School of Computing, Engineering, and the Built Environment
Edinburgh Napier University

PHD STUDENT PROJECT

Application instructions:
Detailed instructions are available at :
https://www.napier.ac.uk/research-and-innovation/doctoral-college/how-to-apply

Prospective candidates are encouraged to contact the Director of Studies (see details below) to discuss the project and their suitability for it.

Project details

Supervisory Team:

- DIRECTOR OF STUDY: Dr Rod Selfridge (Email: r.selfridge@napier.ac.uk)
- 2ND SUPERVISOR: Dr Callum Goddard

Subject Group: Applied Informatics

Research Areas: Computer Science - Other

Project Title: DDSP in Procedural Audio Sound Effects

Project description:

This aim of this research is to improve the audio quality of synthesized procedural audio sound effects, optimising parameters through the use of differential digital signal processing (DDSP) techniques. Physically inspired synthesis techniques often used for procedural audio sound effects [1], where basic knowledge of the sound producing process and behaviour modelling are integrated within the synthesis process. Previous research has incorporated deeper knowledge of the physical processes to improve the quality of the sounds synthesised, but it is still possible for listeners to identify synthesised sounds when compared to the recorded samples [2].

DDSP covers a number of techniques where signal processors are integrated within neural networks [3]. Through backpropagation of loss functions, the signal processors can be optimised for specific synthesis models.

One drawback of physically inspired procedural models is that potential critical aspects of the physical process as well as the behaviour model that controls the sound synthesis process can be missed. By training the parameters of the

synthesis models using DDSP, based on pre-recorded samples, it should be possible to capture missing elements of the models, (behaviour etc), and apply these to new synthesis models. Similar separation of the sounds generated by a musical instrument has been carried out in [4] where the performance data is preserved while the timbre.

The use of DDSP and neural networks for the purposes of sound effects is an ongoing area of research. DDSP has more recently be used to generate sound effects [5] or inspired vocalisation synthesis techniques [6], and different neural synthesis approaches to foley have also been explored [7, 8, 9, 10]. This research looks to build on this body of research, using DDSP to control new physically inspired sound effect models, to improve behaviour and plausibility, and ultimately the quality of synthesised sound effects.

References:

[1] Farnell, A. (2010). Designing sound. Mit Press.

[2] Selfridge, R., Moffat, D., Avital, E. J., & Reiss, J. D. (2018). Creating real-time aeroacoustic sound effects using physically informed models. Journal of the Audio Engineering Society, 66(7/8), 594-607.

[3] Hayes, B., Shier, J., Fazekas, G., McPherson, A., & Saitis, C. (2023). A Review of Differentiable Digital Signal Processing for Music & Speech Synthesis. arXiv preprint arXiv:2308.15422

[4] Dai, S., Zhang, Z., & Xia, G. G. (2018). Music style transfer: A position paper. arXiv preprint arXiv:1803.06841

[5] Barahona-Ríos, A., & Collins, T. (2023). NoiseBandNet: Controllable Time-Varying Neural Synthesis of Sound Effects Using Filterbanks. arXiv preprint arXiv:2307.08007.

[6] Hagiwara, M., Cusimano, M., & Liu, J. Y. (2022). Modeling Animal Vocalizations through Synthesizers. arXiv preprint arXiv:2210.10857.

[7] Andreu, S., & Aylagas, M. V. (2022, October). Neural synthesis of sound effects using flow-based deep generative models. In Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (Vol. 18, No. 1, pp. 2-9).

[8] Comunità, M., Phan, H., & Reiss, J. D. (2021). Neural synthesis of footsteps sound effects with generative adversarial networks. arXiv preprint arXiv:2110.09605

[9] Chung, Y., Lee, J., & Nam, J. (2023). Foley sound synthesis in waveform domain with diffusion model. Tech. Rep., June.

[10] Liu, Y., & Jin, C. (2023). Conditional Sound Effects Generation with Regularized WGAN.

Candidate characteristics

Education:

A first-class honours degree, or a distinction at master level, or equivalent achievements in Computer Science, Sound and Music Computing, Music Technology, Sound Design, Acoustics or Artificial Intelligence

Subject knowledge:

- Digital Signal Processing
- Sound synthesis techniques
- Artificial Intelligence

Essential attributes:

- Experience of fundamental digital signal processing techniques
- Basic understanding of sound synthesis principles
- Good written and oral communication skills
- Strong motivation, with evidence of independent research skills relevant to the project
- Good time management
- Experience of audio programming fundamentals
- Experience of python programming language, command line and linux computing environments