

École doctorale EDITE

Informatique Télécommunications et Électronique

Sujet de thèse - campagne 2023

Laboratoire : Équipe Espaces Acoustiques et Cognitifs

Établissement de rattachement : STMS IRCAM – CNRS – Sorbonne Université (UMR 9912)

Titre de la thèse : **Apprentissage machine pour les simulations d'acoustique de salle en réalité virtuelle et augmentée**

Directeur de thèse : Jean-Louis GIAVITTO (DR, CNRS, STMS UMR 9912)

Mail de contact : Jean-Louis.Giavitto@ircam.fr

Codirecteur : Markus NOISTERNIG (CR, IRCAM, STMS UMR 9912)

Collaborations dans le cadre de la thèse :

Lauri SAVIOJA (Pr, Université Aalto, Helsinki, Finlande)

Olivier Warusfel (CR, IRCAM, STMS UMR 9912)

Rattachement à un programme : Thématique IA-BD

Résumé du sujet

Ce projet de thèse vise à explorer l'application des méthodes d'apprentissage profond pour les simulations d'acoustique des salles dans les environnements de réalité virtuelle et augmentée (RV/RA). Les objectifs sont d'étudier la faisabilité de cette utilisation, de concevoir et développer des modèles pour la simulation d'acoustique des salles, d'évaluer et optimiser leurs performances pour les simulations en temps réel, et de les valider par des études utilisateurs dans des environnements RV/RA. La recherche inclura une revue de littérature, la collecte et la préparation de données, la conception et le développement de modèles utilisant différentes architectures de réseaux de neurones, et leur évaluation objective (qualité d'estimation) et subjective (études utilisateurs). Les premières étapes comprendront l'étude de l'interpolation des réponses impulsionnelles spatiales des salles dans des environnements dynamiques où la source ou l'auditeur est en mouvement, ainsi que l'amélioration de la précision des réponses impulsionnelles spatiales mesurées et simulées.

Mots-clés : Apprentissage en profondeur, Intelligence artificielle, Simulation acoustique des salles, Réalité virtuelle et augmentée.

École doctorale EDITE

Informatique Télécommunications et Électronique

Sujet développé

Contexte et Motivation

La réalité virtuelle et augmentée (VR/AR) a le potentiel pour révolutionner la façon dont nous expérimentons et interagissons avec notre environnement. La simulation de l'acoustique des salles est cruciale pour améliorer de manière significative l'expérience immersive des utilisateurs dans ces environnements. Les méthodes classiques de simulation requièrent la résolution d'équations mathématiques complexes qui décrivent la propagation des ondes sonores dans une salle (Savioja & Svensson, 2015 ; Prinn, 2023). Cependant, ces méthodes reposent sur des calculs intensifs qui les rendent actuellement impropres aux applications VR/AR qui nécessitent des performances de simulation en temps réel. Des solutions hybrides ont été proposées, telles que la combinaison des méthodes d'acoustique géométrique (méthode des sources-images, lancer de rayons, radiosité...) avec des modèles de réverbération basés sur le traitement du signal (Noisternig et al., 2008 ; Poirier-Quinot et al., 2017). Au cours de ces dernières années, les techniques d'apprentissage en profondeur ont montré un grand potentiel dans diverses applications audio, telles que la reconnaissance vocale (Aldarmaki et al., 2022), l'amélioration de la parole (Das et al., 2020), la séparation des sources audio et la transcription musicale automatique (Moyisis et al., 2023), ainsi qu'en acoustique (Bianco et al., 2021), en simulations d'acoustique des salles (Yu et al., 2021) et, en plus générale, en audio spatial (Lovedee-Turner, 2021 ; Cobos et al., 2022).

Ce projet de thèse vise à exploiter les techniques d'apprentissage en profondeur pour les simulations de l'acoustique des salles dans les environnements VR/AR, afin de dépasser les limites des méthodes traditionnelles telles que les méthodes géométriques ou les méthodes basées sur des formalismes ondulatoires. Les méthodes d'apprentissage peuvent être exploitées à différents stades de la simulation : en s'intéressant par exemple à la modélisation de phénomènes acoustiques spécifiques tels que la diffraction et la diffusion ou, à un niveau plus avancé, à l'augmentation des réponses impulsionnelles simulées sommairement (à partir de modèles simplifiés) pour en dériver des réponses plus réalistes. Ces méthodes permettront ainsi de concevoir des environnements immersifs plus réalistes, qui simulent de manière plus précise des situations réelles.

La simulation acoustique des salles s'avère utile pour de nombreux domaines pratiques, tels que l'acoustique architecturale et l'ingénierie du son. En utilisant des techniques d'apprentissage profond, nous visons à créer des modèles de simulation d'acoustique des salles en temps réel avec une plus grande précision et une plus grande efficacité que les méthodes de simulation traditionnelles. De plus, les modèles créés pourraient trouver leur application au-delà des environnements VR/AR, comme la téléphonie mains libres et les systèmes de communication embarqués dans les voitures.

Objectives

Les principaux objectifs sont :

1. Concevoir et développer des modèles basés sur les méthodes d'apprentissage profond pour les simulations acoustique des salles en (temps réel) dans des environnements VR/AR.
2. Évaluer les performances des modèles développés et les comparer aux méthodes de simulation traditionnelles en termes de précision et d'efficacité de calcul.
3. Optimiser les modèles développés pour les simulations en temps réel.
4. Valider les modèles développés à travers des études d'utilisateurs dans des environnements VR/AR.

École doctorale EDITE

Informatique Télécommunications et Électronique

Méthodologie

Le projet proposé se décomposera selon les étapes suivantes :

1. Revue de la littérature : une revue détaillée de la littérature existante relative aux simulations en temps réel de l'acoustique des salles et aux techniques d'apprentissage profond pour les simulations sera effectuée. A titre d'exemple : Lovedee-Turner, 2021 ; Yu et al., 2021 ; Cobos et al., 2022.
2. Collecte et préparation des données : il s'agit de collecter un grand ensemble de données de simulations acoustiques des salles pour entraîner et tester les modèles d'apprentissage profond. Ces données incluront des réponses impulsionnelles spatiales de la salle de concert à acoustique variable de l'Ircam, simulées à l'aide des méthodes FDTD (Tang et al., 2022) et calibrées avec des mesures acoustiques. L'ensemble des données couvrira différentes conditions acoustiques (en changeant notamment la taille et la forme de la salle ainsi que des matériaux de surface) assurant ainsi une représentativité précise et réaliste de l'acoustique pour entraîner et valider les modèles développés.
3. Conception et développement des modèles basés sur l'apprentissage profond en utilisant diverses architectures de réseaux neuronaux telles que les réseaux neuronaux convolutionnels/déconvolutionnels (CNN/DNN). Ces modèles intégreront les récentes avancées des techniques d'apprentissage profond telles que les réseaux antagonistes génératifs (GAN) (Mukherjee et al., 2022) et les réseaux de convolution de graphes (GCN) (Zhang et al., 2019).
4. Évaluation des modèles : les modèles développés seront évalués à l'aide de diverses métriques telles que l'erreur quadratique moyenne (MSE) et le rapport signal sur bruit (SNR). Les résultats seront comparés aux simulations traditionnelles en termes de précision et d'efficacité de calcul.
5. Optimisation des modèles : les modèles développés seront optimisés pour les simulations en temps réel en réduisant leur complexité de calcul et leur utilisation de mémoire. Les techniques d'optimisation pertinentes incluent la taille des modèles et la quantification (Liang et al., 2021) et la compression de modèles (Li et al., 2023).
6. Étude utilisateur : une étude utilisateur sera réalisée dans un environnement VR afin de valider les modèles développés. Cette étude inclura des métriques d'évaluation subjectives telles que la qualité audio perçue et la sensation de *présence*, ainsi que des métriques objectives telles que la précision de localisation (Brinkman & Weinzierl, 2023 ; de Villiers Bosman et al., 2023).

Résultats attendus

Les résultats attendus de ce projet sont les suivants :

1. Un modèle utilisant l'apprentissage profond pour les simulations acoustiques des salles dans les environnements de réalité virtuelle.
2. Une comparaison entre le modèle développé et les modèles traditionnelles de simulations en termes d'exactitude et d'efficacité de calcul.
3. Un modèle optimisé pour les simulations acoustiques des salles en temps réel.
4. Un modèle validé par une étude utilisateur dans des environnements VR.

Profil recherché

La candidate ou le candidat devra faire preuve d'une grande curiosité scientifique, de dynamisme, créativité et de compétences en résolution de problèmes complexes. Elle ou il doit être capable de travailler de manière autonome au sein d'une équipe multidisciplinaire. Il ou elle possèdera un diplôme de Master en Traitement du Signal Numérique audio ou en Intelligence Artificielle. Il n'est pas attendu que la candidate ou le candidat possède l'ensemble des compétences pour ce sujet multidisciplinaire, toutefois un intérêt particulier pour

École doctorale EDITE

Informatique Télécommunications et Électronique

l'acoustique des salles et la perception est requis ainsi qu'une forte capacité pour appréhender ce nouveau domaine de recherche.

Bibliographie

1. Aldarmaki, H., Ullah A., Ram, S., Zaki, N., Unsupervised Automatic Speech Recognition: A review, *Speech Communication*, 139 (2022), 76-91, <https://doi.org/10.1016/j.specom.2022.02.005>
2. Michael J. Bianco, Peter Gerstoft, James Traer, Emma Ozanich, Marie A. Roch, Sharon Gannot, and Charles-Alban Deledalle, "Machine learning in acoustics: Theory and applications", *The Journal of the Acoustical Society of America* 146, 3590-3628 (2019) <https://doi.org/10.1121/1.5133944>
3. Brinkmann, F., Weinzierl, S. (2023). Audio Quality Assessment for Virtual Reality. In: Geronazzo, M., Serafin, S. (eds) *Sonic Interactions in Virtual Environments*. Human-Computer Interaction Series. Springer, Cham. https://doi.org/10.1007/978-3-031-04021-4_5
4. Cobos, M., Ahrens, J., Kowalczyk, K. *et al.* An overview of machine learning and other data-based methods for spatial audio capture, processing, and reproduction. *J Audio Speech Music Proc.*, 10 (2022). <https://doi.org/10.1186/s13636-022-00242-x>
5. Das, N., Chakraborty, S., Chaki, J., Padhy, N., Dey, N., Fundamentals, present and future perspectives of speech enhancement. *Int. J. Speech Technol.* 21:4, (2021), 883–901, <https://doi.org/10.1007/s10772-020-09674-2>
6. de Villiers Bosman, I., Buruk, O., Jørgensen, K., Hamari, J. (2023) The effect of audio on the experience in virtual reality: a scoping review, *Behaviour & Information Technology*, <https://doi.org/10.1080/0144929X.2022.2158371>
7. Li Z, Li H, Meng L. Model Compression for Deep Neural Networks: A Survey. *Computers*. 2023; 12(3):60. <https://doi.org/10.3390/computers12030060>
8. Liang, T., Glossner, J., Wang, L., Shi, S., Zhang, X., Pruning and quantization for deep neural network acceleration: A survey, *Neurocomputing* (2021), 461:21, 370-403, <https://doi.org/10.1016/j.neucom.2021.07.045>
9. Lovedee-Turner, M., Application of Machine Learning for the Spatial Analysis of Binaural Room Impulse Responses. *Applied Sciences Special Edition on Sound and Music Computing*, 8 (2018), <https://doi.org/10.3390/app8010105>
10. Mukherjee, S., Agarwala, A., Agarwala, R. (2022). A Review on Generative Adversarial Networks. In: Satyanarayana, C., Gao, XZ., Ting, CY., Muppalaneni, N.B. (eds) *Machine Learning and Internet of Things for Societal Issues*. Advanced Technologies and Societal Change. Springer, Singapore. https://doi.org/10.1007/978-981-16-5090-1_1
11. Morgenstern, H., Rafaely, B. & Noisternig, M. Design framework for spherical microphone and loudspeaker arrays in a multiple-input multiple-output system. *J Acoust Soc Am* 141, 2024–2038 (2017). <https://doi.org/10.1121/1.4978660>
12. Moysis, L., Iliadis, L. A., Sotiroudis, S. P., Boursianis, A. D., Papadopoulou, M. S., Kokkinidis, K-I. D., Volos, C., Sarigiannidis, P., Nikolaidis, S., Goudos, S. K., "Music Deep Learning: Deep Learning Methods for Music Signal Processing—A Review of the State-of-the-Art," in *IEEE Access*, vol. 11, pp. 17031-17052, 2023, <https://doi.org/10.1109/ACCESS.2023.3244620>
13. Noisternig, M., Katz, B. F. G., Siltanen, S., & Savioja, L. (2008). Framework for Real-Time Auralization in Architectural Acoustics. *Acta Acustica united with Acustica*, 94(6), 1000-1015 <https://doi.org/10.3813/AAA.918116>
14. Poirier-Quinot, D., Noisternig, M. & Katz, B. F. EVERTims: Open source framework for real-time auralization in VR. in 1–5 (Audio Mostly, 2017). <https://doi.org/10.1145/3123514.3123559>
15. Prinn AG. A Review of Finite Element Methods for Room Acoustics. *Acoustics*. 2023; 5(2):367-395. <https://doi.org/10.3390/acoustics5020022>
16. Savioja, L., and Svensson, U. P. Overview of geometrical room acoustic modeling techniques. *J Acoust Soc Am* 138, 708–730 (2015), <https://doi.org/10.1121/1.4926438>
17. Tang, Z., Aralikatti, R., Ratnarajah, A.J., and Manocha, D., "GWA: A Large High-Quality Acoustic Dataset for Audio Processing", *ACM SIGGRAPH Conference Proceedings* (2022), Association for Computing Machinery, New York, NY, USA, Article 36, 1–9. <https://doi.org/10.1145/3528233.3530731>
18. Yu, W., & Kleijn, W. B. (2021). Room Acoustical Parameter Estimation from Room Impulse Responses Using Deep Neural Networks. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 29, 436 - 447. <https://doi.org/10.1109/TASLP.2020.3043115>
19. Zhang, S., Tong, H., Xu, J. et al. Graph convolutional networks: a comprehensive review. *Comput Soc Netw* 6, 11 (2019). <https://doi.org/10.1186/s40649-019-0069-y>

École doctorale EDITE Informatique Télécommunications et Électronique

PhD Thesis Offer -2023

Department: Acoustics and Cognitive Spaces Group

Hosting Institution: STMS IRCAM – CNRS – Sorbonne Université (UMR 9912)

Thesis Title: **Machine Learning for Room Acoustics Simulations in Virtual and Augmented Reality**

Doctoral supervisor: Jean-Louis GIAVITTO (DR, CNRS, STMS UMR 9912)

Email-Address: Jean-Louis.Giavitto@ircam.fr

Co-supervisor: Markus NOISTERNIG (CR, IRCAM, STMS UMR 9912)

Collaborations within the scope of the thesis:

Lauri SAVIOJA (Professor, Université Aalto, Helsinki, Finland)

Benoît ALARY (CR, IRCAM, STMS UMR 9912)

Olivier WARUSFEL (CR, IRCAM, STMS UMR 9912)

Affiliation with a program: IA-BD

Summary of the topic

In this proposed PhD project, we aim to explore the application of deep learning methods for room acoustics simulations in virtual and augmented reality (VR/AR) environments. The objectives of this project are to study the feasibility of using deep learning for room acoustics simulations, design and develop models that can simulate room acoustics, evaluate the performance of the developed models, optimize the models for real-time simulations, and validate them through user studies in VR/AR environments. The project will involve a literature review, data collection and preparation, design and development of models using different neural network architectures, and validation of the models through objective evaluation and user studies. The research will begin by studying the interpolation of spatial room impulse responses in dynamic environments where either the source or listener is in motion, as well as enhancing the accuracy of measured and simulated spatial room impulse responses.

Key words: Deep Learning, Artificial Intelligence, Room Acoustics, Virtual Reality, Augmented Reality

École doctorale EDITE

Informatique Télécommunications et Électronique

Thesis Topic

Context and Motivation

Virtual and augmented reality (VR/AR) has the potential to revolutionize the way we experience and interact with our environments. Room acoustics simulations are an important aspect of VR/AR environments, as they can significantly enhance the immersive experience of users. Traditional methods for room acoustics simulations involve solving complex mathematical equations that describe the behavior of sound waves in a room (Savioja & Svensson, 2015; Prinn, 2023). Real-time simulations in VR applications cannot use such computationally intensive methods. Hybrid solutions, such as combining signal-processing-based reverberation models with geometrical acoustics simulations, have been suggested (Noisternig et al., 2008; Poirier-Quinot et al., 2017). In recent years, deep learning techniques have shown great potential for various audio applications, including speech recognition (Aldarmaki et al., 2022), speech enhancement (Das et al., 2020), audio source separation and automatic music transcription (Moyisis et al., 2023), and have been applied to acoustics (Bianco et al., 2021), spatial audio (Lovedee-Turner, 2021; Cobos et al., 2022) and room acoustics simulations (Yu et al, 2021).

In this context, this proposed PhD project aims to explore the use of deep learning techniques for real-time room acoustics simulations in VR/AR. The main motivation behind this project is to address the limitations of traditional methods for room acoustics simulations (i.e., geometrical methods or wave-based methods), and to design and develop models that can simulate room acoustics in real-time. Deep learning can be used in various stages of acoustic simulations, including the low-level modeling of specific phenomena such as diffraction and diffusion. At a higher level, deep learning can augment room impulse responses estimated by basic room simulation methods in order to obtain a more realistic response. Therefore, the successful accomplishment of this project will enable the creation of VR/AR environments that are more realistic and immersive, providing a better simulation of real-world scenarios.

Room acoustics simulations can have numerous practical applications, such as in architectural acoustics and audio engineering. By using deep learning techniques, we hope to create models that can perform real-time simulations with greater accuracy and efficiency than traditional methods. Additionally, the development of these models could have implications beyond VR/AR, such as hands-free telephony and car in-vehicle communication systems.

Objectives

However, creating realistic room acoustics simulations in real-time is a challenging task that requires significant computational resources. In the proposed PhD project, we aim to explore the use of deep learning techniques for real-time room acoustics simulations in VR/AR environments.

The main objectives of this PhD project are as follows:

1. To design and develop deep learning-based models for real-time room acoustics simulations in VR/AR environments.
2. To evaluate the performance of the developed models and compare them with traditional Finite Difference Time Domain (FDTD) simulations in terms of accuracy and computational efficiency.
3. To optimize the developed models for real-time simulations.
4. To validate the developed models through user studies in VR environments.

École doctorale EDITE

Informatique Télécommunications et Électronique

Methodology

The proposed project will be divided into the following stages:

1. **Literature Review:** A comprehensive review of existing literature related to real-time room acoustics simulations and deep learning techniques for simulations will be conducted. This includes examining recent papers such as (Lovedee-Turner, 2021; Yu et al, 2021; Cobos et al., 2022).
2. **Data Collection and Preparation:** We will collect an extensive dataset of room acoustics simulations to train and test our deep learning models. The dataset will consist of spatial room impulse response data for IRCAM's variable acoustics concert hall, generated using wave-based/FDTD methods (Tang et al., 2022), and calibrated with room acoustics measurements. This high-quality dataset will cover various acoustic conditions, including room sizes, shapes, and surface materials, ensuring accurate and representative real-world acoustics for model training and validation.
3. **Model Design and Development:** Deep learning-based models will be designed and developed using various neural network architectures such as Convolutional/Deconvolutional Neural Networks (CNNs/DNNs). These models will incorporate recent advancements in deep learning techniques such as Generative Adversarial Networks (GANs) (Mukherjee et al., 2022) and graph convolutional networks (GCNs) (Zhang et al., 2019).
4. **Model Evaluation:** The developed models will be evaluated using various metrics such as Mean Squared Error (MSE) and Signal-to-Noise Ratio (SNR). The results will be compared with traditional simulations in terms of accuracy and computational efficiency.
5. **Model Optimization:** The developed models will be optimized for real-time simulations by reducing their computational complexity and memory usage. Relevant optimization techniques include pruning and quantization (Liang et al., 2021) and model compression (Li et al., 2023).
6. **User Study:** A user study will be conducted in a VR environment to validate the developed models. This will include subjective evaluation metrics such as perceived audio quality and presence, as well as objective metrics such as localization accuracy (Brinkman & Weinzierl, 2023; de Villiers Bosman et al, 2023).

Expected Outcomes

The expected outcomes of this proposed PhD project are as follows:

1. A deep learning-based model for real-time room acoustics simulations in VR environments.
2. A comparative analysis of the developed model with traditional FDTD simulations in terms of accuracy and computational efficiency.
3. An optimized model for real-time simulations.
4. A validated model through a user study in VR environments.

Conclusion

In conclusion, this proposed PhD project aims to investigate the use of deep learning techniques for real-time room acoustics simulations in virtual and augmented reality environments. By exploring the potential of machine learning to enhance the accuracy and efficiency of room acoustics simulations, this project seeks to contribute to the development of immersive and realistic VR and AR experiences. The findings of this research could have important implications for various applications, including architectural design, concert hall acoustics, and spatial audio for gaming and entertainment. Overall, this PhD project presents an exciting opportunity to advance the state of the art in the field of acoustic simulations and to explore the possibilities of machine learning in virtual and augmented reality.

École doctorale EDITE

Informatique Télécommunications et Électronique

Candidate Profile

The ideal candidate for this PhD project must possess a master's degree in electrical engineering, computer science, or a related field and have a strong background in audio signal processing, acoustics, and machine learning. They should be proficient in programming with Python and have experience working with deep learning frameworks such as TensorFlow or PyTorch. Furthermore, they should demonstrate the ability to work independently, take initiative, and function well as part of a team. While not expected to possess all the necessary skills for this multidisciplinary topic, the candidate should have effective communication skills, strong analytical abilities, and excellent problem-solving skills.

Literature

1. Aldarmaki, H., Ullah A., Ram, S., Zaki, N., Unsupervised Automatic Speech Recognition: A review, *Speech Communication*, 139 (2022), 76-91, <https://doi.org/10.1016/j.specom.2022.02.005>
2. Michael J. Bianco, Peter Gerstoft, James Traer, Emma Ozanich, Marie A. Roch, Sharon Gannot, and Charles-Alban Deledalle, "Machine learning in acoustics: Theory and applications", *The Journal of the Acoustical Society of America* 146, 3590-3628 (2019) <https://doi.org/10.1121/1.5133944>
3. Brinkmann, F., Weinzierl, S. (2023). Audio Quality Assessment for Virtual Reality. In: Geronazzo, M., Serafin, S. (eds) *Sonic Interactions in Virtual Environments*. Human-Computer Interaction Series. Springer, Cham. https://doi.org/10.1007/978-3-031-04021-4_5
4. Cobos, M., Ahrens, J., Kowalczyk, K. *et al.* An overview of machine learning and other data-based methods for spatial audio capture, processing, and reproduction. *J Audio Speech Music Proc.*, 10 (2022). <https://doi.org/10.1186/s13636-022-00242-x>
5. Das, N., Chakraborty, S., Chaki, J., Padhy, N., Dey, N., Fundamentals, present and future perspectives of speech enhancement. *Int. J. Speech Technol.* 21:4, (2021), 883–901, <https://doi.org/10.1007/s10772-020-09674-2>
6. de Villiers Bosman, I., Buruk, O., Jørgensen, K, Hamari, J, (2023) The effect of audio on the experience in virtual reality: a scoping review, *Behaviour & Information Technology*, <https://doi.org/10.1080/0144929X.2022.2158371>
7. Li Z, Li H, Meng L. Model Compression for Deep Neural Networks: A Survey. *Computers*. 2023; 12(3):60. <https://doi.org/10.3390/computers12030060>
8. Liang, T., Glossner, J., Wang, L., Shi, S., Zhang, X., Pruning and quantization for deep neural network acceleration: A survey, *Neurocomputing* (2021), 461:21, 370-403, <https://doi.org/10.1016/j.neucom.2021.07.045>
9. Lovedee-Turner, M., Application of Machine Learning for the Spatial Analysis of Binaural Room Impulse Responses. *Applied Sciences Special Edition on Sound and Music Computing*, 8 (2018), <https://doi.org/10.3390/app8010105>
10. Mukherjee, S., Agarwala, A., Agarwala, R. (2022). A Review on Generative Adversarial Networks. In: Satyanarayana, C., Gao, XZ., Ting, CY., Muppalaneni, N.B. (eds) *Machine Learning and Internet of Things for Societal Issues*. Advanced Technologies and Societal Change. Springer, Singapore. https://doi.org/10.1007/978-981-16-5090-1_1
11. Morgenstern, H., Rafaely, B. & Noisternig, M. Design framework for spherical microphone and loudspeaker arrays in a multiple-input multiple-output system. *J Acoust Soc Am* 141, 2024–2038 (2017). <https://doi.org/10.1121/1.4978660>
12. Moysis, L., Iliadis, L. A., Sotiroudis, S. P., Boursianis, A. D., Papadopoulou, M. S., Kokkinidis, K-I. D., Volos, C., Sarigiannidis, P., Nikolaidis, S., Goudos, S. K., "Music Deep Learning: Deep Learning Methods for Music Signal Processing—A Review of the State-of-the-Art," in *IEEE Access*, vol. 11, pp. 17031-17052, 2023, <https://doi.org/10.1109/ACCESS.2023.3244620>
13. Noisternig, M., Katz, B. F. G., Siltanen, S., & Savioja, L. (2008). Framework for Real-Time Auralization in Architectural Acoustics. *Acta Acustica united with Acustica*, 94(6), 1000-1015 <https://doi.org/10.3813/AAA.918116>
14. Poirier-Quinot, D., Noisternig, M. & Katz, B. F. EVERTims: Open source framework for real-time auralization in VR. in 1–5 (Audio Mostly, 2017). <https://doi.org/10.1145/3123514.3123559>
15. Prinn AG. A Review of Finite Element Methods for Room Acoustics. *Acoustics*. 2023; 5(2):367-395. <https://doi.org/10.3390/acoustics5020022>
16. Savioja, L., and Svensson, U. P. Overview of geometrical room acoustic modeling techniques. *J Acoust Soc Am* 138, 708–730 (2015), <https://doi.org/10.1121/1.4926438>

École doctorale EDITE Informatique Télécommunications et Électronique

17. Tang, Z., Aralikatti, R., Ratnarajah, A.J., and Manocha, D., "GWA: A Large High-Quality Acoustic Dataset for Audio Processing", ACM SIGGRAPH Conference Proceedings (2022), Association for Computing Machinery, New York, NY, USA, Article 36, 1–9. <https://doi.org/10.1145/3528233.3530731>
18. Yu, W., & Kleijn, W. B. (2021). Room Acoustical Parameter Estimation from Room Impulse Responses Using Deep Neural Networks. IEEE/ACM Transactions on Audio Speech and Language Processing, 29, 436 - 447. <https://doi.org/10.1109/TASLP.2020.3043115>
19. Zhang, S., Tong, H., Xu, J. et al. Graph convolutional networks: a comprehensive review. Comput Soc Netw 6, 11 (2019). <https://doi.org/10.1186/s40649-019-0069-y>