PRELIMINARY GUIDELINES ON THE SONIFICATION OF VISUAL ARTWORKS: LINKING MUSIC, SONIFICATION & VISUAL ARTS

Chihab Nadri, Chairunisa Anaya, Shan Yuan, and Myounghoon Jeon

Mind Music Machine Lab

Virginia Polytechnic Institute and State University,

Department of Industrial and Systems Engineering,

1185 Perry Street Blacksburg, VA 24061 USA

{cnadri, danaya14, shany9, myounghoonjeon}@vt.edu

ABSTRACT

Sonification and data processing algorithms have advanced over the years to reach practical applications in our everyday life. Similarly, image processing techniques have improved over time. While a number of image sonification methods have already been developed, few have delved into potential synergies through the combined use of multiple data and image processing techniques. Additionally, little has been done on the use of image sonification for artworks, as most research has been focused on the transcription of visual data for people with visual impairments. Our goal is to sonify paintings reflecting their art style and genre to improve the experience of both sighted and visually impaired individuals. To this end, we have designed initial sonifications for paintings of abstractionism and realism, and conducted interviews with visual and auditory experts to improve our mappings. We believe the recommendations and design directions we have received will help develop a multidimensional sonification algorithm that can better transcribe visual art into appropriate music.

1. INTRODUCTION

The relationship between visual art, music, and technology has frequently been close, as their combined use resulted in masterpieces in the past [1]. With the increasing amount and interest in sonification techniques [2] and image processing techniques using machine learning, transcribing visual experiences into appropriate auditory experiences is now a real possibility. However, the use of these advances in technology has been concentrated on the sonification of images in general and increasing accessibility through short sound feedback. Based on that, we decided to design a sonification algorithm tailored to transcribing visual artworks and appropriately conveying its wider cognitive and emotional message, with the objective of establishing guidelines for the sonification of artworks based on their characteristics (art style, mood, elements…etc.). Additionally, we have conducted interviews with experts from the fields of sonification, visual arts, and music in order to ascertain the main objectives our sonification algorithm should accomplish, as well as learn from their expertise. We expect this study contributes to the development of sonification algorithms in providing the most effective image and data processing techniques suited for the sonification of different artworks.

1.1. Related Works

Research on visual graphics sonification has focused on providing individuals with visual impairments more ways and tools to experience visuals [3]–[5], and highlight useful data processing techniques used to accurately aid the transcription process, such as shape and edge detection machine learning algorithms. In fact, machine learning algorithms have played an increasing role in image processing tasks [6]. Much research has been done concerning saliency detection and salient region cropping in images [7], [8], as saliency provides a good indicator for the importance and relevance of specific areas of the image. Additionally, saliency is an image parameter that plays an important role on eye fixations and visual perception [9], [10], which machine learning algorithms take into account when going through training datasets.

Machine learning has also been used in the arts to classify artworks and identify artistic styles [11], [12]. Such algorithms can play an important role in streamlining classification tasks for galleries and art directories with large collections of artworks and prove practical when implementing sonification techniques. On the other hand, machine learning can also play a role for the composition of music, adapting and learning from performers' musical genre and style to identify and create similar pieces [13].

Aesthetic research on visitor experience at art galleries [14] has revealed visitor patterns of short yet often repetitive viewing of the same paintings and artworks, which can serve as performance criteria for the effectiveness of different visual artwork design experiences. Related research on sound and color mappings has shown that strong associations between music and appearance can exist for nonsynesthetic individuals [15]. Strong associations were also found between music-color mappings and emotions, as different emotions can be evoked through them [16], and the study also found other musical parameters like tempo and mode can play an important yet unclear role.

Prior research on photographic sonification has also yielded several sonification methods and algorithms, adhering to a musical approach focused on musical structure [17] or a naturalistic one following viewing tendencies [18]. Additionally, research on visual saliency in paintings has yielded several algorithms that can imitate human gaze perception and fixations, regardless of art movement the visual artwork belongs to [19]. Indeed, psychological experiments and other saliency-based algorithms have highlighted how the key to understanding art is the

identification of the perceptual process, and how salient regions can correlate with art movements [20].

1.2. Initial Plans and Mappings

Our sonification would apply to visual artworks from different art styles, with abstract paintings yielding a different sound output than that produced from realistic paintings (Figure 1). The sonification algorithm would use mapping strategies between a variety of visual and musical parameters in order to compose and differentiate sound outputs for different artworks (Table 1). While a variety of mapping strategies already exist [2], we planned on conducting experiments to evaluate the significance of each mapping pair.

Table 1: Initial mapping pairing strategies considered between visual and auditory parameters

Visual Parameters \\ Auditory Variables	Pitch	Tempo	Mode	Timbre	Loudness	Musical Composition
Hue	X					
Brightness		X			X	
Saliency	X		X			
Size	X					
Art Style				X		X

Following the proposed experiments, data processing techniques, including shape detection and sentiment analysis machine learning algorithms, would supplement the sonification program with further accuracy. Later trials and implementation for art galleries would be tested using the completed algorithm, which should accomplish goals set at the start of the project.

Performance criteria currently considered for the success of the algorithm are split between musical criteria and semantic ones, listed below:

1. Sonification sound organization: the ability of the sound output to be arranged in a musical way, as opposed to a random succession of notes that are off-key.
2. Sound quality: how pleasant the sonification sounds to participants, and how likely they would want to listen to the music produced.
3. Mood and emotion: feeling conveyed by the sonification. Learning about the emotional meaning of each piece as identified by participants can allow us to validate our experimental mappings and uncover unexpected pairings.
4. Matching quality: the ability of the sonification to match the painting's cognitive and emotional message and reflect the scene presented by the painting. Measuring this parameter provides a clear, albeit subjective, scale on the sonification's success at one of its primary goals (carrying similar meaning to the artwork).

Due to the inherent subjectivity in art appreciation, performance criteria were also dependent on the user.

However, through the expert interview process, more parameters would be identified and used for the evaluation of the sonification algorithm and its successive iterations.
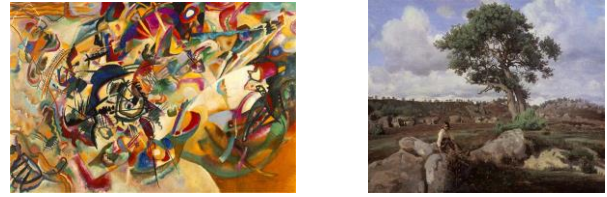

Figure 1: Composition VII, 1913 - Wassily Kandinsky (left), The Raging One, c.1830 - Camille Corot (right), two sample artworks belonging to abstractionism and realism respectively, the two art styles considered for initial experimental tests.

2.　METHODS

2.1. Technical Details

Our program is an application to transcribe digital images of different artworks into MIDI files. We have used JythonMusic[21], a python-based environment for music creation that can also use Java libraries. Given the image provided, our software creates midi files that map image parameters to musical characteristics, creating the music piece.
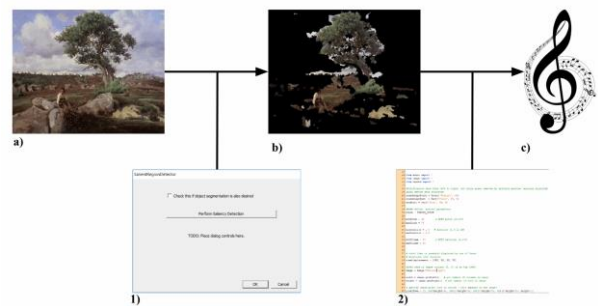

Figure 2: Sample sonification process using saliency mapping, running an image a) through the saliency segmentation software 1) before using JythonMusic code 2) to turn it into music c). The Raging One, c.1830 - Camille Corot.

Over the course of the expert interviews, the saliency and segmentation model created by Achanta et al. was used [7] as an example of a saliency algorithm.

2.2. Participants

As the interview process is still ongoing, five experts, from fields of interest in our study (psychoacoustics, visual arts, sonification), participated in our interview separately. Professors in their field, all had a lengthy experience in their field, with an average experience of 19.4 years.

2.3. Procedure

Experts were first asked questions relevant to their field, such as their previous work and projects. Then, the interviewer asked questions to find out which visual, musical, or sound parameter each expert uses to convey meaning and

accomplish their work objectives, as well as performance measures used in their respective fields. Questions regarding the potential implementation of artwork sonification at art venues were asked next, aiming to discover similarities and differences in perspective between each field. Finally, experts provided advice and observations based on the experimental design, paintings, sonification outputs, or machine learning techniques considered by the team, such as saliency detection.

3. RESULTS

Expert responses varied according to fields of study, with intragroup similarities being found, as well as general trends concerning performance measures shared between all experts (Table 2). Visual arts experts alongside a psychoacoustics expert expressed a preference for the sonification output to carry limited meaning at first. From the perspective of psychoacoustics, following a gradually more complex sound methodology and protocol is essential to reach a complete algorithm. While visual arts experts shared comparable views, not overshadowing the experience of viewing the painting was their main reason. Simply capturing the general feeling of the painting was deemed enough, although greater information would prove beneficial depending on the target audience. Indeed, every visual art expert independently concluded that multiple sonification outputs for any single painting, different in terms of breadth of cognitive and emotional information carried over, would offer the most flexibility for subjective art appreciation at art venues and potential accessibility issues for people with visual impairments.

Expert views on challenges facing the project were less uniform and heavily shaped by their experience. While sonification and psychoacoustics experts emphasized sound quality and methodology respectively, visual arts experts' main priority was to respect the individuality involved in art interpretation. This manifested itself in concerns over matching the painting and not limiting the artwork to a single interpretation.

Important parameters used by experts in their work included brush movement, color choice, geometric shapes, and depicted objects and scenes present in the artwork for visual arts experts. Pitch, timbre cutoff frequency, tempo, and stereo

pan were some of the musical parameters mentioned by musical experts, which helped verify mapping strategies used in earlier sonification projects [2].

Concerning relevant data processing techniques, experts held mixed views on saliency detection. For sound experts (psychoacoustics and sonification), saliency detection would constitute a novel way to determine the temporal flow of the artwork. Using the most salient regions detected first, the sonification algorithm could imitate naturalistic viewing patterns and focus on composing salient elements in detail. For visual arts experts, the importance of the background in setting the mood of a painting meant that saliency detection should be used as an additional parameter when sonifying the entire painting. Additionally, sentiment analysis of an artwork's description (as present in art venues) was perceived as an efficient way to capture information on the painting's background, which can then be implemented into the sonification output (e.g. location of artist at time, state of unrest at location… etc.). Since artworks descriptions usually include the art style present in the artwork, using machine learning algorithms to classify paintings was deemed unnecessary.

Experts agreed researchers should make use of visitor viewing patterns as an objective performance measure. Specifically, visual arts experts considered the number of times a participant reviews a visual artwork more important than viewing length. Another suggestion was to test participants' understanding of the artworks' intended message as a measure of success.

Lastly, experts' opinions on differences that should be present between the sonification of realistic art and abstract art followed similar reasoning, although each expert provided different guidelines (Table 3). According to experts, the sonification of realist art needs to incorporate or convey the scene presented in the artwork by adding iconic sounds for objects in the painting, although sound selection would become a challenge. Also, saliency in realism should be limited, as backgrounds carry important meaning necessary for interpretation. As for abstract art, brush movements, geometric shapes used, and color choice should have a large impact on the sonification output. The use of saliency in abstract art could include the sonification of salient regions only, although the preferred approach would be to alternate sound output between different salient areas. Common amongst both styles was the need for relevant instruments to

Table 2: Initial Expert Feedback during Sonification Project Interview

Experts	Experience	Parameters to convey meaning	Art venue visitor approaches	Challenges
Sonification	23yrs	Pitch, amplitude, timbre cutoff frequency, performance frequency, stereo pan, counterpoint. Respecting painting shape and viewing pattern	Consistent sound quality for painting, i.e. a musical masterpiece for a great work of visual artwork	Sound quality
Psychoacoustics	11yrs	Speech intelligibility, tempo, color contrast for abstract art. Painting elements for realist art	Simple sonification output that can become more complex later, focusing on carrying over a few meanings while letting visitors complete understanding with visual data	Sound protocol and methodology
Visual Arts	20yrs	Brush movement and complementary color choices for abstract art. Objects for realistic paintings	Offering multiple sonification results. Output can be different from painting, focus on offering different modality without minding all musical qualities	Respecting differences in interpretation
Visual Arts	13yrs	Appropriate choice of music and art, choice remains arbitrary, especially for VR projects	Capturing general idea of painting would be fine, emotional ellicitation and clues on painting elements	Matching painting
Visual Arts	30yrs	Color brightness, lines and shapes (geometric or real objects)	Offering multiple sonification results. The more text, information, and meaning provided to visitors the better. Amount of times one views artwork more impactful than length of time viewed	Music not overpowering visual experience

be used according to the artwork's background.

Table 3: Expert consensus on sonification differences between realism and abstractionism

Realism	Abstractionism
Proper representation of painting elements through iconic sounds (e.g. wind blowing)	Color choice, brush movement, and shapes used as most important visual parameters that express artist's mood
Limited scope of saliency (used for object identification, not impacting breadth of painting elements)	Saliency can be used to reduce amount of visual data sonified, but preference on alternating between salient regions in a general sonification of the work better
Instruments/timbre used relevant for time period/geographical location/state of the artist	**Instruments/timbre used relevant for time period/geographical location/state of the artist**

4. CONCLUSION & FUTURE WORK

In this exploratory study, we interviewed experts from the fields of music, sound, visual arts, and sonification to establish design directions for a sonification algorithm that can appropriately transcribe visual artworks into music.

Through the interviews, initial mappings and feedback were gathered, and preliminary guidelines on the sonification of realistic art and abstract art were determined.

Additional expert interviews will be conducted, gathering more data and validating early findings. Future experiments will take the findings of this study into account, adjust the sonification algorithm for recommended techniques and changes experts pointed out, and include additional performance measures found in the study.

5. REFERENCES

[1] B. N. Walker and C. M. Bruce, "ICaD 2013 of biocybernetics A case study of the Accessible Aquarium Project," pp. 39–44, 2013.

[2] G. Dubus and R. Bresin, "A systematic review of mapping strategies for the sonification of physical quantities," *PLoS One*, vol. 8, no. 12, 2013.

[3] M. Jeon, R. J. Winton, J.-B. Yim, C. M. Bruce, and B. N. Walker, "Aquarium fugue: interactive sonification for children and visually impaired audience in informal learning environments," *Proc. 18th Int. Conf. Audit. Disp. (ICAD 2012)*, pp. 246–247, 2012.

[4] S. Cavaco, J. T. Henriques, M. Mengucci, N. Correia, and F. Medeiros, "Color Sonification for the Visually Impaired," *Procedia Technol.*, vol. 9, pp. 1048–1057, 2013.

[5] T. Yoshida, K. M. Kitani, S. Belongie, and K. Schlei, "EdgeSonic : Image Feature Sonification for the Visually Impaired Categories and Subject Descriptors," *Image Rochester NY*, pp. 1–4, 2011.

[6] S. van der Walt *et al.*, "scikit-image: image processing in Python," *PeerJ*, vol. 2, p. e453, 2014.

[7] R. Achantay, S. Hemamiz, F. Estraday, and S. Susstrunky, "Frequency-tuned salient region detection," *IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, no. Ic, p. 1597 { 1604, 2009.

[8] L. Zhang, Y. Gao, R. Ji, Y. Xia, Q. Dai, and X. Li, "Actively learning human gaze shifting paths for semantics-aware photo cropping," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2235–2245, 2014.

[9] I. Fuchs, U. Ansorge, C. Redies, and H. Leder, "Salience in Paintings: Bottom-Up Influences on Eye Fixations," *Cognit. Comput.*, vol. 3, no. 1, pp. 25–36, 2011.

[10] U. Leonards and W. Singer, "Conjunctions of colour, luminance and orientation: The role of colour and luminance contrast on saliency and proximity grouping in texture segregation," *Spat. Vis.*, vol. 13, no. 1, pp. 87–105, 2000.

[11] Y. Bar, N. Levy, and L. Wolf, "Classification of artistic styles using binarized features derived from a deep neural network," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8925, pp. 71–84, 2015.

[12] B. Saleh and A. Elgammal, "Large-scale Classification of Fine-Art Paintings: Learning The Right Metric on The Right Feature," 2015.

[13] R. De Prisco, D. Malandrino, G. Zaccagnino, R. Zaccagnino, and R. Zizza, "A Kind of Bio-inspired Learning of mUsic stylE," in *Computational Intelligence in Music, Sound, Art and Design*, 2017, pp. 97–113.

[14] C. C. Carbon, "Art perception in the museum: How we spend time and space in art exhibitions," *Iperception.*, vol. 8, no. 1, 2017.

[15] S. E. Palmer, K. B. Schloss, Z. Xu, and L. R. Prado-Leon, "Music-color associations are mediated by emotion," *Proc. Natl. Acad. Sci.*, vol. 110, no. 22, pp. 8836–8841, 2013.

[16] T. Tsang and K. B. Schloss, "Associations between Color and Music are Mediated by Emotion and Influenced by Tempo," *Yale Rev. Undergrad. Res. Psychol.*, pp. 82–93, 2010.

[17] N. Rönnberg and J. Löwgren, "Photone: Exploring Modal Synergy in Photographic Images and Music," no. Icad, pp. 73–79, 2018.

[18] A. Polo and X. Sevillano, "Musical Vision: an interactive bio-inspired sonification tool to convert images into music," *J. Multimodal User Interfaces*, no. i, 2018.

[19] R. G. Condorovici, R. Vrânceanu, and C. Vertan, "Saliency Map Retrieval for Artistic Paintings Inspired from Human Understanding," *Proc. SPAMEC*, pp. 101–104, 2011.

[20] B. Stoica, L. Florea, A. Badeanu, A. Racoviteanu, I. Felea, and C. Florea, "Visual saliency analysis in paintings," *ISSCS 2017 - Int. Symp. Signals, Circuits Syst.*, 2017.

[21] B. Manaris, B. Stevens, and A. R. Brown, "JythonMusic: An environment for teaching algorithmic music composition, dynamic coding and musical performativity," *J. Music. Technol. Educ.*, vol. 9, no. 1, pp. 33–56, 2016.