RESEARCH PROPOSAL

Advik Rai
Independent Research II
2024-25

Dr. Alexander Pantelyat and Dr. Kyurim Kang, JHU Center for Music and Medicine, Baltimore

Title: Algorithmic Arias

Introduction and Overview of Research

This research aims to explore the intersection of music and technology by developing a system that personalizes music therapy through Affective Algorithmic Composition (AAC). Music has more pronounced effects on listeners the more it is tailored to them, and the precision of algorithms can make this neurological effect even stronger (Huang & Lin, 2013; Woods et al., 2019). By analyzing real-time emotional data from facial expressions, the system will generate music tailored to an individual's emotional state, offering more precise and adaptive therapy than traditional methods. The goal is to improve therapeutic outcomes, especially for patients with conditions like autism, Alzheimer's, and Parkinson's disease. We hope to create a working prototype of the system and gather preliminary data on its effectiveness with human trials.

Background and Rationale

Music has long been recognized for its therapeutic potential, with plenty research highlighting its impact on emotional and psychological well-being. Music therapy has been successfully used to treat various mental health and neurological conditions, including anxiety, depression, and Alzheimer's disease (Chanda & Levitin, 2013). There is also a growing use of music therapy in clinical settings overall, such as during surgery and in emergency rooms. However, traditional music therapy sessions are often limited in their ability to dynamically adjust to the patient's changing emotional state in real-time, often relying on pre-existing pieces and limited knowledge of the patient. That, and the cost of individual sessions can be prohibitively high, making it difficult to scale the benefits of music therapy for widespread use (Romo & Gifford, 2007).

With advancements in artificial intelligence and computing, there is a growing interest in developing systems for automated music therapy. Much work has been done on that frontier, but the literature draws blanks when it comes to personalized music. Affective Algorithmic Composition (AAC), which uses algorithms to generate music based on emotional inputs, may allow for continuous, real-time adjustments in response to a person's mental state (Williams et al., 2017) and map out exactly which components of music yield measurable outcomes for the patient. Systems that use facial expression recognition (FER) could generate music tailored to an individual's emotional state with this information, creating a unique therapeutic experience – unlike traditional methods that require cumbersome electroencephalograms (EEG) or other invasive measures, AAC with FER can provide immediate feedback and tailored musical compositions without interference or discomfort for the patient (Bahreini et al., 2019). Recent advancements in facial recognition technology have made it possible to read and analyze emotional expressions with high accuracy (Weth et al., 2015) and in speeds in the range of a few

milliseconds (Belaiche et al., 2020): this gives an opportunity to personalize music therapy at a level of precision not previously possible.

Music's power lies in its ability to evoke and regulate emotions, and using it in a personalized context can very greatly amplify its therapeutic effects. Patients with conditions such as autism and Alzheimer's often experience emotional dysregulation and heightened sensitivity to stimuli, making them ideal candidates for this form of therapy. For instance, personalized music interventions have been shown to reduce agitation and improve communication in Alzheimer's patients (Gerdner, 2012) – but while traditional music therapy has been successful, its limitations in real-time personalization may restrict its effectiveness. This project seeks to not only remedy that but further elevate its potency and ease of use – one goal is to create a very user-friendly application that can be set and used by laymen not trained in music theory or programming, etc.

Research Methodology

Research Question:

How can Affective Algorithmic Composition using Facial Expression Recognition technology personalize music therapy in real-time to improve emotional regulation in patients with mental health and neurological conditions?

Hypothesis:

The integration of AAC and FER together will (1) be feasible and (2) allow for highly personalized music therapy experiences by tailoring musical elements to a person's emotional state in real time. This will lead to more effective therapeutic outcomes, such as reduced anxiety and agitation.

Research Design Model:

The study will use a mixed-method design with both quantitative and qualitative data collection, written in a Data Collection notebook. That will allow for a comprehensive analysis of the system's impact by gathering objective biometric data alongside subjective participant feedback.

- Quantitative: Data from the facial expression recognition (FER) systems will be collected. Tools like EEGs or fMRIs, if accessible to the study, could be employed to assess neurological responses to personalized music therapy sessions.
- Qualitative: Post-session interviews and self-reported mood assessments will capture the participants' subjective experiences during and after the therapy sessions.

Variables involved:

- Independent: Affective Algorithmic Composition (AAC) – The algorithm generating music tailored to the detected emotional states of participants in real-time.
- Dependent:
    - Emotional state changes: Measured using FER data (emotional expression categorization, valence, arousal, etc.) and self-reported mood assessments.
    - Biometric responses: Heart rate variability, EEG readings, or other relevant physiological indicators (depending on availability of tools such as EEGs or fMRIs).
    - Self-reported experiences from participants.
- Control:
    - Time of day and environmental factors: These will be controlled to avoid external factors influencing emotional or physiological responses.

Data collection:

- **Facial Expression Recognition (FER) Technology**: This technology will analyze real-time emotional expressions throughout the music therapy sessions. This non-invasive method allows for continuous tracking of emotional changes during therapy.
- **Self-Reported Questionnaires and Interviews**: After each session, participants will complete a mood assessment questionnaire, and selected participants will be interviewed to gain deeper insights into their subjective experience with the AAC system.

Based on my virtual meetings in the summer with Dr. Alex Pantelyat of JHU's Center for Music and Medicine, I learned that conducting human trials with actual hospital patients will be a logistical and legal nightmare. But on the bright side, Dr. Pantelyat mentioned the possibility of actual faculty/members of the Center of Music and Medicine -- well informed individuals who are not ill – taking part in trialing my prototype, should it be created in time. Though not patients, the CMM members are well-versed in the therapeutic applications of music and may provide valuable insights during trials. Their feedback will contribute to the development of the AAC system for future clinical use. They also have special hospital equipment they could use to take specialized measurements like EEGs and fMRIs.

Primary research:
The research will generate new data through human trials – FER will analyze the participants and generate image data, while the AAC component will generate audio data. In my previous literature review, I examined the impact of music therapy on emotional regulation and the emerging role of tech in therapy. This study builds on those findings by applying the theoretical frameworks and insights from existing research into a practical system.

Product Objectives

Product/Outcome:

By the end of this study, I aim to create a prototype system that integrates Affective Algorithmic Composition (AAC) with Facial Expression Recognition (FER) technology for real-time, personalized music therapy. This system will be capable of generating music based on the user's emotional state, detected through facial expressions. The prototype will be accompanied by a detailed research paper outlining the methods, findings, and implications for the future of music therapy. The paper might include a comparison between the biometric and self-reported data collected during the trials.

Target Audience:

The target audience of the product includes:

- Researchers and clinicians in the fields of music therapy, psychology, and neurology, especially those interested in integrating new technology into therapeutic practices.
- Faculty and members of the JHU Center for Music and Medicine, who may trial the system and provide expert feedback.
- Broader audiences in health tech and AI development, who are exploring applications of artificial intelligence in personalized medicine.

One audience group comprises the researchers, and the other would be the people most impacted, in hospital and care settings – seniors who are losing faculties, etc.

Methods of Communication:

- A research paper will be shared with relevant academic journals and stakeholders in the music therapy community.
- A demonstration of the AAC system prototype could be done during a seminar or workshop, likely at JHU's Center for Music and Medicine, to gather feedback from experts. Possibly at the library or even at school.
- A website, a video or even short-form content could be created to explain how the system works and its potential.

Timeline

Date/Complete	Item to be Completed	Details
☐ ##/##/####		
Quarter 1		
~~**9/20/2024**~~	~~Finalize research proposal~~	
	~~Playlist component visible on published website~~	
	Draft and obtain approval for interview questions	
10/1/2024	Contact participants for interviews	Reach out to potential participants from the JHU Center for Music and Medicine and other places
~~**10/10/2024**~~	~~First interview conducted~~	
~~**10/20/2024**~~	~~Second interview conducted~~	Have meetings not quite interviews
~~**11/01/2024**~~	~~Transcribe and analyze meetings and discussions~~	~~Create written transcripts of both interviews for data analysis.~~
11/01/2024	Data collection notebook setup	Establish the format and structure of the data collection notebook to record

		facial expression and mood data.
Quarter 2		
11/29/2024	Final touch-ups of original 2023-24 IR paper	Build on last year's original survey methods and results with new technologies, procedures and processes by utilizing any relevant data and research material.
12/06/2024	Continue working on the above and focus on integrating new Python music model into the existing code	Also document any and all challenges (something I realized I never was quite focusing on until now) which would be immensely beneficial for the methods section
12/13/2024	Consult with Dr. Kang about paper publication, etc. Write more detailed and intuitive explanations for the technology behind the product	Explore options for publishing a methods paper and preprints, potentially on ArXiv, BioArxiv
12/20/2024	Methods paper drafted	Investigate the process of uploading a preprint to ArXiv and others, ensuring drafts adhere to guidelines
1/05/2025	Begin AAC prototype testing	Initiate the first round of testing with the AAC system using participants from JHU, collect initial FER data. (This is very hard)
1/10/2025	Data collection	Enter data in notebook
1/20/2025	Continue data collection	Review FER data and qualitative feedback Uncertainties in testing

1/20/2025	Analyze that data EEG/additional measurements	Coordinate with the JHU team, maybe possibility of EEG or other biometric data collection.
Quarter 3/4:		
1/30/2025	Data collection (more)	Continue testing, collect further data, and note any required system adjustments in the notebook.
2/05/2025	More analysis	See what the data means
2/10/2025	Draft results section	Start drafting the "Results" section of paper based on the first half of collected data.
2/20/2025	Final data collection	Complete all remaining data collection
	Finish data collection notebook	Make sure all data and observations are entered into the notebook and ready for full analysis.
2/15/2025	Final data analysis	Conduct in-depth analysis of collected data, writing final conclusions and p-hacking when necessary (don't do this).
2/15/2025	Synthesis paper draft	First paper draft
3/01/2025	Product draft	
3/15/2025	Finalize synthesis paper	See feedback, revise, etc
4/01/2025	Journaling	See what I wrote in the journaling
5/01/2025	Final productions	Successfully present the completed research project, including findings and

		demonstrations of the functioning product, to the most relevant audiences Preprints and study papers submitted to multiple journals

References:

Bahreini, Kiavash, et al. "A Fuzzy Logic Approach to Reliable Real-time Recognition of Facial Emotions." Multimedia Tools and Applications, vol. 78, no. 14, Feb. 2019, pp. 18943–66. https://doi.org/10.1007/s11042-019-7250-z.

Belaiche, Reda, et al. "Cost-Effective CNNs for Real-Time Micro-Expression Recognition." Applied Sciences, vol. 10, no. 14, July 2020, p. 4959. https://doi.org/10.3390/app10144959.

Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in Cognitive Sciences,* 17(4), 179–193. https://doi.org/10.1016/j.tics.2013.02.007

Gerdner, L. A. (2012). Individualized music for dementia: Evolution and application of evidence-based protocol. World Journal of Psychiatry, 2(2), 26. https://doi.org/10.5498/wjp.v2.i2.26

Huang, C-F. & Lin, E-J. (2013). AN EMOTION-BASED METHOD TO PERFORM ALGORITHMIC COMPOSITION. Proceedings of the 3rd International Conference on Music & Emotion (ICME3), Jyväskylä, Finland, 11th - 15th June 2013. https://jyx.jyu.fi/handle/123456789/41590#

Romo, R., & Gifford, L. (2007). A cost-benefit analysis of music therapy in a home hospice. Nursing economic$, 25(6), 353–358. https://pubmed.ncbi.nlm.nih.gov/18240837/

Weth, Karim, et al. "Investigating Emotional Responses to Self-selected Sad Music via Self-report and Automated Facial Analysis." Musicae Scientiae, vol. 19, no. 4, Sept. 2015, pp. 412–32. https://doi.org/10.1177/1029864915606796.

Williams, D., Kirke, A., Miranda, E., Daly, I., Hwang, F., Weaver, J., & Nasuto, S. (2017, May). Affective Calibration of Musical Featuresets in an Emotionally Intelligent Music Composition System. ACM Transactions on Applied Perception, Volume 14, Issue 3, Article No.17, pp 1–13. https://dl.acm.org/doi/10.1145/3059005

Woods, K.J., Hewett, A., Spencer, A., Morillon, B., & Loui, P. (2019, July). Modulation in background music influences sustained attention. arXiv: Neurons and Cognition. https://www.semanticscholar.org/paper/Modulation-in-background-music-influences-sustained-Woods-Hewett/8cd6d439f59a7587ab0d5fd901407814a5b00e20