software is open-source and completely developed in Pure Data, therefore open to contributions. Its latest version as well as its recent applications in artistic fields are described in this paper. Potential implementations include smartphone apps and games for blind users, image monitoring and color/sound therapy.

Keywords: Audiovisual systems, data sonification, image analysis, sensory substitution, art technology, synesthesia.

1 Introduction

This research follows a project[1] that looked for digital means to provide, through sound, information on surrounding colors and light to blind people [1]. It improved a software tool able to convert a digital image into a characteristic sound composition. Empirical tests and experiments with artists unveiled its potential as a performance art instrument. Synesthesia is a neurological phenomenon in which stimulation of one sensory or cognitive pathway leads to automatic, involuntary experiences in a second sensory or cognitive pathway. It is apparently associated to a sense of wellness by synesthetes, people who report such experiences. Historically the relation between visual and auditive realms, color and sound, has inspired synesthetes [2] and non-synesthetes [3], composers [4] and scientists [5]. The evolution of digital technology has made available a great variety of software and devices to link sound to image for two main purposes: to aid blind people [1,6-14] or people with vision flaws [15-16] and for pure artistic pleasure [17-24]. Data Sonification studies how to exploit the sense of hearing to analyze information from data, allowing a different perception of these than through common visual methods (graphs, histograms, etc.) [25-26]. Images are a particular data set, therefore the application described here is an Image Sonification tool. Research on the relationship between the physics of light and sound, and a search for an informative and attractive output sound has led this study to its actual state. The first known machine for the real-time performance of color 'graphics' was Louis-Bertrand Castel's "Clavecin oculaire" (1734). Many machines have since followed, to produce either compositions of animated colors, dubbed color

[1] See-Through-Sound, funded by FCT, the Portuguese Fundation for Science and Technology (UTA-EXP/MAI/0025/2009)

music, or moving compositions of color and form, like "Lumia" by Thomas Wilfred, the developer of the "Clavilux" organ (1922). In the digital era we find several attempts to explore the generation of audio by visual data; from very simple to highly sophisticated software like *Metasynth* [18] which reads images from left to right, maps the colors by stereo placement (pan) and brightness to volume. It allows to choose which instrument to play. In *Coagula* [19] images are played from left to right as if they were spectrograms of an audio signal. In *Audio Paint* [20] each line of the picture is an oscillator, and the taller the picture, the higher the frequency resolution. While the vertical position of a pixel determines its frequency, its horizontal position corresponds to its time offset. In 2012 a new sensory substitution device, the *EyeMusic* [14], was introduced to the public. It conveys color information by using different musical instruments for each of the five colors: white, blue, red, green and yellow (black is represented by silence). An auditory cue (beep) is sounded at the beginning of each left-to-right scan of the image. The higher musical notes on a pentatonic scale represent pixels that are located higher on the y-axis of an image. The time offset after the cue indicates the x-axis location of the pixel. Many other software applications have been released as research projects or independent productions, as discussed in [6], like the *Voice* and *Kromophone*. Given the current context we implement a novel and alternative mapping of the light spectrum into sound, to convert images into informative audio.

2 Software upgrades and applications

The principal purpose of the application, called SonarX, is real-time sonification of digital images (live or recorded video, pictures). This functionality was explored for blind users in [1,6]. Later it has been used in applications on visual art and performance. The user controlled parameters, like scale, ground note, scan/loop speed and others, allow to customize the audio output. Like the majority of applications previously described, it also scans the image to convert its information into sound with a vertical top-to-bottom scan, not left-to-right. The vertical scan is in accordance with the way our brain processes an image at first glance: the act of human vision grabs information from the whole horizon at once, identifying homogeneous regions [27]. In mappings of image into sound, vertical position is usually mapped to pitch. Color, when analyzed, is frequently mapped to stereo placement (pan) of the sound. As shown in [1,6] the audio print of a digital image is here considered as the composition of the sounds of all its pixels. The horizontal position and color information (Hue, Saturation and Value) for each pixel is mapped respectively to Pan (stereo placement), Pitch, Timbre and Loudness. The vertical position is implied in the time sequence of events generated by a frame scan. An optional sound cue at the beginning of the scan indicates the top of the image. These settings were used in the research with blind users [1,6], while the applications to performance and visual artwork originated several upgrades.

Scale - The light spectrum is mapped to a "just intonation" scale spanning two octaves, the lower one for the cool colors (hues from blue-green through blue-violet) and the higher for the warm colors (red through yellow, browns and tans included). This mapping was experimented during the sonifications of visual artworks by Sylvia Carolinne De Andueza which resulted in an audio-video installation of sonificated pieces at CCJF (Centro Cultural Justiça Federal), Rio De Janeiro in March-May 2014 [28]. It can be chosen (between other mappings which are still available as options (continuous octave, two chromatic octaves, an Indian raga, among others). It ususally results in harmonious sound outputs as it associates the intervals of hues in a scene to sound intervals which minimize the beatings of overtones.

White threshold – Our mapping plays hues, so all the grays, from black to white, are meant to be silent. Usually gray in digital images is not "real" gray (zero saturation), therefore a customizable interval of saturation was introduced to threshold the gray pixels. This is useful to silence a white background, like in [28].

Timbre - The timbre of sound is used to convey information about saturation of color. There are three available options: 1 - the sound is a pure tone mixed with white noise inversely proportional to

saturation (more gray means more white noise with a volume proportional to its brightness); 2 - the sound varies from pure tone to a saw tooth wave proportionally to saturation; 3 - the sound varies from pure tone to a "vocal waveform" proportionally to saturation. The association of white noise to grays has shown to have consistent synesthetic impact during the Festival Bang Awards (Torres Vedras, Portugal, May 2014) where the sonification of the scene was the soundtrack of the digital dance performance Senses Places [29-30] and at exhibition "Entre o 6 e 8A" (AMAC - Auditorio Municipal Augusto Cabrita, Barreiro) [31]. In the first, multicolored avatars projected onto hanging tulles served as a perfect input imagery to create responsive sounds which were complemented with live singing and percussion, enhancing the contact/contrast between virtual and real world. In the second, paintings sonifications of pieces created particular ambient music for the opening of the exhibition.

Resize – The recent visual art sonification experiences showed that the coarse horizontal resolution attenuated the impact of significant visual details present in the scene. After the last publication [6] the image was resized horizontally from 12 to 15 in order to have a central pixel, but actually, the image is resized to 45 pixels. These contain the average values of 45 segments of horizontal lines of the image, and control the parameters of 45 sound generators. The sound produced is richer in harmonics and can capture smaller details in the sonificated image. it has been useful at the performance of live painting "De Cor&Som" [32], which consisted in creating a visual piece from scratch while musicians (voice, guitar, clarinet and percussions) played to the sounds generated by the visual artwork. The visual artist, the musicians and the audience were involved into a synesthetic environment that influenced the act of the creation of a visual product with an extra audio significance.

Scan - The scan line moves vertically on the image at a constant speed, producing a regular "loop". It now can also be manually controlled by the user with a slider control. When the scan line is static it creates a continuous sound. While it is useful to detect objects and to verify the details of a sound generated by an image It has also been used in performance controlled by a smartphone inclination.

4. Conclusions and Future Work

This software was created to provide means for sensory substitution to aid blind people, and has been used afterwards to create environments where music and light were linked together and people could experience synesthesia. The simple image analysis system provides robustness and a quick sound response to any input image: a body moving in front of the camera or colors being painted on a canvas. Thus the effect during a performance is evident and attracts the attention even to who is unaware of the installation. It provides means to explore synesthesia as a therapeutic condition [33], mixing color and sound therapy. In the attempt to keep a coherent link between the physics of sound and light (the range of light frequencies corresponds approximately to an octave $4\text{-}8\times10^{14}$ Hz) the initial aim was to map the light spectrum to a sound octave and maintain all the nuances of color and position in this process. This is still an available option but the Hue and the Pan are now by default discretized for functional reasons. The main focus is actually on improving GUI in order to stimulate contribution by other artists/scientists in an open-source framework. Introducing motion detection to control the scan speed proportionally to amount of movement in the image, more advanced image processing algorithms, Midi and other types of audio output are also being considered.

Acknowledgements. Barreiro municipality for hosting our performance at AMAC, FCT (science and technology Portuguese foundation) for funding the project 'SeeThroughSound' (UTA-EXP/MAI/0025/2009).

References

1. **Cavaco, S., Henrique, J.T., Mengucci, M., Correia, N., Medeiros, F.**. *Color sonification for the visually impaired*, in Procedia Technology, M. M. Cruz-Cunha, J. Varajão, H. Krcmar and R. Martinho (Eds.), Elsevier, volume 9, pp. 1048-1057, 2013.
2. **Hahl-Koch**. Jelena. *Kandinsky*. New York: Rizzoli, 1993.
3. **Mondrian**. *Broadway Boogie Woogie.* New York: MOMA, 1942.
4. **Bowers, Faubion**. *Scriabin, a Biography*. New York:: Dover Publications, 1996.
5. **Peacock, Kenneth**. *Instruments to Perform Color-Music: Two Centuries of Technological Experimentation*, *Leonardo* 21, No. 4, pp. 397-406. 1988.
6. **Cavaco, S., Mengucci, M., Henrique, J.T., Correia, N., Medeiros, F.**. *From Pixels to Pitches: Unveiling the World of Color for the Blind*, in Proceeding of the IEEE International Conference on Serious Games and Applications for Health (SeGAH), 2013.
7. **Meijer, P.**. *An Experimental System for Auditory Image Representations*. IEEE Transactions Biomedical Engineering, vol. 39, pp. 112-121, 1992.
8. **Auvray, M., Hanneton, S., O'Regan, J.K.**. *Learning to perceive with a visuo-auditory substitution system: Localization and object recognition with the vOICe.*In Perception, vol. 36, pp. 416–430, 2007.
9. **Bach-y-Rita, P., Kercel, S.W.**. *Sensory substitution and the human-machine interface*. Trends in Cognitive Neuroscience, vol. 7(12), pp. 541–546, 2003.
10. **Auvray, M., Hanneton, S., Lenay, C., O'Regan, K.**. *There is something out there: distal attribution in sensory substitution, twenty years later*, Journal of Integrative Neuroscience 4, pp. 505-21, 2005.
11. **Capelle, C., Trullemans, C., Arno, P., Veraart, C.**. *A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution,* IEEE Transactions Biomedical Engineering, vol.45, pp. 1279-1293, 1998.
12. **Nielson, L., Mahowald, M., Mead, C.**. *SeeHear*, in Analog VLSI and Neural Systems, by C. Mead, Reading: Addison-Wesley, chapter 13, 207–227, 1989.
13. **Bologna, G., Deville, B., Pun, T.**. "Sonification of Color and Depth in a Mobility Aid for Blind People", in Proceedings of the International Conference on Auditory Displays (ICAD), 2010.
14. **Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S., Amedi, A.**. *EyeMusic: Introducing a "visual" colorful experience for the blind using auditory sensory substitution*. Restorative neurology and neuroscience, vol.32 (2):, pp.247–257, 2014.
15. **Hardisson, Neil**. *Listening to Colors: Life with Extra Senses*, Lecture at the Linda Hall Library of Science, Engineering & Technology, 2013. https://vimeo.com/79003132. Accessed on October 27, 2014.
16. **Brooks, Richard**. *Colour-blind artist learns to paint by hearing*, The Sunday Times, February 24, 2008.
17. **Payling, D., Mills, S., Howle, T.**. *HueMusic – creating timbral soundscapes from colored pictures*, Proceedings of International Conference on Auditory Displays (ICAD), 2007.
18. http://www.uisoftware.com/MetaSynth/index.php, accessed on October 27, 2014.
19. http://www.abc.se/~re/Coagula/Coagula.html, accessed on October 27, 2014.
20. http://www.nicolasfournel.com/audiopaint.htm, accessed October 27, 2014.
21. **Magnusson, Thor**. *Improvising with the threnoscope: integrating code, hardware, GUI, network, and graphic scores*, NIME 2014 (New Interfaces for Musical Expression), Goldsmiths College, Lewisham Way, London: United Kingdom, 2014.
22. **Collins, N., McLean, A., Rohrhuber, J., Ward. A.**. *Live coding in laptop performance*. Organised Sound, 8, 3. pp.321-330, 2003.
23. **Magnusson, Thor**. *An Epistemic Dimension Space for Musical Devices*, Proceedings of NIME. pp.43-46, 2010.
24. **Baldan, S., et al.**. *Musica sull'acqua, a motion tracking based sonification of an aquarium in real time*, Proceedings of the 9th Sound and Music Computing Conference, Copenhagen, Denmark, 2012.
25. **Robert L. Alexander**. 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, vol.119, 7, pp. 5259–5271, 2014.

26. Feder, Toni. *Shh. Listen to the Data*, Physics Today (Physics Today) 65, no. 5, pp.20-21, 2012.
27. Marr, David. *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*, MIT Press, 2010.
28. http://www.sylviacarolinne.com/#!passage-sound/c22i7, accessed on October 28, 2014.
29. Valverde, I., Cochrane, Todd. *Innovative Dance-Technology Educational Practices within Senses Places,* Procedia Technology, vol.13, pp. 122–129, 2014.
30. http://sensesplaces.org/, accessed on October 27, 2014.
31. http://youtu.be/Dr9QO5FN8K8, accessed on October 27, 2014.
32. http://barreiroweb.com/bweb/?p=12084, accessed on October 27, 2014.
33. http://www.theguardian.com/science/2014/apr/27/benefit-synaesthesia-brain-injury-mental-decline accessed on October 27, 2014.