

Offre de stage 2024/2025

Titre	Classical GAS - Classical methods for Generative Audio Synthesis
Niveau	M2
Date de début/ fin	Février 2025-Aout 2025
Ville, Pays	Annecy, France
Laboratoire	Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance - LISTIC
Description du sujet	<p>Keywords: Audio synthesis, C/C++ programming, generative models, frugal AI, deep learning,</p> <p>Context: Signal-based audio synthesis methods like the ones described in [1] or [2] have seen a regain of interest recently after the introduction of the Differential Digital Signal Processing (DDSP) package [3]. It uses a neural architecture to produce control parameters for a signal-based synthesis method. The original DDSP implementation performs successfully in various audio synthesis tasks (synthesis, transfer, ...) using a Variational Autoencoder and a Sine plus noise generative model. It offers an expressive generative model without compromising explainability and control, as well as an opportunity for real time sound synthesis.</p> <p>In contrast, diffusion models lack interpretability and intuitive control, need an intensive training and suffer from a long inference time. Yet, the training of a generative model remains challenging due to the unstable gradients produced by signal-based audio synthesis methods. Notably, the success of such a training of a DDSP-inspired generative model is impacted by (i) the signal-based synthesis method, (ii) the control neural architecture, and (iii) the training methodology.</p> <p>In [4], a GAN-based synthesizer (StyleWaveGAN, SWG) has shown promising results on percussion synthesis with improved expressivity in control, opening an opportunity for instrument synthesis using DDSP-inspired models.</p> <p>Project summary : During this internship, a study of a DDSP-inspired audio generative model will be performed. After a comprehensive study of DDSP-inspired architectures on piano and drums datasets [6, 7], the students will extend the work proposed in [4] to more complex signal-based synthesizers. In particular, the scalability of controls to achieve a perceptive control method similar to [5] will be studied. The internship will result in the implementation of the resulting synthesizer as a virtual instrument on a Digital Audio Workstation (DAW) or on an embedded system.</p> <p>Candidate profile : She/he should be enrolled in a M2 or engineer diploma in one or more of the following fields: signal and image processing, computer science, embedded systems. The candidate should have strong programming skills as well as good writing and oral communication skills. A strong interest and/or experience with audio signal processing will be appreciated.</p> <p>Environment The position can be started anytime from February 2025, and the duration is up to 6 months. The candidate will be based in Annecy. This internship will be hosted in the LISTIC laboratory, with regular meetings and exchanges with researchers from the project.</p> <p>Application procedure Send a detailed CV and motivation letter to antoine.lavault@univ-smb.fr and yassine.mhiri@univ-smb.fr</p> <p>References [1] Serra, Xavier. "Musical Sound Modeling with Sinusoids plus Noise." (1997). [2] Chowning, John. "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation." <i>Journal of The Audio Engineering Society</i> 21 (1973): 526-534. [3] Jesse Engel, Lamtharn (Hanoi) Hantrakul, Chenjie Gu, Adam Roberts. "DDSP: Differentiable Digital Signal Processing." <i>International Conference on Learning Representations</i>. 2020. [4] Antoine Lavault. Generative Adversarial Networks for Synthesis and Control of Drum Sounds, Sorbonne Université, 2023. [5] Antoine Lavault, Axel Roebel, Matthieu Voiry. STYLEWAVEGAN: STYLE-BASED SYNTHESIS OF DRUM SOUNDS WITH EXTENSIVE CONTROLS USING GENERATIVE ADVERSARIAL NETWORKS. 19th Sound and Music Computing Conference, 2022</p>

	<p>[6] Gillet, Olivier and Gaël Richard. "ENST-Drums: an extensive audio-visual database for drum signals processing." <i>International Society for Music Information Retrieval Conference</i> (2006).</p> <p>[7]Curtis Hawthorne, Andriy Stasyuk, Adam Roberts, Ian Simon, Cheng-Zhi Anna Huang, Sander Dieleman, Erich Elsen, Jesse Engel, and Douglas Eck. "Enabling Factorized Piano Music Modeling and Generation with the MAESTRO Dataset.", <i>International Conference on Learning Representations</i>, 2019.</p> <p>[8]Renault, Lenny et al. "DDSP-Piano: A Neural Sound Synthesizer Informed by Instrument Knowledge." <i>Journal of the Audio Engineering Society</i> (2023): n. pag.</p>
Compétences requises	Deep Learning and associated frameworks (TF or Torch), basic signal processing. Skills in signal-based audio synthesis methods is appreciated.
Gratification	Selon législation en vigueur
Tuteurs / Contacts	Antoine Lavault – antoine.lavault@univ-smb.fr Yassine Mhiri – yassine.mhiri@univ-smb.fr