Topic: Smart EQ: Personalising Audio with Context-aware AI using Listener Preferences and Psychological Factors

Supervisors: Dr. Charalampos Saitis, Dr. George Fazekas
in collaboration with Yamaha

Abstract: There is evidence that production techniques like equalisation (EQ) can significantly influence a listener's emotional response to music. In professional music production, precise EQ adjustments are used to evoke specific emotional effects. However, in consumer audio, users often struggle to achieve the desired sound quality, leading to a diminished music listening experience and a pervasive sense of frustration. Ideal EQ settings are highly personal and context-dependent, with no universal solution. Many users find it challenging to achieve the sound they desire due to diverse individual preferences and situational contexts. Some prioritise content, while others focus on sound quality. Psychological factors such as introversion versus extroversion and individual sensitivity also influence preferred EQ settings. These preferences fluctuate with mood, environment, and specific listening contexts, adding another layer of complexity.

The goal of this PhD is to create a model predicting ideal EQ settings for individuals based on their unique preferences and contextual factors. This involves conducting experiments with a diverse group of listeners to gather data on EQ preferences across various contexts. Semantic descriptors related to EQ settings will be collected via crowdsourcing, along with psychological traits and contextual information. Using this rich dataset, a predictive model will be developed with state-of-the-art machine learning techniques, incorporating insights from psychological and contextual factors to enhance accuracy. Validation will be an iterative process involving rigorous testing and refinement. Collaboration with Yamaha will aim to integrate the model into a consumer audio system capable of real-time adaptive EQ adjustments, followed by user testing to evaluate and refine the system. Research by Dourou (2022) found that listeners with low arousal levels prefer EQ settings that boost lower frequencies, aligning with observations in music production and highlighting the need for personalised EQ. Stables et al. (2016) provide a framework for understanding terms and processes to achieve desired timbral effects, informing context-aware EQ development. Further inspiration comes from multimodal models for music and language (Manco et al., 2022) and crowdsourcing for semantic descriptors (Cartwright and Pardo, 2014) as well as other works listed in the references below.

The ideal candidate will have an interest in music technology, music emotion research, signal processing, machine learning as well as the latest deep learning techniques. Basic understanding of music theory is useful but not essential.

The PhD is expected to start at the end of September 2024, or as soon as possible afterwards.

Interested candidates should reach out to the supervisors by email: Charalampos Saitis c.saitis@qmul.ac.uk; George Fazekas george.fazekas@qmul.ac.uk.

To apply, complete the online application form which can be found at:
https://mysis.qmul.ac.uk/urd/sits.urd/run/siw_ipp_lgn.login?process=siw_ipp_app&code1=RFQM-G4ZG-09&code2=0006

Application deadline: 26th Aug. 2024

Further details about the application process and requirements can be found on the AIM CDT website: https://www.aim.qmul.ac.uk/apply/

References:

Dourou et. al. (2022) The Influence of Listeners' Mood on Equalization-Based Listening Experience, MDPI Acoustics 2022, 4(3), 746-763

Stables et al. (2016) Semantic Description of Timbral Transformations in Music Production, Proc. 24th ACM international conference on Multimedia, pp. 337–341

Manco et.al. (2022) Contrastive Audio-Language Learning for Music, ISMIR 2022

Cartwright and Pardo (2013) Social-EQ: Crowdsourcing an Equalization Descriptor Map, ISMIR 2013

S. Yang, et. al. (2023a). Examining Emotion Perception Agreement in Live Music Performance, IEEE Trans. on Affective Computing, vol. 14, no. 2, pp. 1442-1460

S. Yang, M. et.al. (2023b). "Do You Hear What I Hear?" in Computer, vol. 56, no. 12, pp. 4-6

Sulem, A., et.al. (2023). Perception of violin performance expression through expressive musical terms. Musicae Scientiae, 27(2), 442-470.

McIntosh, T., et. al. (2023). Affective Conditional Modifiers in Adaptive Video Game Music, Proc. Audio Mostly.

Y. Frachi, G. et.al. (2023). Affective gaming using adaptive speed controlled by biofeedback. In Proc. ICMI.

Hays et.al. (2022). Disembodied Timbres: a Study on Semantically Prompted FM Synthesis, J. Audio Eng. Soc., vol. 70, no. 5, pp. 373–391

Hays et.al. (2022). timbre.fun: A gamified interactive system for crowdsourcing a timbre semantic vocabulary, 24th International Congress on Acoustics

October 24 to 28, 2022 in Gyeongju, Korea

Preniqi et.al. (2023) Soundscapes of morality: Linking music preferences and moral values through lyrics and audio, PLoS ONE 18(11): e0294402.