Rohan Raj Upadhyay

Professor Parrish

ARTS 1010

02 October 2022

Programmatically Turning *Van Gogh's Starry Night* into music

Sound and colors have objective and subjective correlations. Subjectively, colors can evoke certain feelings and moods, which we can correlate to a specific type of music. Objectively however, certain notes and colors have similar characteristics in terms of their physical properties. Both sound and light are waves, and both have wavelengths and frequencies. Certain notes can be corresponded to a particular color through a formulaic relationship. For example, the note "A4" has a frequency of 440Hz (Goss). When scaled up by 40 octaves, its frequency reaches to 483.8 THz, which corresponds to the color we consider "orange". This paper examines the history of attempts to objectively correlate sound with colors and, based on the sound-color correspondence, attempts to frame a programmatic method to "compose" a piece of painting by decoding audio data and mapping them to colors. Then, the framework will be used to turn Van Gogh's famous painting, *The Starry Night,* into a piece of generative music. The goal of the paper is to provide an objective corroboration about the relationship between music and visual arts, and how they might be tied to a uni-directional subjective experience of visual arts and music.

THESIS: Because of the physical similarities and objective relationships between sound and light, it is possible to derive a programmatic method to turn visual arts into music.

I. INTRODUCTION

A. Physical relationship between sound and color

B. History of attempts at sound-color correspondence

II. METHODOLOGY

A. Finalizing sound-color mapping parameters

B. Visual decoding of *The Starry Night* using programming

C. Description of the composing process and framework

III. CONCLUSIONS

A. Illustration of musical composition of *The Starry Night* based on objective parameters

B. Don McLean and *The Starry Night*

Physical relationship between sound and color

Before we dive into the subjective intertwinings between sound and color, it's imperative to know how they're interrelated in nature. Sound is a mechanical wave, and travels at a rate of 331.29 meters per second in air (Wong, 3). Light is an electromagnetic wave, and travels at a rate of approximately 300,000,000 meters per second (Schaefer), which is 1 million times more than sound. What makes their physical properties intersect is that they both exhibit wavelike nature. More specifically, both sound and light have frequencies and amplitudes. What we are more interested in, is their frequency relationship. Since light moves significantly faster than sound, it possesses a higher frequency or "cycles per second". The standard notes in the Western music system have a fixed frequency. Thus, scaling each note by a multiple can yield us a frequency equivalent to light in the visible range, which in turn can be corresponded to a color. This shows that music (sound) and visual arts (light) share common traits physically, which is an important establishment for what we are going to derive in this paper, a programmatic method to change a piece of music into visual form based on this sound-color relation. One important distinction between the nature of sound waves and light waves is that sound waves are discrete while light waves are continuous.

History of attempts of sound-color correspondence

The question about the relationship between sound and color in terms of their physical properties arose during the times of Plato. Plato's idea of the eight concentric circles was that each eighth note was a repeat of the first, thus forming a "music of spheres" (Goldsmith). He assigned each note a color, thus forming not just a musical sphere but also a sphere of colors. Similarly, Newton, who was the first person to see the seven colors of light via a prism, came up with a cyclical diagram for the colors. He then assigned each color a note arbitrarily. A music-color wheel was thus formed. Aristotle's idea of a linear arrangement of colors also related color with music. In his linear arrangement, the ends of assigned black and white colors and the colors in between were placed based on how well they blended with each other. Each color was given a musical note. Even though these historical attempts don't provide us with a definitive answer to the question of the exact, quantifiable relationship between sound and color, they do hint us that sound and color have some correlations. In today's times, it's possible to measure the frequency of waves reliably, which further helps to eliminate the speculative aspect of sound-color correspondence of these attempts.

Finalizing sound-color parameters

The sound-color correspondence has different measures for correlation. One of the most objective ones is the formulaic relationship between their frequencies. Since both sound and color are waves, they possess frequencies of a particular number. By scaling up the frequency of sound by a certain multiple, we get the frequency of light in the visible range. For this, it's useful to use the concept of octaves. Octaves are an interval where the frequency of one note is twice than that of the same note in one octave lower. Essentially, scaling up octaves refers to scaling up the frequencies of notes by a multiple of 2. This allows us to mathematically relate sound and light waves. The color we will be using to find the correlation is the color with the following quantitative values: (255, 99, 0) in RGB, (0, 156, 255) in CMY, and the color we generally call "orange". It has a frequency of exactly 483.8Thz and a wavelength of 619.7nm (Goldsmith). Because sound waves have lower frequencies than light waves, we'll be scaling down the octave so as to correspond the color with a note in the standard musical system. Mathematically, scaling down the light frequency by 40 octaves is $\frac{483.8\,THz}{2^{40}} = 440.0\text{Hz}$. The resultant frequency, 440.0Hz is exactly the frequency of the A note in the 4th octave in the standard musical notation. This, in terms of physical properties, entails that the A note and the color orange are 40 octaves apart. Consequently, we're able to derive the note-color relationship for all the twelve standard musical notes. This further allows us to use those related parameters to programmatically convert pixel values to notes, in other words, visual element to musical element.

Visual decoding of *The Starry Night* based on objective parameters

Visual decoding is a process where we'll extract pixel data from an image of the painting of The Starry Night. Initially, the pixel data is to be extracted. The data contains color values of a specific pixel. The color values are stored in RGB format. RGB format is then converted to the CIE L*a*b* (Riemersma) format for reasons explained in the next section. For an example, a set of the first five pixels of the image in the first row would be shown as (255, 23, 2), (31, 0, 288), (55, 21, 0), (200, 255, 1), and (211, 2, 5). These values refer to the color of the pixels, which then will be converted into musical notes. However, a standard 1920x1080 image contains 2,073,600 pixels, which would be converted to notes. Even if the notes are to be played for a mere 0.05 seconds, the composition will last for 28.8 hours. To fix the length issue, we'll be using a resizing function. It uses a scale factor value to resize the image's width and height. For example, if the scale factor is set to 0.05, the width and height of the image would now be 96x10.8. This resizing is done using compression and averaging the pixel values, meaning for every certain number of pixels squeezed into one pixel, the average value of the color of the 10 pixels is computed and set to the final pixel.

Description of the composition process and the framework

We'll now programmatically decode the visual, pixel data from a 1920x1080p image of *The Starry Night* by Van Gogh. I used Python's image processing library Cv2 to decode the pixel data. The following steps were taken: go through each pixel starting from the top left to the right of each subsequent pixel. Then extract the color data from each pixel in RGB format. Because of the discrete nature of musical notes, but the continuous nature of colors, it's not viable to map every possible color to a note. Thus, it was required to find the "color distance" between a color and the nearest mappable color to a note. For example, a different shade of orange would be equated to a color orange that could be mapped to the note A4, as shown in the previous section. Human eyes perceive color differently than the RGB pixel arrangement. This is fixed by using a different color-space and a low-cost approximation algorithm for color distance calculation, the weighted Euclidean distance formula (Riemersma). After finding the nearest mappable color to a pixel color, the color is mapped to a note, and finally, the note is played. For example, if the first pixel of the image has RGB values (255, 0, 0) which is pure red, there's no mapping value to a note for this color. Thus, the program will calculate the nearest color to that color using the color distance function. In this case, the nearest mappable color would be (238, 0, 0) with the minimum color distance of 237.92. This nearest color is equivalent to the note A#4. In this way, the pure red color is converted to the nearest mappable color and then is converted into a note. This, in essence, makes the spectrum of colors to a discrete 12-colors set of data.

Illustration of the generated musical composition of *The Starry Night*

For the illustration, I'll be listing out the pixel data with the nearest colors in RGB format, and the corresponding notes of the first 10 pixels of the image. The scale factor used is 0.01.

Pixel #1: (87, 0, 158) Light Purple, corresponding note: F5

Pixel #2: (255, 236, 0) Yellow, corresponding note: B4b

Pixel #3: (0, 255, 232) Turquoise, corresponding note: D5b

Pixel #4: (82, 0, 0) Brown, corresponding note: F4

Pixel #5: (87, 0, 158) Light Purple, corresponding note: F5

Pixel #6: (0, 124, 255) Blue, corresponding note: D5

Pixel #7: (87, 0, 158) Purple, corresponding note: F5

Pixel #8: (153, 255, 0) Green, corresponding note: B5

Pixel #9: (87, 0, 158) Light Purple, corresponding note: F5

Pixel #10: (87, 0, 158) Light Purple, corresponding note: F5

Since the image is resized down by 100 times, the resultant image essentially highlights the main colors in the painting. In *The Starry Night*, the sky's colored using shades of blue and purple. The yellow color is the color of the moon and dispersing moonlight. The dark brown colors are the colors of the tall foreground bush. The green color is of the small bushes around the tall bush. All of these colors create an assortment of musical notes, based on the 40-octave relationship. Thus, the painting's color values create a musical piece using objective mapping values.

Don McLean and *The Starry Night*

Don McLean is an American songwriter and a singer. The lyrics of his musical piece “Starry, Starry Night” is directly inspired from Van Gogh’s *The Starry Night*. While we mainly looked at the objective derivation of a musical piece from a painting, Don McLean’s Starry Starry Night is a perfect example of how paintings can evoke emotions, and thoughts that can lead to a musical composition. One of the few lines in the song is “You took your life, as lovers often do, but I could’ve told you Vincent, this world was never meant for one as beautiful as you.”, and this implies that the song is meant to be a homage to the artist. The lyrics in the song specifically contrast the nature of life with the contrasting dark blue and yellow tones in the painting. The generative music, created using the colors of the painting, and the lyrical and musical poetry by Don McLean have one similarity in that both converge to represent a piece of visual art. Music and art, in essence, work to depict the microcosm of what we call “art”. Subjective inspirations from paintings like Don McLean’s “Starry Starry Night” is arguably just as valid of a musical piece as the generative music created from the physical, objective properties of the painting itself in the sense that both compositions are open to interpretations. In the end, that’s what art is all about – an openness to countless interpretations.

Works Cited

Goss, Clint. “The Color of Sound.” *The Color of Sound*, 3 Feb. 2017, https://www.flutopedia.com/sound_color.htm.

Wong, George SK. "Speed of sound in standard air." *The Journal of the Acoustical Society of America* 79.5 (1986): 1359-1366.

Schaefer, Bradley E. "Severe limits on variations of the speed of light with frequency." *Physical Review Letters* 82.25 (1999): 4964.

Goldsmith, Joe “An investigation into the relationship between sound and color” *Relationship between sound and color*, 2 Jun. 2019, https://www.people.vcu.edu/~djbromle/color-theory/color01/Relationship-color-sound-joe_goldsmith.html

Riemersma, Thiadmer. *Colour Metric*. https://www.compuphase.com/cmetric.htm

Annotated Bibliography

Caivano, Jos Luis. “Color and sound: Physical and psychophysical relations.” *Color Research & Application* 19.2 (1994): 126-133.

This article intends to relate properties of sound with the properties of light or color. Specifically, the article relates the luminosity of color to the loudness of sound, saturation of color with the timbre of sound, and size of color with the duration of sound.

Cho, Jun Dong, et al. "Sound coding color to improve artwork appreciation by people with visual impairments." *Electronics* 9.11 (2020): 1981.

This paper talks about how visual mapping of sound’s three characteristics: timbre, intensity and pitch to a color’s hue, chroma and value can help in creating tactile pictograms, which can aid visually impaired people to better appreciate visual arts.

Sultzbaugh, J. "Chromoacoustics: The science of sound and color." *The rose+ croix journal* 6 (2009): 94-132.

This presentation aims to share findings from decades-long attempts to find the best method to translate music into chromatic visual displays. It also talks about how “chromoacoustics”, the visualization of music, can be applied in therapy and other fields.

Lee, Yong, Chung-Heon Lee, and Jun Dong Cho. "3D Sound Coding Color for the Visually Impaired." *Electronics* 10.9 (2021): 1037.

This paper goes on detail about a method to convey color and depth coding to the visually impaired based on alternative senses such as hearing and touch. The proposed model is claimed to work on 3D and 2.5D projections.

Bernard, Jonathan W. "Messiaen's synaesthesia: The correspondence between color and sound structure in his music." *Music Perception* 4.1 (1986): 41-68.

This paper attempts to tabulate possible correlations between sound and color based on Olivier Messiaen's published works. Olivier Messiaen's works from 1929 to 1974 show a high degree of internal consistency in choosing colors for his labels since he had colored-hearing synaesthesia.

Kim, Sung-Ill, and Jin-Seung Jung. "A Basic Study on the System of Converting Color Image into Sound." *Journal of the Korean Institute of Intelligent Systems* 20.2 (2010): 251-256.

This paper describes a method to convert color images into sound based on the likelihood in the physical frequency information between light and sound. The author also talks about how this basic conversion method could find applications in making intelligent robots that are capable of emulating human synesthetic skills.

Mengucci, Michele, et al. "From color to sound: Assessing the surrounding environment." *Proceedings of the Conference on Digital Arts and New Media (ARTECH)*. 2012.

In this paper, a method to convert color information of an image to sound by corresponding properties of color with sound, is described. After converting, a sonic print of the image is also created, which is useful for visually-impaired people to infer and assess the world of color and luminosity.

GAY III, JOHN FLOYD. *The correlation of sound and color: Three major metaphysical sources*. University of Missouri-Kansas City, 1972.

In this dissertation, the author emphasizes the importance of the combination of sound and music, also called synesthesia. Then the paper talks about the major metaphysical sources of synesthesia in length.

Iwai, Daisuke, et al. "Approach to non-verbal mapping between sound and color." *Proceedings of the 41st SICE Annual Conference. SICE 2002.*. Vol. 1. IEEE, 2002.

Here, an attempt is made to derive a mapping rule between sound and images from synesthetes. Then the mapping rule is applied to real-life people and is confirmed whether it is acceptable by them.

Chen, Chin-Han, et al. "Emotion-based music visualization using photos." *International Conference on Multimedia Modeling*. Springer, Berlin, Heidelberg, 2008.

This paper comes up with a method to visualize music not just by using the low-level attributes of the sound but also by analyzing the emotions intended in the music. Through subjective evaluation, this method of music visualization supposedly enriches the users' listening experience.

Smith, Sean M., and Glen N. Williams. "A visualization of music." *Proceedings. Visualization'97 (Cat. No. 97CB36155)*. IEEE, 1997.

In this paper, a method to visualize music in 3D space is discussed. Moreover, it also goes in detail about alternative ways to achieve the same task and talks about the applications of visualizing music, specifically in music-reading.

Bain, Matthew N. *Real time music visualization: A study in the visual extension of music*. Diss. The Ohio State University, 2008.

Here, an application of visualizing music is described. It talks about how visualizing music can be considered an extension to live musical performances by the use of artist's sensibilities and computer's generative capabilities.

De Mantaras, Ramon Lopez, and Josep Lluis Arcos. "AI and music: From composition to expressive performance." *AI magazine* 23.3 (2002): 43-43.

In this article, the author talks about the three major types of music systems based on AI: compositional, improvisational, and performance systems. Then it aims to address the limitations of an AI to generate music. Finally, it reports on a music system called SAXEX, which is capable of generating high-quality performances of jazz music.

Roads, Curtis. "Research in music and artificial intelligence." *ACM Computing Surveys (CSUR)* 17.2 (1985): 163-190.

This paper addresses the four important areas in the music field that requires AI's need: composition, performance, musical theory, and digital sound processing. Then it surveys recent work involving AI and music, and finally examines how techniques of AI planning can be used to enhance AI-based music systems.

Roads, Curtis. "Artificial intelligence and music." *Computer Music Journal* 4.2 (1980): 13-25.

This book addresses projects working toward musical intelligence. Furthermore, it talks about applied AI methodologies.

Zulić, Harun. "How AI can change/improve/influence music composition, performance and education: three case studies." *INSAM Journal of Contemporary Music, Art and Technology*1.2 (2019): 100-114.

This paper talks about how AI approaches to generative music possess potentials in changing the standard paradigm of music composition, performance and education. It goes through three impactful case studies where AI is seen to influence the very fields in music.

Cosenza, Glenda. "Play me a picture, paint me a song: Integrating music learning with visual art." *General Music Today* 19.2 (2006): 7-11.

Here, the author aims to synthesize parts of visual arts with music and parts of music with visual arts through an experimental approach. The "Play me a picture, paint me a song" is an activity where people analyze a picture to come up with music and also interpret a song to come up with a drawing.

Kaeppler, Adrienne L. "The look of music, the sound of dance: Music as a visual art." *Visual Anthropology* 8.2-4 (1996): 133-153.

Here, the author talks about the visual dimension of music and how it should be studied as a visual art to better understand the artistic and esthetic construction of society.

Boden, Margaret A., and Ernest A. Edmonds. "What is generative art?." *Digital Creativity* 20.1-2 (2009): 21-46.

In this article, the author explains about the various forms of generative art, also known as computer art and asks whether the appropriate aesthetic criteria and the locus of creativity are the same in each case.

Dorin, Alan, et al. "A framework for understanding generative art." *Digital Creativity* 23.3-4 (2012): 239-259.

In this paper, the need for a standard framework to understand generative art is emphasized. Moreover, it points out the inadequacies of the existing ideas from kinetic art and other alternative frameworks.

Galanos, Theodoros, Antonios Liapis, and Georgios N. Yannakakis. "AffectGAN: Affect-based generative art driven by semantics." *2021 9th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 2021.

In this paper, a new method for generating arts based on prompts is introduced that uses semantic encoding and natural language processing methods. It also showcases a sample of 50 generated images that correctly portray the intended emotion.