



Laboratoire
Sciences et technologies
de la musique et du son
(STMS) UMR 9912

ircam
Centre
Pompidou



 SORBONNE
UNIVERSITÉ



 MINISTÈRE
DE LA CULTURE
*Liberté
Égalité
Fraternité*

Faits marquants de la période en cours

◆ L'excellence du laboratoire s'est traduite depuis 2021 par l'obtention de plusieurs prix

- ◆ *Best Paper award SMC-2022*, décerné à **Claire Richards** (doctorante PDS), 1^{re} auteure de l'article « Designing the balance between sound and touch: Methods for multimodal composition »;
- ◆ *Prix Jeune Chercheuse/Chercheur de la Journée Science et Musique 2021*, à **Benjamin Matuszewski**, (ISMM), pour ses travaux « Systèmes musicaux distribués et technologies Web : vers de nouveaux paradigmes d'interactions musicales collectives » (prix décerné en partenariat avec l'AFIM);
- ◆ *Prix Sacem du meilleur documentaire musical au Festival de Lussas 2021* pour le film *À la recherche de la musique de l'Antiquité* d'Arte auquel les chercheurs **Olivier Warusfel** (EAC) et **René Caussé** (S3AM) ont participé. Le film a été également nommé au festival Pariscience 2021 pour la sélection Grand Public - Compétition Science Télévision.

◆ Organisation et collaboration à de nombreux événements publics

- ◆ **Forum Vertigo**: « Les nouvelles modalités de la présence artistique dans la crise Covid », 3 mars 2021 : cette édition s'est tenue en ligne depuis l'Ircam autour de tables rondes sur les thèmes « Création collaborative » et « Le spectacle vivant à l'ère de la pandémie »;
- ◆ **Connecter les mondes**, 9-11 mars 2022 : organisé dans le cadre du cycle « Mutations/Créations » en lien avec le Centre Pompidou et son exposition « Réseaux-Mondes », cette rencontre a réuni une quarantaine d'artistes et chercheurs autour des multiples facettes des liens entre réseaux et création artistique;
- ◆ **Créations artistiques**: opéra *Like Flesh* (2021) de **Sivan Eldar** qui intègre des technologies de l'équipe RepMus et a reçu le Prix Fedora opéra; *Music of choices* (2021) d'**Alexandros Markeas** et *Ex Machina* (2022) de **Steve Lehman** et **Frédéric Maurin** par l'Orchestre National de Jazz. Utilisation de la technologie de générativité interactive DYCI2 développée par l'équipe RepMus; *Convergence* (2021) d'Alexander Schubert intégrant la technologie RAVE à base de deep learning pour la conversion de timbres du projet ACID-TEAM de l'équipe RepMus et qui a obtenu le *Prix Golden Nica d'Ars Electronica*;
- ◆ Inauguration en 2021 de la Philharmonie des enfants. L'un de ses dispositifs phares, *Maestra, maestro !*, développé par l'équipe ISMM avec Ircam Amplify met des enfants en situation de direction d'orchestre;
- ◆ **Ateliers Forum** réunissant la communauté internationale des utilisateurs des logiciels de l'Ircam à Paris (2021 et 2022), Montréal (2021) et New York (2022);

- ◆ Journées finales du projet **ELEMENT**, 13-14 octobre 2022;
- ◆ « **Hôtel du temps** » de **Thierry Ardisson** sur France 2 (2022): interview de Dalida post-mortem avec une synthèse vocale produite par l'équipe Analyse/synthèse en lien avec Ircam Amplify;
- ◆ **Colloque « 1722–2022: trois siècles du Traité de Jean-Philippe Rameau. La musique-science devant la question de l'harmonique »**, Paris, 7-9 novembre 2022 – coorganisé avec l'IreMus;
- ◆ **Séminaires Sheng** (orgue à bouche), avec le Collegium Musicæ (Sorbonne Université);
- ◆ Plusieurs webinaires dont celui organisé conjointement avec l'Ambassade de France à Atlanta et nos partenaires américains (CNMAT à l'université de Berkeley, GeorgiaTech d'Atlanta et l'université de New York) dans le cadre de la semaine culturelle France-Atlanta.

◆ Édition

- ◆ G. Aperghis et N. Donin (2022), *Conversation imagée (2019-2021)*. Éditions de la Philharmonie;
- ◆ H. Scurto, B.V. Kerrebroeck, B. Caramiaux, F. Bevilacqua (2021). « **Designing deep reinforcement learning for human parameter exploration** », *ACM Transactions on Computer-Human Interaction (TOCHI)*, 28(1), 1-35;
- ◆ P. Lascabettes, C. Agon, M. Andreatta, I. Bloch (2022), « **Computational Analysis of Musical Structures based on Morphological Filters** », in *Computer Science book series (LNAI, volume 1326)*.

◆ Actions internationales

- ◆ Accord de coopération (2020-2027) avec le C-LAB pour la **constitution du Taiwan Sound Lab**;
- ◆ Démarrage du **réseau international ACTOR** sur l'orchestration avec l'université McGill pour lequel l'Ircam est partenaire principal avec 19 institutions;
- ◆ Convention de résidence artistique avec **GeorgiaTech, Atlanta**.

◆ Conseil scientifique

Le conseil scientifique s'est thématiqué en 2021 sur deux axes forts : Spatialisation ; Son/Musique et Santé.

La recherche musicale et ses ressources artistiques ont intégré l'UMR.

Projets dans l'UMR STMS 2021 — 2023

ERC (Conseil européen de la recherche) et autres projets européens

ERC REACH Raising Co-Creativity in Cyber-Human Musicianship, European Research Council Advanced Grant, porté par Gérard Assayag, 2021-2025. Le projet REACH vise à comprendre, modéliser et développer la co-créativité musicale entre les humains et les machines par le biais d'interactions improvisées, permettant aux musiciens de tout niveau de formation de développer leurs compétences et d'accroître leur potentiel créatif individuel et social.

ERC COSMOS Computational Shaping and Modeling of Musical, European Research Council Advanced Grant, porté par Elaine Chew, 2019-2024. COSMOS propose un programme intégré de recherche complétant la compréhension de l'acte interprétatif chez le musicien et de la perception de cette interprétation chez l'auditeur.

ERC POC HEART.FM Maximizing the Therapeutic Potential of Music through Tailored Therapy with Physiological Feedback in Cardiovascular Disease, European Research Council Proof of Concept, porté par Elaine Chew, 2020-2022. Heart.FM est un projet « preuve de concept » issu de Cosmos. Il vise à développer des applications de thérapie musicale personnalisée, informées par un retour d'information physiologique.

ERC POC ACTIVATE Augmenting the Value of Conversations with Voice Transformations, 2020-2021 porté par Jean-Julien Aucouturier. Ce projet « preuve de concept » a pour objectif de valoriser des technologies de transformation temps réel de voix issues du projet ERC CREAM (2014-2019), dans des applications de contrôle des caractéristiques émotionnelles dans la conversation parlée.

Horizon 2020 AI4Media constituant un laboratoire d'excellence distribué sur les applications des techniques de deep learning au secteur des médias et dans lequel l'équipe AnaSyn développe une nouvelle génération de synthétiseur sonore.

Horizon Europe DAFNE+ proposant, en lien avec le Forum Ircam, une plateforme à base de blockchain, NFT et sous la forme de DAO (distributed autonomous organization) destinée à la création, la diffusion et la monétisation de contenus artistiques.

Participation de STMS au projet européen **SOUND.AI** (coordonné par Sorbonne Université) et organisant un réseau européen de doctorants autour des techniques d'intelligence artificielle.

7 nouveaux projets ANR

Exvoices – Lecture Virtuelle d'Histoires pour Enfants : Aspects Expressifs et Cognitifs de la Voix de Synthèse (2022-2026), équipe AnaSyn

BRUEL – ElaBoRation d'Une méthodologie d'EvaLuation des systèmes d'identification par la voix (2022-2026), équipe AnaSyn

Aqua-Rius – Analyse de la Qualité Audio pour Représenter, Indexer et Unifier les Signaux (2022-2026), équipe AnaSyn

DOTS – Distributed Musical Things for Collective Interactions (2022-2025), JCJC équipe ISMM

INSPECTSYN – Speech-in-noise INTElligibility Deficits : Designing Psychophysical and Electrophysiological Markers of SPECTral-Shape Coding Sensitive to Cochlear SYNaptopathy (2022-2025), JCJC équipe PDS

DeTox – Lutte contre les vidéos hyper-truquées de personnalités françaises (2022-2024), équipe AnaSyn

AVATARS – Artificial Voice production: control of bio-inspired port-HamilToniAn numerical and mechatronic modelS (2023-2027), équipe S3AM

Les chiffres clés

7

équipes de recherche sur les différentes disciplines scientifiques en rapport avec la musique : acoustique, traitement du signal, modélisation, contrôle, informatique, psychologie, musicologie.

105

collaborateurs (chercheurs, ingénieurs, techniciens, administratifs), 30 stagiaires et chercheurs invités par an.

140

publications par an, dont 60 dans des revues et chapitres d'ouvrages.

40

collaborations annuelles avec des laboratoires et organismes publics internationaux.

20

environnements logiciels développés et diffusés auprès de **30 000** utilisateurs professionnels dans le monde via le Forum Ircam (groupe d'utilisateurs).

Le mot de la direction

Les activités de recherche accueillies à l'Ircam s'inscrivent dans le cadre de l'unité mixte de recherche (UMR 9912) *Sciences et technologies de la musique et du son* (STMS), associant aux côtés de l'Ircam le CNRS, Sorbonne Université et le ministère de la Culture.

Le contrat quinquennal 2019-2023 prolongé d'un an suite à la crise COVID vise à répondre à plusieurs défis : l'intégration de la recherche artistique dans les structures universitaires et la montée en puissance des thématiques art-science (avec, par exemple, la création des doctorats en arts) ; la lisibilité et l'accroissement de l'attractivité de l'unité au sein de l'éco-système de recherche et d'innovation français et international.

L'activité de recherche à STMS est portée par sept équipes et se distribue sur trois axes structurants : l'atelier du son, le corps musicien et les dynamiques créatives.

L'UMR s'appuie pour les aspects contractuels et d'innovation sur le département IMR de l'Ircam (Innovation et Moyens de la recherche, sous la responsabilité d'Hugues Vinet) avec en particulier la création en 2019 par l'Ircam d'un nouvel instrument de valorisation, la filiale Ircam Amplify, qui permet une diffusion plus large des technologies issues du laboratoire.

Enfin, j'ai le plaisir de souligner que depuis le début du contrat quinquennal, STMS a participé à 16 projets européens dont 5 ERC, ainsi que 18 projets ANR, dont 53% en tant que coordinateur.

Brigitte d'Andréa-Novel

Directrice de l'UMR 9912, octobre 2022

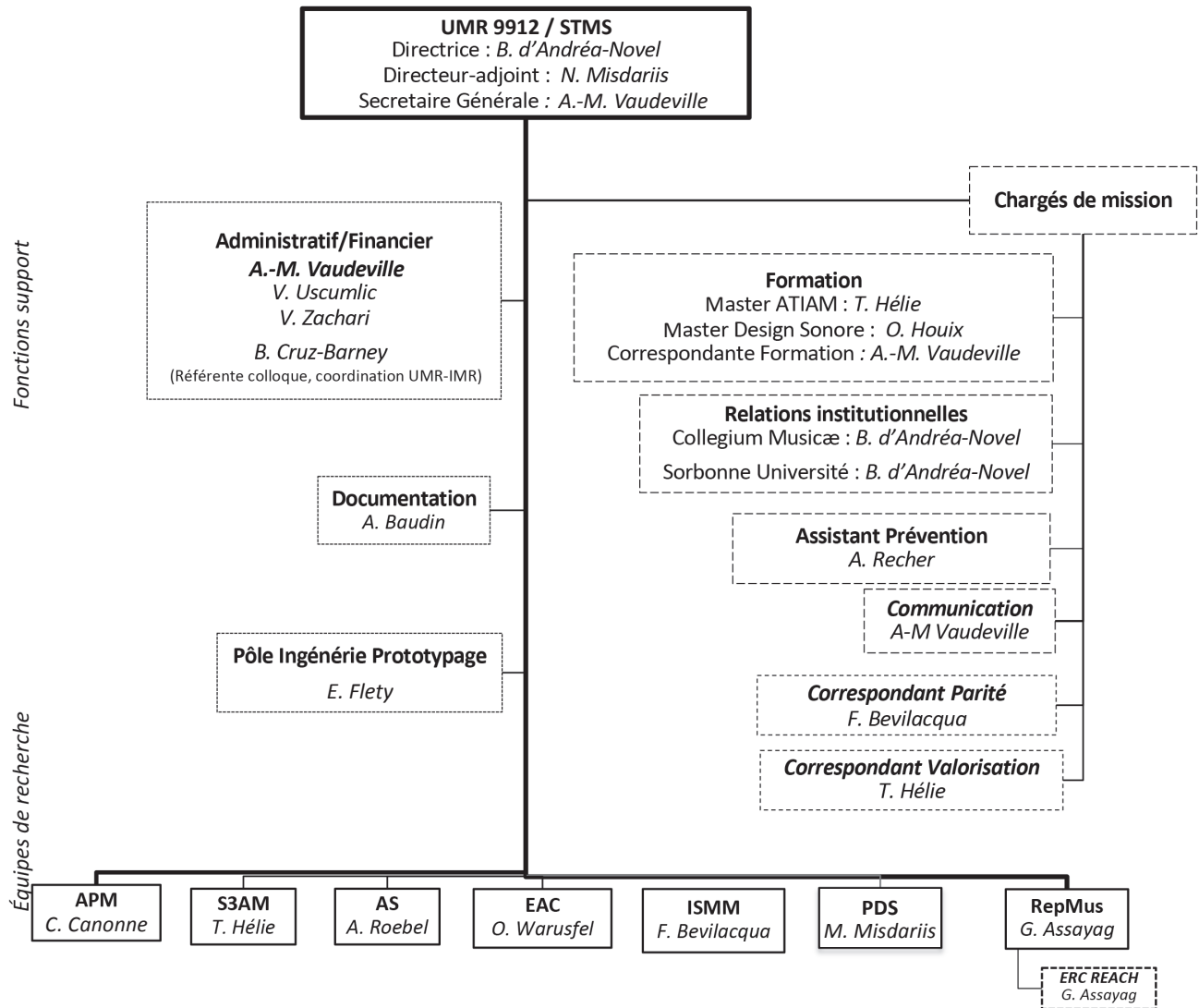
Sommaire

7	Le laboratoire
8	Organigramme
9	Missions
10	Organisation et enjeux scientifiques
13	Liens institutionnels
14	Innovation et moyens de la recherche
15	Collaborations de recherche et partenariats récents
17	Les équipes
18	Systèmes et signaux sonores : audio/acoustique, instruments (S3AM)
19	Espaces acoustiques et cognitifs (EAC)
20	Perception et design sonores (PDS)
21	Analyse et synthèse des sons (AS)
22	Interaction son musique mouvement (ISMM)
23	Représentations musicales (RepMus)
24	Analyse des pratiques musicales (APM)
25	Les projets
26	C1. L'atelier du son
38	C2. Le corps musicien
46	C3. Les dynamiques créatives
63	Les logiciels
64	ADMix Tools
65	Antescofo
66	CataRT
67	DYCI2
68	IrcamLab TS
69	Max
70	Modalys
71	MuBu pour Max
71	OMax & co
72	OpenMusic
73	Orchid*
74	OSCar
75	Panoramix
76	SkataRT
76	SOMAX2
77	Spat~
78	SpeaK
79	Suivi de geste et de formes temporelles
79	SuperVP, Trax et bibliothèques
80	The Snail-Absolute Tuning
81	Le Forum Ircam
82	Les équipes

1

Le laboratoire

Organigramme



Missions

Depuis sa création en 1977, l'Ircam a pour mission fondamentale de susciter une interaction féconde entre recherche scientifique, développement technologique et création musicale contemporaine. Cette articulation constitue le principal axe structurant de l'ensemble de ses activités. L'un des enjeux majeurs est de contribuer, par les apports des sciences et techniques, au renouvellement de l'expression musicale. Réciproquement, les problèmes spécifiques posés par la création contemporaine donnent lieu à des avancées scientifiques originales, tant théoriques, méthodologiques, qu'appliquées, dont la portée dépasse largement le seul domaine musical. Cette dynamique de recherche originale, tournée vers la création artistique dans toute l'exigence et la sensibilité qui s'y manifestent pour des modes de représentation et de manipulation élaborés du sonore et du musical, suscite des innovations scientifiques et technologiques dont les applications se déploient dans des secteurs d'activité de plus en plus larges : professionnels de la musique et du son, industrie, enseignement, public mélomane... Cette médiation entre recherche et création musicales comporte en particulier le développement d'outils logiciels et technologiques pour les musiciens (compositeurs, interprètes, musicologues), à partir de modèles et prototypes élaborés par des équipes de recherche travaillant dans les différents domaines en rapport avec la musique : informatique (langages, interfaces homme-machine, temps réel, bases de données), traitement du signal et automatique, acoustique, perception et psychologie cognitive de l'audition, musicologie...

Les travaux reposent ainsi sur l'articulation de deux types d'activités complémentaires, la recherche et le développement, réalisées dans le cadre de l'unité mixte de recherche Sciences et technologies de la musique et du son (UMR 9912 – STMS) soutenue par le CNRS, le ministère de la Culture et Sorbonne Université.

La **recherche**, source d'innovation, vise l'élaboration de connaissances en rapport avec les problématiques musicales. Interdisciplinaire par nature, elle inscrit son activité sous la forme de nombreuses collaborations avec des laboratoires français et étrangers, avec des organismes d'enseignement supérieur et avec des partenaires institutionnels et privés. L'accueil d'élèves-chercheurs et ingénieurs dans le

cadre de thèses de doctorat, de stages de master et d'écoles d'ingénieurs contribue à la formation par la recherche. Les compétences ainsi développées trouvent de nombreuses applications au-delà des problématiques musicales, et font l'objet de projets réalisés en collaboration avec des industriels ou dans le cadre de programmes nationaux, européens et internationaux.

Le **développement** effectue l'adaptation des connaissances, modèles et prototypes issus de la recherche sous la forme d'environnements logiciels et d'outils technologiques. Les principales applications visent la réalisation d'outils pour la création musicale, à travers la mise en œuvre d'environnements ouverts et programmables, afin de pouvoir répondre à des approches esthétiques très diverses, et d'intégrer les modèles issus des travaux de recherche au fur et à mesure de leur avancement. Des versions simplifiées des logiciels, destinées à des cibles d'utilisateurs plus larges, rendent accessible la technologie produite dans le laboratoire sous forme de collections de produits : Ircam Tools (professionnels de l'audio), modules Max for Live (musique électronique live), etc. Le Forum Ircam, club d'utilisateurs des logiciels, favorise leur diffusion auprès d'une communauté internationale de professionnels de la musique et du son (compositeurs, artistes multimédias et plasticiens, ingénieurs du son, sound designers, chercheurs, enseignants...), évaluée à plus de 30 000 utilisateurs depuis sa création. La filiale Ircam Amplify, fondée en 2019, commercialise les logiciels et connaissances issus du laboratoire sous forme de lignes de produits, de licences et de prestations dans différents secteurs d'activités.

Organisation et enjeux scientifiques

Pour répondre à cette mission, le laboratoire est structuré en **sept équipes** qui se caractérisent par leurs ancrages théoriques et les objets étudiés :

- **S3AM** (Systèmes et signaux sonores : audio/acoustique, instruments) mobilise mathématiques appliquées (signal, système, contrôle, géométrie différentielle, analyse numérique), physique et mécanique pour explorer, comprendre, reproduire ou inventer des systèmes multiphysiques produisant des signaux sonores (instruments de musique, voix) ;
- **EAC** (Espaces acoustiques et cognitifs) s'appuie sur l'acoustique, le traitement du signal, la psychologie cognitive de l'audition et les neurosciences pour la captation, l'analyse et la reproduction de champs sonores et l'étude et la compréhension de la perception spatiale du son en contexte d'interaction multisensorielle ;
- **PDS** (Perception et design sonores) adresse la perception des sons vocaux, musicaux et environnementaux, leur traitement émotionnel et les applications en design sonore à travers la psychoacoustique, la psychologie, la cognition musicale et les neurosciences ;
- **AS** (Analyse et synthèse des sons) étudie le signal audionumérique en utilisant des théories et techniques de traitement du signal, statistique, apprentissage et reconnaissance des formes pour en extraire toutes sortes de caractéristiques, pour le synthétiser selon des propriétés spécifiées, ou encore pour le transformer en fonction de besoins compositionnels ou autre.
- **ISMM** (Interaction son musique mouvement) focalise ses recherches sur l'interaction « incarnée » (embodied interaction) entre humains et médias sonores et musicaux : captation du geste, programmation réactive et temps réel, apprentissage, interface tangible et multimodale, techniques de synthèses sonores ;
- **RepMus** (Représentations musicales) travaille sur les structures formelles de la musique et les environnements créatifs pour la composition, l'analyse et l'interaction musicien-machine à partir d'architectures et d'approches langages, de formalismes algébriques, et de techniques d'IA telles que l'optimisation et l'apprentissage ;
- **APM** (Analyse des pratiques musicales) développe une musicologie interdisciplinaire des œuvres et des pratiques musicales contemporaines, en considérant les points de vue historique, sociologique, anthropologique, psychologique, philosophique et en privilégiant des approches « outillées », souvent en lien avec les autres équipes de l'UMR.

La recherche scientifique au sein de l'UMR s'appuie sur ces 7 équipes articulées autour de **3 champs d'interaction** :

- **L'atelier du son ;**
- **Le corps musicien ;**
- **Les dynamiques créatives.**

C1. L'atelier du son

Le son est ici compris comme un phénomène physique (mécanique, acoustique) mais aussi comme une information numérique engageant des techniques mathématiques et informatiques (traitement du signal) pour traiter des flux, des contenus et des scènes sonores, musicales, ou multimédia.

Le signal audionumérique est abordé sous ses deux aspects de source et de champs. En tant que source, nous le caractérisons ou bien nous le synthétisons, avec des contraintes de précision, de finesse et de qualité à même de répondre à l'exigence de la création contemporaine. Nos recherches se focalisent donc sur les spécificités du signal musical et nous amènent à développer des modèles plus précis en les informant par la physique, la biomécanique, la perception et la cognition, ou encore des données musicales symboliques... L'extraction de descriptions plus symboliques, sémantiquement pertinentes, et le contrôle des modèles par des données qualitatives de haut niveau (par exemple, la prosodie dans le cas de la synthèse de la voix chantée) sont deux des objectifs du laboratoire dans ce domaine. En tant que champs, nos études visent à comprendre ses représentations spatiales et les coupler avec les représentations temps/fréquence, à capter ce champ avec une grande résolution, et à le reconstruire de manière réaliste.

C2. Le corps musicien

Le domaine visé est celui de la musique et du son perçus et produits par des humains qui interagissent avec un environnement, une machine ou d'autres êtres humains. Le sonore ne se réduit plus à une vibration, mais devient un véhicule permettant l'interaction, qui est produit par un corps biologique et qui agit sur les individus et les comportements. Ce domaine est abordé en se positionnant sur quatre problématiques : le geste, la voix, la perception spatiale du son et son intégration multisensorielle, en étudiant les liens entre contenus sonores ou musicaux et valence émotionnelle.

C3. Les dynamiques créatives

Si la musique est un phénomène produit et perçu par un corps qui tire parti des objets physiques et qui organise un monde sonore, elle est aussi pensée, imaginée – ou rêvée – et parfois formalisée avant d'être réalisée. La question de la créativité émerge aujourd'hui comme une problématique scientifique¹ et comme un paradigme de travail utile pour la conception de systèmes intelligents. Contrairement à ce que pourrait laisser penser une analyse superficielle des nouvelles directions prises par l'IA², le défi ne nous semble pas tant être celui de l'*imitation* que celui du *compagnonnage* : quelles sont les idées et les outils nécessaires pour inspirer, guider et assister l'artiste dans son processus de création et faire ainsi de la machine un partenaire ?

Nous opposons donc ici une approche qui vise à remplacer l'homme en rendant la machine autonome et celle qui vise à augmenter les capacités humaines. On peut alors parler d'*intelligence augmentée*.

Dans le domaine artistique, la première approche peut éventuellement amener à éclairer les mécanismes qui sont remplacés, mais la seconde ouvre quant à elle des dimensions créatives inédites susceptibles de renouveler profondément la relation science-musique.

S'appuyant sur un de nos atouts majeurs – notre intimité avec la création musicale –, les travaux de STMS visent à éclairer et documenter le processus créatif chez des compositeurs ; à formaliser, analyser, calculer et produire des structures symboliques musicales en interaction avec les visées des créateurs ; à développer de nouveaux outils d'écriture adressant les enjeux contemporains de la notation ; et à appréhender les dispositifs assistant la création collective.

Ces trois champs assurent ainsi une lisibilité de notre recherche et couvrent l'ensemble de nos activités, puisqu'ils adressent le monde sonore et musical dans sa dimension physique et numérique, sa perception et sa production par un sujet humain, en lien avec les enjeux de création et de créativité.

Parmi ces sujets d'étude, mentionnons quelques focus ou sujets que nous souhaitons investiguer avec une attention plus particulière dans les prochaines années :

- La modélisation multiphysique
- Le geste expressif expert
- Les CP&HS (Cyber Physical & Human Systems)
- L'apport des neurosciences, sciences cognitives et les liens avec la santé
- IA et créativité musicale
- Les nouvelles dimensions de la notation musicale

Malgré la diversité des travaux menés, le laboratoire possède une identité et une visibilité fortes à la fois du point de vue de ses objets d'études et de son positionnement scientifique.

1. Axe Son-Musique Santé

À côté de ces 3 champs de recherche, l'axe transversal Son-Musique Santé regroupe les recherches du laboratoire liées au bien-être et à la santé et est organisé en 3 activités principales :

- Prévention et promotion de la santé et du mieux-être.
- Interventions dans la prise en charge du patient.
- Étude des mécanismes de la perception de la musique et du son, notamment en cas de dysfonctionnement.

Articulation recherche-crédation

L'adossement des thématiques de recherche à la création musicale contemporaine constitue une interaction féconde³ et permet une grande diffusion des résultats de recherche. L'interaction recherche-crédation est soutenue par plusieurs dispositifs récurrents. Nous en citerons quatre.

- Les équipes de l'UMR participent aux *productions artistiques* de l'Ircam, soit directement, soit à travers les réalisateurs en informatique musicale (RIM) qui servent de courroie de transmission entre les artistes et les prototypes et les outils élaborés dans le laboratoire. Leur rôle est essentiel pour adapter les temps longs de la recherche au temps nécessairement plus limité des productions artistiques.
- Les projets de résidence en recherche artistique permettent aux équipes d'interagir avec un créateur qui propose un projet de recherche et vient en résidence à l'Ircam pour une durée d'un à trois mois. Ces projets, détachés du contexte et des contraintes de production, permettent de développer un travail en amont. Depuis 2012, une soixantaine de résidences ont été sélectionnées par un appel international annuel via la plateforme du réseau ULYSSES⁴. La plupart des artistes en résidence sont étrangers et la moitié des projets impliquent au moins deux équipes.
- Apparue en 2013, le dispositif des *thèses en art* permet de développer une interaction sur un temps long avec un artiste qui poursuit un travail de thèse à plein temps. Ce type de doctorat permet d'articuler finement pratique artistique et approche théorique d'une part, recherche et création d'autre part.

1. Voir les nouvelles conférences internationales « Digital Intelligence (#DI) », « Musical MetaCreation (MUME) », « International Conference on Computational Creativity (ICCC) » auxquelles STMS participe régulièrement depuis leur fondation, ou encore l'initiative européenne FET « Creative ICT » qui a financé cinq projets en 2016.

2. Comme le projet **Magenta** développé par Google : « [how to develop] algorithms that can learn how to generate art and music. »

3. Ce modèle d'interaction, dans lequel l'Ircam a été pionnier, est aujourd'hui reconnu et s'impose avec par exemple le projet H2020-ICT **Vertigo** (Adding socio-economic value to industry through the integration of artists in research and open innovation processes) qui a débuté fin 2016 porté par l'Ircam : vertigo.ircam.fr/

4. www.ulysses-network.eu

2. Un aller-retour entre applications finalisées et réflexions théoriques

Un des principes qui caractérise la stratégie scientifique de l'UMR est de toujours valider les réflexions théoriques par des développements qui sont utilisés en dehors des équipes et, à l'inverse, d'étayer les développements applicatifs par une réflexion épistémologique ou théorique. Dans cet aller-retour, l'application logicielle ou matérielle agit dans le laboratoire comme un indispensable dispositif expérimental. Les prototypes développés par les chercheurs le sont souvent dans le cadre de projets collaboratifs, ce qui entraîne leur utilisation par des partenaires extérieurs et aussi dans les productions artistiques, confrontant les résultats du laboratoire aux exigences du compositeur et à son oreille aguerrie. Au-delà de l'expérimentation à travers les projets artistiques propre à l'Ircam, les logiciels atteignent aussi un public plus large à travers le *Forum Ircam*¹ qui fait l'objet d'une coordination transversale à l'Ircam (plus de 30 000 membres professionnels dans le monde).

3. Un très large ancrage national et international

Notons que le laboratoire est fortement ancré dans le paysage national et international de la recherche au plus haut niveau, dans toutes les communautés scientifiques auxquelles sont rattachées les différentes équipes de recherche. Ce point fait l'objet du paragraphe suivant.

1. forumnet.ircam.fr

Liens institutionnels

L'UMR s'appuie sur les liens naturels et diversifiés entre ses équipes de recherche mais tire aussi parti de l'ancrage de ses membres au sein d'une communauté régionale, nationale et internationale. En attestent en particulier le grand nombre d'invitations des chercheurs du laboratoire (séminaires, conférenciers invités), leur participation dans les comités de programme et les conférences et colloques qu'ils ont initiés ou organisés.

Au niveau régional, le laboratoire est partenaire du parcours de master de recherche (M2) **ATIAM** (Acoustique, traitement du signal et informatique appliqués à la musique, parcours Systèmes et applications réparties), accueilli et coordonné par l'Ircam dans le cadre du master Sciences et technologie de Sorbonne Université, en collaboration avec Télécom Paris. Le laboratoire est aussi partenaire du master **Design Sonore** coorganisé avec l'EPCC – École supérieure d'art et de design Tours Angers Le Mans, l'université du Maine et l'ENSCI – Les Ateliers. Ces deux formations n'ont pas d'équivalent sur le territoire français. Au niveau international, les cursus comparables sont plutôt orientés « Music Technology » et, inscrits dans des facultés de musique, s'adressent plutôt à des étudiants dont la formation initiale est la musique. Les membres du laboratoire interviennent aussi ponctuellement dans plusieurs autres masters (Master spécialisé AIMOVE avec Mines Paris, Nantes, Grenoble, Strasbourg, Marseille).

STMS est un laboratoire d'accueil des écoles doctorales de Sorbonne Université dans ses domaines de compétence, en particulier EDITE (École doctorale d'informatique, télécommunications et électronique de Paris), SMAER (Sciences mécaniques, acoustique, électronique et robotique) et ED3C (Cerveau, cognition, comportement). Dans le cadre de Sorbonne Université, une nouvelle filière doctorante destinée aux compositeurs et intitulée Doctorat de musique – Recherche en composition, a vu le jour à la rentrée universitaire 2014 mais n'a pas proposé de contrat doctoral ces deux dernières années.

L'Ircam est membre fondateur du **Collegium Musicæ**, institut qui mène une action structurante au sein de Sorbonne Université. Cinq axes thématiques sont couverts : Improvisation, Apprentissage et IA ; Instrument d'hier au service de celui de demain ; Les lieux de production musicale ; Musique et médecine ; La construction des savoirs musicaux.

Au niveau international, l'UMR entretient des liens réguliers avec les centres internationaux de recherche en informatique musicale dont voici une liste des principaux par équipe. internationaux de recherche en informatique musicale dont voici une liste des principaux par équipe.

S3AM

Athena Research Center (Grèce)
Centre hospitalier universitaire de Liège (Belgique)
Université de Nouvelle-Galles du Sud, UNSW de Sydney (Australie)

A/S

Institute for Digital Media Technology, Fraunhofer (Allemagne)
Center for Research in New Music (CeReNeM), université d'Huddersfield (Royaume-Uni)

Media Integration and Communication Center, Université Aristote de Thessalonique (Grèce)
Université de Florence (Italie)

RepMus

CIRMMT Centre for Interdisciplinary Research in Music Media and Technology, université McGill (Canada)
CNMAT Center for New Music and Audio Technologies, Université de Californie (États-Unis)
CCRMA Center for Computer Research in Music and Acoustics (États-Unis)
Haute Ecole de Musique de Genève (Suisse)
STEGI Onassis cultural Center, université d'Athènes (Grèce)
GEIDA University of the Arts, université de Tokyo (Japon)
Université catholique de Louvain (Belgique)

EAC

Aalto University Helsinki (Finlande)
Acoustics research Institute of the Austrian Academy of Sciences - ARI-ÖAW (Autriche)
Ecole supérieure polytechnique de Rhénanie-Westphalie (Allemagne)
Institut polytechnique Rensselaer (États-Unis)
Université de New York (États-Unis)

ISMM

Department of Computer Science and Engineering, Université Carlos III de Madrid, (Espagne)
Future Narratives Lab, Université Nationale Yang Ming Chiao Tung Taiwan)
Université de l'Arizona(USA)
HKU, University of the Arts, Utrecht (Pays-Bas)
CNMAT Center for New Music and Audio Technologies, université de Californie (États-Unis)

PDS

Watanabe Katsumi Laboratory, Department of Intermedia Art and Science, Faculty of Science and Engineering, université Waseda de Tokyo (Japon)
Faculty of Industrial Design Engineering, Université technologique de Delft (Pays-Bas)
CIRMMT Centre for Interdisciplinary Research in Music Media and Technology, École de musique Schulich, université McGill (Canada)
Université de Gand (Belgique)
Université de Mayence (Allemagne)
École polytechnique fédérale de Lausanne (Suisse)
Royal Institute of Technology of Stockholm (Suède)

APM

CIRMMT Centre for Interdisciplinary Research in Music Media and Technology, École de musique Schulich, université McGill (Canada)
Social Mind and Body Group, Central European University (Autriche)
Haute École de musique – Conservatoire de Genève (Suisse)
Department of Music, New York University (États-Unis)

Innovation et Moyens de la recherche

Au cœur d'enjeux sociétaux et économiques croisant culture et technologies de l'information, les recherches accueillies à l'Ircam se présentent dans le paysage international de la recherche comme pôle de référence interdisciplinaire autour des sciences et technologies du son et de la musique et s'exposent en permanence aux nouveaux besoins et usages de la société. Cette médiation prend diverses formes en fonction des cibles et sphères d'activité concernées et est portée par le **département IMR – Innovation et Moyens de la recherche** de l'Ircam qui est également en charge de la tutelle Ircam de l'UMR STMS et des infrastructures web et informatiques de l'Ircam.

La spécificité de l'Ircam est de réunir un large spectre de compétences scientifiques et technologiques, qui trouvent des applications dans de nombreux secteurs d'activité, centrés sur la production et la diffusion de la musique, du son, des arts numériques, des industries culturelles et rayonnant dans tous les domaines faisant intervenir la dimension sonore : recommandation musicale, traitement de la voix, design sonore (automobile, transports, environnement urbain), interfaces multimodales, réalité virtuelle et augmentée, simulation, etc. Fort de son savoir-faire de plusieurs décennies sur le management de l'innovation et de la créativité collective, l'Ircam intervient également dans différents contextes pour l'application de son modèle d'innovation technologique en lien avec la création artistique.

Celui-ci a notamment été étendu à tous les champs technologiques et artistiques dans le cadre de l'initiative STARTS¹ (Science, Technology and the Arts) de la Commission européenne : après avoir coordonné le projet pilote européen VERTIGO-STARTS Residencies organisant 45 résidences d'artistes en lien avec des projets de recherche technologique dans toute l'Europe, l'Ircam participe au projet MediaFutures appliquant la méthodologie de STARTS Residencies pour la préfiguration de médias innovants faisant appel à de grandes bases de données informationnelles.

Ce modèle s'expose dans le cadre du **Forum annuel Vertigo**², symposium interdisciplinaire organisé au Centre Pompidou dans le cadre de l'événement Mutations Création, associant artistes, chercheurs, ingénieurs et acteurs de l'innovation autour des enjeux actuels de la création artistique dans son rapport aux sciences et technologies.

L'importance croissante des industries culturelles et créatives (ICC) donne lieu à de nouveaux programmes destinés à soutenir l'innovation et la souveraineté nationale et européenne dans ce secteur. L'Ircam y est présent dans le cadre des appels du programme France 2030 consacrés aux ICC en tant que chef de file du nouveau projet Continuum visant une expérience d'immersion sonore augmentée du spectacle vivant à la fois en salle et à travers des dispositifs de réalité virtuelle. Au niveau européen dans le cadre du nouveau programme-cadre Horizon Europe, l'Ircam est l'un des principaux partenaires du projet DAFNE+ expérimentant de nouveaux modes de diffusion et de valorisation des productions artistiques à travers une plateforme blockchain/ NFT.

Spécificité de l'Ircam, la culture de **développement technologique professionnel** intégré à ses équipes de recherche aboutit à la production finalisée de modules fonctionnels optimisés, directement transférables. Plus globalement, cette activité de développement se traduit par diverses formes de réalisation adaptées aux différentes catégories de cibles visées. Plus d'une dizaine d'environnements logiciels sont ainsi développés au sein des équipes (cf. Partie 4), perfectionnés de manière incrémentale en fonction des dernières avancées de la recherche, utilisés par les autres départements de l'Ircam pour leurs activités de production artistique et de formation (compositeurs, réalisateurs en informatique musicale, ingénieurs du son, instrumentistes, danseurs...) et diffusés dans le cadre du Forum Ircam auprès d'une communauté internationale de plus de 30 000 utilisateurs professionnels. Des applications simplifiées dans leur usage en sont dérivées sous forme de produits logiciels dans le cadre de différentes collections commerciales : Ircam Lab avec Plugivity, Ircam Tools avec Flux., IrcamMax avec Ableton.

L'Ircam est agréé par le ministère de la Recherche et de l'Enseignement supérieur au titre du Crédit Impôt Recherche et les prestations effectuées par ses équipes pour le compte de partenaires privés bénéficient d'avantages fiscaux. Les liens avec l'industrie s'inscrivent dans de nombreux cadres de collaboration : projets de R&D nationaux et européens (une vingtaine en cours, dont 40% en tant que chef de file), prestations de recherche, co-encadrement de thèses CIFRE, accompagnement à la création de jeunes pousses (Phonotonic, Mogeos, Niland (rachetée par Spotify en 2017), Antescofo et HyVibe parmi les plus récentes), prises de participation et cessions de licences. La fondation en 2019 d'**Ircam Amplify**³, filiale de l'Ircam avec le soutien de la Banque des Territoires et d'investisseurs privés, constitue un nouveau momentum d'intensification de la diffusion commerciale des technologies et expertises des recherches accueillies à l'Ircam.

1 starts.eu

2 vertigo.ircam.fr

3 ircamamplify.com

Collaborations de recherche et partenariats récents

3D for All Kft (Hongrie)	France Télévisions	Orbe
A-Volute	Fratelli Piacenza SpA (Italie)	OSU Ecce Terra
Ableton (Allemagne)	Fraunhofer (Allemagne)	Parisson
Acapela	GENESIS (Genesis Acoustics)	Parrot
Arkamys	Goldsmiths' College (Royaume-Uni)	Pathé
Arte	Grame	Philharmonie de Paris
Artipolis	Haute École de musique de Genève (Suisse)	Phonotonic
Artshare (Belgique)	HearDis! Corporate Sound GmbH (Allemagne)	Playground (Suède)
Athena Research Center (Grèce)	I3S	pluX - Wireless Biosignals SA
Athens Technology Center (Grèce)	ID Scenes	PSA Peugeot Citroën
b<->com	Inova+ (Portugal)	Qwant
Bass Nation (Finlande)	Inria	Radio France
Bayerischer Rundfunk (Allemagne)	Inserm-Sorbonne Université-faculté de médecine	Reactable Systems (Espagne)
BBC (Royaume-Uni)	Institut Jean Le Rond d'Alembert (Sorbonne Université)	Renault
Bmat (Espagne)	Instituts de recherche technologique	Sigma-Orionis
Buffet-Crampon	INTEGRAL Markt – und Meinungsforschungsges.m.b.H. (Allemagne)	Sky-Deutschland (Allemagne)
Cabrilog (France)	IreMus	SNCF
Centre de psychiatrie et neurosciences (Inserm-université Paris Descartes)	ISAE SUPAERO	Somethin'Else Sound Directions Ltd (Royaume-Uni)
Centre de recherche de l'Institut du cerveau et de la moëlle épinière (UMR 7225)	ISIR	Soundtrack Your Brand (Suède)
Centre hospitalier universitaire de Nice	IUAV (Italie)	StreetLab
Centre interfacultaire des sciences affectives (Suisse)	Kainos (Royaume-Uni)	Stromatolite Innovation Lab (Royaume-Uni)
CHU La Conception-Marseille	Kantar Media	Stupeflix
CHU-Liège (Belgique)	KTH (Suède)	Supelec
CNES	Laboratoire de neurophysique et physiologie (UMR 8119, université Paris-Descartes)	Technicolor
CNMAT-UC Berkeley	Laboratoire des sciences, des procédés et des matériaux de l'université Paris-13	Technische Universität Berlin (Allemagne)
CNRS Telecom ParisTech	LaBRI	Technische Universität Wien (Autriche)
Conservatoire national supérieur de musique et de danse de Paris	LAGEP	Testaluna (Italie)
CTEL	Lambde Limited (Roli)	Trinity College Dublin (Irlande)
Culture Tech	LATMOS	Trinnov-Audio
Dassault Systèmes	Libelium Comunicaciones (Espagne)	Ubisoft
Deezer	LIMSI	UC Limburg (Belgique)
Deutsches Forschungszentrum für Künstliche Intelligenz (Allemagne)	LMA	Universal (Royaume Uni)
Dualo	Lovemonk S.L. (Espagne)	Universal Music Publishing Classical
Dubbing Brothers	LPL	Université Ben Gurion (Israël)
École de musique Schulich (Canada)	LRI	Université de Bretagne occidentale
École polytechnique fédérale de Lausanne (Suisse)	Magix (Allemagne)	Université de Fribourg (Suisse)
ECR-Musée de la musique	Maha	Université de Gênes (Italie)
EHESS	Makemusic	Université de Thessalonique (Grèce)
elephantcandy (Pays-Bas)	Mathematics for more (Espagne)	Université de Zaragoza (Espagne)
Ellinogermaniki Agogi (Grèce)	Mogees	Université Kiel (Allemagne)
EMI	Musée du Quai Branly – Jacques-Chirac	Université McGill (Canada)
EMPAC (États-Unis)	Musimap (Belgique)	Université Pompeu Fabra (Espagne)
Ensad	Niland	UVI Sounds and Software
esba TALM	No Design	Velti (Grèce)
Eurecom Sophia Antipolis	Nokia (Finlande)	Vi-live
Eurescom	Novelab	Vizion'r
FINCONS SpA	Novespace	Xtranormal
Flux::	Open University (Royaume-Uni)	
Flying Eye (Allemagne)	Orange	

2

—
Les équipes
—

Équipe Systèmes et signaux sonores : audio/acoustique, instruments (S3AM)

Responsable : Thomas Hélie

Activités

—

L'équipe Systèmes et signaux sonores : Audio/acoustique, instruments élabore des outils théoriques, technologiques et expérimentaux portant sur les systèmes multiphysiques et les signaux sonores qu'ils produisent. Elle s'intéresse à explorer, comprendre, reproduire avec réalisme ou inventer des objets sonores en audio, en acoustique, avec une focalisation sur les instruments de musique jusqu'à la production de la voix.

Plus précisément, les objectifs sont de modéliser, simuler, identifier et optimiser ces systèmes (voix, musicien/instrument, haut-parleurs, effets électroniques audio, etc.), avec pour particularité de s'appuyer sur la physique pour faire émerger des structures intrinsèques et en bénéficier. L'équipe élabore des méthodes d'analyse, transformation, contrôle et simulation de sons et des outils d'aide à la conception, dans des paradigmes virtuels, réels ou hybrides.

Dans cette démarche, l'approche globale «systèmes et signaux» apporte une synergie dès la conception des outils, en connectant (et non simplement en juxtaposant) plusieurs disciplines et champs scientifiques : physique, théorie des systèmes et du contrôle, géométrie différentielle, analyse numérique, traitement du signal, informatique, électronique, mécatronique et robotique.

Les applications visées concernent les domaines scientifiques, artistiques, pédagogiques et de la santé.

Thématiques et projets associés

—

- Modélisation physique des instruments de musique et de la voix
- Synthèse sonore par modélisation physique
- Modélisation de systèmes physiques audio et de haut-parleurs
- Plateformes expérimentales robotisées : archet robotisé, bouche artificielle robotisée pour le jeu des cuivres, appareil vocal robotisé à l'échelle 1:1
- Identification de systèmes non linéaires
- Contrôle de systèmes non linéaires et contrôle d'instruments et d'objets du patrimoine
- Instruments augmentés, instruments hybrides
- Lutherie augmentée
- Projet ATRIM (Analyseur temps réel haute précision de justesse et de timbre pour instruments musicaux)
- Projet AVATARS (Artificial Voice production: control of bio-inspired port-hAmilToniAn numeRical and mechatronic models)
- Projet Bass Holograms in Pune¹
- Projet Accordéon augmenté²
- Projet « The nonlinear spring »³
- Projet Infidhem (Systèmes interconnectés de dimension infinie pour les milieux hétérogènes)

- Projet Ondes Martenot
- Projet Sheng⁴
- Projet Organologie des harpes pharaoniques

Domaines de compétence

—

Acoustique, mécanique, systèmes non linéaires, automatique et contrôle, théorie du signal, géométrie différentielle, analyse numérique, expérimentation, mécatronique, lutherie réelle et virtuelle, synthèse sonore.

Collaborations

—

Athena-RIC (Grèce), Buffet-Crampon, Cabrilog SAS, C2RMF-Louvre, C2RMF et Louvre- Lens, CHU Liège, Collegium Musicae, CYU-Cergy Université, GdR-GDM, École Centrale Lille, GIPSA-lab, HISOMA, IJLRA-Sorbonne Université, Institut National du Patrimoine, INRAE, INRIA-Lille, INRIA-Bordeaux, IREMUS, ISAE-SUPAÉRO, LAGEP-Université Lyon-1, LaSiE-Université de La Rochelle, LEOPOLY (Hongrie), LJLL-Sorbonne Université, LMA-CNRS, LMD-ENS, LPL, Musée de la musique (Cité de la musique-Philharmonie de Paris), Université Technologique de Troyes, UVI.

Équipe

—

Chercheurs : B. d'Andréa-Novel (Sorbonne Université), H. Boutin (Sorbonne Université), R. Caussé (émérite), T. Hélie (CNRS), D. Roze (CNRS), M. Wijnand – **Ingénieur :** R. Piéchaud – **Doctorants :** L. Forma, V. Martos, T. Risse, A. Thibault, C. Voisembert – **Résidences/invités/collaborations longues :** P. Bloland, R. Causton, S. Emerit, A. Falaize, J.-L. Le Carrou, F. Krouchi, E. Rouhaud, N. Schütz, F. Silva, J.-E. Sotty, H. P. Stubbe Teglbjærg.

1 <https://forum.ircam.fr/article/robert-piechaud-ircam/> <https://pune.afindia.org/events/bass-holograms-a-musical-encounter-between-the-fkbass-an-augmented-bass-with-integrated-technology-and-the-rudra-veena/>

2 Résidence artistique : <https://www.ircam.fr/article/jean-etienne-sotty-lacordeon-augmente>

3 Projet artistique sur le couplage non linéaire entre instruments par contrôle actif, en collaboration avec H.-P. Stubbe Teglbjærg, soutenu par le Ministère de la Culture du Danemark.

4 <https://anr.fr/Projet-ANR-15-CE23-0007>

Équipe Espaces acoustiques et cognitifs (EAC)

Responsable : Olivier Warusfel

Activités

—

L'activité de recherche et de développement de l'équipe Espaces acoustiques et cognitifs est consacrée à la reproduction, à l'analyse/synthèse et à la perception de scènes sonores. Les disciplines scientifiques de l'équipe sont le traitement du signal et l'acoustique pour l'élaboration de techniques de reproduction audio spatialisée et de méthodes d'analyse/synthèse du champ sonore. Parallèlement, l'équipe consacre un important volet d'études cognitives sur l'intégration multisensorielle pour un développement raisonné de nouvelles médiations sonores basées sur l'interaction corps/audition/espace. Les activités de recherche scientifique décrites ci-dessous s'articulent avec une activité de développement de bibliothèques logicielles. Ces développements consignent le savoir-faire de l'équipe, soutiennent son activité de recherche théorique et expérimentale et sont le vecteur majeur de notre relation avec la création musicale et d'autres secteurs applicatifs.

Les travaux concernant les techniques de spatialisation se concentrent sur les modèles basés sur un formalisme physique du champ sonore. L'objectif principal est le développement d'un cadre formel d'analyse/synthèse du champ sonore exploitant des réponses impulsionnelles spatialisées (SRIR pour Spatial Room Impulse Response). Les SRIRs sont généralement mesurées par des réseaux sphériques comportant plusieurs dizaines de transducteurs (microphones et/ou haut-parleurs). L'application principale concerne le développement de réverbérateurs à convolution exploitant ces SRIRs à haute résolution spatiale afin de reproduire fidèlement la complexité du champ sonore.

La technique de spatialisation binaurale sur casque retient également notre attention. L'évolution des pratiques d'écoute et la démocratisation des applications interactives tendent à privilégier l'écoute sur casque à travers l'usage des smartphones. Grâce à sa capacité d'immersion sonore, l'écoute binaurale devient le premier vecteur d'écoute tridimensionnelle. Basée sur l'exploitation des fonctions de transfert d'oreille (HRTFs) elle reste à ce jour la seule approche assurant une reconstruction exacte et dynamique des indices responsables de la localisation auditive. Elle s'impose comme un outil de référence pour la recherche expérimentale liée à la cognition spatiale en contexte multisensoriel et pour les applications de réalité virtuelle.

Ces techniques de spatialisation audio 3D, associées aux dispositifs de captation des mouvements de l'interprète ou de l'auditeur dans l'espace, constituent une base organologique essentielle pour aborder questions « d'interaction musicale, sonore et multimédia ». Parallèlement, elles nourrissent une recherche sur les mécanismes cognitifs liés à la sensation d'espace, notamment sur la nécessaire coordination entre les différentes modalités sensorielles (audition, vision, proprioception, motricité, ...) pour la perception et la représentation de l'espace. Nous tentons de mettre à jour l'influence des différents indices acoustiques (localisation, distance, réverbération...) utilisés par le système nerveux central de l'homme sur l'intégration des informations sensorielles et leur interaction avec les processus émotionnels.

Sur le plan musical, notre ambition est de fournir des modèles et des outils permettant aux compositeurs d'intégrer la mise en espace des sons depuis le stade de la composition jusqu'à la situation de concert, contribuant ainsi à élever la spatialisation au statut de paramètre d'écriture musicale. Plus généralement, dans le domaine artistique, ces recherches s'appliquent également à la postproduction, aux installations sonores interactives et à la danse à travers les enjeux de l'interaction son/espace/corps.

Thématiques et projets associés

—

- **Spatialisation sonore**: Réverbération hybride et réponses impulsionnelles spatialisées (SRIR); Analyse-synthèse de SRIR; Réverbération hybride et réponses impulsionnelles spatialisées; Synthèse de champs sonores par réseaux à haute densité spatiale; Système WFS et HOA de l'Espace de projection; Écoute binaurale; CONTINUUM, Spatialisation distribuée; Réalité augmentée et Projet HAIKUS
- **Fondements cognitifs**: Intégration multisensorielle et émotion; Musique et plasticité cérébrale; Perception de la distance en Réalité Augmentée
- **Création / Médiation**: Auralisation de salle Projet RASPUTIN; Composition Urbaine et Paysagère; Étude acoustique du temple de Dendara; Synthèse de la directivité par corpus
- **Logiciels**: Spatialisateur; OSCar; Panoramix; ADMix Tools

Collaborations

—

Amadeus (FR), AALTO (FIN), ARI-ÖAW (Wien, AUT), CNSMDP (FR), Greenwich Univ. (GB), Flux:: (FR), HEGP (FR), HISOMA (FR), IFAO (FR), IRBA (FR), LAM-IJLRA (FR), LORIA (FR), McGill (CAN), MPIA-IJLRA (FR), NOTAM (NW), NYU (US), Radio France (FR), RPI (Troy, US), RWTH Aachen (D), Univ. Lille (FR).

Équipe

—

Chercheurs: B. Alary, M. Noisternig, I. Viaud-Delmon (CNRS), O. Warusfel
Ingénieurs: T. Carpentier (CNRS), J. C. Vincent (CNRS) – **Doctorants**: M. Gospodarek (NYU) – **Résidences/invités/collaborations longues**: L. Hobeika, N. Schütz, C. Suied, M. Taffou.

Équipe Perception et design sonores (PDS)

Responsable : Nicolas Misdariis

Activités

—

Le projet scientifique de l'équipe Perception et design sonores combine des recherches en perception et cognition sonores avec des recherches et des applications en design sonore. Il met en œuvre des relations art/science, d'une part, en alimentant la recherche musicale et sonore avec des questionnements scientifiques d'ordre perceptif et cognitif (axe 'du son au sujet'), et d'autre part, en ancrant les problématiques de la création sonore appliquée dans le champ de la recherche en design (axe 'du sujet au son').

La composante applicative permet, quant à elle, de développer des actions de type recherche-projet ou recherche-action, dans le cadre de thèses ou de collaborations industrielles (Renault, Krug, SNCF...) qui allient « savoir » scientifique et technologique des chercheurs.e.s de l'équipe et « savoir-faire » artistique des compositeurs et compositrices systématiquement associés.e.s (A. Cera, S. Gaxie, A. Sigman, R. Rivas, N. Schütz...).

Ce cadre de travail est enfin complété par une dimension pédagogique essentielle qui, grâce à son implication dans le DNSEP Design Sonore de l'École supérieure d'art et design TALM-Le Mans, permet à l'équipe de contribuer activement à la formation et la structuration de la discipline. Cette implication passe notamment par l'encadrement d'un workshop applicatif annuel qui rassemble des étudiants en design sonore et en design autour d'un cas concret apporté par un partenaire industriel ou institutionnel (Ville du Mans, Région Ile-de-France, RATP, Maison de la Radio, Hôpital Sainte-Anne...).

Les recherches menées sur l'axe 'du son au sujet' portent en premier lieu – et historiquement – sur les sons environnementaux, objet singulier qui permet d'accéder à plusieurs niveaux de représentation cognitive, en lien avec des modes d'écoute distincts (écoute réduite, causale, sémantique). Elles sont mises en œuvre à différentes échelles temporelles (de l'événement à la scène sonore), à différents degrés de complexité (du son pur au son complexe), et visent autant la caractérisation de mécanismes perceptifs de bas-niveau (sonie, saillance auditive...), que la compréhension de processus cognitifs de plus haut niveau impliquant la reconnaissance, l'identification, la catégorisation, la mémorisation ou le traitement émotionnel de l'information sonore. Sur les dimensions mémorielles et émotionnelles, les recherches de l'équipe se déplacent sur le terrain de la perception multimodale en s'intéressant notamment aux relations entre la modalité auditive et olfactive (thèse CIFRE en cours de montage avec IFF).

Les recherches menées sur l'axe 'du sujet au son' étendent la catégorie de sons environnementaux à celle d'artefacts sonores – issus d'une transformation humaine, d'origine artificielle – et posent la question de leur conception (design) mais aussi de leur réception (perception). Elles s'intéressent donc globalement à la discipline du design sonore, considérée comme un objet d'étude et de recherche à part entière, en l'intégrant dans le champ de la recherche en design, cadre conceptuel plus large et plus établi d'un point de vue historique et épistémologique. Cette volonté de faire interagir le design sonore avec le design, et d'en acquérir des connaissances, méthodologies ou outils propres, prend le parti de s'inspirer de l'approche des sciences du design formalisée notamment par Nigel Cross (2006). Cette approche se déploie sur des dimensions relatives aux acteurs ('people'), procédures ('process') et produits ('products') de la discipline; elle se complète, en outre, par une prise en

compte spécifique des modalités de réception des 'objets' (artefacts) de design sonore – qu'ils soient de nature tangible, digitale ou spatiale –, ce qui induit notamment la problématique de l'évaluation perceptuelle. Enfin, l'équipe est partie prenante dans l'axe thématique – et transverse au laboratoire – « Son, Musique / Santé », à travers des projets sur l'utilisation du son et de la musique dans la réduction de l'anxiété (projet Psyson), l'utilisation du feedback vocal dans des protocoles de psychothérapie (thèse N. Guerouaou), ou bien le design de dispositifs audio-tactiles pour l'écoute par voie extra-tympanique (thèse C. Richards).

Thématiques et projets associés :

—

- Perception, représentation et description des sons : HearDevComp
- Sonie multi-source
- Imitation vocale et identification
- Acoustique et sémantique du timbre : Speak
- Perception / cognition des sons complexes
- Perception locale / globale
- Cognition et émotion vocales et musicales
- Cognition sociale et identité vocale : CREAM ; ACTIVATE
- Corrélation inverse et prosodie : SEPIA
- Cognition et créativité musicale : GRIAMI
- Design et interaction sonores
- Design sonore multimodale, sonification
- Design sonore et santé : Psy-son
- Outils d'aide à la conception sonore : Speak ; SkataRT

Domaines de compétence

—

Psychologie expérimentale, analyse de données (numériques, verbales), acoustique, psychoacoustique, psychologie et neurosciences cognitives, design sonore, design.

Collaborations

—

LMA (CNRS), Ecole Centrale Nantes (LS2N), Sorbonne Université (Institut Jean le Rond d'Alembert), Université de Strasbourg, EPFL (Lausanne), TU Delft (Faculté Industrial Design Engin.), ENAC, ENS Ulm (LSP), CAP (Université Lyon 1), CHU Paris Psychiatrie & Neurosciences (Neuro Sainte-Anne et Lab-ah), CHU Lille, ENAC, Université Waseda de Tokyo, McGill University (Canada), KTH (Suède), ZHdK (Suisse), ENS Paris Saclay (CRD), ENSCI Les Ateliers, Music Unit, Groupe Renault, Actronika, McGill Univ., CIRMMT /Schulich School of Music, Université de Gand, Université de Mayence.

Équipe

—

Chercheurs : M. Ayari (Université de Strasbourg), O. Houix (ESAD TALM Le Mans), N. Misdariis, E. Ponsot (CNRS), P. Susini – **Post-doctorants :** E. Frid – **Doctorants :** C. Richards, T. Souaille (co-dir. LS2N), N. Guerouaou (co-dir. CHU Lille), B. Bouvier, M. Duroyon, V. Fraisse (co-tutelle McGill Univ.) – **Résidences/collaborations artistiques :** A. Cera, R. Rivas, N. Schütz.

Équipe Analyse et synthèse des sons (AS)

Responsable : Axel Roebel

Activités

—

L'équipe Analyse et synthèse des sons effectue des recherches et des développements en analyse, transformation et synthèse des signaux sonores.

L'analyse de sons comprend les méthodes permettant l'extraction ou la structuration automatique de divers types d'informations provenant du signal, comme la fréquence fondamentale ou les évolutions spectrales déterminant la hauteur et le timbre du son perçu.

Des informations non strictement musicales sont également prises en compte et intéressent des domaines tels que l'acoustique industrielle, le design sonore et le multimédia. On peut citer, en particulier, l'indexation automatique des enregistrements sonores. Les méthodes utilisées reposent sur le traitement du signal, l'analyse statistique, les techniques d'apprentissage et l'apprentissage profond, la reconnaissance des formes mais aussi sur la connaissance de la perception auditive.

Les techniques de transformation et la synthèse des sons sont d'abord conçues pour répondre aux demandes des musiciens pour la création de nouveaux sons et de nouvelles musiques. Un exemple typique en est la synthèse d'un chœur virtuel par un ordinateur sur la scène d'un opéra. Ces travaux trouvent également de nombreuses applications dans des domaines tels que la téléphonie mobile, les jeux vidéo, le cinéma ou la réalité virtuelle en général. Analyse et synthèse reposent sur la conception, d'une part, de modèles de signaux (modélisation des effets des sons produits en termes de signaux) et d'autre part, des modèles sous forme de réseaux profonds appris sur des grandes bases de données qui souvent sont générés ou augmentés à l'aide de modèles de signaux. Ces modèles sont réalisés sous la forme de bibliothèques et logiciels (Mac OS x, Windows et Linux), spécifiquement conçues à l'intention d'utilisateurs professionnels ou non, musiciens, mais aussi ingénieurs du son, acousticiens et amateurs.

Domaines de compétence

—

Traitement du signal, statistiques, théorie de l'information, techniques d'apprentissage, analyse numérique, modélisation.

Thématiques et projets associés :

—

- Modèles de signaux : Traitements par vocodeur
- Caractérisation des sons : Indexation automatique ; AI4Media
- Analyse, caractérisation et transformation des enregistrements musicaux : AQUA-RIUS
- Analyse, transformation et synthèse de la voix, conversion des locuteurs : projets theVoice ; ARS
- Analyse de scènes sonores : projet ROUTE
- Logiciels : IRCAMLab TS ; ISiS Singing Synthesis

Collaborations

—

Aristotle University of Thessaloniki (Greece), Athens Technology Center (Greece), bmat (Espagne), Centre for Research and Technology Hellas (Greece), Flux::Audio, France Télécom R&D / Orange, Fraunhofer Gesellschaft ICMT (Allemagne), IBISC/Université d'Evry, ISIR/Sorbonne Université, LAM/Sorbonne Université, LIA (Université d'Avignon), MICC Univ. Florence (Italie), PSA (Paris), Sony Music France, Stupeflix, UMG, Univ. Huddersfield (Royaume-Uni), Uni. Lumière Lyon 2.

Équipe

—

Chercheurs : A. Cohen-Hadria (Sorbonne Université), F. Bous, G. Doras, R. Mignot, N. Obin (Sorbonne Université), A. Roebel
Ingénieurs : F. Cornu
Doctorants : M. Farès, A. Lavault, L. Renault, C. Le Moine Veillon, L. Salais, Y. Teytaut.

Équipe Interaction son musique mouvement (ISMM)

Responsable : Frédéric Bevilacqua

Activités

—

L'équipe Interaction son musique mouvement (précédemment : Interactions musicales temps réel) mène des recherches et des développements sur les systèmes interactifs dédiés à la musique et au spectacle vivant.

Nos travaux concernent toute la chaîne du processus interactif, comprenant la captation et l'analyse des gestes et des sons, les outils de gestion de l'interaction et de la synchronisation, ainsi que des techniques de synthèse et traitement sonore temps réel. Ces recherches, et leurs développements informatiques associés (MuBu for Max, CataRT, Soundworks et librairies Web), sont généralement réalisés dans le cadre de projets interdisciplinaires, intégrant scientifiques, artistes, pédagogues, designers et trouvent des applications dans des projets de création artistique, de pédagogie musicale, d'apprentissage du mouvement, ou encore dans des domaines médicaux tels que la rééducation guidée par le son et la musique.

Thématiques et projets associés

—

- Modélisation des gestes et des mouvements : ce thème regroupe le développement de modèles basés sur des études expérimentales, concernant généralement l'interaction entre son et mouvement, allant du jeu d'instrumentistes aux mouvements dansés, (cf. Analyse et reconnaissance du geste ; Instruments augmentés
- Synthèse et traitement sonore interactif : ce thème regroupe essentiellement des méthodes de synthèse et traitement sonore basées sur des sons enregistrés ou de large corpus sonores (cf. Synthèse concaténative par corpus ; AQUA-RIUS
- Systèmes interactifs sonores fondés sur le geste et nouveaux instruments : ce thème concerne le design et le développement d'environnements sonores interactifs utilisant gestes, mouvements et toucher. L'apprentissage machine interactif est l'un des outils développés spécifiquement (cf. ELEMENT ; Systèmes distribués sonores interactifs ; Rapid-Mix ; Skat-VG
- Interaction musicale collective et systèmes distribués : ce thème aborde les questions d'interactions musicales allant de quelques utilisateurs à des centaines. Il concerne plus particulièrement le développement d'un environnement Web mêlant ordinateurs, smartphones et/ou systèmes embarqués, pour explorer de nouvelles possibilités d'interactions expressives et synchronisées (cf. Systèmes distribués, DOTS)

Projets nationaux et européens

—

- AQUA-RIUS
- ELEMENT
- DOTS
- SkAT-VG

Domaines de compétence

—

Systèmes interactifs sonores, interaction humain-machine, captation du mouvement, modélisation du son et du geste, analyse et synthèse sonore temps réel, modélisation statistique et apprentissage machine interactif, traitement de signal, systèmes interactifs distribués.

Collaborations

—

Arizona State University (États-Unis), Atelier des feuillantines, BEK (Norvège), CNMAT Berkeley (États-Unis), ENSAD, ENSCI, GRAME, HKU (Pays-Bas), Hôpital Pitié-Salpêtrière, IEM (Autriche), ISIR-CNRS et LIB Sorbonne Université, Little Heart Movement, Mines-Paris Tech, Mogeas (Royaume-Uni/Italie), No Design, Motion Bank (Allemagne), université Pompeu Fabra (Espagne), UserStudio, université de Genève (Suisse), LIMSI-CNRS université Paris-Saclay, LRI-CNRS université Paris-Saclay, GeorgiaTech (USA), Legacy Lab (Tawain), NOTAM (Norvège), Orbe.mobi, PLUX (Portugal), ReactTable Systems (Espagne), UCL (Royaume-Uni), Univers Sons/Ultimate Sound bank, université Paris 8, Université Nanterre, Universidad Carlos III Madrid (Espagne), université de Gênes (Italie), université McGill (Canada), TU Berlin (Allemagne), ZhDK (Suisse).

Équipe

—

Chercheurs et développeurs : F. Bevilacqua, R. Borghesi, B. Matuszewski, D. Schwarz, – **Doctorantes :** A. Golvet, V. Paredes., I. Peyre, M. Voillot – **Résidence/invités/colloborations longues :** B. Caramiaux (CNRS), J. Françoise (CNRS), S. Fdili Aloui, A. W. Liu, M. Suarez Cifuentes, M. A. Magalhaes, L. Bianchi.

Équipe Représentations musicales (RepMus)

Responsable : Gérard Assayag

Activités

—

L'équipe Représentations musicales travaille sur les structures formelles de la musique et les environnements créatifs pour la composition et l'interaction musicales.

Ces travaux mènent à des applications dans les domaines de la composition assistée par ordinateur (CAO), de la performance, de l'improvisation, de l'interprétation et de la musicologie computationnelle. La réflexion sur les représentations de haut niveau des concepts et des structures musicales, appuyée sur les langages informatiques originaux développés par l'équipe, débouche sur l'implantation de modèles qui peuvent se tourner vers l'analyse musicale comme vers la création, vers la composition comme la performance et l'improvisation.

Sur le versant musicologique, les outils de représentation et de modélisation permettent une approche véritablement expérimentale qui dynamise de manière significative cette discipline.

Sur le versant création, l'objectif est de concevoir des compagnons musicaux qui interagissent avec les compositeurs, musiciens, ingénieurs du son... dans toutes les phases du workflow musical. Les logiciels développés ont pu être diffusés vers une communauté importante de musiciens, concrétisant des formes de pensée originales liées à cette caractéristique particulière des supports informatiques qu'ils peuvent représenter (et exécuter) à la fois la partition finale, ses divers niveaux d'élaboration formelle, ses générateurs algorithmiques, ses réalisations sonores, et qu'ils permettent d'interagir dans le vif de la performance et notamment de l'improvisation.

L'équipe intègre l'interaction symbolique et la créativité artificielle à travers ses travaux sur la modélisation de l'improvisation et l'intégration de nouvelles formes compositionnelles ouvertes et dynamiques. Ces recherches permettent des avancées en intelligence artificielle, avec des modèles d'écoute, d'apprentissage génératif, de synchronisation, et posent les fondements de nouvelles technologies d'agents créatifs qui peuvent devenir des compagnons musicaux dotés de musicalité artificielle (« machine musicianship »). Cela préfigure la dynamique de coopération inhérente aux réseaux cyber-humains et l'émergence de formes de co-créativité homme-machine.

L'équipe a une longue histoire de collaboration intensive avec des compositeurs et musiciens, qu'ils soient internes ou externes à l'Ircam. Les trois tomes de l'ouvrage *OM Composer's Book* archivent ces travaux et garantissent leur diffusion internationale et leur pérennité.

Thématiques et projet associés

—

- Composition assistée : Composition assistée par ordinateur, logiciel OpenMusic
- Orchestration : Orchestration assistée par ordinateur ; projet ACTOR ; logiciel Orchid*
- Contrôle de la synthèse et de la spatialisation, écriture du temps : Composition assistée par ordinateur ; Écriture du temps synchrone ; logiciel OpenMusic ; Antescofo

- Mathématique et musique
- Langages informatiques pour la création : logiciels OpenMusic ; Antescofo
- Dynamiques de l'interaction improvisée, co-créativité : projets DYCI2 ; MERCI ; REACH ; Logiciels Somax2 et DYCI2,
- Intelligence Artificielle Créative : projet ACIDS
- Études sur les structures musicales dans la performance : projet COSMOS
- Études en lien avec la thérapeutique : projet HEART.FM

Domaine de compétence

—

{Composition, analyse, performance, improvisation, orchestration} assistés par ordinateur, musicologie computationnelle, intelligence artificielle créative, langages informatiques musicaux, mathématiques musicales, langages temps réel synchrones, notations exécutables, architectures d'interaction, (co-)créativité computationnelle.

Collaborations

—

Bergen Center for Electronic Arts (NO), CIRMMT/ McGill University (CA), City University London, Conservatoire national supérieur de musique et de danse de Paris, Columbia New York, CNMAT/UC Berkeley, Electronic Music Foundation, Gmem, Grame Lyon, HEM Genève, École normale supérieure, EsMuC Barcelone, Harvard University, Inria, Sorbonne Université, université de Strasbourg, université Paris-Sud Orsay, UCSD San Diego, université de Yale, université du Minnesota, Barts Heart Centre, University College London, GREAM Strasbourg, Institut Pasteur, Jardin des Sciences Strasbourg, université de Padoue, EHES, HyVibe, université de Tokyo, Geidai University of the Arts, université de Kyoto, Metropolitan University Tokyo, STEGI Onassis Cultural Center, université d'Athènes, université de Pennsylvanie, Université catholique de Louvain.

Équipe

—

Chercheurs : C. Agon (PR Sorbonne Université), G. Assayag (DR Ircam), E. Chew (DR CNRS), P. Esling (MC Sorbonne Université), J.-L. Giavitto (DR CNRS), K. Haddad (CR Ircam, compositeur), M. Malt (CR Ircam), G. Bloch (MC U. Strasbourg, chercheur associé), M. Andreatta (DR CNRS, U. Strasbourg, associé) – **Postdocs, ingénieurs :** Joakim Borg (Ing. Reach) – **Chercheurs invités, résidences :** C. Malherbe (compositeur associé), M. Stroppa (compositeur associé) – Joëlle Léandre (résidence), Hervé Sellin (résidence), Rémi Fox (résidence) – **Doctorants :** C. Douwes, D. A. Bedoya, P. Lascabettes, N. Devis, G. Romero, Giovanni Bindi

Équipe Analyse des pratiques musicales (APM)

Responsable : Clément Canonne

Activités

—

L'équipe Analyse des Pratiques Musicales mène des recherches interdisciplinaires sur le monde de la musique – de la production (composition, improvisation, interprétation), à la réception (écoute, concert, critique, discours scientifique et technique), en passant par les multiples objets techniques qui le constituent (instruments, enregistrements, algorithmes).

Pour ce faire, elle développe une approche à la fois empirique (en empruntant ses outils aux sciences sociales comme aux sciences cognitives), intégrative (en connectant les objets, pratiques, et discours musicaux à des questions théoriques d'ordre général), et contextuelle (en abordant les pratiques considérées en tant que complexités dynamiques et socialement situées).

Les travaux de l'équipe sont diffusés par des publications et colloques au sein des différentes communautés scientifiques concernées (musicologie, histoire, philosophie, esthétique, sciences cognitives, sociologie, critique génétique, anthropologie, étude des sciences et des techniques); ils font aussi l'objet d'autres formes de dissémination: courts-métrages documentaires, enseignements universitaires et auprès d'un plus large public, développements informatiques, conférences-performances, enregistrements, etc.

Thématiques et projets associés

—

- Analyse des processus de création
- Pratiques musicales et action collective
- Histoire des pratiques musicales et scientifiques et de leurs interactions
- Musique mixte et électroacoustique
- Performance studies
- Histoire et esthétique des avant-gardes musicales
- Sound Studies
- Musicologie & SHS
- Esthétique philosophique
- Temporalité et spatialité musicales de la conscience morbide
- Dire la musique
- Instruments of Improvisation
- Histoire de l'acoustique musicale en France
- Accorder le monde

Collaborations

—

Centre de recherches sur les arts et le langage (EHESS, CNRS), CIRMMT/ McGill University (Canada), Social Mind and Body Group/ Central European University (Autriche), Conservatoire national supérieur de musique et de danse de Paris, Faculté de musique de l'université d'Oxford, Haute École de musique/conservatoire de Genève, Institut des textes et manuscrits modernes (CNRS-ENS Ulm), OICRM/université de Montréal, New-York University/Department of Music (États-Unis)

Équipe

—

Chercheurs : C. Canonne (CNRS), F.-X. Féron (CNRS), L. Feynerou (CNRS), P. Saint-Germier (CNRS) – **Chercheurs associés :** Alain Bonardi

3



Les projets



C1. l'atelier du son

AI4Media



Équipe concernée : Analyse et synthèse des sons

Financement : H2020, Commission européenne

Période de réalisation : septembre 2020 – août 2024

Partenaires extérieurs : Coordinateur CERTH pour 30 partenaires

Face à l'évolution rapide du monde des médias et de l'information, le projet AI4Media a pour but de développer des outils d'intelligence artificielle pour les médias, le multimédia et l'audiovisuel. Par exemple, afin d'aider nos démocraties européennes à lutter contre la désinformation, le projet vise à proposer des analyses de médias pour la détection de dernières nouvelles, la détection de fausses informations, ou la validation d'information pour les journalistes. Également, le projet s'intéresse à des applications artistiques et d'aide à la création de jeux vidéo ; le travail de l'équipe Analyse-Synthèse se place dans ce contexte. Grâce à des techniques avancées d'intelligence artificielle, nous développerons une nouvelle approche de synthèse sonore réaliste d'instruments de musique. Cette méthode rendra possible la production de contenus musicaux de qualité à partir de partitions au format MIDI. Cela servira d'une part à des applications artistiques ou de jeux vidéo, et d'autre part à l'apprentissage supervisé de modèles d'analyse automatiques d'enregistrements audio : estimation du tempo, détection de la clef, reconnaissance des accords, transcription automatique, séparation de sources.

Analyse-synthèse de SRIR

Équipe concernée : Espaces acoustiques et cognitifs

Collaborations extérieures : université Aalto (Helsinki, FIN).

L'analyse et le traitement de réponses impulsionnelles spatiales (SRIR) est une thématique centrale de l'équipe. Le travail se focalise sur la modélisation et la resynthèse du champ de réverbération tardif. Dans ce contexte, un processus de débruitage robuste a été développé afin de traiter la base de données de SRIR de l'équipe. Quand la queue de réverbération tardive présente des propriétés de champ diffus (incohérence des ondes planes, distribution spatiale de puissance isotrope), le débruitage par resynthèse peut être réalisé dans le domaine des harmoniques sphériques, représentation spatiale à la base du format HOA. Le domaine des harmoniques sphériques préserve la nature spatiale incohérente du champ diffus (grâce à l'orthogonalité de la base de décomposition). La décroissance exponentielle de puissance est alors modélisée composante par composante et fréquence par fréquence, opération au cours de laquelle le seuil de bruit est simultanément détecté. Une prolongation de la queue de réverbération peut enfin être synthétisée et utilisée pour remplacer ce seuil de bruit. Dans le cadre d'une collaboration avec l'université Aalto une étude est consacrée à l'analyse de la directivité du champ tardif dans le cas où celui-ci n'est pas entièrement diffus, i.e. où les ondes planes sont incohérentes mais la distribution spatiale de puissance est anisotrope. Dans ce cas, il n'est plus possible de resynthétiser la queue de réverbération en agissant seulement sur l'enveloppe de puissance dans le domaine des harmoniques sphériques, et une représentation directionnelle de l'IR est nécessaire (décomposition en ondes planes). Le nombre de directions sur lesquelles l'IR peut être décomposée est limité par le nombre d'harmoniques sphériques puisque l'indépendance linéaire des signaux générés doit être garantie. La répartition de ces directions sur la sphère est choisie afin de maximiser leur différence angulaire, pour que les signaux soient les plus indépendants possibles, et minimiser la variance de la directivité totale, pour qu'aucun sous-espace de la sphère ne soit privilégié par rapport à d'autres. L'opération de débruitage est appliquée direction par direction, préservant à la fois l'incohérence spatiale et la directivité du champ tardif. Ce projet se poursuit par l'extraction de descripteurs des SRIRs permettant de concevoir un répertoire de manipulations intéressantes dans le cadre d'applications artistiques exploitant des SRIRs (cf. Composition urbaine et paysagère)

AQUA-RIUS

Analyse de la qualité audio pour représenter, indexer et unifier les signaux

—

Équipes concernées : Analyse et synthèse des sons,

Interaction son musique mouvement

Financement : ANR CE23

Période de réalisation : janvier 2023 - juin 2026

Partenaires extérieurs : Coordinateur IBISC/Université d'Evry,

Télécom-Paris

Le projet AQUA-RIUS d'une durée de 42 mois, à compter du 1^{er} janvier 2023, a pour objet l'étude, l'analyse automatique et le traitement de la qualité audio dans les enregistrements audio. Dans le cadre du projet, la qualité audio est définie comme étant l'ensemble des caractéristiques sonores d'un enregistrement musical ne dépendant pas des sources sonores. Cela dépend donc des effets du mixage (tels que: filtrage, compression de dynamique, réverbération) et les dégradations (par exemple: ajout de bruit, codage audio avec pertes, saturation). Ce projet a pour but, d'une part, l'analyse et la description de la qualité audio et des dégradations (d'un point de vue signal et d'un point de vue psycho-acoustique) et, d'autre part, la simulation et la modification de la qualité audio, incluant par exemple le remixage d'un enregistrement. Profitant des récentes avancées en apprentissage automatique, l'apprentissage profond sera pleinement exploité, avec notamment le développement d'une librairie différentiable de traitements audio, et l'apprentissage de réseaux génératifs. Les partenaires du projet sont: l'IBISC/Université d'Evry (Dominique Fourer le coordinateur du projet, Hichem Maaref), l'Ircam (Rémi Mignot, et Diemo Schwarz), et Télécom Paris (Geoffroy Peeters).

ATRIM

Analyseur temps réel haute précision de justesse et de timbre pour instruments musicaux

—

Équipe concernée : Systèmes et signaux sonores :

audio/acoustique, instruments

Financement : ANR-Plan de Relance

Période de réalisation : mai 2022 – avril 2024

Partenaire : Buffet-Crampon

Le projet ATRIM vise à développer une analyse en temps réel de haute précision de la hauteur et du timbre des instruments de musique à vent. L'approche consiste à concevoir et à explorer des rendus visuels réactifs des sons produits par les musiciens, sous des formes qui fournissent un retour informatif précis pendant la performance musicale.

Un premier objectif est de développer des outils fiables pour les musiciens testeurs, à la fin du cycle de production des instruments de musique.

Un deuxième objectif est de concevoir un retour visuel pour aider les musiciens à ajuster leur contrôle moteur sur leurs propres cibles d'intonation, de vibrato, de glissando, de timbre, etc.

À cette fin, le travail principal est consacré à l'élaboration de nouveaux indicateurs et de rendus visuels basés sur des estimations spectrales précises utilisées dans l'analyseur Snail, ainsi qu'à des interfaces graphiques conçues en collaboration avec des musiciens experts.

AVATARS

Artificial Voice production: control of bio-inspired port-hAmilToniAn numeRical and mechatronic modelS

—
Équipe concernée: Systèmes et signaux sonores: audio/acoustique, instruments

Financement: Projet ANR

Période de réalisation: mars 2023 – mars 2027

Partenaires: GIPSA-LAB (Grenoble), LMA (Marseille), CHU de Liège (Belgique)

Ce projet étudie la production de la voix humaine et les régimes oscillatoires résultant de l'articulation du larynx et du conduit vocal (impliqués dans les registres de voix de poitrine, de fausset, etc.). Il se fonde sur la modélisation physique et la théorie des systèmes non linéaires pour concevoir des simulations, observateurs-contrôleurs et analyses de bifurcations d'appareils vocaux. Sa première spécificité est d'élaborer ces outils sous l'hypothèse systématique de bilans de puissance garantis (écartant les difficultés d'instabilités non physiques). La seconde est de les compléter, nourrir et évaluer par des mesures in vivo et sur un ensemble cohérent de bancs mécatroniques (contrôle de larynx ex vivo, biomimétiques ou standardisés, et de conduits vocaux artificiels). L'objectif est de comprendre, reproduire et analyser scientifiquement la voix (sur des cas sains ou pathologiques) grâce à des avatars virtuels et robotiques, avec de possibles retombées artistiques et médicales.

BRUEL

—
Équipe concernée: Analyse et synthèse des sons

Financement: Projet ANR

Période de réalisation: octobre 2022 – octobre 2026

Coordinateur: Laboratoire d'informatique d'Avignon (LIA)

Partenaires: Commissariat à l'énergie atomique et aux énergies alternatives (CEA), Eurecom Sophia Antipolis, Service national de police scientifique

Le projet BRUEL concerne l'évaluation/certification des systèmes d'identification par la voix face aux attaques adverses.

En effet, les systèmes de reconnaissance automatique du locuteur sont vulnérables non seulement à la parole produite artificiellement par synthèse vocale, mais aussi à d'autres formes d'attaques telles que la conversion d'identité vocale et la relecture. Les artefacts créés lors de la création ou la manipulation de ces attaques frauduleuses constituent les marques laissées dans le signal par les algorithmes de synthèse vocale permettant ainsi de distinguer la voix réelle originale d'une voix usurpée. Dans ces conditions, la détection de l'usurpation d'identité requiert d'évaluer les contre-mesures d'usurpation d'identité en même temps que les systèmes de reconnaissance du locuteur. Le projet BRUEL ambitionne de proposer la première méthodologie d'évaluation/certification des systèmes d'identification par la voix basée sur une approche Critères Communs.

Continuum

—
Équipes concernées : Direction Innovation et moyens de la recherche, Espaces acoustiques et cognitifs, Direction Création Production
Programme : PIA4-ICC,
appel « Expérience augmentée du spectacle vivant »
Période de réalisation : octobre 2022 - septembre 2025
Partenaires : Amadeus, VRtuoz

Continuum désigne une nouvelle conception de la production et de la diffusion du spectacle vivant augmenté dans ses dimensions sonores : un continuum entre scènes réelles et virtuelles, entre artiste et spectateur-visiteur-auditeur d'aujourd'hui et entre innovation technologique française et productions culturelles aux formats innovants. Porteur d'un nouveau standard qualitatif de l'immersion sonore, ce programme à la pointe de l'état de l'art de la recherche technologique s'attache au développement d'une chaîne de production complète, de la captation à la restitution finale, permettant de créer et de transmettre aux spectateurs un contenu spatialisé sur différentes plateformes d'écoute et d'interactions individualisées et de le diffuser dans des lieux de différentes configurations (salles de spectacle, espaces publics ou privés). Les fonctions visées sont expérimentées et validées à travers un ensemble de créations remarquables et leur commercialisation concourt à leur démocratisation.

Le projet Continuum est coordonné par l'Ircam, en collaboration avec les sociétés Amadeus et VRtuoz. Il est soutenu par l'État dans le cadre du dispositif « Expérience augmentée du spectacle vivant » de la filière des industries culturelles et créatives (ICC) de France 2030, opérée par la Caisse des Dépôts.



Écoute binaurale

—
Équipe concernée : Espaces acoustiques et cognitifs

Les appareils personnels (smartphones, tablettes) concentrent de multiples fonctions et représentent le principal vecteur de diffusion de contenus musicaux. À mesure que l'audience dérive vers ces nouveaux équipements, l'écoute binaurale sur casque d'écoute gagne du terrain et symbolise la notion d'expérience personnelle en donnant accès en théorie à la reproduction de scènes tridimensionnelles. Cette reproduction binaurale est basée sur le filtrage dynamique de la source sonore par les fonctions de transfert de la tête (Head-Related Transfer Function, HRTF) préalablement mesurées sur la tête d'un auditeur ou d'un mannequin. Cependant, la dépendance individuelle de ces HRTFs a jusqu'à présent limité sa diffusion grand public.

À l'occasion du projet BiLi (financement FUI), l'environnement de mesures des HRTFs dans la chambre anéchoïque de l'Ircam a fait l'objet d'une mise à jour matérielle et logicielle permettant d'augmenter considérablement la résolution spatiale par rapport aux bases de données disponibles. Le nouvel échantillonnage spatial (1680 directions) autorise une décomposition en harmoniques sphériques d'ordre élevé, utile pour diverses phases d'exploitation (interpolation spatiale, transcodage HOA / binaural).

Le partage des bases de données de HRTFs au sein de la communauté internationale a motivé la démarche de standardisation d'un format d'échange. Issu d'une collaboration internationale, le format SOFA (Spatially Oriented Format for Acoustics), approuvé par l'instance de standardisation de l'Audio Engineering Society, permet le stockage de données acoustiques spatiales telles que des HRTFs ou les SRIRs. L'Ircam a mis en place un serveur OPenDAP (Open-source Project for Network Data Access Protocol) qui héberge différentes bases de données HRTFs au format SOFA et auquel des applications clientes (applications web, Matlab, etc.) peuvent adresser des requêtes de téléchargement (une HRTF spécifique, une tête complète, etc.).

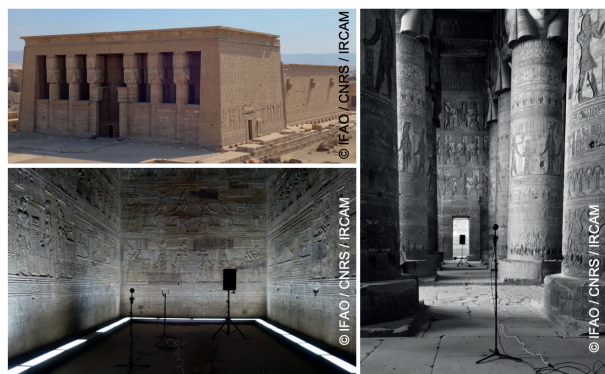
Le travail sur l'individualisation des HRTFs se poursuit, notamment sous forme d'estimation des HRTFs ne requérant ni mesures acoustiques ni mesures morphologiques. Certaines méthodes exploitent les bases de données désormais disponibles pour guider l'utilisateur dans la sélection du jeu de HRTFs le plus approprié (projet RASPUTIN). Une nouvelle voie proposée s'appuie sur le recours aux méthodes d'apprentissage profond (projet HAIKUS) et l'analyse en aveugle d'enregistrements binauraux effectués sur l'auditeur en conditions non contrôlées (milieu réverbérant, signaux quelconques, sources et auditeur en mouvement).

Étude acoustique du temple de Dendara

Équipe concernée : Espaces acoustiques et cognitifs

Partenaires extérieurs : HISOMA, IFAO

Le CNRS et l'IFAO (Institut français d'archéologie orientale) ont entrepris, en collaboration avec l'Ircam, l'étude archéo-acoustique du temple principal de Dendara (projet dirigé par Sibylle Emerit, chercheuse du laboratoire CNRS HISOMA et ancien membre de l'IFAO). Situé en Haute-Égypte, et construit entre la fin de l'époque ptolémaïque et le début de l'époque romaine, ce temple est dédié à Hathor, déesse de la musique, de l'amour et de l'ivresse. Les études épigraphiques et iconographiques attestent de l'importance de l'activité sonore dans ce temple que ce soit au travers des mentions textuelles ou par l'omniprésence de représentations de divinités ou de danseurs et musiciens vénérant Hathor par leurs chants et manifestations instrumentales (tambourins, harpes, sistres, luths) qui ornent les parois et les colonnes de différents espaces. C'est la première fois qu'une telle étude, au carrefour de l'anthropologie, de l'archéologie, de l'architecture, de la musique et de l'acoustique, est consacrée à un temple égyptien. L'enjeu est la compréhension de la manière dont les anciens percevaient et interprétaient les phénomènes sonores par l'analyse du vocabulaire consacré aux sons, par l'étude des instruments de musique et l'observation des sites dans lesquels se déroulaient des manifestations riches en événements sonores. Les études architecturales se concentrent généralement sur la dimension visuelle du bâti et plus rarement sur sa dimension acoustique à l'exception des théâtres antiques grecs et romains. Dans ce projet, la mission de l'Ircam est de mener une caractérisation acoustique des différents espaces du temple par la mesure de réponses impulsionnelles. Ces réponses, enregistrées par un réseau microphonique à haute-résolution spatiale (microphone EM32 de MHAcoustics®) seront utilisées ultérieurement pour calibrer un modèle numérique 3D, en cours de réalisation, qui permettra de compléter l'étude par la prise en compte d'éléments architecturaux aujourd'hui disparus (portes du sanctuaire et des différentes chapelles, cour péristyle).



Étude acoustique du temple de Dendara. Mesure dans le Naos (en bas) et le Pronaos (à droite).

FuturePulse



Équipe concernée : Analyse et synthèse des sons

Financement : programme H2020 ICT-19-2017

Période de réalisation : septembre 2017 – août 2020

Partenaires : Bmat (coordinateur, Espagne), Athens Technology Center (Grèce), Musimap (Belgique), Playground (Suède), Bass Nation (Finlande), Soundtrack Your Brand (Suède)

L'industrie musicale connaît actuellement une importante mutation liée à la montée en charge du streaming, des médias sociaux et des technologies de convergence. La disponibilité de très grands catalogues fait des fonctions de recommandation et de découverte des éléments clés de compétitivité entre acteurs commerciaux, tandis que l'accès en continu à de multiples sources de consommation musicale donne naissance à des pratiques dynamiques, caractérisées par une grande diversité de goûts et une volatilité des préférences en fonction du contexte d'écoute. Afin de répondre aux besoins de plus en plus complexes de l'écosystème musical, le projet FuturePulse s'attache au développement d'une plateforme technologique de marketing analytique destinée à trois secteurs de l'industrie musicale à fort impact :

- l'édition phonographique ;
- la production de concerts ;
- les plateformes de diffusion en ligne.

L'enjeu du projet est d'aider les acteurs du domaine à tirer parti d'une variété de données, allant des contenus de diffusion (télévision, radio) et de streaming aux statistiques de ventes et aux flux de discussions, d'interactions et de contenus sur les médias sociaux, au moyen d'analyses élaborées et de services de modélisation prédictive pour prendre des décisions commerciales éclairées, mieux comprendre leur public et les tendances futures de la musique et aller vers une meilleure efficacité et rentabilité de la distribution musicale.

L'Ircam intervient dans le projet par l'adaptation et le perfectionnement de ses technologies d'indexation, permettant l'extraction automatisée d'informations à partir d'enregistrements musicaux.

Infidhem

—
Systèmes interconnectés de dimension infinie
pour les milieux hétérogènes

Équipe concernée : Systèmes et signaux sonores : audio/acoustique,
instruments

Financement : ANR/DFG

Période de réalisation : février 2017 – décembre 2020

Partenaires : Deutsche Forschungsgemeinschaft, ISAE SUPAERO
(Toulouse), LAGEP (Lyon), Kiel University

Motivés par les progrès technologiques récents dans les domaines de la mécanique, de l'aéronautique, des systèmes énergétiques et de l'ingénierie chimique et des nouveaux outils de calcul, l'analyse et le contrôle des systèmes à dimension infinie sont devenus un sujet d'intérêt majeur ces dernières décennies. Les concepts de base de la théorie des systèmes classiques ont été progressivement généralisés à des systèmes à dimension infinie avec des contributions issues de la communauté mathématique et de la communauté des ingénieurs. Plus récemment, les scientifiques se sont intéressés à la compréhension de systèmes composés de sous-systèmes à paramètres distribués (décrits par des systèmes d'équations aux dérivées partielles, des EDP) qui interagissent dans les réseaux. La majeure partie de la littérature existante sur la modélisation, l'analyse et le contrôle de systèmes, traite des réseaux de systèmes homogènes tels que les formes de barres élastiques ou les propriétés de conductivité thermique des mousses métalliques.

INSPECTSYN

Speech-in-noise INtelligibility Deficits: Designing Psychophysical and Electrophysiological Markers of SPECTral-Shape Coding Sensitive to Cochlear SYNaptopathy

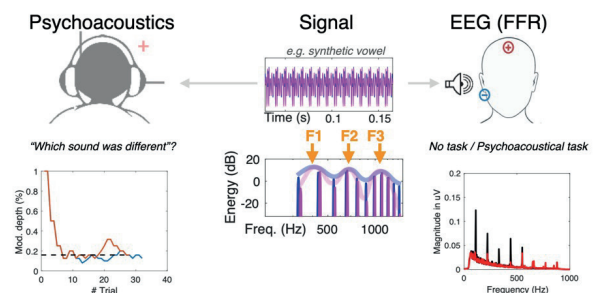
—
Équipe concernée : Perception et design sonores

Financement : ANR JCJC

Période de réalisation : 2023 – 2026

Partenaires : ENS, Université de Rochester, Institut de l'Audition

On estime aujourd'hui que plus de 10% des personnes dont l'audiogramme est cliniquement normal souffrent de troubles importants de la compréhension de la parole dans le bruit (PdB), par exemple en présence d'autres locuteurs (écoute 'cocktail party'). En particulier, la synaptopathie – la perte de synapses reliant la cochlée au nerf auditif causée par le vieillissement ou l'exposition au bruit – est considérée comme un facteur important contribuant à ce problème. Les récentes propositions d'évaluation électrophysiologiques de la synaptopathie chez l'homme se sont limitées à considérer son impact sur le codage neural de l'enveloppe temporelle du signal acoustique, et il reste encore difficile d'évaluer dans quelle mesure cette pathologie contribue aux difficultés individuelles de compréhension de la PdB. En raison de son rôle pivot dans la compréhension de la PdB (Lorenzi et al., PNAS, 2006), nous testerons ici l'hypothèse que l'évaluation de la fidélité de codage neural de la structure temporelle fine (TFS) du signal audio (qui n'est représentée par le système auditif que jusqu'à ~ 2 kHz) devrait offrir une mesure plus directe pour quantifier l'impact de la synaptopathie sur les déficits d'intelligibilité de la PdB. Le projet INSPECTSYN se concentrera sur les mécanismes d'encodage neural de l'enveloppe spectrale de signaux stationnaires basses fréquences (un modèle des voyelles dans la parole) et leurs déficits potentiels, afin de développer et de tester de nouveaux outils pour évaluer l'impact de la synaptopathie sur le codage de la TFS. Ce projet repose sur une approche multidisciplinaire intégrée combinant modélisation computationnelle, psychophysique et électrophysiologie, qui sera ici déployée sur différents groupes d'auditeurs, jeunes/âgés, avec/sans pertes audiométriques. Globalement, ce projet vise à faire progresser les tentatives actuelles de caractérisation d'une pathologie auditive qui reste non détectable par les tests audiométriques standards, et conduire ainsi à proposer de nouveaux outils audiolinguistiques permettant de mieux caractériser la santé auditive humaine.



Instruments augmentés

Équipes concernées : Interaction son musique mouvement,
Systèmes et signaux sonores : audio/acoustique, instruments

Les instruments augmentés correspondent à des instruments acoustiques auxquels sont intégrés des capteurs, afin de transmettre en temps réel des paramètres gestuels. Ces instruments sont spécialement bien adaptés au contexte de musique mixte acoustique et électronique.

Ces travaux s'effectuent en étroite relation avec des compositeurs et interprètes dont le but est d'intégrer ces technologies dans leurs œuvres.

Ces développements s'inscrivent également dans le cadre de recherche sur le geste instrumental.

Ce projet avait été initié par le développement du « violon augmenté », qui s'est fortement développé, en raison d'un intérêt accru des compositeurs. Nous travaillons désormais sur tous les instruments du quatuor à cordes, ainsi que sur certains instruments de percussions et le piano. Diverses techniques informatiques sont également développées pour analyser, reconnaître et suivre les gestes instrumentaux. Par exemple, un vocabulaire d'éléments musicaux comme des modes de jeux ou des phrases musicales peut être défini par le compositeur, et servir de base pour l'interaction avec des processus sonores comme la synthèse ou la spatialisation.

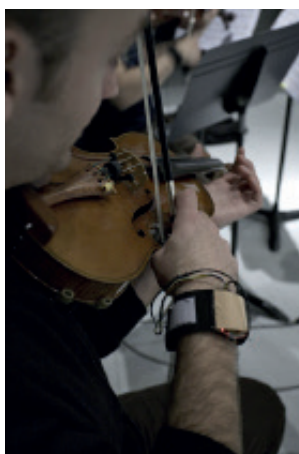
Finalement, il est à noter que les instruments augmentés sont utilisés dans des créations musicales mais également dans le cadre d'applications pédagogiques.

Lutherie augmentée

Équipe concernée : Systèmes et signaux sonores : audio/acoustique, instruments

Le projet « Lutherie augmentée » est au cœur des recherches sur les instruments de musique. Il a produit des prototypes d'instruments ou parties d'instruments utilisables en concert. On peut citer par exemple l'élargissement de la gamme de sourdines des cuivres, l'étude d'un dispositif d'accord automatique pour la timbale et la réalisation d'un bec de clarinette à volume variable, ou encore un archet et un bec instrumentés par plusieurs capteurs.

De plus, l'Ircam conçoit et possède un instrumentarium de SmartInstruments (guitares, quatuor à cordes, clarinette basse, sourdines de cuivres) munis de capteurs et d'actionneurs dont les propriétés acoustiques sont modifiables par l'instrumentiste ou le compositeur, et qui diffusent des sons d'origines diverses sans enceinte extérieure. Cela a amené à la réalisation de COALA, système embarqué de contrôle actif à très faible latence pour SmartInstruments.



© F. Kleinfeim

Violon augmenté.

Ondes Martenot

—
Équipe concernée : Systèmes et signaux sonores :
audio/acoustique, instruments
Financement : Collegium Musicæ
Partenaire : ECR-Musée de la musique

Une trentaine d'années après l'arrêt de leur production, les Ondes Martenot, présentées au public en 1928, sont encore jouées par de prestigieux interprètes et leur répertoire continue de s'élargir aussi bien dans le contexte de la musique savante que populaire. Les premières versions de l'instrument sont, pour la plupart, conservées dans les musées, dont les positionnements déontologiques et les contraintes techniques de la conservation-restauration interdisent aujourd'hui d'espérer faire aisément entendre leur son. Afin de conserver ces instruments historiques et de répondre à la demande des compositeurs et musiciens, ce projet vise à élaborer des fac-similés reproduisant au mieux les caractéristiques de chaque instrument, unique, à l'aide des technologies actuelles.

Un projet antérieur au sein de l'équipe S3AM a mené à la reproduction numérique d'un circuit électronique d'époque (oscillateur hétérodyne, mélangeur et amplificateurs), embarqué dans un processeur numérique temps réel et connecté à une interface de contrôle utilisateur : clavier MIDI et capteur de position résistif en ruban pour la hauteur, touche de contrôle de la pression de la main (Touché, Expressive E) pour le niveau sonore.

Le projet actuel, en collaboration avec l'équipe Conservation Recherche du Musée de la musique, Paris, et financé par l'institut Collegium Musicae de l'Alliance Sorbonne-Université, se concentre désormais sur les « diffuseurs » de l'Onde Martenot, transducteurs conçus par Maurice Martenot pour convertir le signal électrique en onde sonore.

Dans un premier temps, une étude musicologique a permis d'étudier les diffuseurs d'Ondes Martenot conservés au Musée de la musique et de formuler des hypothèses sur leur ordre d'apparition, à partir de leur documentation et des archives disponibles.

La seconde partie du projet en cours, consiste à mesurer certaines caractéristiques électro-mécano-acoustiques sur un diffuseur de type A-B datant de 1937 : impédance d'entrée et directivité dans les limites du domaine linéaire, ainsi que les produits de distorsion. Leur analyse comparée avec celles d'autres haut-parleurs d'époque et actuels visent à proposer de nouvelles hypothèses sur les choix de facture de Maurice Martenot.

Un modèle mathématique sera ensuite établi et informé par ces mesures. Son rayonnement sera simulé et comparé à celui du diffuseur A-B original. La dernière partie du projet vise à élaborer un fac-similé du diffuseur A-B permettant à un interprète de recréer le son sans utiliser les diffuseurs d'époque devenus fragiles. Cette dernière étape vise à reproduire le rayonnement du diffuseur A-B original avec un haut-parleur actuel, en modifiant la vibration de sa membrane par contrôle actif.

Organologie des harpes pharaoniques

—
Équipe concernée : Systèmes et signaux sonores :
audio/acoustique, instruments
Période : 2017-2022
Collaborations : HiSoMA, IJLRDA-LAM et Susanna Schulz (Berlin)

Lors de fouilles dans les années 2002-2005 dans la nécropole de Dra' Abou el-Naga, située sur la rive ouest de Louxor, les archéologues ont trouvé trois harpes. L'une d'entre elles, dans un état de conservation exceptionnel, a incité à mettre en place une étude interdisciplinaire pour essayer de tirer le maximum d'informations de l'objet, dont une analyse sur ses propriétés acoustiques. La qualité sonore de cette harpe de l'Égypte ancienne ne pouvant pas être appréhendée à partir du vestige archéologique, une copie de l'instrument a été réalisée par une luthière à partir d'un relevé archéologique précis des dimensions de l'original et de l'identification des bois utilisés. L'analyse vibratoire et acoustique de cette copie a permis, d'une part, de mieux saisir l'expertise des artisans qui ont conçu cet instrument et, d'autre part, d'ouvrir une réflexion sur l'accordage inconnu de l'instrument et sur la position de jeu que l'on voit dans les scènes de musiques représentées dans les tombes.

Réverbération hybride et réponses impulsionnelles spatialisées

Équipe concernée : Espaces acoustiques et cognitifs

La plupart des réverbérateurs artificiels actuels s'appuient soit sur des réseaux de retards rebouclés (Feedback Delay Networks FDN), soit sur des moteurs de convolution. Ces approches se distinguent par le coût de calcul, l'«authenticité» du rendu, la flexibilité d'usage et l'adaptabilité vis-à-vis du dispositif de restitution. L'équipe EAC explore depuis plusieurs années une approche dite «hybride». À partir des réponses impulsionnelles mesurées (Impulse Responses IR), les réflexions précoces sont restituées par convolution, tandis que la partie tardive de la réverbération est synthétisée par un FDN qui approxime le relief de décroissance énergétique (Energy Decay Relief EDR) de l'IR. La transition entre les deux modes s'opère au temps de mélange, et le FDN est calibré de sorte à assurer la continuité de l'EDR, garantissant ainsi un rendu perceptivement indiscernable de l'IR originelle. Cette technique hybride a été étendue à des IR spatialisées (SpatialRoom Impulse Responses SRIR), mesurées avec des microphones sphériques (cf. ci-contre), de sorte à restituer la distribution spatiale des réflexions dans la salle.

Depuis plusieurs années, l'équipe EAC et ses collaborateurs (artistes en résidence, ingénieurs du son, etc.) ont collecté des réponses impulsionnelles directionnelles (DRIR) dans une variété d'espaces (plein air, salles de concert, églises, site archéologique, etc.). Initiée dans le cadre de la résidence de recherche artistique de Pedro Garcia-Velazquez et Augustin Muller, la bibliothèque comprend, à ce jour, une quarantaine d'espaces acoustiques. En vue de sa mise à disposition pour le département production et, à terme, pour le public, les différentes campagnes de mesures ont été annotées (positions des sources et récepteurs, typologie architecturale, etc.) et uniformisées (format de stockage, système de coordonnées, nommage des fichiers, etc.). Un démonstrateur a été développé pour 'naviguer' dans ce catalogue raisonné, représenter schématiquement la scène sonore et afficher les différents champs d'annotation associés.

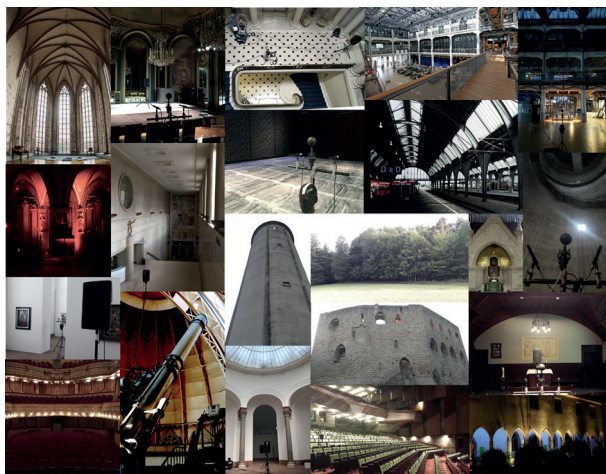
SHENG – orgue à bouche

Équipes concernées : Analyse des pratiques musicales, Systèmes et signaux sonores : audio/acoustique, instruments

Période : 2019-2021

Collaborations : Collegium Musicæ, Iremus, TPMC

Le sheng (orgue à bouche) est un instrument riche d'une longue histoire dont le répertoire se partage entre la tradition et le contemporain. Pour accueillir et encourager de prochaines créations, une équipe de chercheurs (France, Allemagne, Autriche, Chine, Taiwan, Japon) travaillent, au sein d'un projet conduit par l'institut Collegium Musicæ de Sorbonne Université, sur l'histoire, les différents modèles d'orgue à bouche, le répertoire, l'étude acoustique, l'analyse gestuelle d'improvisation, la combinaison de doigtés, la notation, et explorer cet instrument historique qui se prête de façon surprenante à la modernité. Toutes ces recherches seront accompagnées par une série de séminaires.



© P. Garcia-Velazquez, A. Muller

Pêle-mêle des sites remarquables ayant fait l'objet d'une prise d'empreinte acoustique 3D.

Synthèse de champs sonores par réseaux à haute densité spatiale

—
 Équipe concernée : Espaces acoustiques et cognitifs
 Collaborations extérieures : EMPAC (Troy, États-Unis).

La recherche sur la synthèse de champ sonore reste un enjeu privilégié pour l'équipe EAC et trouve son application naturelle dans les productions musicales. Elle représente également un excellent vecteur de valorisation de nos développements et de collaboration avec des institutions de recherche et de création musicale prestigieuses.

Dans le cadre d'une collaboration de recherche avec le Rensselaer Polytechnic Institute, la salle de concert du Curtis R. Priem Experimental Media and Performing Arts Center (EMPAC) a été équipée d'un ambitieux dispositif de spatialisation sonore. Ce nouveau système est constitué d'un réseau de 512 haut-parleurs (16 modules de 31 haut-parleurs associés à un subwoofer) pour la diffusion en WFS (Wave Field Synthesis) et d'un dôme de 100 haut-parleurs pour une diffusion en HOA (High Order Ambisonics).

L'espacement des transducteurs est de 5,8cm. Différents agencements géométriques des modules ont été testés, chacun permettant de tirer parti de façon optimale des différentes méthodes de synthèse de champ sonore. Par exemple, la superposition des modules permet de constituer une antenne rectiligne de 15m pour la diffusion en WFS avec des transducteurs distribués en quinconce et espacés de 2,9 cm. La fréquence de repliement spatial est ainsi rejetée au-delà du spectre audible et permet d'obtenir une reconstruction du champ sonore saisissante.

Disposés sous forme d'une ceinture polygonale, ces modules permettent, cette fois, de constituer un dispositif de reproduction HOA 2D à l'ordre N=250. Les haut-parleurs constituant le dispositif sont contrôlés indépendamment par une ferme d'ordinateurs gérant en temps réel la spatialisation des sources sonores. L'un des enjeux majeurs de cette recherche est la mise en pratique du contrôle de champ sonore multizone.



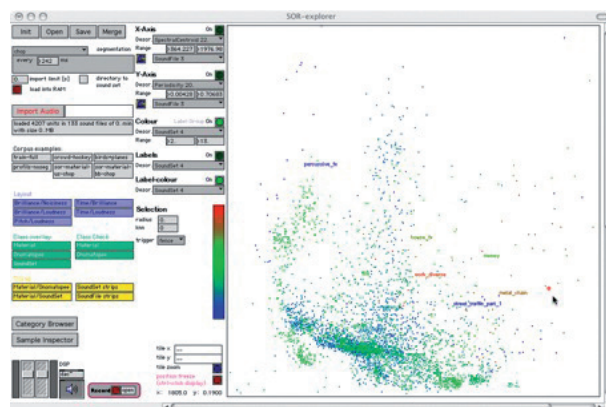
© M. Noisternig

Dispositif de restitution sonore WFS – HOA du Curtis R. Priem Experimental Media and Performing Arts Center au Rensselaer Polytechnic Institute de Troy (US). Le dispositif est constitué de 16 antennes linéaires de haut-parleurs espacés de 5,8 cm, ici configurées pour une diffusion HOA bidimensionnelle.

Synthèse concaténative par corpus

—
 Équipe concernée : Interaction son musique mouvement

La synthèse concaténative par corpus utilise une base de données de sons enregistrés, segmentés et indexés par leurs caractéristiques sonores, telles que leur hauteur, ou leur timbre. Cette base nommée corpus, est exploitée par un algorithme de sélection d'unités qui choisit les segments du corpus qui conviennent le mieux pour la séquence musicale que l'on souhaite synthétiser par concaténation. La sélection est fondée sur les descripteurs sonores qui caractérisent les enregistrements, obtenus par analyse du signal et correspondant par exemple à la hauteur, l'énergie, la brillance, ou la rugosité. Les méthodes de synthèse musicale habituelles sont fondées sur un modèle du signal sonore, mais il est très difficile d'établir un modèle qui préserverait la totalité des détails et de la finesse du son. En revanche, la synthèse concaténative, qui utilise des enregistrements réels, préserve ces détails. La mise en œuvre de la nouvelle approche de synthèse sonore concaténative par corpus en temps réel permet une exploration interactive d'une base sonore et une composition granulaire ciblée par des caractéristiques sonores précises, et permet aux compositeurs et musiciens d'atteindre de nouvelles sonorités. Si la position cible de la synthèse est obtenue par analyse d'un signal audio en entrée, on parle alors d'*audio mosaicing*. La synthèse concaténative par corpus et l'audio mosaicing par similarité spectrale sont réalisés par la bibliothèque MuBu de modules optimisés pour Max, et dans le système CataRT, qui permet, via l'affichage d'une projection 2D de l'espace des descripteurs, une navigation simple avec la souris, par des contrôleurs externes, ou par l'analyse d'un signal audio. CataRT, comme bibliothèque de modules pour Max, comme device pour Ableton Live, ou comme application indépendante, est utilisé dans des contextes musicaux de composition, de performance et d'installation sonore variés.



Exemple de visualisation de synthèse concaténative.

Système WFS et ambisonique à l'Espace de projection

Équipe concernée : Espaces acoustiques et cognitifs

Financement : programme Sesame (Conseil général d'Île-de-France), CNRS, Ircam, ANR, Sorbonne Université

La technique Wave Field Synthesis (WFS) désigne un procédé de reproduction holophonique qui permet, par analogie avec les hologrammes visuels, de capter ou synthétiser une scène sonore en préservant les informations spatiales de distance et de direction des sources qui la composent. Cette approche, initiée par l'université de Delft, dépasse les limites des systèmes conventionnels en termes de fidélité de reproduction sur une zone d'écoute étendue. Tandis que les techniques stéréophoniques conventionnelles (stéréo, 5.1) s'apparentent au trompe-l'œil et ne peuvent ainsi être appréciées que depuis le centre du dispositif, l'holophonie a l'ambition de reconstruire un champ sonore dans lequel les auditeurs peuvent se déplacer en gardant une perception cohérente de la localisation des sources.

La technique Ambisonics consiste à représenter le champ sonore comme une distribution angulaire de la pression acoustique exprimée sur une base de fonctions périodiques de l'espace, les harmoniques sphériques. La captation et la reproduction reposent sur des réseaux de transducteurs concentriques (sphères de microphones ou réseaux de haut-parleurs) qui échantillonnent l'espace et sur des opérations d'encodage et décodage permettant de passer du domaine des transducteurs au domaine des harmoniques sphériques. Plus l'ordre de décomposition en harmonique sphérique est élevé, meilleure est la résolution spatiale. On parle alors de Higher Order Ambisonics (HOA). Le formalisme Ambisonics est très puissant et permet une certaine indépendance entre le système de reproduction et le format d'encodage/décodage. Il permet en outre de réaliser de manière efficace des opérations de rotation d'une scène sonore 2D ou 3D ou encore d'opérer des formations de voies (microphone virtuel directif).

Grâce au soutien de la région Île-de-France, du CNRS et de Sorbonne Université, l'Ircam a fait l'acquisition d'un système de diffusion WFS et HOA installé dans l'Espace de projection. Le système inauguré en 2012 est composé, d'une part, d'une couronne horizontale de 264 haut-parleurs régulièrement répartis autour de la scène et du public pour la diffusion en WFS et, d'autre part, d'un dôme de 75 haut-parleurs pour une diffusion tridimensionnelle en mode HOA. Cet équipement a pour double vocation de susciter l'exploration de nouvelles modalités de spatialisation pour la création musicale et de permettre la mise en œuvre d'expériences à caractère scientifique consacrées à la réalité virtuelle et la cognition spatiale. Une partie du dispositif est mobile en sorte de permettre son installation hors les murs. Il a été exploité depuis dans le cadre de nombreuses créations musicales, représentations théâtrales et installations sonores in situ ou hors les murs (*Operspective Hölderlin* de P. Schoeller; *Le Père* de M. Jarrell; *Mimesis* de M. Garcia-Vitoria; *La Tragédie du Roi Richard II* d'après W. Shakespeare et mise en scène J.-B. Sastre, Cour d'honneur du Palais des papes à Avignon; *Mon Cœur parle tout seul* de D. Ghisi et D. Janne-teau; *Disenchanted Island* de O. Neuwirth et T. Rosner; *Pocket of Spaces* de N. Barrett et The OpenEnded Group).

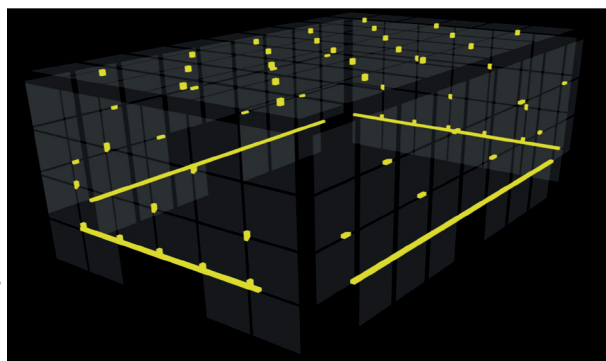


Schéma d'implantation du système de diffusion dans l'Espace de projection.



Le dispositif HOA installé dans la Grande salle du Centre Pompidou pour la pièce *Pocket of Spaces* de N. Barrett et The Open Ended Group.

Traitements par vocodeur

—
Équipe concernée : Analyse et synthèse des sons

La recherche sur les vocodeurs pour la transformation de la musique et de la parole est un des sujets clés de l'équipe. Au début (depuis 1995) les recherches étaient concentrées sur le vocodeur de phase, l'une des techniques les plus performantes pour l'analyse et la transformation des sons. Ces recherches sont à la base du logiciel SuperVP, qui permet par exemple de transposer, d'étirer ou de raccourcir des sons, de les filtrer pratiquement sans limitation, etc. Pour la parole également, la qualité sonore des signaux transformés atteint un excellent niveau. De très nombreuses améliorations et extensions y ont été apportées. Citons par exemple :

- L'estimation d'enveloppe spectrale par «true envelope» ;
- La transposition avec préservation d'enveloppe spectrale ;
- La transposition de la voix avec modèle «shape invariant» ;
- La synthèse croisée généralisée qui permet de synthétiser des sons hybrides ;
- La détection et le traitement séparé des zones temps-fréquence sinusoidales, non-sinusoidales et transitoires.

Ces différents modules d'analyse, de synthèse et de traitement sont utilisés dans plusieurs logiciels commerciaux. À partir de 2009 un nouveau type de vocodeur, fondé sur la gestion des impulsions glottique a été envisagé. Issu de ses recherches, le vocodeur PaN «Pulse and Noise» a été intégré dans le logiciel de synthèse de chant ISIS (disponible via le Forum de l'Ircam).

Depuis 2017, le succès des techniques d'apprentissage profond dans le domaine de la synthèse de la parole a motivé une nouvelle direction de recherche sur la transformation de la parole avec des *vocodeurs neuro-naux*, qui actuellement est en cours d'étude notamment dans le projet ARS.

C2. Le corps musicien

ACTIVATE

Augmenting the value of conversations with voice technologies

Projet ERC

Équipe concernée : Perception et design sonores

Financement : ERC Proof of Concept

Période de réalisation : 2020 – 2022

Partenaires extérieurs : Altavoce SAS, CNRS Innovation

Le projet ACTIVATE a pour objectif de valoriser des technologies de transformation temps réel de voix issues du projet ERC CREAM, en augmentant la valeur économique de conversations parlées en permettant de contrôler leurs caractéristiques émotionnelles. Par exemple, dans le contexte d'une conversation de relation-client au téléphone, l'algorithmie de la transformation de voix peut rendre la voix d'un client énervé 10 % moins agressive, réduisant ainsi la fatigue émotionnelle des employés à la fin de la journée, ou rendre la voix d'un opérateur 10 % plus digne de confiance, augmentant ainsi la satisfaction client après l'appel. Le projet ACTIVATE est une collaboration entre chercheurs et business developers et va réaliser (1) des analyses de marché et des entretiens avec des acteurs du monde socio-économique pour identifier des situations de conversation dans lesquelles de telles technologies peuvent être testées, (2) identifier quels indicateurs de performance post-conversation sont pertinents pour le marché, (3) mesurer l'impact des transformations sur ces indicateurs, soit en grandeur nature, soit dans un environnement de test simulé en laboratoire et (4) utiliser ces résultats pour estimer la valeur d'un éventuel produit, qui sera valorisé par la startup partenaire Altavoce (www.altavoce.tech).

Les résultats du projet ERC CREAM ouvrent des perspectives socio-économiques, que le projet ERC ACTIVATE se donne comme objectif de valoriser.

ARS

Analyse et transformation de style de chant

ars.ircam.fr

Équipe concernée : Analyse et synthèse des sons

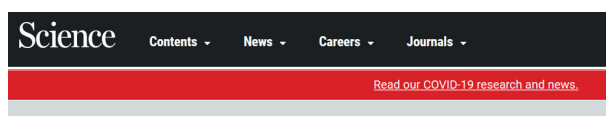
Financement : ANR Révolution numérique : rapports aux savoirs et à la culture (CE38)

Période de réalisation : janvier 2020 – juin 2023

Partenaires extérieurs : Institut d'Alembert (Sorbonne Université), Passages XX-XXI (université de Lyon), Flux::audio.

Le style de chant est un attribut important de l'interprétation chantée, qui affecte profondément la perception et la réception subjectives de la musique. Récemment, des musicologues se sont lancés dans une étude des styles de chant dans le domaine de la musique populaire, un domaine où, dans le même temps, les effets numériques appliqués au chant sont devenus si répandus que la vaste majorité des musiques populaires ont aujourd'hui recours à au moins un des plugins disponibles sur le marché. Dans ce contexte, le projet ARS aspire à instaurer une collaboration, bénéfique pour tous, entre les musicologies étudiant l'interprétation chantée (qui relèvent des humanités), les ingénieurs travaillant à la transformation de la voix chantée (sciences numériques) et les développeurs spécialistes des logiciels destinés aux professionnels de l'audio.

Le projet ARS a deux objectifs principaux : le premier est de mettre au point de nouveaux moyens d'étude des styles de chant dans le domaine de la performance musicale. Le deuxième est de développer des algorithmes de modification expressive du style de chant de haute qualité. Parmi les nombreuses caractéristiques stylistiques relatives au chant, le projet ARS se concentrera sur ceux qui ont un lien direct avec le style personnel, c'est-à-dire sur les effets stylistiques auxquels un chanteur peut choisir d'avoir recours au cours de sa performance dans le but de créer un effet interactif. Cela inclut les modifications de l'intonation et des contours d'intensité de même que des caractéristiques vocales telles que : la rugosité, le grognement, le souffle, le belting, ou le formant du chanteur.



SHARE



ISTOCKPHOTO/ STEVE DEBENPORT

Want to sound like someone people can trust? This new software could help

By Tania Rabesandratana | Mar. 30, 2018, 8:00 AM

COSMOS

Computational Shaping and Modeling of Musical Structures

cosmos.cnrs.fr

—

Équipe concernée : Représentations musicales

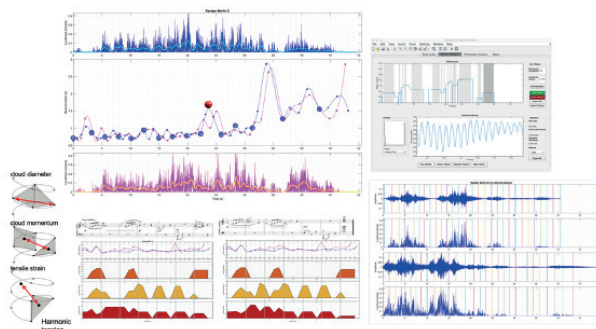
Financement : European Research Council Advanced Grant d'Elaine Chew (2019 – 2024)

La performance en tant qu'acte créatif a été largement reconnue, mais la nature de cette créativité reste insaisissable. Dans le projet Cosmos, nous estimons que cette créativité réside dans la création et la mise en forme des structures musicales, et que cette pensée créative prend la forme d'une activité de « résolution de problème ». Cosmos vise à transformer la compréhension commune de l'expérience humaine de la musique interprétée, qui constitue presque toute la musique que nous entendons, et de la créativité inhérente à l'interprétation musicale, qui conditionne la façon dont la musique est produite.

Cosmos utilise la science des données, l'optimisation, l'analyse des données et la science citoyenne pour étudier les structures musicales telles qu'elles sont expérimentées et créées dans la performance. Le projet étudie des sources inhabituelles comme les données cardiaques rythmiques.

Les thèmes de recherche généraux du projet Cosmos sont les suivants : i) trouver de nouvelles façons de représenter et d'explorer les variations dans les performances enregistrées et les séquences variant dans le temps ; ii) étudier le mode de pensée des volontaires (science citoyenne) pour la recherche sur la performance musicale en se concentrant sur les structures expérimentées et la résolution de problèmes ; iii) créer des environnements expérimentaux pour la création de telles structures ; iv) créer des cadres théoriques pour découvrir les raisons qui expliquent les structures perçues et réalisées ; v) encourager la participation communautaire en formant des experts (notamment des musiciens retraités) pour fournir des commentaires sur les structures, afin que le public comprenne mieux la création dans la performance musicale.

La plateforme de science citoyenne développée dans Cosmos permettra au grand public d'annoter et de déconstruire les structures créées dans les performances musicales, et d'explorer l'espace des interprétations musicales possibles. Les outils et techniques mathématiques permettant de caractériser les variations et les motifs des musiques interprétées et des séquences électrocardiographiques élargiront les outils mis à la disposition des chercheurs en information musicale, des mathématiciens, des musicologues, des pédagogues de la musique, des psychologues de la musique et des musicothérapeutes. Les applications de la plateforme et des outils Cosmos comprendront des projets artistiques, des événements d'engagement du public et la communication des sciences de la musique.



ELEMENT

Stimuler l'apprentissage de mouvements dans les interactions humain-machine

element-project.ircam.fr

—

Équipe concernée : Interaction, son, musique, mouvement

Financement : ANR, projet AAP

Période de réalisation : novembre 2018 – octobre 2021

Coordinateur : Ircam

Partenaires : LRI, LIMSI-CNRS

La plupart des interactions basées sur le mouvement proposent des interfaces « intuitives » et des vocabulaires gestuels triviaux. Bien que ceux-ci facilitent l'adoption, ils limitent néanmoins les possibilités d'interactions plus complexes, expressives et véritablement incarnées. Nous proposons de passer des notions d'intuitivité à des notions d'« apprenabilité ». Notre projet aborde à la fois des problèmes méthodologiques et de modélisation computationnelle.

Premièrement, nous devons élaborer des méthodes pour concevoir des vocabulaires de mouvements, lesquelles seraient faciles à apprendre et à composer afin de créer des phrases de mouvement riches et expressives. Deuxièmement, nous devons concevoir des modes computationnels capables d'analyser les mouvements des utilisateurs en temps réel pour fournir divers mécanismes de feedback et de guidage multimodaux (par exemple visuels et auditifs).

Trois questions de recherche fondamentale : 1) Comment concevoir des mouvements et des gestes, formés de composants faciles à apprendre, tout en permettant des techniques d'interactions complexes au-delà des simples commandes ? 2) Comment rendre compte de l'apprentissage sensori-moteur avec des modélisations computationnelles du mouvement et de l'interaction ? 3) Comment optimiser des systèmes de feedback et guidages informatique afin de faciliter l'acquisition de compétences ?

L'objectif à long terme est de favoriser l'innovation dans l'interaction multimodale, de la communication non verbale à l'interaction avec médias numériques dans des applications créatives.

HAIKUS

Réalité augmentée et apprentissage artificiel

—

Équipe concernée : Espaces acoustiques et cognitifs

Financement : ANR AAP19

Période de réalisation : décembre 2019 – mai 2023

Partenaires extérieurs : MPIA (CNRS-SU), LORIA (INRIA, Univ. Lorraine)

Les applications de réalité augmentée auditive reçoivent un intérêt croissant dans différents domaines applicatifs comme la création artistique, la médiation culturelle, la communication ou le divertissement. L'audition joue un rôle majeur dans la sensation d'immersion de l'utilisateur et la qualité d'expérience ressentie. Intégrer des événements sonores virtuels préenregistrés ou générés automatiquement dans l'environnement réel de l'utilisateur peut procurer une sensation interactive attrayante et peut être exploitée par exemple pour les jeux vidéo, les audio-guides ou de nouvelles formes de dramatiques radio. L'enjeu majeur des applications de réalité augmentée auditive réside dans la capacité à intégrer ces événements sonores sans hiatus perceptif, c'est-à-dire avec un rendu spatial s'adaptant en permanence aux conditions acoustiques de l'environnement réel, par exemple en fonction du mouvement des sources ou lors de la déambulation à travers différents espaces. L'objectif du projet HAIKUS est l'exploitation conjointe de méthodes d'apprentissage artificiel (AA) et de traitement du signal audio afin de résoudre les problèmes acoustiques rencontrés dans le cadre des applications de réalité augmentée. Les méthodes d'AA sont notamment appliquées pour l'identification automatique du canal acoustique entre les sources et l'auditeur. L'intégration sans hiatus perceptif des sons virtuels dans l'environnement réel nécessite l'estimation des caractéristiques acoustiques de la salle ou du site de sorte à adapter automatiquement le traitement de réverbération appliqué aux sources virtuelles. L'enjeu est alors l'estimation en aveugle des paramètres acoustiques (temps de réverbération, rapport son direct / son réverbéré) ou des caractéristiques architecturales de la salle (volume et forme de la salle, absorption des parois) à partir de la simple observation des signaux audio réverbérés émanant des sources réelles présentes dans la salle. L'adhésion de l'utilisateur à la scène auditive augmentée repose sur une évolution des indices acoustiques réaliste et congruente avec son déplacement dans la scène et le mouvement des sources virtuelles. Ceci nécessite l'inférence de règles plausibles de modification des paramètres de spatialisation ou la mise en œuvre de techniques d'interpolation de réponses impulsives de salle en fonction des mouvements relatifs des sources et de l'auditeur. Les scènes sonores virtuelles interactives sont typiquement rendues en mode binaural sur casque d'écoute. Il est établi qu'une restitution binaurale convaincante requiert l'usage d'HRTFs individuelles qui devraient être idéalement mesurées pour chaque utilisateur en chambre anéchoïque avec des signaux parfaitement calibrés. On se propose d'estimer en aveugle les HRTFs du participant à partir de l'observation des signaux binauraux captés dans les environnements quotidiens et de manière non supervisée (sources quelconques et en mouvement).

HEARDEVCOMP

Développement biologique de la perception de la parole : inférence computationnelle du codage auditif par données psychoacoustique et électrophysiologique

—

Équipe concernée : Perception et design sonores

Financement : CNRS MITI 80PRIME

Période de réalisation : 2023

Partenaires extérieurs : INCC (Université de Paris)

L'apprentissage du langage chez l'homme repose sur la mise en place d'un processus biologique d'extraction et de traitement sélectif des paramètres acoustiques contenus dans les signaux de parole, en particulier des modulations d'amplitude et de fréquence. Si les études précédentes montrent que l'efficacité du traitement auditif de ces indices, déjà en place dès la naissance, s'améliore jusqu'à l'âge adulte, les étapes de ce développement restent encore mal caractérisées. Dans ce projet, nous proposons une nouvelle combinaison interdisciplinaire de méthodes de mesure de l'audition couplée à des modèles computationnels pour accéder aux paramètres de cette trajectoire développementale, depuis la petite enfance jusqu'à l'âge adulte. Chaque individu testé sera soumis à une batterie de tests auditifs visant à caractériser l'encodage auditif physiologique ainsi que ses capacités de perception de la parole. La caractérisation de l'encodage auditif associera réponses comportementales (psychophysique, adaptée chez le bébé) et neurophysiologiques (EEG) de signaux de synthèse reproduisant les modulations temps-fréquence de la parole. Les capacités de perception de la parole seront évaluées par des tests subjectifs de mesure d'intelligibilité de signaux de parole réels dans des conditions d'écoute naturelle difficiles (ex. écoute spatialisée 'cocktail party').

L'objectif initial de ce projet est d'établir une preuve de concept de l'utilisation de modèles computationnels du système auditif pour exploiter des données comportementales et neurophysiologiques récoltées chez le nourrisson et chez l'adulte afin d'inférer les paramètres latents du codage auditif pour ces deux populations, en relation avec les capacités de perception de la parole observées au quotidien. À terme, ce travail vise à développer de nouveaux biomarqueurs du développement typique et atypique de la parole et du langage, permettant de mieux appréhender l'impact des surdités infantiles et adultes sur le développement auditif et la perception de la parole au quotidien.

HEART.FM

Maximizing the Therapeutic Potential of Music through Tailored Therapy with Physiological Feedback in Cardiovascular Disease

Équipe concernée: Représentations musicales

Financement: European Research Council Proof of Concept Grant

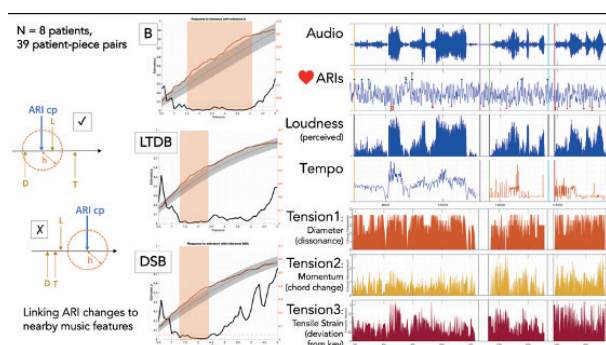
Période de réalisation: décembre 2020 - mai 2022

Partenaires: University College London, CNRS Innovation

Heart.FM est un projet de preuve de concept lié à Cosmos.

Les maladies cardiovasculaires sont la première cause de décès dans le monde, et l'hypertension est le principal facteur de risque médical. La réduction de la pression artérielle chez les patients hypertendus a des effets cardiovasculaires positifs, et l'écoute de musique offre des moyens non pharmacologiques de réguler le rythme cardiaque et la pression artérielle. Actuellement, la manière la plus efficace d'administrer des interventions musicales ayant des effets durables est à forte intensité de main-d'œuvre et repose sur les intuitions des thérapeutes professionnels, ce qui la rend non extensible à de vastes cohortes. HEART.FM est une initiative de création d'applications visant à démocratiser l'accès à des interventions personnalisées de thérapie musicale, éclairées par un retour d'information physiologique.

L'objectif est de préparer HEART.FM à un déploiement à grande échelle dans le cadre d'essais de recherche contrôlés et randomisés, afin d'établir des relations de cause à effet entre la musique et l'hypertension pour les personnes en bonne santé ainsi que pour les patients hypertendus. HEART.FM profitera à la communauté des chercheurs en offrant un portail à forte fréquentation pour la collecte de données de recherche sur les maladies cardiovasculaires et les thérapies préventives basées sur la musique. Le PoC explorera des solutions de commercialisation et de marché dans les secteurs de la musique et de la médecine.



Intégration multisensorielle et émotion

Équipe concernée: Espaces acoustiques et cognitifs

Au cours des vingt dernières années, de nouveaux types de thérapie par exposition ayant recours aux technologies de réalité virtuelle se sont développés pour le traitement et la réhabilitation des troubles émotionnels, et notamment des phobies. La réalité virtuelle (RV) offre un contrôle très riche de la présentation sensorielle et spatiale des stimulations ce qui permet de mieux contrôler l'impact émotionnel des situations anxiogènes auxquelles le patient est exposé. Cependant, afin d'exploiter ces avantages uniques de la RV, il est nécessaire d'approfondir les connaissances sur les liens entre les caractéristiques de présentation des stimuli et leur effet émotionnel.

Nous avons exploré ces liens dans le cadre de foules virtuelles. La peur de la foule est un symptôme retrouvé dans plusieurs troubles émotionnels (agoraphobie, phobie sociale, peur de tomber...) et présente des composantes à la fois sensorielles (auditives et visuelles) et spatiales, ce qui fait de la foule un stimulus idéal pour nos études.

Une étude a été menée sur deux groupes de participants (sensibles ou non à la peur de la foule), immergés dans un environnement visuo-auditif contenant des foules virtuelles, et qui devaient indiquer l'intensité de leur inconfort. La présentation sensorielle des foules était soit visuelle, soit auditive, soit visuelle et auditive. Les résultats indiquent que la présentation visuo-auditive amplifie le ressenti négatif chez les participants sensibles à la peur de la foule. Cependant, cet effet n'est observé que lorsque la foule est proche du participant. Ce résultat révèle l'interaction des caractéristiques spatiale et sensorielle du stimulus anxiogène sur l'impact émotionnel.

Cette dépendance spatiale a motivé l'initiation d'un volet de recherches consacrées à l'étude de l'espace péri-personnel. L'homme ne perçoit pas l'espace de manière homogène: le cerveau code l'espace proche du corps différemment de l'espace lointain. L'espace péri-personnel (proche du corps) est codé par des neurones multisensoriels. Nous cherchons à mettre en évidence la plasticité des limites de cet espace en fonction de la nature du son et du contexte émotionnel ou social.



Dispositif de RV dans le studio 4 de l'Ircam. La stéréo visuelle est projetée sur un grand écran. La mise à jour des informations visuelles et auditives est réalisée grâce à un système de suivi de position

MoVE

Modélisation générative des attitudes de la voix et application à un agent conversationnel expressif

Équipe concernée : Analyse et synthèse des sons

Financement : Allocation doctorale 2019 « Paris Region PhD » financée par la Région Île-de-France

Période de réalisation : septembre 2019 – septembre 2022

Partenaires extérieurs : PSA Peugeot Citroën

Dans un contexte où les assistants personnels et les interactions avec les machines deviennent une réalité quotidienne pour l'être humain, la voix s'est imposée comme une modalité privilégiée de l'interaction avec la machine. En particulier, la synthèse vocale a réalisé d'énormes progrès ces dernières années, notamment par l'utilisation de l'apprentissage profond et de grandes bases de données multi-locuteurs. Les limitations principales sont d'une part une faible expressivité : le comportement de l'agent est encore souvent monomodale (voix, comme les assistants Alexa ou Google Home) et demeure très monotone, ce qui limite grandement l'acceptabilité, la durée et la qualité de l'interaction; et d'autre part le comportement de l'agent est peu ou pas adapté à l'interlocuteur ou à la situation d'interaction, ce qui diminue sa compréhension de l'information et son temps de réaction à l'information transmise.

Le projet MoVE développera des algorithmes d'apprentissage neuro-naux permettant d'adapter le style de parole d'une voix de synthèse pour l'adapter à une situation d'interaction spécifique, avec par exemple un focus sur des attitudes de la voix de synthèse (cordiale, souriante, autoritaire, etc.). La meilleure adaptation du style de la voix permettra d'améliorer la compréhension de l'information communiquée par l'agent et de réduire le temps de réaction de l'humain aux informations communiquées (par exemple en situation d'urgence).

Musique et plasticité cérébrale

Équipe concernée : Espaces acoustiques et cognitifs

Partenaires extérieurs : université de Lille

Ce travail, mené dans le cadre de la collaboration avec l'université de Lille, se concentre sur les effets cognitifs de la musique dans le vieillissement pathologique. La maladie d'Alzheimer est une maladie neurodégénérative qui affecte le fonctionnement cognitif, émotionnel et social des individus. Compte tenu de l'efficacité limitée des approches pharmacologiques, des traitements non médicamenteux et en particulier des interventions musicales sont fréquemment proposés pour surmonter ces difficultés. Cependant, les facteurs qui expliquent leur efficacité ne sont pas clairement identifiés. Le but de ce projet est d'évaluer les facteurs influençant l'engagement socio-émotionnel et moteur de patients atteints de la maladie d'Alzheimer pendant une activité musicale, ainsi que leurs capacités à synchroniser leurs mouvements avec le rythme musical. Pour cela, des patients atteints de la maladie d'Alzheimer et des participants contrôles appariés ont réalisé une tâche de synchronisation sensorimotrice, consistant à taper avec la main au rythme d'un métronome ou de musique en présence d'un musicien. Le musicien réalisait la tâche de synchronisation avec le patient. Il pouvait être physiquement présent en face du patient (condition live), ou virtuellement présent à travers une vidéo préenregistrée et projetée en taille réelle en face du patient (condition vidéo). Les résultats montrent l'importance du contexte musical et des interactions sociales sur les performances de synchronisation, mais également sur l'implication sociale, émotionnelle et motrice des participants atteints de la maladie d'Alzheimer. Ces résultats ouvrent des perspectives thérapeutiques prometteuses, avec notamment la possibilité de compléter les prises en charges actuelles par des activités musicales régulières grâce à l'utilisation d'enregistrements vidéo et de la technologie mobile.



Dispositif expérimental.

[À gauche] Le participant tape sur la tablette en réponse au rythme des séquences auditives (musique ou métronome). En face, le musicien tape également sur une tablette avec sa main au rythme des séquences auditives.

[À droite] Manipulation de la présence sociale : la performance du musicien pouvait être préenregistrée et diffusée sur un écran à taille réelle (en haut), ou être *live* (en bas)

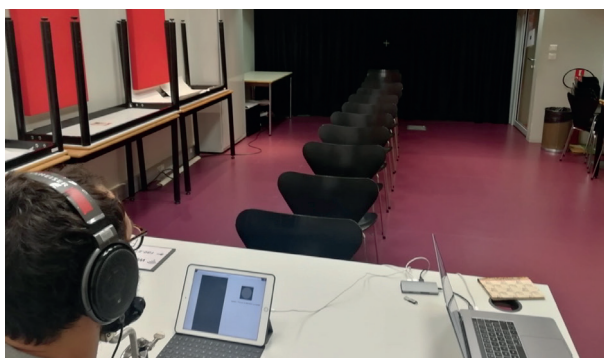
Perception de la distance auditive en réalité augmentée

Équipe concernée : Espaces acoustiques et cognitifs

Ce travail s'inscrit dans le cadre d'une étude plus vaste sur la modélisation et la perception spatiale dans le contexte de la réalité augmentée auditive. Dans un tel contexte, où l'environnement réel est enrichi par un ensemble d'événements sonores virtuels, une problématique importante est de garantir la cohérence de la perception spatiale de ces événements sonores virtuels et des sources sonores réelles ou ancrées visuelles présentes dans la salle. Assurer cette cohérence nécessite, en particulier, le choix d'un modèle de gestion de l'effet de distance sonore et l'accès à des informations sur les propriétés acoustiques de la salle.

L'objectif d'une première étude était d'évaluer les performances d'un modèle, informé par la donnée d'une unique réponse impulsionnelle directionnelle mesurée dans le lieu réel, pour contrôler la distance apparente d'une source virtuelle. Au sein de ce modèle, l'effet de distance sonore est rendu par la modification de l'énergie de différents segments temporels de la réponse impulsionnelle de base. Les réponses impulsionnelles ainsi extrapolées sont ensuite exploitées dans un moteur de rendu par convolution pour placer la source virtuelle à différentes distances de l'auditeur.

L'évaluation des performances du modèle a fait l'objet d'un protocole de test perceptif dans lequel le participant était placé dans une situation de réalité augmentée sonore dans laquelle il devait juger la distance d'une source sonore virtuelle en présence d'une série d'ancres visuelles (chaises). Les performances du modèle étaient menées par comparaison avec des réponses impulsionnelles binaurales de référence, mesurées dans la salle de test pour chaque distance testée. L'analyse statistique des résultats, collectés sur 40 participants au cours de deux sessions, démontre les performances équivalentes, d'un point de vue perceptif, de ce modèle face aux mesures de référence. Les critères jugés dans l'analyse sont la distance moyenne perçue, la variabilité inter-sujets et intra-sujet.



© Ircam-Centre Pompidou

Protocole de test pour l'évaluation perceptive de modèles de distance auditive dans un cadre de réalité augmentée

Psy-Son

Médias collaboratifs situés
cosima.ircam.fr

Équipe concernée : Perception et design sonores

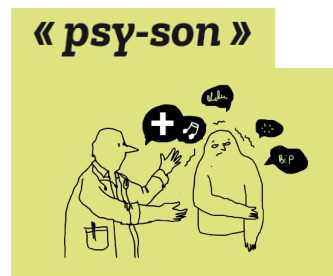
Financement : PHRIP

Période de réalisation : janvier 2023 – janvier 2025

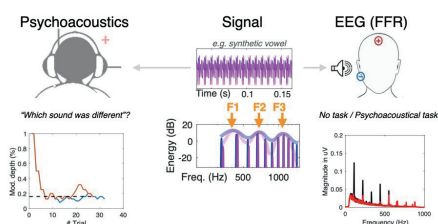
Partenaires extérieurs : GHU Paris Psychiatrie & Neurosciences (Lab-ah), CRD ENS Paris Saclay (ENSCI – Les Ateliers)

Le projet *Psy-Son* s'appuie sur le concept d'enveloppe sonore, questionne l'efficacité d'un espace temporel et sonore dans la prise en charge infirmière des états de « crise » de patients hospitalisés en psychiatrie, et vise la recherche et le développement d'un dispositif multimodal et multi-sensoriel d'écoute sonore et musicale de modulation de l'anxiété. Il s'articule suivant quatre axes distincts de recherche-action : la forme des dispositifs mobiliers et spatiaux favorisant une écoute sonore et musicale qualitative et multi-sensorielle ; le contenu sonore et musical incarnant l'adéquation avec les états émotionnels du patient ; l'interface de contrôle du dispositif proposant des alternatives à l'écran pour favoriser la perception multi-sensorielle ; les scénarios d'usage et de mise en œuvre, permettant l'évaluation du dispositif tant du côté soignant que patient.

La contribution de l'Ircam dans le projet se concentre sur les questions du contenu et d'interface de médiation qui s'incarnent dans le concept d'entretien musico-soignant. Il s'agit d'un dialogue, ou d'une série de dialogues, entre le soignant et le patient dont l'objectif est, dans un premier temps, de définir conjointement le profil musical et sonore – la biographie sensorielle – du patient, et dans un second temps, de proposer au patient une offre musicale et sonore (*playlist*) en adéquation avec son état et à même de réduire son niveau d'anxiété, au moment voulu. La compréhension, la conception et la personnalisation de l'univers sonore et musicale du patient à des fins thérapeutiques constituent l'une des clés de voûte du projet *Psy-Son*, représentent un véritable enjeu de recherche en design sonore appliqué au monde médical et clinique, et s'appuient sur les compétences croisées dans les domaines des sciences et technologies du son et ceux du soin et de la thérapie.



© Lab-ah / GHU Paris Psy. & Neurosc

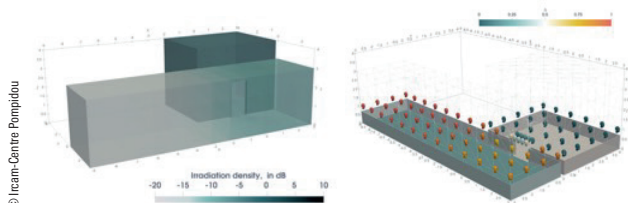


RASPUTIN

—
 Équipe concernée : Espaces acoustiques et cognitifs
 Financement : ANR, programme AAP 2018
 Période de réalisation : novembre 2018 – avril 2022
 Partenaires extérieurs : IJLRA (Sorbonne Université / CNRS),
 université Paris Descartes, Novelab

RASPUTIN est un projet de recherche fondamentale collaboratif à l'intersection des « Sciences et Technologies du Numérique » et de la « Psychologie ». Il vise à réduire la complexité cognitive de la navigation pour les non-voyants dans un nouvel environnement intérieur, et ce, grâce à des simulations numériques et des explorations en réalité virtuelle en audio 3D. Ces explorations permettent d'élaborer mentalement une carte cognitive de ces environnements encore jamais physiquement explorés.

L'enjeu technique principal du projet est de générer un rendu audio réaliste d'un édifice (auralisation) en temps réel en s'appuyant, d'une part, sur la librairie de spatialisation Spat~ et, d'autre part, sur la librairie Evertims dédiée à la simulation de la propagation acoustique dans un modèle numérique décrivant les paramètres architecturaux de l'édifice. La topologie des lieux d'intérêts nous a poussés à examiner comment les volumes couplés peuvent être simulés et auralisés de manière efficace. Tout particulièrement, les propriétés de la réverbération tardive dans de tels lieux rendent l'auralisation délicate car elles impliquent l'utilisation de réverbérateurs à décroissance multi-exponentielle. Cette étude nous a donc amenés à proposer un tel réverbérateur, basé sur un réseau de retards rebouclés (FDN Feedback Delay Network), dont les caractéristiques de décroissance peuvent être contrôlées par des paramètres physiques traduisant les effets de couplage entre les volumes. À cet effet, un modèle de radiosité acoustique a été proposé afin de rendre compte de l'hétérogénéité et de l'anisotropie de l'intensité acoustique réverbérée au sein de tels volumes et de sa dépendance en fonction de l'emplacement des sources et récepteurs. Le modèle permet notamment de contrôler la manière dont la puissance de la source est répartie entre les différentes entrées du réverbérateur selon sa position et son orientation dans le volume. De même, il est possible d'estimer comment les différentes sorties du réverbérateur doivent être pondérées pour obtenir le signal du récepteur.



Modélisation de la distribution de l'intensité acoustique au sein du volume.
 À gauche : distribution de la densité d'irradiation aux parois dans le cas de deux volumes couplés en régime de décroissance libre. De larges variations sont observables, notamment à proximité de l'ouverture. À droite : évolution du paramètre régissant la pondération relative des sorties du réverbérateur à décroissance multi-exponentielle en fonction de l'emplacement du récepteur. Des cartographies semblables peuvent être établies pour contrôler la répartition de la puissance d'une source vers les entrées du réverbérateur en fonction de son emplacement.

REFLETS

**Rétroaction émotionnelle faciale et linguistique,
 et états de stress post-traumatique**
 Projet ANR « AAP GÉNÉRIQUE 2017 »

—
 Équipe concernée : Perception et design sonores
 Financement : ANR, programme blanc « Technologies pour la santé »
 Période de réalisation : octobre 2017 – septembre 2021
 Partenaires extérieurs : CentraleSupélec (Rennes, France –
 coordinateur), UMR 8257 Cognac-G (CNRS Paris-Descartes/Service
 de Santé des Armées), Hôpital militaire Percy et Institut de recherche
 biomédicale des Armées (France), Dynamixyz (France), HumanEvo
 (France), Chanel (France)

Le projet REFLETS (Rétroaction émotionnelle faciale et linguistique et états de stress post-traumatique) vise l'amélioration de la prise en charge des personnes souffrant de troubles de stress post-traumatique (TSPT) via un dispositif technologique agissant sur leur capacité à percevoir et réguler leurs propres émotions. Le projet est basé sur de récentes technologies développées à l'Ircam sur la manipulation temps réel des indices du sourire dans la voix parlée, et à CentraleSupélec sur la manipulation vidéo des indices visuels du sourire sur les visages.

Les efforts de REFLETS se concentrent dans trois domaines disciplinaires : du point de vue des sciences de l'information, le projet vise à développer une technologie de « miroir émotionnel » dans lequel un participant peut se voir et s'entendre parler, avec un ton émotionnel manipulé algorithmiquement (en pratique, se voir et s'entendre sourire alors que l'on s'exprime avec un ton neutre). Du point de vue de la psychologie et neurosciences cognitives, le projet vise à étudier les mécanismes de perception de soi et de métacognition impliqués dans ce paradigme de faux feedback, des travaux préliminaires ayant déjà établi que le fait de s'entendre parler avec un ton de voix plus joyeux a un effet positif sur les émotions du locuteur (Aucouturier et al., PNAS 2016). Enfin, du point de vue clinique, le projet vise à tester l'impact thérapeutique d'un tel dispositif dans la prise en charge des déficits de perception de soi et de régulation émotionnelle (alexithymie) dans les troubles post-traumatiques. Le projet permet la mise en place d'essais cliniques sur des patients TSPT, notamment depuis 2019 dans le cadre d'une nouvelle collaboration avec le CHU de Lille.



SEPIA

Sensory and Emotional Processing in Autism spectrum disorders

Projet ANR Programme blanc

—

Équipe concernée: Perception et design sonores

Financement: ANR Programme Blanc (Neurosciences intégratives et cognitives)

Période de réalisation: 2020 – 2024

Partenaires extérieurs: Unité Imagerie et Cerveau (iBrain, Inserm U1253), université de Tours

Le projet SEPIA combine l'expertise de l'équipe PDS en matière de techniques innovantes en psychoacoustique et celle de l'unité Imagerie et Cerveau de l'Inserm/université de Tours (Dr Marie Gomot) dans l'étude neuroscientifique des troubles du spectre autistique (TSA).

Le diagnostic de l'autisme requière la présence conjointe de troubles de la communication sociale et de comportements répétitifs et restreints. Pourtant, peu d'études sur le TSA intègrent les domaines socio-émotionnel et perceptif. Ainsi, il reste encore à établir si les difficultés des TSA sont liées à des déficits émotionnels spécifiques, à des particularités sensorielles qui seraient encore plus marquées pour des stimuli sociaux, ou aux deux.

En utilisant le sourire vocal comme modèle, chaque étape majeure de la boucle Perception-Représentation-Action sera explorée chez les mêmes patients, enfants et adultes. Au niveau perceptif, les corrélats neuroaux de l'encodage de la régularité auditive aux stimuli émotionnels seront étudiés en électrophysiologie (EEG), tandis que des trains de voix répétées souriantes ou neutres seront présentés aux participants. Les représentations mentales du sourire auditif prototypique et le mimétisme moteur et autonome seront étudiés grâce au recueil de données comportementales (reverse corrélations) et physiologiques (électromyographie faciale, pupilométrie) respectivement, alors que les participants écoutent et jugent des expressions vocales souriantes ou non.

Ce projet permettra de dissocier les processus perceptifs et émotionnels, et de déterminer à quel niveau les processus physiopathologiques peuvent intervenir, offrant ainsi un nouvel éclairage mécanistique sur les difficultés socio-émotionnelles des TSA.

theVOICE

Design de voix pour l'industrie créative

—

Équipe concernée: Analyse et synthèse des sons

Financement: ANR, programme Société de l'information et de la communication (DS07)

Période de réalisation: janvier 2018 – juin 2021

Partenaires extérieurs: LIA (Laboratoire d'informatique d'Avignon), Dubbing Brothers

Le projet theVoice s'attaque à la création de voix pour la production de contenu dans le secteur de l'industrie créative (films, séries, documentaires), secteur très important en termes de potentiel industriel mais extrêmement exigeant en termes de qualité. Le projet s'appuie sur un constat simple: la production de voix demeure exclusivement effectuée par des opérateurs humains dans un secteur quasi exclusivement numérique. Les objectifs scientifiques et technologiques du projet visent à modéliser la « palette vocale » d'un acteur pour permettre la recommandation de voix par similarité, et la création de voix artificielles capables de reproduire l'identité vocale d'un acteur. Le projet créera une rupture des usages par la réalisation et l'industrialisation de nouvelles technologies pour la création de contenus vocaux naturels et expressifs. Le consortium, porté par un acteur majeur du secteur de l'industrie de la création de contenus numérique et constitué de laboratoires de recherches reconnus, ambitionne de consolider une position d'excellence de la recherche et des technologies numériques made in France et la promotion de la culture française à travers le monde.



C3. Les dynamiques créatives

ACIDS

Artificial Creative Intelligence and Data Science

acids.ircam.fr

Équipe concernée : Représentations musicales

L'objectif du projet Artificial Creative Intelligence and Data Science (ACIDS) est de modéliser la créativité musicale en développant des modèles d'intelligence artificielle et d'apprentissage machine innovants et de fournir des outils d'exploration intuitive de la créativité. Le projet fournit une ample activité théorique, de modélisation et d'expérimentation d'outils.

L'étude de la créativité dans des situations interactives humain – IA est déterminante pour la compréhension des interactions « symbiotiques ». Le fait de disposer de modèles d'intelligence artificielle capables de démontrer des comportements créatifs pourrait donner naissance à une toute nouvelle catégorie de systèmes d'apprentissage créatif génériques.

Le temps est l'essence même de la musique, et pourtant c'est une donnée complexe, à plusieurs échelles et à plusieurs facettes. C'est pourquoi la musique doit être examinée à des granularités temporelles variables, car une multitude d'échelles de temps coexistent (de l'identité de notes individuelles jusqu'à la structure de morceaux entiers). Nous introduisons donc l'idée d'un apprentissage de la granularité temporelle profonde qui pourrait permettre de trouver non seulement les caractéristiques saillantes d'un ensemble de données, mais aussi l'échelle de temps à laquelle elles se comportent le mieux.

Nous avons par exemple récemment (projet ACTOR) mis au point le premier système de Live Orchestral Piano par apprentissage automatique de répertoires piano/orchestre, permettant de composer de la musique avec un orchestre classique en temps réel en jouant simplement sur un clavier MIDI. En observant la corrélation entre partitions de piano et orchestrations correspondantes historiques, nous pourrions déduire la connaissance spectrale des compositeurs. Les modèles probabilistes que nous étudions sont des réseaux de neurones avec des structures conditionnelles et temporelles.

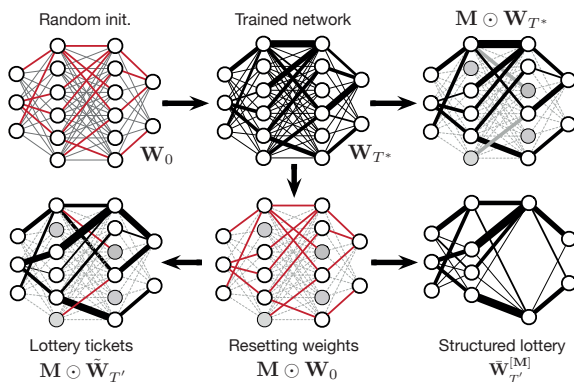
ACIDS encourage un modèle interactif et centré sur l'utilisateur, où nous visons à placer l'humain au centre. Par exemple (collaboration avec les projets REACH et MERCI de l'équipe), les interactions collectives humain machine, notamment l'improvisation nous intéressent comme modèle

général des interactions humaines dans lesquelles la décision, l'initiative et la coopération sont à l'œuvre, et constituent un pont d'observation et de contrôle idéal pour comprendre et modéliser l'interaction symbiotique en général.

Alors que la plupart des recherches actuelles tentent de surpasser les approches précédentes en utilisant des modèles plus complexes et plus lourds, nous encourageons le besoin d'un modèle simple et contrôlable. En outre, nous sommes convaincus que les modèles véritablement intelligents devraient pouvoir apprendre et généraliser à partir de petites quantités de données. L'une des idées clés d'ACIDS est basée sur l'hypothèse de la multiplicité. Ce concept affirme que des informations très complexes dans leur forme originale pourraient se trouver dans un espace plus simple et organisé. Nous cherchons donc à modéliser cette sémantique de haut niveau de la musique par la notion d'espaces latents. Cela pourrait permettre de comprendre les caractéristiques complexes de la musique mais aussi de produire des paramètres de contrôle compréhensibles.

Les espaces de timbre ont été utilisés pour étudier les relations entre les différents timbres instrumentaux, sur la base de notations perceptuelles. Cependant, ils ne permettent qu'une interprétation limitée, aucune capacité de génération et aucune généralisation. Nous étudions dans ACIDS les auto-encodeurs variationnels (VAE) qui peuvent pallier ces limitations, en régularisant leur espace latent pendant la formation afin de s'assurer que l'espace latent de l'audio suit la même topologie que celle de l'espace de timbre perceptif. Ainsi, nous faisons le pont entre l'analyse, la perception et la synthèse audio en un seul système. Les synthétiseurs de son sont omniprésents dans la musique et ils définissent même maintenant entièrement de nouveaux genres musicaux. Cependant, leur complexité et leurs ensembles de paramètres les rendent difficiles à maîtriser. Nous avons créé un modèle probabiliste génératif innovant qui apprend une correspondance inversible entre un espace auditif latent continu des capacités audio d'un synthétiseur et l'espace de ses paramètres. Nous abordons cette tâche en utilisant des auto-encodeurs variationnels et en normalisant les flux. Grâce à ce nouveau modèle d'apprentissage, nous pouvons apprendre les principales macro-commandes d'un synthétiseur, ce qui nous permet de voyager à travers sa multitude de sons organisés, d'effectuer des inférences de paramètres à partir de l'audio pour contrôler le synthétiseur avec notre voix, et même d'aborder l'apprentissage de la dimension sémantique où nous trouvons comment les commandes s'adaptent à des concepts sémantiques donnés, le tout dans un modèle unique.

Les travaux d'ACIDS sont régulièrement utilisés en création contemporaine, par exemple dans la production de la voix d'un chanteur d'opéra synthétique (*La fabrique des monstres*, Ghisi / Peyret) ou des collaborations avec le compositeur Alexandre Schubert sur apprentissage et captation gestuelle.



ACTOR

Analysis, Creation, and Teaching of Orchestration

—

Équipes concernées : Représentations musicales,

Analyse des pratiques musicales

Financement : Conseil de recherches en sciences humaines du Canada (SSHRC, 2019 – 2026)

Partenariat : 20+ institutions américaines et européennes

Le projet ACTOR est une collaboration interdisciplinaire comprenant plus de vingt organisations académiques, industrielles et artistiques travaillant ensemble pour développer une approche de l'orchestration du XXI^e siècle.

Le partenariat ACTOR vise à stimuler l'intérêt pour le timbre et l'orchestration ainsi que leur utilité musicale au sein du milieu académique, des pratiques musicales et du grand public en sollicitant la participation d'artistes, de chercheurs en sciences humaines et de scientifiques de renommée mondiale. Ce partenariat met en relation les pratiques et la pédagogie de l'orchestration nord-américaine et européenne, et encourage le développement de nouveaux outils numériques stimulant la créativité lors de l'apprentissage, de la création et de l'étude de l'orchestration dans les musiques de concert, populaires, de film et de jeu vidéo. En plus il cherche à sensibiliser les auditeurs plus jeunes aux merveilles et à la complexité d'une musique de haute qualité.

ACTOR a pour but de transformer la recherche en musique à la fois grâce à l'application de nouveaux outils analytiques, de manière à mettre en lumière des phénomènes mystérieux qui n'ont pas encore été théorisés sur plus de quatre siècles de musique, ainsi qu'à travers le développement d'outils d'analyse basés sur le son, qui peuvent être appliqués à des musiques non notées ou à des enregistrements de musique écrite.

ACTOR vise une transformation des méthodes d'investigation dans les domaines de la musicologie, de la théorie musicale, de la psychologie de la musique, de l'étude des musiques populaires et de l'ethnomusicologie, où le rôle du timbre en musique a été soit ignoré, soit rarement abordé de manière systématique.

ACTOR enrichira l'enseignement aux compositeurs, aux arrangeurs et aux orchestrateurs en leur fournissant des outils technologiques pour l'apprentissage des liens entre les représentations symboliques présentes dans la partition et le résultat sonore correspondant.

Enfin, ACTOR permettra la création d'outils stimulant la créativité musicale en lien avec le timbre et l'orchestration dans une variété de genres musicaux, à la fois sur scène et à la maison, à travers le développement d'environnements d'orchestration assistée par ordinateur qui intègrent des techniques de pointe en traitement du signal et en apprentissage machine. Afin d'atteindre ces objectifs, les activités en lien avec ACTOR sont structurées en trois axes de recherche : Analyse (Analysis), Développement d'outils technologiques (Technological Tool Development) et Innovation (Innovative Outputs), le tout à travers plusieurs domaines d'expertise. L'équipe est responsable de l'axe 2 et met en œuvre une série de recherches en Intelligence artificielle et machine learning autour de l'analyse et de la génération de l'orchestration, comme le « Live Orchestral Piano » qui arrange automatiquement des séquences pianistiques pour grand orchestre et synthétise le résultat, ou qui peut à l'inverse réduire une partition d'orchestre.

L'agentivité algorithmique

—

Équipe concernée : Analyse des pratiques musicales

Début du projet : 01/01/2021

L'enjeu philosophique soulevé par la révolution numérique est celui d'une possible déshumanisation de l'agentivité. Les progrès de l'intelligence artificielle conduiraient à une nouvelle forme non-humaine d'agentivité. L'opacité qui entoure l'emploi de techniques d'apprentissage profond, ferait alors courir le risque d'une incommensurabilité entre l'agentivité humaine et l'agentivité algorithmique, et par voie de conséquence conduit à l'implosion du concept même d'agentivité. Selon une hypothèse alternative, l'emploi massif d'intelligences artificielles ne conduit pas à une déshumanisation de l'agentivité, mais plutôt à de nouvelles formes de distribution de l'agentivité, entre agents humains et non-humains faisant jouer un rôle prédominant à des artefacts informationnels. La nouveauté ne serait pas tant une déshumanisation, qu'une distribution plus poussée de l'agentivité.

L'IA musicale fournit une collection d'exemples plus ou moins spectaculaires d'algorithmes capables de composer ou d'improviser de la musique comme (ou presque comme) le font des compositeurs et des improvisateurs humains. Ce genre d'exemples joue souvent un rôle central chez ceux qui mobilisent le thème de la déshumanisation, soit pour alerter contre la menace de notre remplacement par des algorithmes, soit pour nous rassurer en montrant que les algorithmes ne sauraient être des « artistes authentiques ». C'est donc un terrain particulièrement utile pour tester l'hypothèse opposée qui est celle de la redistribution. Comme la sociologie de l'art l'a bien mis en évidence, la création artistique est une entreprise éminemment collective, de sorte que l'idée d'une mise en concurrence de Ludwig van Beethoven par un algorithme comme sujet individuel et autonome de création musicale se fonde sur la prémisse erronée selon laquelle l'activité de composer de la musique est une activité homogène d'un seul et unique agent. Au contraire, de nombreuses capacités et ressources hétérogènes sont mises en jeu, et il est banal de remarquer que certaines d'entre elles (réaliser une partie d'accompagnement étant donné une basse chiffrée, par exemple) peuvent être confiées à un élève ou, pourquoi pas, à un algorithme convenablement programmé. Ce qui est beaucoup moins clair, de ce point de vue, c'est la manière dont le complexe d'activités qui sous-tendent un processus créatif peut être décomposé et redistribué entre des agents humains et des artefacts informationnels.

Le projet se dédouble en un versant théorique, et un versant empirique. Sur le plan théorique, la difficulté est de parvenir à penser la composition d'une agentivité créative à partir de composants non agentiels ou non créatifs. L'hypothèse directrice est que ce problème peut être résolu au sein d'une théorie de l'agentivité créative distribuée. Sur le plan empirique, il s'agit d'ouvrir la boîte noire des algorithmes créatifs, par des études de terrains focalisée sur des projets d'IA musicales, incluant en particulier le projet OMax et ses dérivés à l'Ircam.

Analyser la flexibilité temporelle dans les pratiques musicales collectives

—
Équipe concernée : Analyse des pratiques musicales
Collaboration : SOMBY (Central European University, Autriche)
Début du projet : 01/01/2021

Ce projet de recherche prolonge sur un nouveau terrain certaines des questions abordées lors du projet ANR « Improvisation musicale et action collective ». L'idée est maintenant de s'intéresser plus spécifiquement aux formes de coordination musicale qui ne se ramènent pas à une simple synchronisation, ou, pour le dire autrement, à tout le champ de la coordination temporelle qui s'étend au-delà de la synchronisation. Dans un grand nombre de pratiques, les musiciens manifestent en effet des comportements temporellement flexibles, qu'il s'agisse d'une flexibilité « verticale » (maintenir des comportements qui sont, au moins dans une certaine mesure, temporellement indépendants les uns des autres) ou d'une flexibilité « horizontale » (rompre les schémas de synchronisation précédemment établis afin de modifier le tempo commun). Ce faisant, ce projet se propose de résoudre en partie la tension qu'il peut y avoir entre, d'un côté, les sciences cognitives de l'action conjointe – qui ont souvent utilisé la musique comme paradigme expérimental mais en l'envisageant systématiquement sous l'angle de la synchronisation – et, de l'autre, la musicologie empirique de la performance musicale – qui, si elle s'est largement penchée sur les multiples formes de flexibilité temporelle (rubato, swing, groove, etc.), n'a pas encore vraiment interrogé les mécanismes qui rendent cette flexibilité possible.

En combinant études de cas (analyses, enquêtes ethnographiques) et expériences de laboratoire, il s'agit d'analyser ces mécanismes et par là même d'éclairer notre disposition à la flexibilité quand nous agissons ensemble. Ce projet se fait en collaboration avec le Social Mind and Body Group de Central European University (dirigé par Natalie Sebanz et Günther Knoblich).

Biais algorithmiques et injustices esthétiques à l'ère des plateformes numériques

—
Équipe concernée : Analyse des pratiques musicales
Début du projet : 01/01/2021

Les plateformes numériques comme google, youtube, ou spotify sélectionnent les signaux qu'elles transmettent à leurs utilisateurs dans le but de maximiser la probabilité que ces utilisateurs interagissent avec eux. Pour cela, elles s'appuient sur le comportement passé des utilisateurs (les signaux qu'ils ont émis et les traces qu'ils ont laissées). Un effet pervers des algorithmes qui ont pour tâche d'effectuer cette sélection est de nourrir les utilisateurs de signaux avec lesquels ils sont déjà familiers, parce qu'ils véhiculent des opinions ou des valeurs qu'ils partagent déjà, et d'exclure les signaux qui véhiculent des opinions et des valeurs contraires ou tout simplement différentes. Ces effets sont connus sous le nom de « chambres d'écho », et de « bulles épistémiques », lorsqu'il est plus spécifiquement question de la dynamique des croyances des utilisateurs ainsi pris au piège.

S'il n'est pas question de nier l'importance des conséquences épistémiques de ces chambres d'écho, il serait réducteur de les considérer uniquement sous l'angle de la dynamique des croyances. Il est plausible les algorithmes de recommandation aient également des conséquences sur la formation des jugements et des préférences esthétiques. L'hypothèse qu'il s'agit d'explorer concerne l'existence de bulles esthétiques, et dans le cas particulier du son, de ce que l'on pourrait appeler des bulles musicales. De la même manière que des bulles épistémiques se forment lorsque des algorithmes ne nous donnent accès qu'à des informations allant dans le sens de ce que nous croyons déjà, les bulles acoustiques se formeraient lorsque ces algorithmes privilégient les contenus pour lesquels nous avons déjà des préférences esthétiques et filtrent les autres. Les auditeurs de musique baroque n'auraient aucune chance de goûter les plaisirs de la musique électronique minimale berlinoise, et les amateurs de K-pop n'auront jamais l'opportunité d'apprécier (ou de détester) le free-jazz japonais.

Cette hypothèse contient un volet descriptif et un volet normatif. Sur le plan descriptif, la question est de savoir si de telles bulles acoustiques existent, et si oui par quels mécanismes. Sur le plan normatif, la question principale est de savoir si la constitution de bulles esthétiques, (au cas où l'hypothèse descriptive serait confirmée), conduit à une catégorie nouvelle d'injustices esthétiques. Les travaux de Miranda Fricker ont popularisé l'idée selon laquelle il existe des injustices *épistémiques* qui peuvent se produire lorsque les témoignages de personnes appartenant à des groupes marginalisés ne reçoivent pas leur juste crédit, ou lorsque des personnes sont empêchées de donner un sens à certains aspects de leur vie sociale en raison de défauts dans les ressources conceptuelles partagées. L'hypothèse qu'il s'agit d'explorer est l'existence d'injustices *esthétiques*, générées par les bulles esthétiques évoquées, et qui consisteraient en des torts causés à des personnes en tant que sujets de capacités esthétiques, qu'il s'agisse de producteurs ou de récepteurs d'expériences esthétiques.

Composer (avec) l'espace

—

Équipe concernée : Analyse des pratiques musicales

Collaboration : CIRMMT (Université McGill, Canada)

Début du projet : 01/01/2021

Au fil des siècles, l'espace est devenu pour beaucoup d'artistes une dimension à part entière occupant une place centrale au sein du processus créatif. Si la notion d'espace en musique peut souvent renvoyer à des concepts abstraits, elle est ici abordée comme une dimension sonore qui est travaillée par les compositeurs et perçue par les auditeurs. Dans une perspective historique, nous nous intéressons à l'émergence du paradigme spatial dans les musiques de tradition « savante » depuis la Renaissance mais ce projet de recherche vise surtout à mieux formaliser la notion d'espace en s'intéressant d'une part, aux écrits théoriques produits à ce sujet au cours des XX et XXI^e siècles et en analysant d'autre part, un corpus varié d'œuvres musicales. La localisation, le déplacement, le rayonnement des sources sonores, la construction de perspectives imaginaires, l'exploration des propriétés physiques des lieux de diffusion sont le siège d'expériences auditives qui n'ont fait l'objet que d'embryonnaires recherches en musicologie et en psychoacoustique. Il s'agira donc ici de croiser ces deux disciplines pour mieux comprendre comment la dimension spatiale est imaginée, verbalisée et intégrée par les compositeurs, implémentée par les ingénieurs du son, les réalisateurs en informatique musicale et les interprètes, et perçue par les auditeurs.

Composition assistée par ordinateur : écriture du son, du temps et de l'espace

Équipe concernée: Représentations musicales

La recherche en composition assistée par ordinateur (CAO) a pour but d'étudier et concevoir des modèles et techniques informatiques adaptés aux processus de création, intégrant paradigmes de calcul, interactions et représentations musicales. Cette démarche met en avant une orientation symbolique s'appuyant sur les langages de programmation pour la création et le traitement des données harmoniques, temporelles, rythmiques ou des autres aspects entrant en jeu dans les processus compositionnels. Nos travaux dans ce domaine s'articulent principalement autour de l'environnement OpenMusic, un langage de programmation visuelle basé sur Common Lisp et dédié à la composition musicale. Cet environnement utilisé par les compositeurs de musique contemporaine depuis une quinzaine d'années, est aujourd'hui considéré comme l'une des principales références dans le domaine de la composition assistée par ordinateur. Enseigné dans les principaux centres d'informatique musicale ainsi que dans plusieurs universités en Europe et dans le monde, il a fait l'objet de milliers de téléchargements.

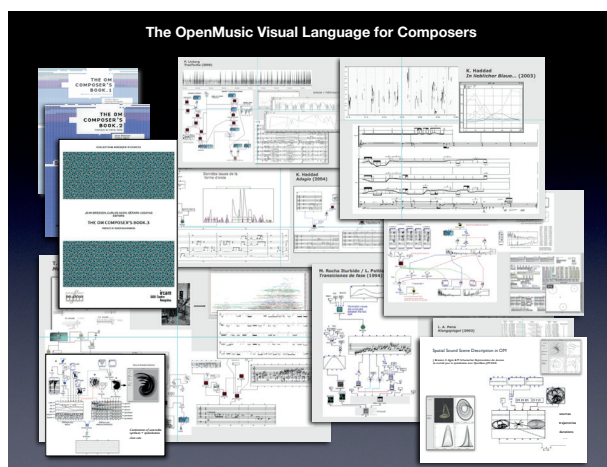
OpenMusic (OM) est un environnement de programmation visuelle pour la composition ou l'analyse musicale assistées par ordinateur. OM offre à l'utilisateur de nombreux modules associés à des fonctions, connectés les uns aux autres pour constituer un programme (ou patch) permettant de générer ou transformer des structures et données musicales. OM propose également de nombreux éditeurs permettant de manipuler ces données, ainsi que des bibliothèques spécialisées dans des domaines tels que l'analyse et la synthèse sonore, les modèles mathématiques, la résolution des problèmes de contraintes, etc. Des interfaces originales comme l'éditeur de maquettes permettent de construire des structures intégrant relations fonctionnelles et temporelles entre les objets musicaux. OpenMusic est utilisé par un grand nombre de compositeurs et de musicologues, et est enseigné dans les principaux centres d'informatique musicale ainsi que dans plusieurs universités en Europe et dans le monde.

Récemment, un nouveau paradigme de calcul et de programmation a été proposé à partir d'OpenMusic, combinant l'approche existante basée sur le style fonctionnel/*demand-driven* à une approche réactive inspirée des systèmes interactifs temps réel (*event-driven*). L'intégration des boucles d'exécution et de calcul des structures musicales permet l'activation de chaînes réactives dans les programmes visuels, et accroît les possibilités d'interaction dans l'environnement de CAO: un changement ou une action de l'utilisateur (événement) dans un programme ou dans les données qui le composent, produit une série de réactions conduisant à sa mise à jour (réévaluation). L'environnement de CAO se trouve alors inséré dans la temporalité d'un système plus large, et potentiellement régi par les événements et interactions produits par ou dans ce système. Cette temporalité peut être celle du processus de composition, ou celle de la performance. Ce projet a produit le logiciel OpenSource OM#, qui est désormais développé indépendamment de l'Ircam.

Les technologies d'analyse, traitement et synthèse du signal sonore permettent d'entrevoir des modalités d'écriture inédites assimilant la création sonore au cœur de la composition musicale. OpenMusic permet l'intégration de telles technologies via un ensemble de bibliothèques spécialisées liant les programmes créés dans l'environnement de CAO aux processus de traitement, de synthèse ou de spatialisation sonore (réalisés notamment par des outils de l'Ircam: SuperVP, Pm2, Chant, Modalys, Spat~, mais également des outils externes comme Csound ou Faust). Ce rapprochement des domaines du son et de la CAO constitue une approche nouvelle pour la représentation et le traitement sonore à travers des programmes et structures de données symboliques de haut niveau.

Développée en collaboration avec le compositeur Marco Stroppa, la bibliothèque OMChroma permet de contrôler des processus de synthèse sonore à l'aide de structures de données matricielles.

Son extension au domaine de la spatialisation, OMPrisma, permet de réaliser des processus de «synthèse sonore spatialisée», faisant intervenir la spatialisation (positions et trajectoires, mais également caractéristiques de salle, orientation ou directivité des sources sonores) au moment de la production des sons. Contrôlés dans OpenMusic grâce à un ensemble d'éditeurs graphiques et d'opérateurs, ces outils offrent une grande richesse dans la spécification conjointe des sons de synthèse et des scènes spatiales. Le projet OM-Chant a récemment permis de remettre sur le devant de la scène la technologie de synthèse par FOFs (fonctions d'ondes formantiques), et de réaliser au cœur des processus de CAO des sons de synthèse inspirés du modèle de production vocale.



Composition urbaine et paysagère

—
Équipes concernées: Espaces acoustiques et cognitifs,
Perception et design sonore

La résidence artistique de Nadine Schütz, porte sur l'intégration des qualités acoustiques et compositions sonores au sein de projets urbains et paysagers. Depuis 2018, ce travail s'est organisé autour de différents cas d'usage, notamment la Place de La Défense, sur laquelle N. Schütz est intervenue en collaboration avec une équipe paysagiste-urbaniste, les Grandes-Serres de Pantin, un site industriel historique sur lequel N. Schütz intervient avec une série d'œuvres évoquant la mémoire sonore industrielle du lieu et, plus récemment, la création de « Niches Acoustiques » sur le parvis du Tribunal de Grande Instance, projet lauréat du Budget Participatif 2019 de la Ville de Paris. Amenée à composer à partir de, avec et dans l'existant, condition inhérente à ce type de projets, N. Schütz interroge l'application d'ambiances et de réponses impulsives (RI) enregistrées en format High Order Ambisonics (HOA), ainsi que les outils de spatialisation associés, lors des différentes étapes des projets artistiques.

Ces enregistrements et mesures permettent de reconstruire au sein du studio un environnement de composition qui porte les traits clés de l'environnement sonore existant du site destinataire de l'œuvre. Ils permettent de simuler l'intégration des nouvelles sources sonores dans ce même contexte. Ils donnent aussi lieu à la création de modèles auditifs qui jouent le rôle de fil conducteur (Leitbild) réveur pour l'étape de création de la matière sonore prenant compte des contraintes techniques du site et du dispositif à mettre en place. Cette recherche implique également l'équipe PDS pour l'analyse des données sonores des lieux mettant en relation des aspects acoustiques et sémantiques ainsi que le développement d'une approche d'évaluation longitudinale, accompagnant la création, la livraison et les premières années de vie d'une installation. Ces études se prolongent par un travail, mené en collaboration avec l'équipe S3AM, sur la complémentarité des approches de modélisation physique et perceptive, notamment pour la préfiguration des projets de création et pour la production des contenus sonores. Cette recherche implique également l'équipe PDS pour l'analyse des données sonores des lieux mettant en relation des aspects acoustiques et sémantiques.



© bureau de Alain Bublax.

Boîte à tubes #1, œuvre sonore et plastique réalisée par Nadine Schütz (2021). Il s'agit d'une archive sonore de l'activité industrielle historique du lieu résultant du travail d'enregistrement et de mesures acoustiques effectuées aux Grandes Serres de Pantin. Le dispositif électroacoustique, logé dans une cabane de contrôle existante située au centre de la grande halle historique, est constitué de transducteurs insérés dans des tubes de différentes dimensions répondant à un accord parfait (la - do - mi).

DAFNE+

dafneplus.eu

Début: 1/7/2022 – Fin: 30/6/2025

Programme: Horizon Europe, appel HORIZON-CL2-2021-HERITAGE-01-03

Decentralized platform for fair creative content distribution empowering creators and communities through new digital distribution models based on digital tokens

—
Équipes:

Direction Innovation et moyens de la recherche, Forum, Web audio musique, Toutes équipes STMS

Partenaires:

Université polytechnique de Madrid (Coordinateur,E), CERTH (GR), Engineering (IT), FabLab Barcelona-Institut d'architecture avancée de Catalogne (E), Netcompany Intrasoft (L), Ecole des arts numériques- Université des médias de Manchester (GB), Synelixis (GR), Université catholique de Louvain (B), Université technique de Tallin (EE)

Le projet DAFNE+ soutenu par le programme Horizon Europe de la Commission européenne, vise à soutenir les créateurs de contenu numérique à travers de nouvelles formes de création, de diffusion et de monétisation de leurs œuvres s'appuyant sur les technologies blockchain. DAFNE+ fournira de nouveaux services et outils permettant la création et la co-création de contenu de manière intuitive et simple sans nécessiter de connaissances techniques sur les blockchains et NFT. Ces outils seront compatibles avec ceux actuellement utilisés par les créateurs de contenu dans leurs studios, afin qu'ils s'intègrent dans les processus de travail existants. DAFNE+ favorisera les communautés de créateurs grâce aux nouvelles organisations autonomes décentralisées (*Decentralized Autonomous Organizations - DAO*) qui se forment autour des NFT. DAFNE+ créera une DAO dans laquelle tous les détenteurs d'un NFT de gouvernance auront un pouvoir de décision sur les règles de la communauté et la plateforme de distribution de contenu qui sera créée. Les recherches de DAFNE+ porteront également sur la définition de nouveaux modèles économiques à travers la distribution de contenus, permettant aux créateurs et aux utilisateurs de monétiser les créations multimédias. Ces nouveaux modèles économiques permettront l'implication de nouveaux acteurs, en tenant compte des besoins de la communauté impliquée, par le biais de la DAO.

DAFNE+ s'adresse à un large éventail de pratiques créatives, notamment les arts visuels, le son et la musique, le design, la mode, l'architecture, les makers, la création amateur, etc.

L'Ircam participe à DAFNE+ et est plus particulièrement en charge de constituer une communauté d'artistes, de designers, de chercheurs/ingénieurs et de collectionneurs/mélobanes autour des technologies du son et de la musique. Pour eux, la plateforme DAFNE+ gèrera deux types de contenus:

- des *composants* - bibliothèques, applications, scripts et bases de données pour le traitement du son et de la musique,
- des *archives* rassemblant les éléments numériques nécessaires à la production et l'exécution d'œuvres (composants, médias, documentation, etc.) dans leurs versions successives.

Dire la Musique

—
Équipe concernée : Analyse des pratiques musicales
Début du projet : 01/01/2021

En regard de l'œuvre et de la création, desquelles ils procèdent souvent, les écrits de compositeurs sont indispensables à la compréhension de la musique du XX^e siècle. Avec Charles Baudelaire naissait en effet l'exigence de l'artiste raisonnant son art et découvrant les lois obscures en vertu desquelles il écrit : « Il est impossible qu'un poète ne contienne pas un critique. » Biographiques, critiques, théoriques, esthétiques, généalogiques, balisant les chemins d'une œuvre encore en devenir ou se livrant à l'introspection analytique, les écrits de compositeurs renouvellent concepts et méthodes des discours sur la musique.

Parfois menacées par la dispersion des archives et la détérioration, sinon la disparition des manuscrits et des tapuscrits en langue originale, ainsi que des bandes magnétiques d'entretiens, les éditions envisagées, volontiers exhaustives, situent chacun des articles et proposent un appareil critique (références historiques, biographiques et bibliographiques, index, index des œuvres, catalogues...), mettant ainsi à la disposition des musiciens, des mélomanes, des institutions et des salles de concert des documents réunis, traduits et annotés.

Après les éditions d'écrits de Jean Barraqué, Giacomo Manzoni, Luigi Nono, Louis Saguer, Salvatore Sciarrino, Frédéric Durieux ou Niccolò Castiglioni, Laurent Feneyrou prépare des volumes consacrés à Franco Donatoni, Giuseppe Sinopoli et Emmanuel Nunes.

DOTS

Objets musicaux distribués pour les interactions collectives

—
Équipe concernée : ISMM
Financement : ANR CE 33
Période de réalisation : janvier 2023 - janvier 2026
Partenaires extérieurs : LISN/Ex-Situ, I3S-Université Côte d'Azur, MSH Paris Nord, ENSAD-Lab

Le développement récent des technologies Web et Internet des objets (IoT) - en démocratisant l'accès et la manipulation des technologies tant aux niveaux logiciels que matériels - a ouvert de nouvelles perspectives aux chercheurs et aux artistes s'appuyant sur des environnements multimédias distribués et s'intéressant aux interactions collaboratives. Cependant, la complexité inhérente à de tels systèmes distribués, en particulier quant aux nécessités d'interactions temps réel, synchronisées et multimodales propres à la musique et aux arts vivants, pose de nouvelles questions sur la conception d'interactions permettant à ces praticiens d'utiliser et surtout, de s'approprier de manière créative ces technologies.

Dans ce contexte, les objectifs du projet Distributed Musical Things for Collective Interactions (DOTS) sont doubles : 1) Comprendre comment notre système expérimental pourrait mieux soutenir le travail créatif, collaboratif et multidisciplinaire dans des situations réelles, en questionnant l'impact des technologies dans de telles tâches. En particulier, notre objectif est de développer un cadre méthodologique destiné à faciliter le processus d'apprentissage des utilisateurs et à soutenir leur appropriation des aspects tant technologiques que de conception. 2) Développer une nouvelle version de notre plateforme pour créer un écosystème cohérent d'outils modulaires, portables, interopérables et à bas coût (ciblant à la fois les plateformes mobiles et embarquées) dédié aux applications créatives et musicales, y compris pédagogiques. En particulier – en tenant compte des spécificités liées aux interactions multimodales continues et synchronisées dans les systèmes distribués – notre objectif est de développer de nouveaux outils et de nouvelles interfaces qui permettent de soutenir le travail créatif (e.g. feedback immédiat, boucles rapides d'essais et d'erreurs) d'utilisateurs ayant différents parcours, compétences et objectifs.

Partant de l'hypothèse que les pratiques artistiques expertes nous fournissent un cadre exemplaire et stimulant pour aborder ces questions de recherche et proposer des technologies innovantes, nous mènerons ce projet selon une méthodologie de recherche-action et en nous appuyant sur le cadre conceptuel du méta-design. Plus précisément - dans une perspective interdisciplinaire centrée sur la pratique mêlant Interaction Homme-Machine, Ubiquitous Computing, Sound and Music Computing, Design et Arts - nous mettrons en œuvre une méthodologie itérative et incrémentale alternant phases de développement et travaux de terrain menés lors de résidences artistiques et d'ateliers menés dans des écoles d'art et des universités. Ainsi, nous allons à la fois 1) développer de nouveaux outils (logiciels et matériels) favorisant l'appropriation et le travail créatif au sein des systèmes distribués musicaux, et 2) produire des études de cas documentées nous permettant de mieux comprendre comment les praticiens experts en arts et en musique interagissent avec les outils technologiques et se les approprient de manière créative. Nous pensons également que les technologies et les méthodologies développées au cours du projet favoriseront la mise en place de nouveaux cadres expérimentaux permettant de soutenir des recherches dans d'autres domaines tels que la pédagogie ou le bien-être et le soin.

(DYCI2) Agents créatifs, interactions improvisées et « meta-composition »

Équipe concernée : Représentations musicales

L'équipe explore le paradigme de la créativité computationnelle à l'aide de dispositifs inspirés par l'intelligence artificielle, dans le sens des nouvelles interactions symboliques musicien-machine ou dans celui de la science des données et l'extraction des connaissances.

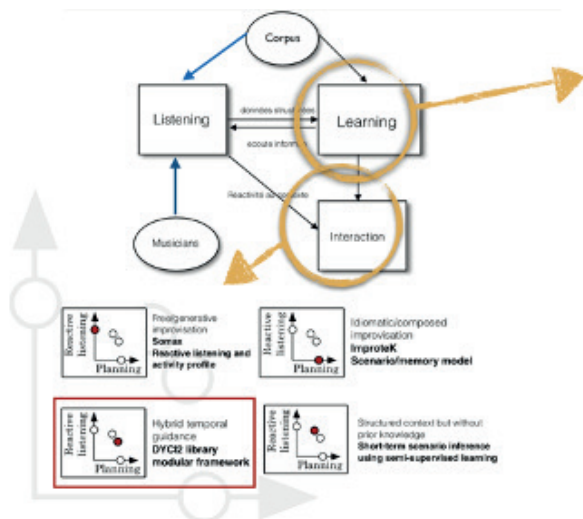
Dans la lignée du logiciel Omax, les recherches sur l'apprentissage et la génération musicale interactive ont abouti à la création de plusieurs paradigmes d'interaction musicien-machine caractérisés par une architecture combinant écoute artificielle du signal, découverte d'un vocabulaire symbolique, apprentissage statistique d'un modèle de séquence et génération de nouvelles séquences musicales par des mécanismes réactifs et/ou de planification (scénario).

Le projet DYCI2 met en avant l'interaction improvisée, à la fois comme modèle anthropologique et cognitif de l'action et de la décision, comme schéma de découverte et d'apprentissage non supervisé, et comme outil discursif pour l'échange humain – artefact numérique, dans une perspective de modélisation du style et de l'interaction.

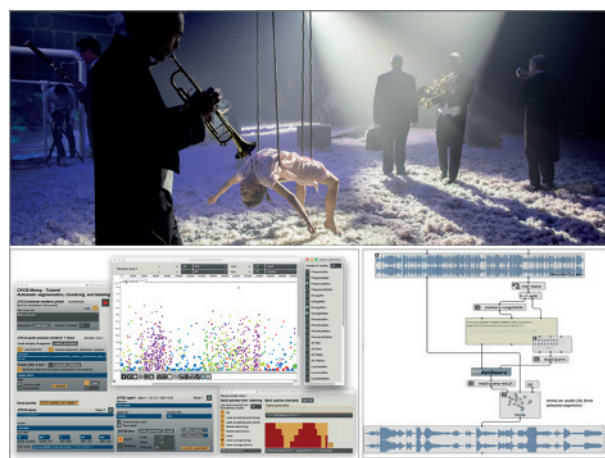
L'objectif est de constituer des agents créatifs autonomes par apprentissage direct résultant d'une exposition au jeu vivant (live) de musiciens humains improvisant, en créant une boucle de rétroaction stylistique par l'exposition simultanée de l'humain aux productions improvisées des artefacts numériques eux-mêmes, donc à partir d'une situation de communication humain-artefact évoluant dans une dynamique complexe de co-créativité.

Ces agents/instruments s'inscrivent également dans une recherche de nouveaux paradigmes génératifs « informés par IA », en travaillant par exemple à associer aux mécanismes génératifs des modules d'extraction et inférence de grille harmonique en temps réel.

Après avoir été le socle de productions d'envergures permettant une validation par des musiciens experts (Pascal Dusapin, Bernard Lubat, Steve Lehman, Rémi Fox, Hervé Sellin, etc.) ainsi que de workshops et festivals (Festival Improtech Paris-Athina et Paris-Philly; MassMoca, Cycling '74, etc.) les thématiques DYCI2 se poursuivent avec le lancement du projet ANR MERCI (Mixed Musical Reality with Creative Instruments) coordonné par G. Assayag avec l'EHESS et la startup HyVibe, ainsi que l'ERC Advanced Grant REACH porté par G. Assayag. La mise en place d'environnements homme-machine puissants et réalistes pour l'improvisation nécessite d'aller au-delà de l'ingénierie logicielle des agents créatifs avec des capacités d'écoute et de génération de signaux audio. Cette ligne de recherche propose de renouveler radicalement le paradigme de l'interaction improvisée homme-machine en établissant un continuum allant de la logique musicale co-créative à une forme d'« interréalité physique » (un schéma de réalité mixte où le monde physique est activement modifié) ancrée dans des instruments acoustiques.



Paradigmes d'interaction musicale humain-machine



« Lullaby Experience » – Librairie DYCI2 – Librairie om-dyci2

© Quentin Chevrier

GRIAMI

Groupe de recherches interdisciplinaires sur l'acte musical et interartistique

—

Équipe concernée : Perception et design sonores

Financement : ITI GRIAMI, université de Strasbourg

Période de réalisation : 2021-2024

Collaborations : ITI GRIAMI / EA 3402, université de Strasbourg.

Cette recherche sur l'acte musical tente de comprendre, d'un côté, comment fonctionne l'activité artistique et l'intelligence musicale dans leur plasticité par la pratique, l'apprentissage et la créativité orale/aurale, et, de l'autre côté, comment un groupe social se donne des règles d'organisation et des cadres pour penser la musique particularisée.

L'oralité est un processus constituant susceptible de conditionner et de créer des formes improvisées et spontanées, que ce soit musicales ou poétiques (Lortat-Jacob). L'immense terrain musical que représente le bassin méditerranéen est un exemple fascinant de transfert d'éléments culturels, mais aussi de construction de spécificités identitaires. La compréhension de ses mécanismes, engageant des processus d'*échange*, d'*interaction*, d'*imitation*, de *mémorisation*, de *reproduction*, implique la construction d'une anthropologie globale du phénomène de l'écoute, pour laquelle la connaissance approfondie des aspects cognitifs, dans le rapport qu'instaure la musique orale avec la mémoire, est essentielle (Chouvel).

Les enquêtes de terrain et les expériences psychologiques se focalisent principalement sur le rôle de l'acculturation et l'implication des schémas cognitifs dans l'écoute de l'auditeur ainsi que sur les rapports complexes du créateur (improvisateur, interprète) à la musique, à la culture commune (l'art de s'émouvoir ensemble, Daring) et aux normes de la tradition ; c'est-à-dire comment, et sur quels critères, l'improvisateur pourrait être porteur d'un don particulier pour le groupe social, et avoir un lien spécifique avec ce qui est hors de l'ordinaire et du commun ? Notre intérêt est porté également sur le vécu et la profondeur sémantique de tout ce que le musicien peut dire sur sa pratique sociale et son art ; verbalisations diverses, formulations métaphoriques, représentations symboliques, etc. Ce sont les expériences esthétiques des sujets-acteurs et les stratégies performatives et cognitives actées qui retiennent notre attention. Autant de références culturelles et psychologiques à recueillir sur le terrain de la performance, à analyser et à rendre compte dans notre « modèle d'écoute culturalisée ».

Instruments of Improvisation

—

Équipe concernée : Analyse des pratiques musicales

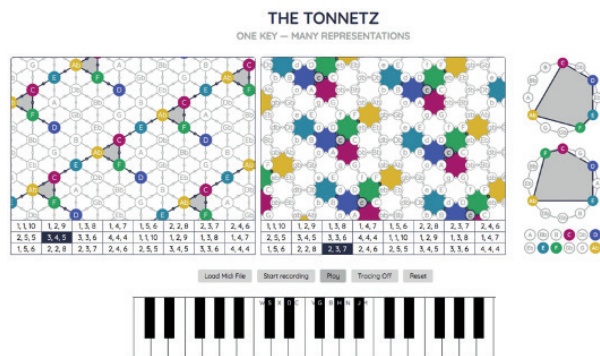
Ce projet prolonge les travaux menés précédemment dans le cadre du projet «Music & Hacking». Il vise à analyser les multiples transformations matérielles, adjonctions ou inventions que les musiciens exercent sur leurs instruments, considérés comme des moyens de singularisation artistique, d'optimisation ergonomique, de circulation institutionnelle ou d'extension cognitive. Près de cinquante entretiens filmés ont été réalisés avec des instrumentistes aux pratiques d'improvisation variées (improvisation libre, jazz, musique baroque, musique traditionnelle). Chaque musicien a été interviewé dans son atelier, instrument « en main », afin de permettre une exploration la plus fine possible de ses pratiques luthériennes. Ce travail aboutira à l'écriture d'un livre, et sera accompagné de la réalisation d'un web-documentaire.

Mathématiques et musique

Équipe concernée : Représentations musicales

Le projet SMIR (« Structural Music Information Research ») prolonge le projet retenu par le CNRS à l'occasion de la création d'un poste de directeur de recherche au sein de l'équipe Représentations musicales en élargissant le spectre d'outils mathématiques au service du musicologue computationnel (des outils issus non seulement de l'algèbre mais aussi de la topologie et de la théorie des catégories). En même temps, il s'attaque aux articulations entre musique savante et popular music (rock, pop, jazz et chanson), dans une dimension à la fois théorique mais aussi pratique et ouverte à la diffusion des savoirs. Le projet bénéficie d'un quadruple ancrage institutionnel au sein de la *Society for Mathematics and Computation in Music*, (société internationale dont l'activité a été fédérée par la revue *Journal of Mathematics and Music*), d'un partenariat avec l'université de Strasbourg (en particulier l'IRMA, le Labex GREAM, l'Idex « Université et Cité » et le Jardin des Sciences) et d'une collaboration avec l'Institut Pasteur (dans le cadre du « Défi - Processus et techniques d'apprentissages 2020 » financé par la Mission pour les initiatives transverses et interdisciplinaires du CNRS) et avec l'université de Padoue (création d'une chaire de Professeur invité pour la nouvelle bourse doctorale « Mathématiques et Musique », 2020-2021). Ce projet alimente, de surcroît, le séminaire MaMuPhi (mathématiques, musique et philosophie), coorganisé avec l'École normale supérieure, ainsi que les deux collections Musique/Sciences (chez Delatour France) et Computational Music Sciences (chez Springer).

Le projet s'appuie sur un environnement expérimental d'aide à l'analyse musicale computationnelle appelé The Tonnnetz, qui s'appuie sur les fonctionnalités du logiciel HexaChord. Tonnnetz est un environnement web permettant de construire des représentations spatiales associées à des progressions harmoniques et de les analyser à travers plusieurs espaces géométriques (les Tonnnetze). L'environnement est couramment utilisé dans des ateliers scolaires et grand public ainsi que dans des projets de diffusion des savoirs (en particulier dans le cadre du spectacle « Math'n Pop »).



MERCI

Mixed Musical Reality with Creative Instruments

Équipe concernée: Représentations musicales

Financement: ANR.

Période: 2020-2022

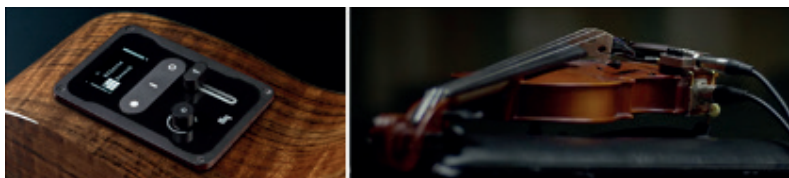
Partenaires: HyVibe, CAMS/EHESS

L'improvisation peut être considérée comme une force motrice majeure dans les interactions humaines, stratégique dans tous les aspects de la communication et de l'action. Dans sa forme la plus élevée, l'improvisation est un mélange d'actions structurées, planifiées, dirigées, et de décisions et déviations locales difficilement prévisibles optimisant l'adaptation au contexte, exprimant de manière unique le moi créatif, et stimulant la coordination et la coopération entre les agents. La mise en place d'environnements homme-machine puissants et réalistes pour l'improvisation nécessite d'aller au-delà de la simple ingénierie logicielle des agents créatifs avec des capacités d'écoute et de génération de signaux audio, comme ce qui a été fait la plupart du temps jusqu'à présent. Dans les nouvelles configurations d'«*interréalité physique*» (un schéma de réalité mixte où le monde physique est activement modifié par l'action humaine), les sujets humains sont immergés et engagés dans des actions tangibles permettant une pleine incarnation (*embodiment*) dans les mondes numérique, physique et social.

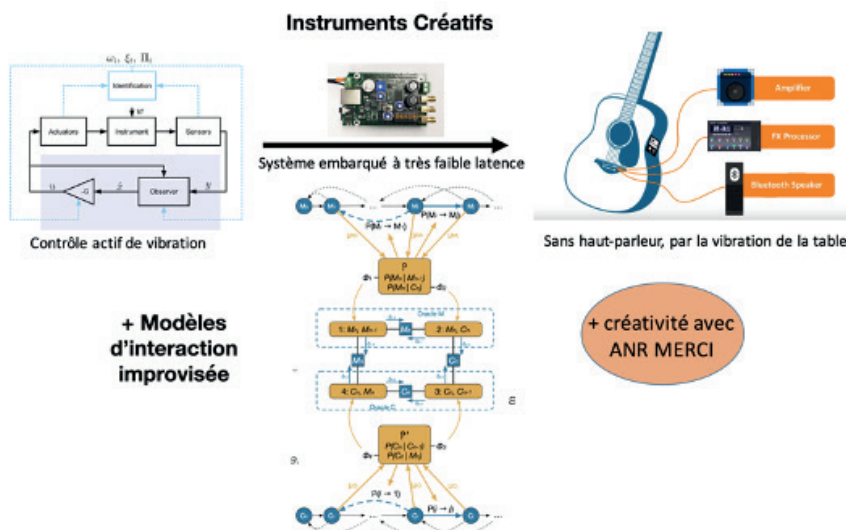
L'objectif principal de ce projet est de créer les conditions scientifiques et technologiques pour des systèmes musicaux à réalité mixte, permettant des interactions improvisées homme-machine, basées sur l'interrelation d'agents numériques créatifs et le contrôle acoustique actif dans les instruments de musique. Nous appelons ces dispositifs de réalité mélangée

des instruments créatifs. L'intégration fonctionnelle de l'intelligence artificielle créative et du contrôle actif de l'acoustique dans le cœur organologique de l'instrument de musique, afin de favoriser des situations d'interréalité physique plausibles, nécessite la synergie de recherches publiques et privées hautement interdisciplinaires, telles qu'elles sont menées par les partenaires. De tels progrès seront susceptibles de perturber les pratiques artistiques et sociales, et d'avoir à terme un impact puissant sur l'industrie musicale ainsi que sur les pratiques musicales amateurs et professionnelles.

Un instrument créatif sera capable d'écouter en permanence la performance du musicien qui en joue, afin de déterminer ses orientations musicales, et de dialoguer de manière créative avec lui dans une expérience inédite de réalité musicale mixte où les sons produits par l'interprète et ceux créés artificiellement se mélangeront dans une polyphonie musicale significative. Cette réalité musicale mixte se déploiera dans les dimensions mélodique, harmonique, rythmique et orchestrale de la musique, avec un compagnon musical créatif situé au cœur même de l'instrument et agissant comme un avatar artistique, un assistant à la création, un partenaire, une incitation nouvelle à apprendre, à pratiquer et à communiquer dans le langage musical.



1/ Le tableau de contrôle de la HyVibe Guitar 2/ Premier prototype de violon SmartInstrument à l'Ircam



MICA

Improvisation musicale et action collective

—

Équipes concernées : Analyse des pratiques musicales, Perception et design sonores, Interaction son musique mouvement
Financement : ANR, programme JCJC
Période de réalisation : mars 2018 – février 2021
Partenaires extérieurs : Sorbonne Université, EHES
Coordinateur : Ircam

Le projet MICA rassemble une équipe pluridisciplinaire, à la croisée de la musicologie, de la philosophie et des sciences cognitives. Son objectif est, premièrement, de développer une nouvelle série de méthodes pour l'analyse des pratiques d'improvisation ; et, deuxièmement, d'utiliser les situations d'improvisation collective libre comme une sorte de « laboratoire » permettant de poser des questions d'ordre général sur l'action conjointe.

Dans un premier temps, l'équipe du projet MICA collectera des descriptions denses et précises d'actions collectivement improvisées et des processus de coordination à l'œuvre dans ce type d'action à la fois complexe et spontanée à travers plusieurs enquêtes ethnographiques longitudinales auprès de différents collectifs d'improvisateurs. L'étude croisée de ces différents terrains permettra d'observer comment le mode opératoire de la création collective se modifie lorsque les conditions de la coordination entre les musiciens varient.

Ce travail ethnographique fournira les bases qui nous permettront de renouveler en profondeur les questions expérimentales traditionnellement posées au sujet de l'action conjointe, en abordant la question de la coordination émergente autrement que par le paradigme de la synchronisation. Il s'agira donc dans un deuxième temps de mettre en place une série de protocoles expérimentaux, inscrits dans la psychologie expérimentale et la cognition sociale, qui auront pour objectif de mettre en lumière certains aspects fondamentaux de l'interaction dyadique improvisée.

Enfin, il s'agira de proposer un travail d'élaboration conceptuelle en confrontant aux cas des actions conjointes improvisées les notions de coordination et de planification telles qu'elles sont généralement définies au sein des théories philosophiques et psychologiques de l'action conjointe. Ainsi, ce sont toutes ces actions collectives à la fois créatives et spontanées – qui occupent une place si centrale dans nos vies – que nous tenterons d'éclairer d'un jour nouveau.

Music & Fun

—

Équipe concernée : Analyse des pratiques musicales
Début du projet : 01/01/2021

S'il existe bien une sous-discipline de la musicologie qui se définit comme « ludo-musicologie », celle-ci traite de fait essentiellement de musique de jeu vidéo ou de la manière dont la musique participe de l'expérience ludique procurée par les jeux vidéo. Le projet Music & Fun vise pour sa part quelque chose de beaucoup plus général, l'idée étant ici de s'intéresser aux œuvres et aux pratiques dont le but est de susciter l'*amusement* (et qu'il convient de bien distinguer du divertissement comme de la joie). Il y a là une des dimensions fondamentales du faire musical : nous musiquons (qu'il s'agisse d'écoute ou de performance) aussi pour nous amuser. Mais si les sciences de la musique se sont penchées depuis longtemps sur d'autres dimensions des pratiques musicales, la condition ludique de la musique reste encore étonnamment largement à questionner.

Ce projet s'articule autour d'un certain nombre de chantiers : entretiens avec des instrumentistes autour des sons qu'ils considèrent comme amusants ; enquête ethnographique sur les productions du label Dur&Doux ; analyse des formes d'humour musical au sein de la culture web ; et expériences de psychologie visant à explorer systématiquement les mécanismes sous-tendant la production et la perception de l'humour musical.

Musique et Maladie

Équipe concernée : Analyse des pratiques musicales

Début du projet : 01/01/2021

1. Le discours médical a souvent abordé le thème de l'usage médical de la musique :

- musique et grossesse, Hippocrate associant déjà embryologie et musique dans Du régime. Quelle musique la femme gravide peut-elle écouter pour le bien-être du fœtus et le sien ?
- musique et petite enfance. Quel soin musical prodiguer à l'enfant, s'il est malade, s'il ne peut se retenir d'uriner au lit, s'il peine à s'endormir, s'il est inquiet, s'il est agité... ?
- musique et alimentation. Qu'écouter pour faciliter la digestion de tel ou tel mets ?
- musique et bain, pour le bien-être – les instituts de beauté actuels n'y échappent pas.
- musique et eros, les médecins de l'école de Montpellier considérant que la musique disposerait à l'activité sexuelle et guérirait l'impuissance. Ailleurs, elle calme les ardeurs trop pressantes, comme chez Athanasius Kircher, où elle sert à reconduire les érotomanes à la tempérance.
- la musique contre l'agitation et la perte du sommeil – ce que l'école pythagoricienne pratiquait déjà, par des chants et mélodies assurant un sommeil calme, aux rêves peu nombreux.
- la musique comme remède contre la maladie ou soutien de la thérapie...

Ce corpus ouvre un vaste champ de pensée entre histoire de la musique et histoire des sciences.

2. Le second axe, plus courant, est celui de la pathographie, visant à décrire les conséquences de la maladie sur la biographie, la pratique et l'esthétique d'un musicien. En un saisissant renversement datant de l'âge moderne, ce musicien, autrefois sorte d'Orphée, exprime, expose son pathos, et devient malade, de plus en plus gravement : Bach, Haendel, Mozart, Beethoven, Schubert, Mendelssohn, Chopin, Schumann, Puccini, Ravel, Bartók, parmi d'autres, permettent d'aborder la chirurgie de l'œil, des tableaux cliniques complexes, deux des grands maux du XIX^e siècle, la syphilis et la tuberculose, une maladie héréditaire, des troubles psychiatriques, le cancer ou le Sida, mais aussi des pathologies construites comme spécifiquement musicales (amusies, hypermusies, « vers auditifs »), ainsi que l'histoire de l'hygiène, de la sexualité, de la pharmacologie, de l'hôpital..., en prêtant attention aux mutations du regard médical.

3. Il s'agit enfin d'étudier comment des œuvres musicales reflètent des pathologies, mentales ou somatiques. Il y a, bien sûr, le cas de musiciens psychotiques, mais aussi les principes d'écriture musicale qui dénotent la psychose à travers les siècles. Il convient également, dans ce contexte, d'étudier les disciplines médicales, en mesurant, par exemple, l'histoire de la pédiatrie à l'aune de l'enfance en musique, l'histoire de la cardiologie à celle du rythme, ou l'histoire de la pneumologie à celle du souffle, en rapportant les maladies aux discours et représentations musicaux, notamment dans le domaine, littéraire et scénique, de l'opéra.

Orchestration assistée par ordinateur

Équipe concernée : Représentations musicales

Ce projet aborde la question de l'orchestration par la recherche automatique d'instrumentations et de superpositions d'instruments approchant, en fonction de différents critères de similarité acoustique, une cible définie par le compositeur. Les recherches actuelles s'attachent à étendre ce paradigme à l'orchestration dynamique, suivant des cibles dont les caractéristiques sonores varient au cours du temps.

Réalisé à la suite du logiciel Orchidée, qui a défini toutes les bases en termes de méthodologies et de fonctionnalités de l'aide à l'orchestration, Orchids est le premier système de la série Orchid* pour l'orchestration assistée par ordinateur et l'optimisation de mélanges de timbres qui permet de gérer les dynamiques continues d'évolution. Il fournit un ensemble d'algorithmes permettant de reconstruire n'importe quelle cible sonore évoluant dans le temps par une combinaison d'instruments ou échantillons, selon un ensemble de critères psycho-acoustiques. Il peut aider les compositeurs à obtenir des couleurs de timbre inouïes en fournissant une multitude de solutions efficaces qui recréent au mieux cette cible sonore. Grâce à un ensemble étendu de fonctionnalités, Orchids peut également reproduire des évolutions et formes abstraites de mouvements spectraux. Ses résultats fournissent des partitions d'orchestre multiples pouvant être organisées de manière intuitive afin d'obtenir rapidement une réalisation d'idées orchestrales et musicales. Ce système fournit plusieurs algorithmes d'approximation permettant d'optimiser conjointement plusieurs propriétés de timbre. Les avantages du système Orchids résident dans le fait que cette approximation peut être faite séparément sur des formes temporelles, valeurs moyennes ou écarts-types (ou toute combinaison des trois) de chaque descripteur psychoacoustique. En outre, les utilisateurs peuvent également définir une déformation temporelle manuelle, et même effectuer une recherche multicible à l'intérieur de multiples segments sonores, offrant ainsi des réalisations de pièces orchestrales complètes en quelques secondes.

La nouvelle version de la série orchid*, orchidea, développée dans l'équipe RepMus en collaboration avec la HEM, est partie d'orchids pour optimiser encore plus l'analyse, la recherche et l'interface, elle est désormais disponible en open-source, et propose des optimisations significatives et une plus grande fiabilité des résultats.

Outre la production des logiciels phares Orchid*, le projet orchestration de l'Ircam suscite de nombreuses collaborations nationales et internationales coordonnées par l'équipe RepMus : le projet ANR MAKIMONO explore les apports de l'intelligence artificielle et du Machine Learning à la synthèse timbrale et orchestrale ; le projet Orchestration financé par le Conseil de recherches en sciences humaines du Canada (SSHRC) avec la Haute École de musique de Genève (Suisse) a permis de poser les bases d'un vaste projet international impliquant 19 partenaires américains et européens, le projet ACTOR, qui explore toutes les facettes de l'orchestration, de la modélisation à la pédagogie et a démarré en 2019.

Philosophie du son numérique

—

Équipe concernée : Analyse des pratiques musicales

Début du projet : 01/01/2021

Les philosophes de la perception auditive ont récemment fait progresser notre compréhension de la nature des sons, compris comme les objets de la perception auditive, mais les débats se sont focalisés principalement sur la nature des sons naturels, résultant d'événements vibratoires dans nos environnements physiques proches, et seulement dans une moindre mesure sur les sons enregistrés et les images sonores. Il est toutefois contestable de ranger tous les sons numériques dans la catégorie des enregistrements, car il existe des sons numériques synthétiques. Il est également discutable de les subsumer sous la catégorie plus large des images sonores dans la mesure où certains événements sonores qui peuplent notre quotidien numérique, tels les alarmes et les notifications, ne visent pas à représenter d'autres sons qu'eux-mêmes, mais à performer des actes d'avertissement ou de notification.

Ce projet se fonde sur la prémisse selon laquelle les sons numériques ont des conditions d'existence et d'identité distinctes de celles des sons naturels, ce qui nécessite de construire une ontologie philosophique distincte pour les sons numériques. Selon l'hypothèse envisagée, les conditions d'identité et d'existence des sons numériques ne sont pas dépendantes d'un événement vibratoire au sein d'un milieu de propagation, mais plutôt d'une structure de données qui joue un rôle analogue à celui d'une notation, et explique la répétabilité et la portabilité des sons numériques. Cette hypothèse a également des conséquences sur les théories de la notation des musiques numériques, et sur la question du réalisme phonographique appliqué aux enregistrements sonores numériques.

RAMHO

Recherche et Acoustique Musicales en France après 1945 : une Histoire Orale

—

Équipe concernée : Analyse des pratiques musicales

Financement : Sorbonne Université, programme Émergences (2021-2023)

C'est au cours de la seconde moitié du XX^e siècle que l'acoustique musicale et la recherche musicale se sont imposées comme des disciplines scientifiques à part entière avec l'émergence, à travers le monde, de laboratoires et de studios s'intéressant à la science des sons au prisme des pratiques musicales. La France a largement contribué à l'institutionnalisation de ces champs de recherche avec, entre autres, la création du Groupe de Recherche Musicale (GRM) en 1958, du Laboratoire d'Acoustique Musicale (LAM) en 1963 ou de l'Institut de Recherche et de Coordination Acoustique/Musique (Ircam) en 1977 mais bien d'autres structures virent le jour jusqu'à la fin des années 1980. Comment sont-elles nées ? Quels étaient les enjeux scientifiques ? De quels moyens humains et matériels disposaient-elles ? Comment se positionnaient-elles les unes par rapport aux autres ? Pour répondre à ces questions, les chercheurs associés au projet RAMHO mènent de nombreux entretiens semi-dirigés auprès des figures historiques de ces différentes institutions françaises et consultent aussi des fonds d'archives publiques ou privés de sorte à retracer de manière précise l'histoire de ces institutions et la carrière scientifique de celles et ceux qui en furent les acteurs principaux. Le projet RAMHO permet ce faisant de nourrir des discussions entre musicologues, acousticiens, informaticiens et historiens des sciences, sur la pluralité des approches scientifiques de la musique, la diversité des outils technologiques mis en jeu et les diverses fonctions assignées à ce qu'on appelle l'acoustique musicale et la recherche musicale. Une base de données sera spécialement développée afin de diffuser l'ensemble des entretiens menés au cours du projet et de partager certains documents d'archive. En vue de la commémoration du cinquantième anniversaire du Centre Georges Pompidou en 2027, la genèse de l'Ircam sera plus particulièrement scrutée.

REACH

Raising Co-Creativity in Cyber-Human Musicianship

Équipes concernées : Représentations musicales

Financement : European Research Council Advanced Grant de Gérard Assayag (2021 – 2025)

Collaborations : CAMS / EHESS, UCSD

L'imbrication de la créativité humaine avec les processus computationnels environnants suscite de nouvelles potentialités d'action conjointe homme-machine et de nouveaux modèles d'interaction. Des formes originales de réalité partagée impliquant des «interactions symbiotiques» se généraliseront dans les arts et l'activité humaine en général. La co-créativité entre les humains et les machines entraînera l'émergence de structures d'information distribuées et performatives impliquant la coopération d'agents artificiels et humains. Cela transformera certainement les ordres culturels connus et aura un impact significatif sur le développement humain. L'apprentissage génératif de représentations symboliques à partir de signaux physiques et humains, et la compréhension des stratégies artistiques et sociales de l'improvisation nous aideront à mieux comprendre cette dynamique de coopération (ou de conflits) inhérente aux réseaux cyber-humains.

Pour cela, le projet REACH vise à comprendre, modéliser et développer la co-créativité musicale entre les humains et les machines par le biais d'interactions improvisées, permettant aux musiciens de tout niveau de formation de développer leurs compétences et d'accroître leur potentiel créatif individuel et social. En effet, l'improvisation est au cœur même de toutes les interactions humaines, et la musique est un terrain fertile pour développer des modèles et des outils de créativité qui peuvent être généralisés à d'autres activités, du fait de la richesse de ses structures et de ses contraintes qui favorisent l'émergence de comportements coopératifs et de cours d'actions fortement intégrés. REACH étudiera la «musicalité partagée» à l'intersection des sphères physique, humaine et numérique, comme archétype d'une intelligence distribuée, et produira des modèles et des outils pour mieux comprendre et encourager la créativité humaine dans un contexte où elle est de plus en plus imbriquée avec le calcul.

REACH part de l'hypothèse que la co-créativité dans les systèmes cyber-humains résulte d'une forme d'émergence de comportements cohérents et de régimes non-linéaire de formation d'événements et de structures, qui conduisent à une riche co-évolution des formes musicales. Ces phénomènes résultent de processus d'apprentissage croisé entre agents impliquant des boucles de rétroaction et des mécanismes de renforcement complexes. REACH étudiera ces mécanismes in vivo et in vitro, et produira des outils créatifs, par la convergence de méthodes issues des recherches en créativité computationnelle interactive, en intelligence artificielle et machine learning, en sciences sociales avec l'anthropologie des pratiques improvisées (collaboration avec le CAMS à l'EHESS), et des systèmes de réalité mixte instrumentaux (collaboration avec la compagnie HyVibe).

Systèmes distribués sonores et interactifs

Équipe concernée : Interaction son, musique, mouvement

Ce projet, qui s'inscrit dans la suite des projets CoSiMa et Wave, concerne le développement d'outils logiciels et théoriques dédiés à la mise en œuvre de systèmes sonores et interactifs distribués sur un grand nombre d'appareils, en s'appuyant en particulier sur les technologies et les nouveaux standards du Web. Ce projet s'attache notamment à envisager et faciliter la création de nouveaux scénarios musicaux allant de la spatialisation distribuée aux interactions collectives. Outre la réalisation de plusieurs pièces, concerts et installations, ce projet a également permis d'ouvrir de nouvelles perspectives de recherches scientifiques, par exemple sur l'étude du mouvement en situation d'interaction collective (cf. EmoDemos).

Un axe important de ce projet s'attache aujourd'hui à proposer un environnement hybride, pouvant mêler ordinateurs, smartphones et/ou systèmes embarqués, permettant d'envisager de nouvelles possibilités d'interactions expressives et synchronisées. Cette démarche s'accompagne également d'un travail important autour de la mise en place d'outils et d'interfaces permettant d'abstraire la complexité de ces systèmes et ainsi de permettre à des artistes et des chercheurs de plus facilement se les approprier et de les manipuler.

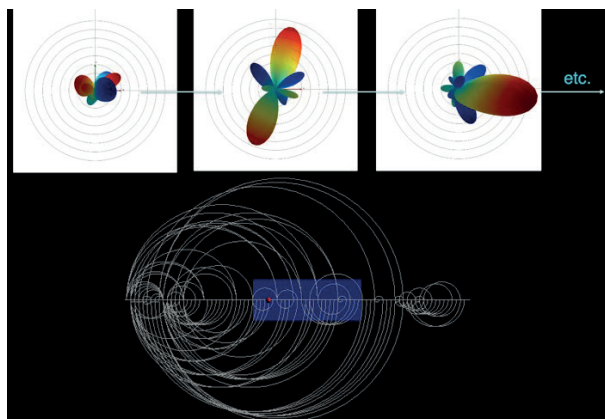
Synthèse de la directivité par corpus

—
Équipe concernée : Espaces acoustiques et cognitifs

La résidence artistique d'Aaron Einbond s'intéresse à la cohabitation entre sons instrumentaux et sons de synthèse dans un espace de diffusion.

Le jeu d'un instrumentiste sur scène est capté et analysé en temps réel : différents descripteurs audio (timbraux) sont calculés et exploités pour produire l'électronique par synthèse concaténative par corpus (réalisée par CataRT-MuBu). Se pose alors la question de la diffusion des échantillons (grains) du corpus. Nous exploitons pour cela un réseau compact de haut-parleurs IKO, qui permet de simuler des diagrammes de rayonnement (décrits par leur représentation en harmoniques sphériques d'ordre 3). Les motifs de rayonnement ici utilisés sont sélectionnés au sein d'une base de données de directivité d'instruments acoustiques (historiques et modernes) mesurée et mise à disposition par TU Berlin : les descripteurs audio (provenant du jeu de l'instrumentiste) servent à sélectionner un (ou plusieurs) instruments dans la base de données TU, afin d'appliquer leur diagramme de directivité aux grains. L'idée sous-jacente n'est pas de reproduire fidèlement la radiation spatiale d'instruments, mais d'insuffler aux sons de synthèse une spatialité « naturelle, plausible », afin que l'électronique fusionne harmonieusement avec des instruments acoustiques présents sur scène.

En outre, nous avons organisé la base de données de diagrammes de rayonnement, en analysant les ressemblances/dissembances des motifs spatiaux à l'aide d'un opérateur de corrélation sphérique. Il est ainsi possible de 'naviguer' de façon cohérente dans l'espace des directivités, et d'interpoler entre les rayonnements de plusieurs instruments. Ainsi, différentes stratégies créatives ont été implémentées pour générer des rayonnements 3D dynamiques, et produire une polyphonie spatiale.



« Improvisation » dans l'espace des directivités

WASABI

Web Audio Semantic Aggregated in the Browser for Indexation

—
Équipes concernées : Analyse des pratiques musicales et Analyse et synthèse des sons, Centre de ressources de l'Ircam
Financement : ANR, programme générique
Période de réalisation : octobre 2016 - mars 2021
Partenaires : I3S (coordinateur), Parisson, Deezer, Radio France

Le projet propose de définir une méthodologie d'optimisation de l'indexation musicale dans un contexte Web et pour de très grands corpus de données par l'utilisation conjointe de métadonnées issues de l'analyse audio, du Web sémantique, de l'analyse en langage naturel des paroles de chansons, puis de confronter cette méthodologie à des cas d'usages en développant des services et des applications originales exploitant les technologies Web audio.

Il s'agit donc d'utiliser conjointement les algorithmes d'extraction d'information musicale et le Web sémantique afin de produire des bases de connaissances musicales plus consistantes

pour les services de streaming et les bases de données musicales. Les services de données du Web sémantique (LastFM, MusicBrainz, DBPedia, etc.) favoriseront l'extraction de données structurées, liant les œuvres à des métadonnées telles que le producteur, le studio d'enregistrement, le compositeur, l'année de diffusion, les thèmes qui y sont abordés, par exemple. Les données en texte libre comme les paroles seront aussi analysées pour déterminer le contexte musical de l'œuvre. Les technologies Web audio permettront enfin d'explorer ces espaces musicaux enrichis par des analyses de type indexation musicale de haut niveau : détection d'émotion, détection de plagiat, détection et caractérisation de voix chantée, détection de structure et séparation de sources.

Il sera proposé une suite de briques logicielles open source et de services en ligne de type « open data » pour :

- la visualisation de métadonnées audio et l'écoute de pistes démixées dans le navigateur en exploitant les dernières technologies issues de la Web Audio API (mixage temps réel, effets audio)
- le traitement automatique de textes de chansons, reconnaissance et liage d'entités nommées, d'annotation et correction collaborative,
- l'accès à un service Web doté d'une API proposant un environnement d'étude de similarités musicales issu des analyses audio et sémantiques.

Ces briques logicielles serviront au développement des démonstrateurs formalisés avec nos partenaires et collaborateurs (journalistes et compositeurs), utilisant le nouveau standard Web Audio API et permettant ainsi le développement d'applications musicales accessibles au grand public depuis un navigateur Web.

4

Les logiciels

ADMix Tools

Conception et développement : équipe Espaces acoustiques et cognitifs

La suite logicielle ADMix Tools, développée dans le cadre du projet européen ORPHEUS (programme H2020), permet l'enregistrement (ADMix Recorder) et la reproduction (ADMix Renderer) de contenus en format objet. Le format objet retenu suit le standard ADM (Audio Definition Model) défini par les instances de normalisation internationales (EBU et ITU). Les contenus ADM prennent la forme de fichiers audio au format .wav (BW64) dont un segment contient les métadonnées structurées et encodées au format <axml>. L'association des pistes audio et des objets référencés dans la section de métadonnées est définie dans un segment de type <chna>.

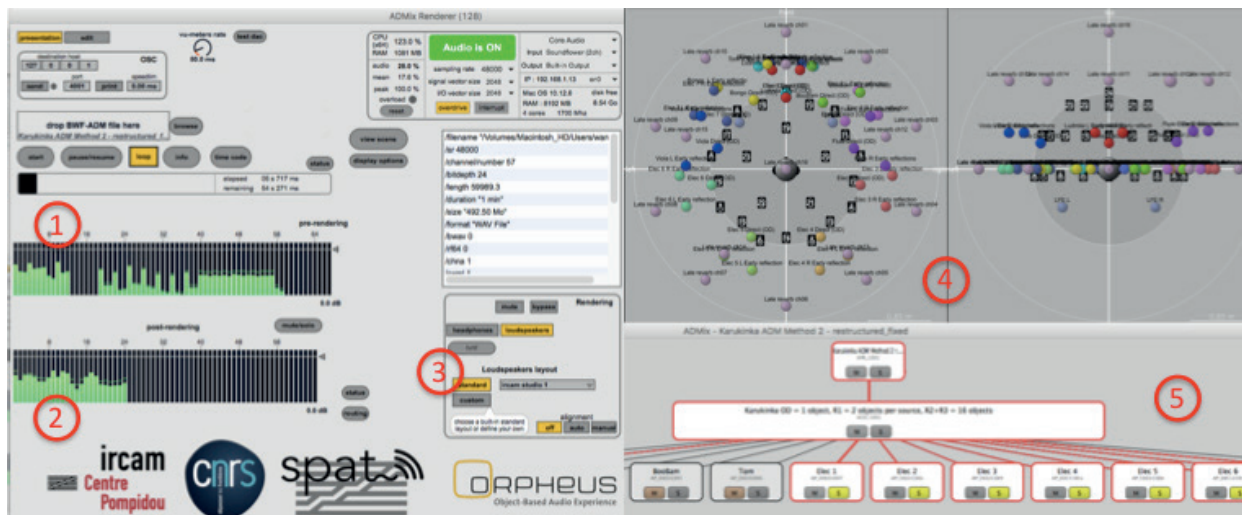
Les éléments constitutifs du fichier peuvent être de différents types, « DirectSpeakers », « HOA », « binaural », « Matrix » ou « Objets ». Les éléments de type « DirectSpeakers » permettent de véhiculer une ambiance ou un premix multicanal et sont définis par leur standard (stéréo, 5.1, etc.). Les éléments de type « HOA » sont définis par leur ordre. Les éléments de type « Objets » sont les plus génériques et sont définis par la position spatiale et le gain de leurs différentes pistes constitutives. Ces positions et gains peuvent varier au cours du temps.

Un fichier ADM peut également contenir plusieurs « programmes » qui exploitent chacun tout ou partie des pistes audio et sont associés à différents jeux de métadonnées. Cette fonctionnalité peut être utilisée pour transmettre un contenu en différentes langues, où seuls les dialogues de la langue sélectionnée seront diffusés tandis que les composantes musicales et sons d'ambiances seront partagés par tous les programmes. Une autre application de ces contenus multiprogrammes peut consister à transmettre une même scène sonore, mixée et restituée selon différents points d'écoute.

En pratique, le module d'enregistrement (ADMix Recorder) s'interface avec toute station de travail audionumérique et permet de traduire et enregistrer les métadonnées ADM sous forme d'automatisation en ayant recours au protocole de communication OSC. Le module de reproduction (ADMix Renderer) est basé sur la bibliothèque du Spat~ et permet ainsi de décoder les contenus sur tout type de système de restitution 2D ou 3D, en particulier le binaural sur casque, VBAP et HOA sur une distribution 2D ou 3D de haut-parleurs.

Configuration

La suite logicielle ADMix Tools, téléchargeable depuis le site du Forum Ircam, est disponible pour les environnements Mac OSX, Windows.



- Fenêtre principale de l'application ADMix Renderer.
- 1) Vu-mètres correspondant aux différents objets constitutifs du fichier audio;
 - 2) Vu-mètres des différents canaux de sortie créés en fonction du mode de rendu;
 - 3) Choix du dispositif de rendu (binaural sur casque ou configurations standard ou ad hoc de haut-parleurs);
 - 4) Visualisation des positions instantanées des différents objets;
 - 5) Graphe de la scène sonore et activation des objets (mute/solo).

Antescofo

<http://antescofo-doc.ircam.fr/>

—

Conception et développement : équipe Représentations musicales et la start-up Antescofo

Antescofo est un système de suivi de partition modulaire et un langage de programmation synchrone temporisé pour la composition musicale. Le module d'écoute effectue une reconnaissance automatique en temps réel de l'interprétation — position dans la partition et tempo — permettant ainsi de synchroniser une performance instrumentale avec celle d'une partition virtuelle informatique qui pilote les traitements audionumériques (synthèses, effets, spatialisation, mixage, etc.) et les dispositifs scéniques. Antescofo permet la description des parties instrumentales et électroniques dans une même partition, grâce à un langage temps réel conçu pour la musique mixte et visant à faciliter l'écriture des processus musicaux temps réel. Après le chargement de la partition, Antescofo est capable de suivre la position et le tempo des musiciens en temps réel et de synchroniser les actions programmées pour la partie d'informatique musicale (déclenchement et contrôle de la partie électronique). Antescofo prend comme entrée un flux d'audio polyphonique. Antescofo est aussi utilisé sans suivi comme séquenceur programmable interactif pour le contrôle en temps réel de processus de synthèse sonore et de spatialisation complexes.

Applications

Œuvres musicales interactives, accompagnement automatique, installation sonore, mapping geste/synthèse sonore, contrôle de la spatialisation, pédagogie musicale.

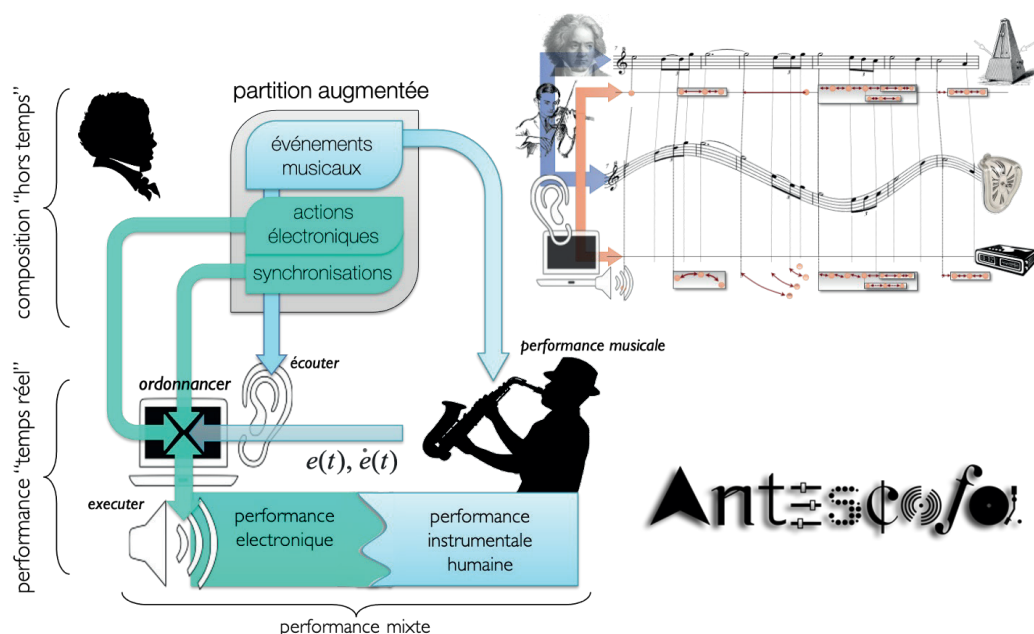
Antescofo a été utilisé dans la réalisation de plus d'une centaine de créations ou de recréations de pièces emblématiques du répertoire contemporain.

Fonctions principales

- Reconnaissance temps réel de la position et tempo des flux audio polyphoniques sur une partition;
- Écriture des processus électroniques en temps relatif, absolu et rationnel;
- Gestion du temps événementiel et chronométrique (durée continue);
- Séquencement interactif des partitions multimédias;
- Traduction de partitions MIDI ou MusicXML;
- Structures de données et librairie de plus de 300 fonctions de base permettant de programmer aisément des traitements algorithmiques;
- Communication OSC, entrée/sortie au format JSON;
- Documentation en-ligne de plus de 400 pages.

Configuration

Antescofo se présente sous la forme d'un objet pour l'environnement Max (Ircam/ Cycling'74) sous systèmes d'exploitation Mac OS X ou Windows XP. Antescofo est également disponible comme objet pour PureData sous Mac OS X, Windows et Linux.



CataRT

Conception et développement : équipe Interaction son musique mouvement

La synthèse concaténative par corpus utilise une base de données de sons enregistrés, et un algorithme de sélection d'unités qui permet de choisir les segments de la base de données, afin de synthétiser par concaténation une séquence musicale. La sélection est fondée sur les caractéristiques de l'enregistrement, qui sont obtenues par analyse du signal et correspondent par exemple à la hauteur, à l'énergie ou au spectre.

La mise en œuvre en temps réel de cette nouvelle approche de la synthèse sonore par le système CataRT permet une exploration interactive et en temps réel d'une base sonore et une composition granulaire ciblée par des caractéristiques sonores précises. Cela permet aux compositeurs et musiciens d'atteindre de nouvelles sonorités, et aux designers sonores de rapidement explorer un corpus sonore constitué par un grand nombre d'enregistrements.

CataRT existe en application standalone, en device pour Ableton Live ou en système modulaire implémenté dans Max avec les extensions MuBu. L'interaction repose sur une interface simple consistant en l'affichage d'une projection 2D de l'espace de descripteurs, et une navigation où les grains sont sélectionnés et joués par proximité géométrique. Cette navigation peut être contrôlée avec la souris, avec des contrôleurs externes (pads multitouch, joysticks, cameras), ou avec l'analyse d'un signal audio.

Applications

Œuvres musicales et multimédias fixes ou interactives, sound design, synthèse de textures environnementales, nouveaux instruments, recherche dans les bases de données sonores.

Fonctions principales

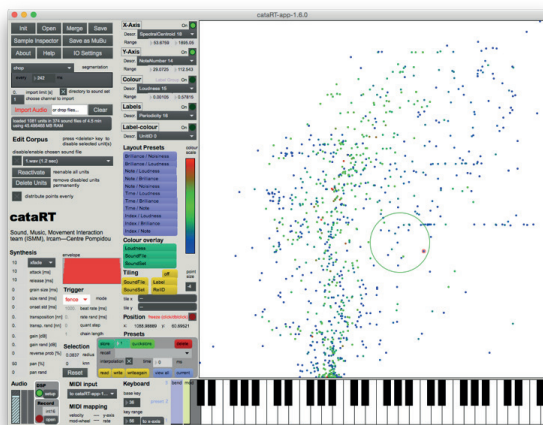
- Importation audio avec segmentation et analyse automatique ;
- Analyse, visualisation, et jeu interactif avec des sons ;
- Exportation des données d'analyse ;
- Composition par navigation interactive dans un espace de descripteurs sonores ;
- Brassages et juxtapositions d'un nombre massif de sons ;
- Traitements granulaires flexibles ;
- Spatialisation granulaire en ambisonique d'ordre supérieur (HOA) ;
- Communication avec séquenceurs MIDI ;
- Recherche interactive efficace d'événements ou singularités dans des nombreux et longs enregistrements, séparation de bruits parasites.

Configuration

- Version application : Mac OS 10 ou supérieur, collection pour Max 7 runtime ou supérieur ;
- Version device Max4Live : Ableton Live 10 ou supérieur ;
- Version modulaire : Mac OS 10.10 ou supérieur ou PC Windows avec Max 7 ou supérieur et package MuBu.

Formats de fichiers audio

- jusqu'à 32-bit/192 kHz entier et flottant : AIFF, AIFC, WAV, FLAC, AAC, MP3 ;
- fichiers SDIF pour l'importation de marqueurs de segmentation d'AudioSculpt (version application) ;
- fichiers MIDI pour l'importation de segmentation et annotation de pitch (version application).
- fichiers Text pour l'importation de segmentation et annotation de labels, par exemple de Audacity (version application).



DYCI2

Conception et développement : équipe Représentations musicales

DYCI2 est une bibliothèque d'agents génératifs pour la performance et l'interaction musicales combinant les approches libres, planifiées et réactives de la génération à partir d'un corpus, ainsi que des modèles de scénarios dynamiques à court terme (« meta-Djing »). DYCI2 consiste en :

- une bibliothèque Python définissant des modèles et des outils pour la génération de séquences créatives à partir de modèles de séquences. Elle implémente plusieurs modèles, heuristiques, stratégies de gestion du temps, et architectures d'agents interactifs ;
- une bibliothèque Max d'agents temps réel interfacée avec la bibliothèque Python

À partir d'un agent embarquant une « mémoire » musicale, DYCI2 propose différentes stratégies génératives.

- Génération libre : Une fois le modèle de mémoire construit, une requête dite « free » entraîne la génération d'une séquence musicale inédite selon la logique temporelle interne du matériau appris, sans autre contrainte de structure ou d'écoute réactive. Cette stratégie est semblable à celle de OMax, légèrement simplifiée
- Guidage par scénario long-terme : La génération peut être guidée par un scénario déterminé par l'utilisateur. Cette séquence symbolique de référence est définie sur le même alphabet que les annotations de la mémoire musicale utilisée (labels d'accords, modes de jeu, etc.) et guide l'improvisation (séquence d'accords, séquence de modes de jeu, etc.). Il s'agit donc de trouver des segments de la mémoire correspondant aux portions successives du scénario à suivre et de les enchaîner de manière créative.

- Guidage par scénario dynamique à court terme : L'implémentation de la réaction en tant que réécriture d'anticipations préalablement générées a permis d'introduire la notion de scénarios dynamiques à court terme. Au cours de la performance, on peut ainsi envoyer des requêtes correspondant à des séquences de labels spécifiant ce que l'agent doit générer et jouer sur-le-champ, tout en maintenant la cohérence avec ce qui vient d'être joué ainsi qu'avec les anticipations préalablement générées. Ce paradigme de guidage introduit un mode de jeu que l'on pourrait qualifier de « meta DJing » ou « DJing d'intentions » : en effet, un opérateur-musicien peut improviser en contrôlant un agent à l'échelle de la narration musicale. Par exemple, on envoie cette instruction : « à partir du temps prochain : générer et jouer une séquence correspondant à la suite d'accords Dm7 G7 CMaj7 » ; ou celle-ci : « maintenant : générer et jouer une séquence partant de grave rugueux pour arriver à aigu brillant ».

Utilise CataRT by MuBu, ISMM



IRCAM Lab TS

Conception et développement : équipe Analyse et synthèse des sons

TS est un logiciel indépendant de transformation et de dilatation temporelle à partir de SuperVP et qui reprend des éléments d'interface d'AudioSculpt 3.

Applications

Composition, design sonore, postproduction, enseignement.

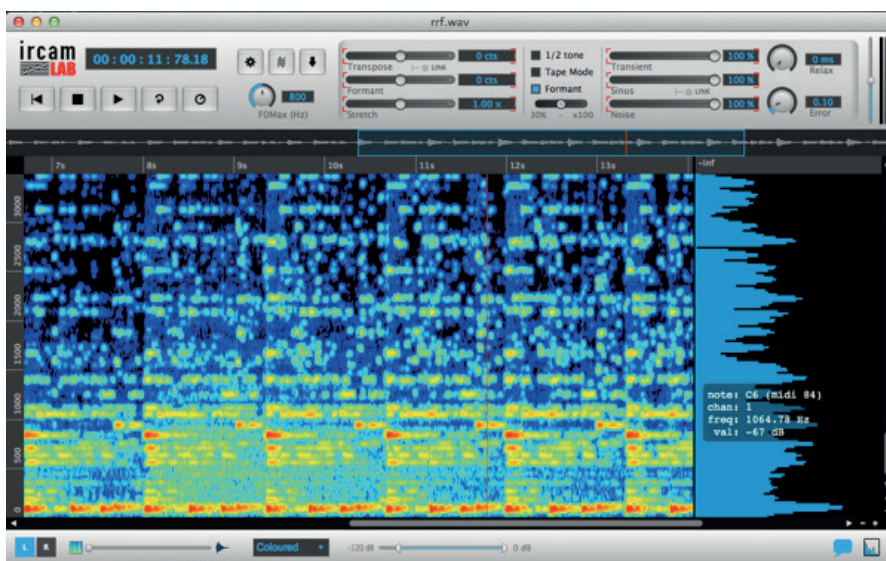
Fonctions principales

- Visualisation :
Affichage du sonogramme, spectre et de la forme d'onde. Zoom interactif sur le signal temporel jusqu'au niveau de l'échantillon, sonogramme linéaire et logarithmique. Analyse FFT, LPC, True Envelope.

- Compression/expansion :
Modification de la longueur du son sans changer sa hauteur ou son timbre et en préservant la qualité des attaques et transitoires. Préréglages définis pour différents modes de transposition avec la possibilité d'ajuster plus finement la F0 max et l'enveloppe spectrale.
- Fonction Remix :
Remixage des transitoires, sinusoides et de la partie bruitée du signal.
- Contrôle :
Lecture audio avec contrôles standard. Mode d'enregistrement des résultats en différé ou en temps réel avec modification des paramètres. Pilotage et automation via contrôleur MIDI possible.

Configuration

- Macintosh sous Mac OSX 10.6 ou supérieur.
- Formats de fichiers audio : AIFF/AIFC, Wav (et autres formats permis par la bibliothèque libsndfile).



Max

Le standard mondial pour l'interactivité sonore temps réel

Max (anciennement Max/MSP) est un environnement visuel pour la programmation d'applications interactives temps réel. C'est actuellement la référence mondiale pour la création d'installations sonores interactives. Max est la combinaison du logiciel Max (Ircam/Cycling'74) pour le contrôle temps réel d'applications musicales et multimédias interactives par MIDI, de MSP, une bibliothèque d'objets pour l'analyse, la synthèse et le traitement du signal audio en temps réel et de Jitter qui est un ensemble d'objets vidéo, matriciels et graphiques 3D pour l'environnement. Max est conçu pour les musiciens, les designers sonores, les enseignants et les chercheurs qui souhaitent développer des programmes interactifs temps réel. Max est développé et édité par la société Cycling'74 (États-Unis) sous licence exclusive de l'Ircam.

Composition

Génération de structures musicales au moyen de modèles mathématiques et aléatoires.

Accompagnement Live

Production d'œuvres mixtes où la partie électronique réagit au jeu de l'instrumentiste (par captation du jeu instrumental par microphone ou par MIDI) ou bien par modification directe du son de l'instrument. Max se comporte comme un séquenceur en contrôlant la diffusion de sons ou de fichiers audio dans le temps. Le contrôle dynamique de programmes complexes pilotant de la musique, de la vidéo ou des effets de lumière est aussi possible.

Postproduction audio et pilotage de périphériques

Contrôle possible de périphériques extérieurs tels que tables de mixage, systèmes « direct-to-disk », synthétiseurs, etc.

Traitement vidéo temps réel

Une bibliothèque d'objets pour le traitement vidéo (Jitter) est incluse avec le logiciel. Elle permet un contrôle de très bas niveau et la programmation de nombreuses applications.

Enseignement

Max est un outil pédagogique largement répandu dans les universités, les écoles de musique et les conservatoires. La documentation intégrée directement dans le logiciel comprend de nombreux tutoriels interactifs pour la compréhension du signal audio, du système MIDI ou encore de la vidéo.

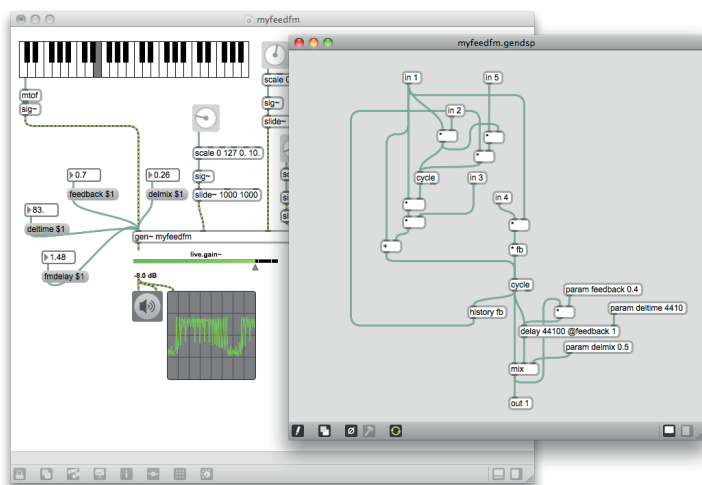
Recherche

Max est utilisé en recherche pour prototyper des algorithmes de synthèse ou de traitement du signal en temps réel. De nouveaux objets externes peuvent être programmés en langage JavaScript, Java et C.

Caractéristiques techniques

Plusieurs centaines d'objets pour la synthèse, le contrôle, le traitement, l'échantillonnage ou l'analyse. Le moteur audio offre le support du multiprocessing, du calcul en 64 bits et apporte de nombreuses améliorations du rendu sonore.

Max intègre un système de génération de code et compilation à la volée à partir de patches, permettant la création rapide de nouveaux traitements performant directement dans l'environnement Max.



Modalys

Conception et développement : équipe S3AM

Modalys est un environnement de lutherie virtuelle permettant de créer des instruments à partir d'objets physiques « simples » tels que cordes, plaques, tubes, membranes, plectres, archets ou marteaux, et de les faire interagir. « Physique » signifie que chaque objet est défini à partir des équations physico-mathématiques qui régissent son comportement, contrairement à une approche de type « sample » par exemple.

Il est également possible de construire des objets aux formes plus complexes par maillage 3D ou par le résultat de mesures réelles ; Modalys se charge alors de leur donner vie et de calculer leurs vibrations sonores en situation de jeu.

En combinant ces différents objets physiques résonnants et en spécifiant le type d'interaction (frappe, frottement, embouchure, anche, etc.), l'utilisateur construit ses propres instruments et choisit le type de jeu et de contrôle. En outre, Max et OpenMusic offrent un dialogue direct avec le moteur de synthèse Modalys via des interfaces graphiques intuitives.

Domaines d'application

- Composition musicale : musique contemporaine, techno, électro, etc. ;
- Design sonore : travail naturel sur la qualité sonore (métallique, boisé, etc.) ;
- Installation sonore en temps réel ou différé ;
- Pédagogie : la théorie modale en pratique ;
- Cinéma & vidéo : habillage sonore ;
- Recherche & développement scientifiques : comparaison des modes déterminés théoriquement et mesurés.

Fonctions principales

Au cœur de Modalys réside un moteur de synthèse permettant de créer et de faire sonner des instruments virtuels à partir d'objets physiques simples ou créés à partir de mesures ou même à partir de maillage 3D. L'utilisateur construit ainsi un instrument et décide ensuite de la manière dont il va en jouer. Pour ce faire, l'utilisateur passe par les étapes suivantes :

- Choix des objets physiques qui vont constituer l'instrument (exemples : tube, corde, membrane, plaque, plectre, archet, marteau, etc.). Ces objets sont simulés avec des paramètres par défaut que l'utilisateur peut modifier ;
- Choix des « accès » aux objets, positions où ces derniers seront mis en interaction ;
- Mise en place des « connexions » correspondant aux modes de jeu tels que frapper, souffler, frotter, etc. ;
- Choix des contrôleurs permettant de modifier en temps réel les paramètres d'un mode de jeu.
- Riche documentation en ligne.

Utilisation

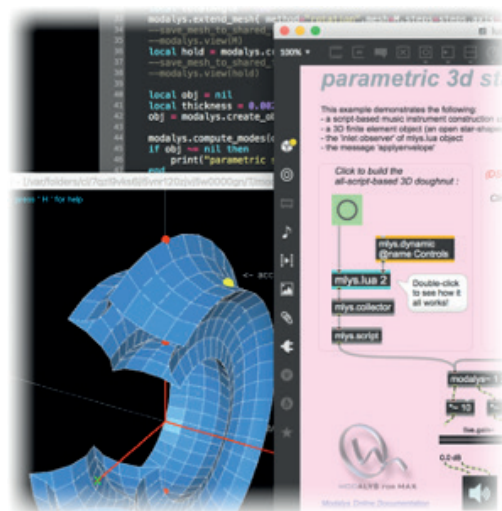
Choisi par de nombreux compositeurs et designers sonores, Modalys for Max est l'incarnation de Modalys dans l'environnement temps réel Max, utilisé pour construire graphiquement des instruments sophistiqués à partir de cordes, plaques, membranes, tubes ou de maillages 3D, avec accès à tous types de matériaux (métal, bois, diamant, etc.), taille (illimitée) et interactions : frappe, pincement, frottement, souffle etc. Cette interface permet le contrôle direct de Modalys en temps réel. Modalys for Max permet aussi de définir un instrument et ses interactions en passant par le langage LUA, avec un accès au contrôle à l'échantillon. Modalys for Max nécessite Max 7 ou supérieur (8 inclus), et est disponible pour Mac OS X (à partir de 10.9) et Windows (7 ou plus récent).

Alternativement, Modalys peut être mis en œuvre à partir de Modalisp, un environnement textuel de programmation, et peut donc être utilisé pour définir les modèles physiques de l'instrument et le faire jouer avec une approche de type script. Dans cette approche, l'utilisateur écrit et exécute des programmes Lisp pour obtenir le résultat musical escompté. Également, le logiciel de programmation musicale de l'Ircam OpenMusic, lui-même basé sur Lisp, permet de contrôler les modèles de Modalys d'une façon à la fois musicale et graphique (contrôle par une partition ou par des éditeurs de courbes).

Enfin, pour une approche scientifique de la synthèse modale, Modalys s'interface aisément avec MatLab grâce à des bibliothèques dédiées.

Configuration

- Macintosh sous Mac OS X 10.9 minimum, natif Intel ou Silicon.
- Windows 7 ou supérieur (Modalys for Max uniquement).



MuBu pour Max

Conception et développement : équipe Interaction son musique mouvement

MuBu (pour «Multi-buffer») est un ensemble de modules Max pour le traitement du signal temps réel multimodal (audio et mouvement), l'apprentissage automatique, et la synthèse sonore interactive.

Le conteneur MuBu permet de stocker, éditer et visualiser divers types de pistes temporelles synchronisées : audio, descripteurs sonores, données de captation gestuelle, marqueurs de segmentation, partitions MIDI. Des représentations musicales symboliques simplifiées des paramètres de contrôle de synthèse et de la spatialisation peuvent être également intégrées. Cela permet de traiter dans Max des structures de données complexes : données audio segmentées et annotées alignées avec des descripteurs audio, données de captation gestuelles annotées, données audio et gestuelles alignées.

MuBu intègre des modules pour l'apprentissage machine interactif, pour la reconnaissance de formes sonores ou gestuelles. Il intègre également un format de plug-in pour le traitement de flux audio ou gestuels, PiPo (Plugin Interface for Processing Objects). Les données dans MuBu sont scriptables grâce à l'intégration du support javascript de Max.

Applications

Œuvres musicales et multimédias interactives, synthèse sonore interactive, musique, danse.

Fonctions principales

- Multi-buffer avec interface graphique ;
- Traitement de flux de données et de données stockées ;
- Extraction de descripteurs sonores ;
- Importation/exportation de fichiers audio, SDF, json, MIDI, MuBu ;
- Sélection d'unité par k-NN ;
- Synthèse concaténative, granulaire et PSOLA ;
- Synthèse additive ;
- Reconnaissance de formes sonores ou gestuelles (GMM, HMM, XMM, GF, DTW) ;
- Apprentissage automatique.

Configuration

MuBu est une librairie d'objets Max pour MacOSX et Windows, distribué via le Forum Ircam et le Package Manager de Max.



OMax & co

Conception et développement : équipe Représentations musicales

OMax est un environnement pour l'improvisation avec ordinateur qui analyse, modélise et réimprovise en temps réel le jeu d'un ou de plusieurs instrumentistes, en audio ou en MIDI. OMax est basé sur une représentation informatique nommée «Oracle des facteurs», un graphe qui interconnecte tous les motifs et leurs variations et fournit donc une carte de navigation dans la structure de la musique, engendrant ainsi un grand nombre de variantes cohérentes stylistiquement. OMax base sa reconnaissance soit sur des notes (suivi de hauteurs), soit sur des timbres (suivi spectral).

Applications

Improvisation avec ordinateur, œuvres musicales interactives, accompagnement automatique, co-créativité homme-machine.

Fonctions principales

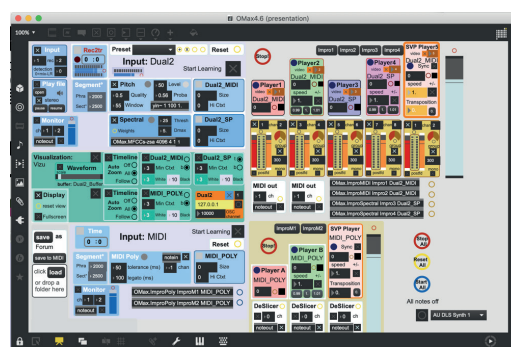
- Architecture modulaire permettant d'allouer des agents improvisateurs, des agents joueurs (players) et de les interconnecter librement (e.g. plusieurs joueurs connectés sur un même improvisateur peuvent créer par exemple un canon avec augmentation rythmique) ;
- Visualisation temps réel interactive du modèle de connaissance, fonctions de sélections multiples et de focalisation opérant sur la représentation visuelle ;
- Écoute artificielle adaptative à partir du signal audio, Modélisation incrémentale du modèle de connaissances représentant la carte des motifs musicaux, Heuristiques de Génération selon des critères de continuité formelle, motivique ou énergétique.

OMax a engendré plusieurs logiciels compagnons : SoMax, ImproteK

- SoMax est un agent autonome capable de générer du contenu musical cohérent en se basant sur une connaissance extraite par apprentissage d'un corpus ;
- SoMax met l'accent sur l'écoute instantanée et le guidage implicite opéré par le musicien humain pour naviguer de manière à la fois réactive et auto-cohérente, de sorte à produire accompagnement, contrepoint ou co-improvisation.
- DYCI2 implémente un concept réactif de micro-scenarios répondants aux interactions avec les musiciens
- ImproteK est un système d'improvisation musicale homme-machine guidé par un scénario temporel formalisé (grille harmonique, composition. Objet d'un partenariat avec l'EHESS, Il est particulièrement efficace dans l'improvisation pulsée et idiomatique (jazz, blues... par exemple) ou la composition.

Configuration

Versions récentes de Mac OSX et de Max.



OpenMusic

Conception et développement : équipe Représentations musicales

OpenMusic (OM) est un environnement de programmation visuelle dédié à la composition et à l'analyse musicale assistées par ordinateur. OpenMusic offre à l'utilisateur la possibilité de programmer graphiquement des procédures de traitement ou génération de données musicales à l'aide de nombreux modules graphiques et fonctions prédéfinies, assemblées en programmes visuels. Les programmes peuvent ensuite s'intégrer les uns dans les autres pour constituer des processus complexes produisant des structures de plus en plus élaborées. OpenMusic est aujourd'hui utilisé par un grand nombre de compositeurs et de musicologues. Il est enseigné dans les principaux centres d'informatique musicale ainsi que dans plusieurs universités en Europe et dans le monde.

Fonctions principales

- Construction graphique de programmes fonctionnels (méthodes d'abstraction/ application, fonctions d'ordre supérieures, structures conditionnelles, itérations...);
- Fonctions de haut niveau pour la création et la transformation des structures musicales (hauteurs, rythmes, etc.);
- Manipulation algorithmique des données aux formats MIDI, audio, SDIF, Communication OSC;
- Éditeurs graphiques : partitions en notation musicale traditionnelle et proportionnelle (avec import et export vers les principaux formats d'échanges), courbes 2D/3D et « breakpoint functions ».
- Outils mathématiques pour l'analyse et la composition ;
- « Maquettes » : éditeur graphique permettant de créer des formes temporelles et hiérarchiques aux programmes visuels.

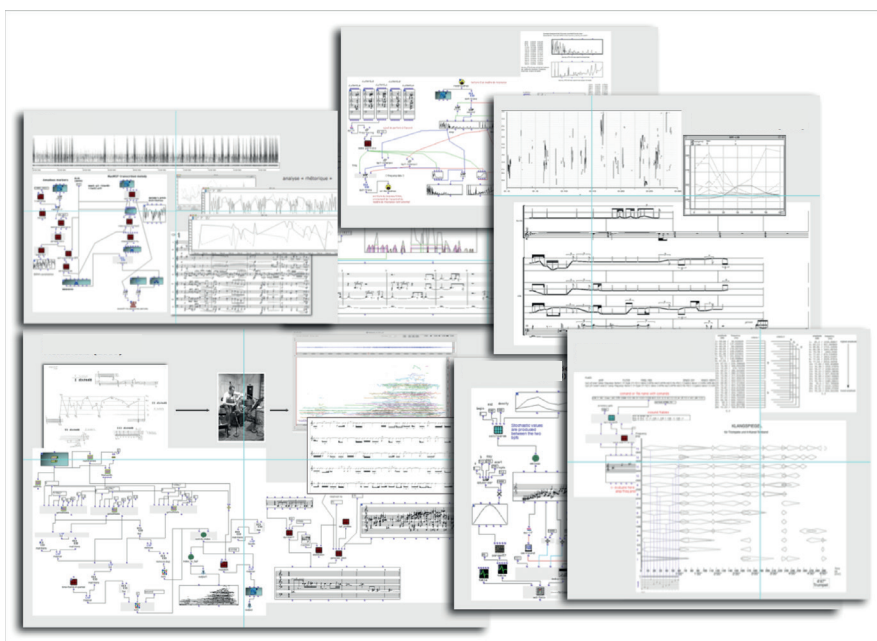
Bibliothèques spécialisées

Musique spectrale, contraintes musicales, algorithmique musicale, structures rythmiques, générateurs probabilistes stochastiques ou chaotiques, synthèse sonore, synthèse spatiale (Spat~), contrôle et traitements sonores (super-vp, additive, chant, csound), etc.

Caractéristiques techniques et configuration

Programmation basée sur le langage Common Lisp / LispWorks (www.lispworks.com/)

Principalement MacOS, versions Windows et Linux possibles.



Le logiciel de composition OpenMusic (Assayag, Agon, Bresson, Haddad)

Orchid*

Conception et développement : équipe Représentations musicales

Le logiciel d'aide à l'orchestration de l'Ircam, développé par l'équipe RepMus, propose des mixtures instrumentales comme solution d'un problème d'orchestration consistant à imiter le mieux possible perceptivement une cible acoustique.

Orchid* s'est décliné en trois instances d'orchidées développées au sein de l'équipe : Orchidee, Orchids et Orchidea.

- Orchid / Orchidée, la première réalisation a fixé les lois du genre en termes de fonctionnalités d'orchestration imitative, et de méthodes avec la prédiction de descripteurs audio des mixtures, l'optimisation multi ou mono objectif par algorithme génétique pour la recherche des solutions orchestrales, l'orchestration statique (cible unique) ou séquentielle (par segmentation) de la cible. Orchidée se déclinait en standalone ou toolbox d'interface Max dialoguant avec le moteur d'orchestration.
- Orchids a étendu le champ en passant de l'optimisation multi-objectif dans un espace de descripteurs à l'optimisation de séries temporelles et ouvert la voie à l'orchestration dynamique continue de morphologies spectrales complexes et évolutives et a proposé une riche interface standalone.
- Orchidea, la version la plus récente, se base comme les autres sur la recherche génétique pour l'optimisation mono ou multi-objectif et la prédiction de descripteurs, et propose aussi, soit une version standalone, soit une toolbox et un GUI en Max. Orchidea affine et optimise l'analyse et la prédiction de descripteurs, la recherche de solutions et la qualité des enchaînements pour l'orchestration séquentielle, et fluidifie l'interface utilisateur. Orchidea est une collaboration Ircam / HEM, qui se continuera par une collaboration Ircam / Berkeley, ainsi qu'avec des développeurs indépendants tel Daniele Ghisi (Bach). De ce fait Orchidea sera open source.

Orchid* est un système complet pour l'orchestration assistée par ordinateur et l'optimisation de mélanges de timbres. Il fournit un ensemble d'algorithmes permettant de reconstruire n'importe quelle cible sonore évoluant dans le temps par une combinaison d'instruments ou échantillons,

selon un ensemble de critères acoustiques et perceptifs. Il peut aider les compositeurs à obtenir des couleurs de timbre inouïes en fournissant une multitude de solutions efficaces qui recréent au mieux cette cible sonore. Grâce à un ensemble étendu de fonctionnalités, Orchid* peut également reproduire des évolutions spectrales en analysant la dynamique temporelle. Ses résultats fournissent des partitions d'orchestre multiples pouvant être organisées de manière intuitive afin d'obtenir rapidement une réalisation d'idées orchestrales et musicales. Orchid* peut utiliser plusieurs bases de données de sons instrumentaux mais peut également être étendu de manière illimitée, y compris avec des sons de synthèse, en important simplement sa propre banque de sons.

Application

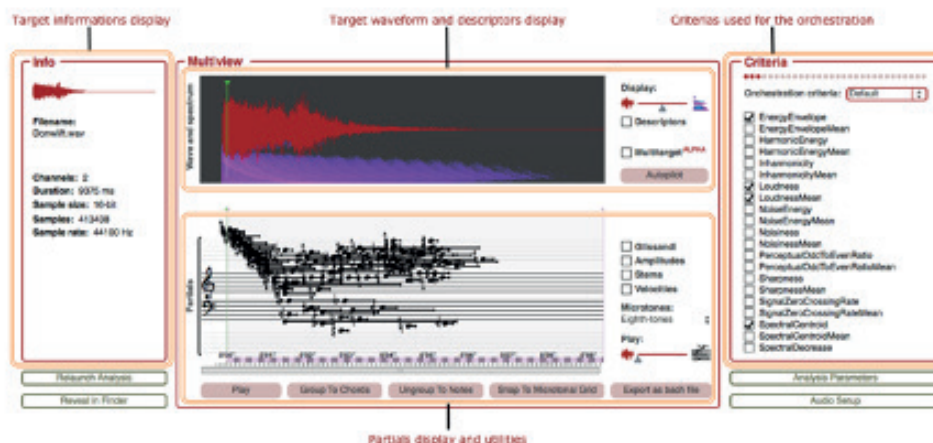
Composition orchestrale, réorchestration, écriture du timbre, pédagogie musicale.

Fonctions principales

- Génération de mélanges et combinaisons orchestrales inouïes par optimisation d'une cible ;
- Possibilité d'écriture du timbre par mélanges orchestraux et optimisation de l'évolution temporelle ;
- Définition de cibles abstraites et d'évolutions temporelles de descripteurs spectraux ;
- Interface simple permettant la spatialisation de l'orchestre ainsi que le travail efficace d'une maquette d'œuvre orchestrale ;
- Extension facile de la base de connaissances et recherches temporelles efficaces ;
- Interaction et exports vers de multiples systèmes de notation musicale ;

Configuration

Orchids est un logiciel standalone marchant sous toutes les versions du système d'exploitation Mac OSX. Orchids est également disponible sous forme de patches Max et de serveur C++.



OSCar

Conception et développement : équipe Espaces acoustiques et cognitifs

OSCar est un plugin permettant de communiquer, en lecture comme en écriture, l'automatisation des paramètres d'une station de travail audio-numérique vers d'autres applications, via le protocole OSC. Son application typique concerne la réalisation de mixages spatialisés orientés objet, indépendamment des contraintes des logiciels hôtes.

Fonctions principales

OSCar s'insère en tant que plugin audio dans le DAW (Digital Audio Workstation, environnement de montage/mixage), sur chacune des pistes à spatialiser. Le plugin OSCar n'affecte aucunement les signaux audio : sa fonction est en revanche d'exposer un certain nombre de paramètres automatisables. Dans un souci de compatibilité avec une vaste gamme de stations audio-numériques, il est préférable de restreindre le nombre de paramètres exposés par piste à 32. Lorsque ces pistes d'automatisation sont activées en lecture, OSCar envoie les données de ces pistes via OSC. De même, si les automatisations sont armées en écriture, OSCar peut recevoir des paquets OSC depuis une application distante, et inscrire ces données dans les pistes du DAW.

Chaque instance d'OSCar possède un identifiant (ID) qui est réglable par l'utilisateur. Cet identifiant peut être utilisé pour formater les messages OSC émis/reçus par OSCar. Leur syntaxe prend par exemple la forme suivante : «/track/ID/NomParamètre ValeurParamètre» par exemple : «/track/3/azimuth 135.0»

Dans un souci de simplicité, les paramètres dans OSCar sont tous des nombres flottants en double précision. Les autres types de données possiblement encapsulables dans un paquet OSC (tels que nombres entiers, chaînes de caractères, etc.) ne sont actuellement pas gérés.

OSCar n'est pas lié à un moteur de spatialisation ou de synthèse en particulier ; autrement dit, les paramètres d'automatisation qu'il expose sont génériques. Par défaut, ces 32 paramètres se nomment param1, param2, etc. Le libellé des paramètres peut ensuite être configuré. Ceci se fait par le biais d'un fichier de configuration (fichier dit de mapping) au format XML qui doit être chargé dans OSCar. (cf. figure).

Le fichier XML de configuration permet de spécifier, pour chacun des paramètres, une mise à l'échelle des valeurs d'automatisation : dans l'environnement du DAW, les pistes d'automatisation sont codées en flottant sur [0 – 1] ; OSCar réalise une mise à l'échelle de [0 – 1] Cette mise à l'échelle affecte les données sortantes et l'opération inverse est appliquée aux données entrantes.

Configuration

OSCar a été développé avec le framework Juce. Juce offre un wrapper de plugin qui permet, à partir d'une base de code commune, de générer des plugins dans différents formats (VST, AudioUnits, etc.). Ainsi, OSCar est disponible sous MacOS et Windows, en format VST3, AU, AAX. Il est distribué via le Forum Ircam.



```
<oscar version="1.0">
  <parameter index="1" name="x" pattern="/source/1/x" min="-20" max="20" default="0"/>
  <parameter index="2" name="y" pattern="/source/1/y" min="-20" max="20" default="0"/>
  <parameter index="3" name="z" pattern="/source/1/z" min="-20" max="20" default="0"/>
  <parameter index="4" name="pres" pattern="/source/1/pres" min="0" max="120" default="40"/>
  <parameter index="5" name="bril" pattern="/source/1/bril" min="0" max="50" default="25"/>
  <parameter index="6" name="warm" pattern="/source/1/warm" min="0" max="50" default="25"/>
</oscar>
```

Vue du plugin OSCar dans ProTools et fichier d'exemple de configuration des paramètres d'automatisation

- 1 Plugin en insert sur la piste 1
2. Liste des paramètres exposés
3. Fenêtre du plugin
4. Pistes d'automatisation.

Panoramix

Conception et développement :
équipe Espaces acoustiques et cognitifs

Panoramix est une station de travail pour le mixage et la postproduction de contenus audio 3D. L'outil propose un environnement cohérent permettant de mixer, réverbérer et spatialiser des matériaux sonores provenant de divers dispositifs micro-phoniques : arbre principal, appoints, captation Ambisonic aux ordres supérieurs. Plusieurs techniques de spatialisation 3D (VBAP, HOA, binaural) peuvent être combinées, et les mixages peuvent être rendus simultanément dans différents formats. Le logiciel présente également les caractéristiques des mixeurs traditionnels (modules d'égalisation et compression dynamique, gestion groupée de paramètres, routage des signaux d'entrée/sortie, etc.), et il est entièrement contrôlable via le protocole Open Sound Control.

Architecture

À l'instar d'une console traditionnelle, Panoramix se présente sous forme de tranches. Les tranches sont classées en deux catégories principales : tranches d'entrée et bus.

Les tranches d'entrée correspondent aux signaux sources issus de la prise de son. Plusieurs formats de tranches d'entrée sont supportés : « Mono » (typiquement un appoint microphonique ou une piste d'électronique devant être spatialisée), « Multi » (une tranche « Multi » est essentiellement une collection de plusieurs tranches mono), « SMA » (employée pour traiter un enregistrement provenant d'un microphone sphérique de type Eigen-Mike®) ou Tree (sert à traiter un arbre microphonique principal).

Trois types de bus sont présents : des bus de panning, des bus de réverbération et un bus « LFE » (« low frequency enhancement »). Les bus de panning jouent le rôle de point de sommation des signaux provenant des tranches d'entrée, régissent la technique de spatialisation (VBAP, HOA, binaural) et réalisent l'encodage/décodage des signaux selon le dispositif de restitution. Les bus de réverbération génèrent et contrôlent les sections tardives/diffuses du réverbérateur artificiel. Dans la version actuelle, un réseau de retards rebouclés est employé.

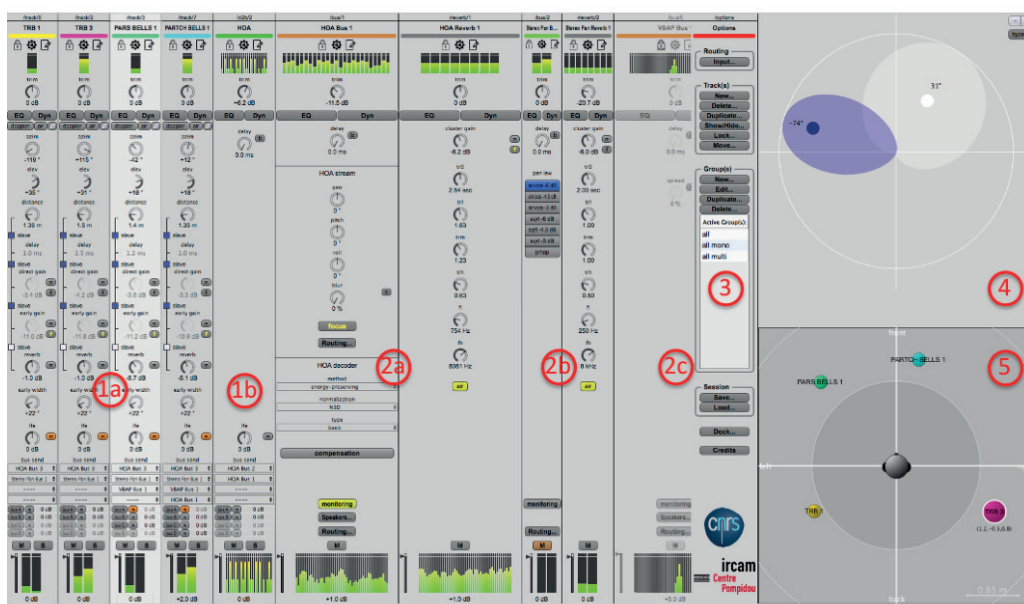
Enfin, la tranche « Master » permet de collecter les signaux de sortie de tous les bus et de les adresser aux sorties physiques de Panoramix. Bien que la session ne comporte qu'une unique tranche Master, il est possible de produire simultanément des mixages dans plusieurs formats.

Configuration

Panoramix est téléchargeable depuis le site du Forum Ircam et disponible pour les environnements MacOS, Windows et Linux.

Fenêtre principale de la station Panoramix.

- 1a) Tranches d'entrée mono ;
- 1b) Tranche d'entrée HOA ;
- 2a) Bus de panning et de réverbération HOA ;
- 2b) Bus stéréo ;
- 2c) Bus VBAP ;
- 3) Options de session (gestion des groupes, import/export de presets, etc.) ;
- 4) Module de formation de voies (focus) en HOA ;
- 5) Interface géométrique de positionnement.



SkataRT

Outil d'esquisses pour la composition musicale et le design sonore

Conception et développement: équipes Interaction Son-Musique-Mouvement, Perception et Design Sonores (collab. MusicUnit)

L'esquisse est une étape fondamentale des activités de composition et de design. L'environnement SkataRT explore la capacité de la voix et du geste à transposer le paradigme «papier / crayon» largement utilisé en création visuelle.

L'environnement SkataRT est construit à l'intersection d'un travail de recherche et développement sur la synthèse sonore concaténative par corpus (CataRT) et d'un projet de recherche européen sur la question de l'imitation vocale comme outil d'esquisse et de prototypage rapide (SkatVG). Il offre donc différents champs d'application et d'usage qui vont de la création musicale au design sonore, en passant par l'exploration de corpus sonores, ou la performance. D'un point de vue technologique, il s'incarne dans un device «Max for Live» développé par Manuel Poletti et Thomas Goepfer (Music Unit), dans le cadre d'une collaboration avec les équipes ISMM et PDS de l'Ircam.

Le cœur du moteur de rendu réside dans un dispositif capable d'interpréter les intentions des utilisateurs exprimées par des imitations vocales et gestuelles, de sélectionner des modules de synthèses sonores appropriés, et de rendre ainsi possible le partage et le raffinement des idées. Il s'articule autour de 2 principes :

- une exploration d'un espace 2D sur lequel des grains de sons sont projetés (CataRt)
- la capture d'une imitation vocale, son analyse et sa modélisation sous la forme d'une esquisse sonore paramétrisable exploitant une synthèse par corpus de type CataRt.

Pour le design sonore, l'outil SkataRT vise à permettre aux concepteurs d'utiliser directement leur voix et leurs mains pour esquisser les caractéristiques sonores d'un objet, facilitant ainsi l'exploitation des possibilités fonctionnelles et esthétiques du son.



SOMAX2

Conception et développement: équipe Représentations musicales

Somax 2 est une application pour l'improvisation et la composition musicale. Elle est implémentée dans Max et est basée sur un modèle génératif utilisant un processus similaire à la synthèse concaténative pour fournir une improvisation stylistiquement cohérente, tout en écoutant et en s'adaptant à un musicien (ou tout autre type de source audio ou MIDI) en temps réel. Le modèle fonctionne dans le domaine symbolique ou audio et est entraîné sur un corpus musical, constitué d'un ou plusieurs fichiers, dans lequel il puise son matériel utilisé pour l'improvisation. Le modèle peut être utilisé avec peu de configuration pour interagir de manière autonome avec un musicien, mais il permet également le contrôle manuel de son processus de génération, servant alors d'instrument qui peut être joué à part entière. Somax2 est également conçu comme une bibliothèque, permettant à l'utilisateur de créer des modèles personnalisés ainsi que de mettre en place des réseaux de modèles et de sources multiples (agents) qui s'écoutent et interagissent les uns avec les autres.

Somax 2 est une version totalement nouvelle du paradigme de co-improvisation réactive Somax, conçu par l'équipe de représentation musicale de l'Ircam, mais qui n'avait pas encore été rendu public. Écrit en Max et Python, il présente une implémentation modulaire concurrente, de multiples agents interconnectables, une nouvelle interface utilisateur avec tutoriels interactifs et documentation, ainsi qu'un certain nombre de nouveaux paramètres puissants pour l'interaction.

Somax 2 est fourni avec un tutoriel interactif donnant une brève introduction aux différents modules de Somax et guidant les premiers pas vers l'interaction avec le modèle.

Il existe également des fichiers d'aide individuels pour chaque objet Max, qui expliquent comment utiliser l'objet, ses contrôles paramétriques et un certain nombre de cas d'utilisation.

Spat~

Conception et développement : équipe Espaces acoustiques et cognitifs

Spat~, le Spatialisateur de l'Ircam, est un logiciel dédié à la spatialisation du son en temps réel. Conçu sous forme d'une bibliothèque, il permet au musicien ou à l'ingénieur du son de contrôler la spatialisation du son pour divers systèmes de restitution. Les applications s'étendent de la reproduction en contexte domestique (casque d'écoute, système stéréophonique ou 5.1) aux situations de concert (distributions 2D ou 3D de haut-parleurs) ou aux installations sonores holophoniques et interactives.

Fonctions principales

- Concerts et spatialisation du son en temps réel
Le compositeur peut associer à chaque événement sonore de la partition un effet de salle ou une donnée de localisation dans l'espace. Le Spat~ peut être contrôlé à partir d'un séquenceur d'un système de suivi de partition, ou à partir de processus de contrôle de haut niveau dédiés à l'écriture musicale (OpenMusic).
- Postproduction sonore
On peut connecter le Spatialisateur à chaque canal d'une table de mixage ou d'un environnement de mixage logiciel pour disposer d'un contrôle intuitif et global des positions de chaque source et de leur effet de salle associé.
- Simulation et réalité virtuelle
Sur casque, l'effet d'immersion obtenu avec le Spat~ repose sur l'utilisation de techniques binaurales, d'autant plus convaincantes si le système est associé à un dispositif de suivi de la position ou de l'orientation de la tête. Les sources sonores gardent ainsi leurs positions spatiales en fonction des commandes de navigation de l'utilisateur.
- Restitution Ambisonique
De nombreux modules sont dédiés à la génération, au traitement, et au décodage de flux HOA (Higher Order Ambisonics): encodeurs pour les microphones sphériques EM32 de MH acoustics® et ZM1 de Zylia®, formation de voies synthétisant des microphones virtuels directifs dans le domaine HOA, déformations non linéaires du champ sonore (compression ou dilatation vers les pôles ou l'équateur), etc.

Restitution holophonique

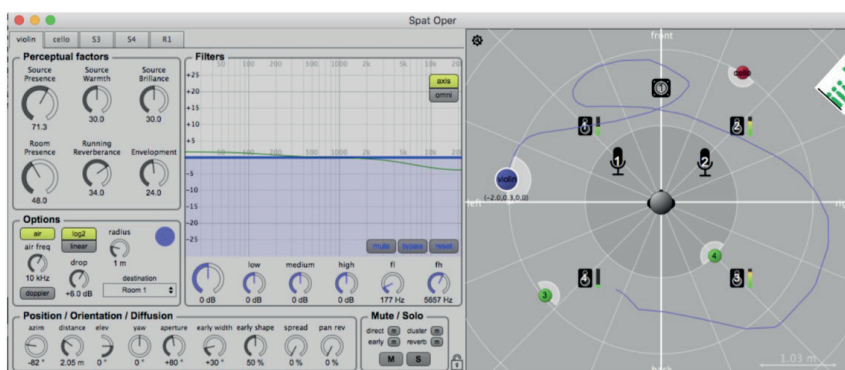
Le Spatialisateur a été adapté pour le mode de restitution holophonique (WFS) qui permet de reconstruire un champ sonore préservant les qualités de restitution spatiale pour un auditoire étendu. Pour les installations sonores interactives ce mode de restitution permet de retrouver les sensations spatiales naturelles au cours de l'exploration de la scène sonore.

Caractéristiques techniques et configuration

- Contrôles perceptifs
Azimut et élévation, distance, présence de la source, présence de la salle, réverbérance précoce et tardive, enveloppement, balances fréquentielles.
- Contrôles DSP de bas niveau
Égalisation, effet Doppler, absorption de l'air, réverbération multicanale, direction de la source en fonction du dispositif de diffusion (binaural, transaural, stéréo, stéréo 3/2, multicanal 4 à 8 HP, VBAP pour un agencement tridimensionnel des haut-parleurs, Ambisonics et Higher Order Ambisonics (HOA), Wave Field Synthesis (WFS).

Configuration

Le Spat~ est une bibliothèque d'objets pour l'environnement Max (Ircam/Opcode/Cycling'74), compatible avec les systèmes d'exploitation MacOS ou Windows. Cette bibliothèque constitue également le cœur de Spat Revolution, une station autonome de mixage multicanal et de spatialisation des sons, développée en collaboration avec la société Flux::. Le Spat~ est également disponible sous forme d'un ensemble de plugins développés en collaboration avec la société Flux::.

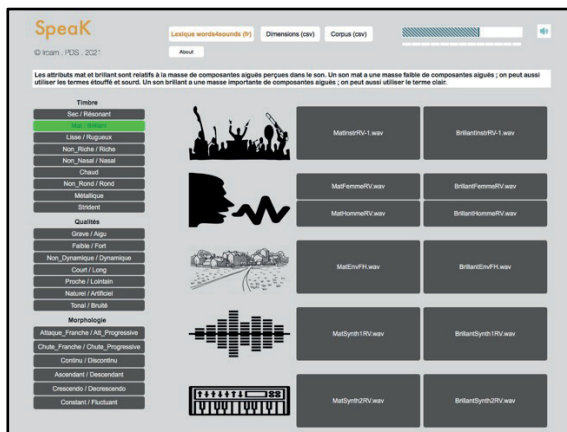
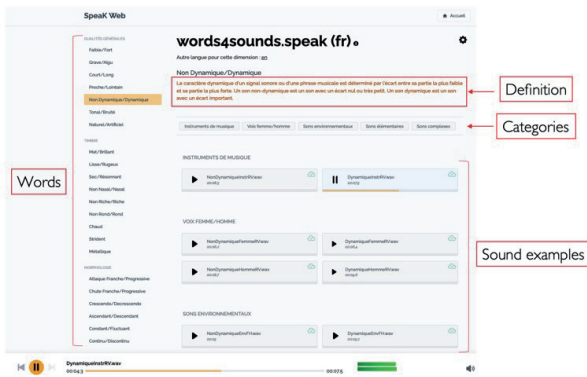


Une vue de l'interface utilisateur du Spat~. À droite : contrôle de la position des sources, de leur trajet (ici pour la source bleue), de leur directivité/ orientation et représentation du système de captation et de reproduction. À gauche : contrôle des paramètres perceptifs de la salle associée à la source (ici la source bleue).

Speak

Lexique sonore pour la description des qualités sonores
 Conception et développement : équipe Perception et Design Sonores
 (collab. IMR /WAM)

L’outil Speak est une interface qui présente un lexique sonore. Les descripteurs sémantiques du lexique s’organisent selon trois grandes classes : timbre (sec/résonnant, mat/brillant, lisse/rugueux, chaud, rond, riche, nasal...), qualités générales (grave/aigu, tonal/bruité, proche/lointain...), et morphologie (attaque, chute, continu/discontinu, constant/fluctuant...). L’outil Speak, dans sa forme actuelle, offre deux principales possibilités. La première est de présenter une définition et un ensemble d’exemples sonores pour chacun des descripteurs du lexique proposé par Carron et al. (Journal of Design Research, 2017). Les exemples sonores ont été produits par l’Ircam (collab. R. Rivas et Ircam PROD) et couvrent plusieurs catégories de sources : instruments de musique, voix femme/homme, sons de synthèse élémentaires et complexes. La deuxième possibilité est de configurer très facilement l’interface pour intégrer de nouveaux descripteurs sémantiques et des sons associés. Ainsi le lexique peut être enrichi ou encore adapté à un projet spécifique. L’outil Speak permet d’avoir un langage commun pour répondre à différents objectifs. Dans le cadre d’un projet en design sonore, il peut être utilisé pour accompagner des séances d’échanges entre différents partenaires. Il a été utilisé à l’Ircam dans des projets de design avec SNCF, Krug et, au sein de plusieurs projets d’Ircam-Amplify – notamment Sorbonne Université (SU). Dans un cadre pédagogique, il pourrait permettre d’apprendre à écouter et à nommer des qualités sonores. Les récents développements de l’outil Speak sont : l’amélioration des définitions de certains descripteurs comme chaud et rond (thèse de Victor Rosi, 2022), le développement d’une version en ligne (Speak web) et le développement de lexiques sonores spécifiques – dont celui dédié à la description des espaces acoustiques.



Suivi de geste et de formes temporelles

Conception et développement : équipe Interaction son musique mouvement

Plusieurs objets Max sont disponibles pour effectuer le suivi de morphologies temporelles à partir de modèles de Markov.

L’objet voicefollower^ indique en temps réel la progression temporelle d’une interprétation, en comparaison avec un enregistrement audio. Cela permet par exemple de synchroniser divers processus sonores ou visuels avec un texte préenregistré. Il a été expérimenté avec succès pour le théâtre avec de la voix parlée ou pour des performances avec de la voix chantée.

L’objet motionfollower^ est similaire au voice follower^ (utilisant une interface proche) mais fonctionne avec des paramètres gestuels, comme ceux issus de capteurs. Il permet par exemple d’effectuer un suivi, dans le cas d’instruments augmentés ou de battue, effectué avec un capteur de mouvement.

L’objet gf (pour gesture follower), intégré dans les modules MuBu, permet des utilisations plus générales de suivi et reconnaissance de formes temporelles sur la base d’exemples enregistrés.

Applications

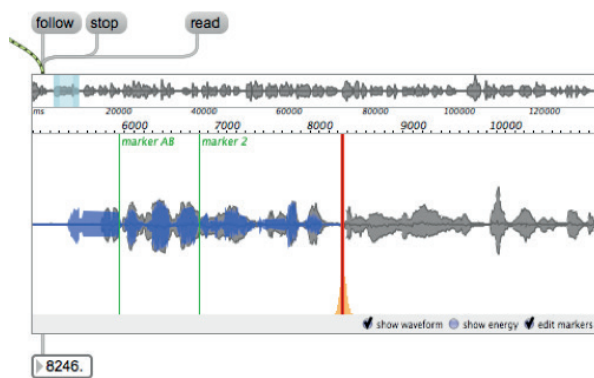
Œuvres musicales et multimédias interactives, musique, danse, contrôle gestuel de la synthèse sonore.

Fonctions principales

- Synchronisation entre gestes et médias numériques ;
- Reconnaissance de gestes en continu ;
- Mesures de similarité entre morphologies temporelles continues.

Configuration

Logiciel Max, MacOSx, Windows



Voice follower~ dans le logiciel Max

SuperVP, TRaX et bibliothèques

Conception et développement : équipe Analyse et synthèse des sons

SuperVP est une bibliothèque de traitement de signal reposant sur un vocodeur de phase perfectionné. Elle permet un grand nombre de transformations du signal avec une très grande qualité sonore (étirement temporel, transposition de la fréquence fondamentale et de l'enveloppe spectrale, débruitage, re-mixage des composantes sinusoïdales, bruitées et transitoires, dilatation de l'enveloppe spectrale, synthèse croisée généralisée, synthèse croisée en mode source et filtre...). Elle donne accès à un vaste jeu de paramètres qui fournissent un contrôle complet, et à grain fin, du résultat d'algorithmes différents. En plus des algorithmes de transformation sonores, la bibliothèque comprend une collection importante d'algorithmes d'analyse du signal (F0, détection des débuts de notes, spectrogramme, spectrogramme réassigné, enveloppe spectrale...).

Un des dérivés de SuperVP est la bibliothèque Voice-Forger qui regroupe une grande part des traitements temps réel de SuperVP et permet un accès de haut niveau (cf. projets : Sample Orchestrator, Vivos, Affective Avatars). Les outils TRaX ont été conçus pour que la plupart des paramètres de la bibliothèque soient condensés en un jeu de paramètres utilisateur de haut niveau qui permettent de contrôler l'application de manière intuitive.

Fonctions principales

Transformation du signal

- Dilatation/compression : avec et sans préservation de transitoires, mode spécialisé (préservation de la forme d'onde) pour la transformation;
- Transposition : transposition indépendante de la fréquence fondamentale et de l'enveloppe spectrale, modification du taux d'échantillonnage (ré-échantillonnage);
- Débruitage : plusieurs motifs de bruit avec interpolation et traitement indépendant du bruit sinusoïdal;
- Re-mixage transitoires/sinusoïdes/bruit : détection des sinusoïdes, bruits, transitoires et contrôle de leurs mixages respectifs;
- Transformation de l'enveloppe : estimation de l'enveloppe spectrale et temporelle et contrôle de leurs mixages respectifs;
- Synthèse croisée : fondée sur le vocodeur de phase ou sur le modèle source-filtre et contrôles variés.

Analyse du signal

- Spectrogramme : FFT à court terme et réassignée;
- Enveloppe spectrale : LPC, « True Envelope »;
- F0 : analyse haute précision à faible coût CPU, paramètres par défaut spécialisés pour différents instruments (exemples : violon, flûte, xylophone, vibraphone...);
- Attaques : algorithme paramétrable pour la détection des attaques de notes et autres transitoires;
- Voisé/non-voisé : fréquence de coupure voisé/non-voisé ou sinusoïdal/non-sinusoïdal.

Caractéristiques techniques et configuration

La bibliothèque peut être contrôlée en utilisant différents niveaux d'abstraction. Elle permet l'entrée/sortie du son par fichiers et par mémoire aussi bien que le contrôle des paramètres en temps réel. La bibliothèque est utilisable à partir d'une large gamme d'environnements. Une application en ligne de commande donne accès à toutes les transformations et analyses sonores disponibles. Elle est utilisée par exemple dans les applications AudioSculpt et OpenMusic. Dans Max, des objets permettent la transformation sonore et l'analyse en temps réel et l'application TRaX repose sur ces objets. La bibliothèque est programmée en C++ en utilisant des dispositifs d'optimisation avancés comme les opérations SIMD (SSE2, SSE, AltiVec) des processeurs actuels. La bibliothèque fonctionne sous les systèmes d'exploitation Mac OSX, Windows XP et Linux, sur les processeurs Intel et PowerPC. Son implémentation multithread permet d'exploiter les capacités de calcul des processeurs multi-cœurs et ordinateurs multiprocesseurs.

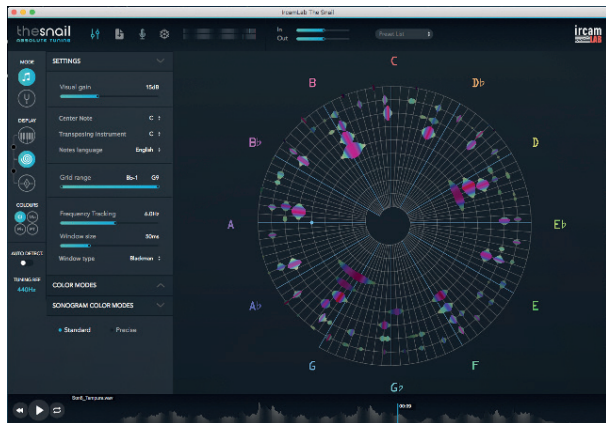
Technologies et partenaires

MakeMusic (SmartMusic), Flux::, Univers Sons, Xtranormal, Voxler, MXP4, etc.



The Snail-Absolute Tuning

Conception et développement : équipe S3AM



The Snail-Absolute Tuning est un analyseur fréquentiel de sons qui repose sur une représentation originale à alignement chromatique. Comparée aux analyseurs standard cette représentation permet de visualiser simplement les zones actives par activation lumineuse (comme sur un spectrogramme) et d'organiser la disposition des fréquences par notes sur une échelle tempérée, accordée sur un diapason.

The Snail intègre une analyse ultrafine et unique des points fréquentiels, technologie brevetée.

Applications

Visualisation musicale du signal, pédagogie, accordage.

Fonctions principales

- Visualisation en temps réel : points de fréquence projetés selon le système tempéré sur la spirale octaviante du Snail pour une « lecture musicale » du son.
- Interface : paramétrage du moteur d'analyse ; réglages des différents modes de visualisation : phase, spectre ou snail (tuner) ; réglage de l'ambitus spectral ; réglage du diapason et de la transposition, etc.

Configuration

- Mac OS X 10.9 ou supérieur, natif sur processeur Intel ou Silicon (ARM). Plug-in VST, Audio Units et AAX (sur ces plateformes).
- iOS 9 ou supérieur (iPhone ou iPad).
- Windows 7 ou supérieur.
- Lecture de fichier audio (tout format, Mac ou Windows uniquement) ou entrée du signal via l'entrée audio de l'équipement utilisé.

Le Forum Ircam

forum.ircam.fr

L'Ircam s'engage au service d'artistes, designers, ingénieurs du son, chercheurs, entrepreneurs, enseignants et étudiants en leur donnant accès à ses technologies les plus récentes et en leur permettant de diffuser leurs propres travaux.

Synthèse sonore, traitement de la voix, composition assistée par ordinateur, interaction temps réel, spatialisation sonore, contrôle gestuel et nouveaux instruments... : des centaines de modules technologiques issus des dernières avancées de la recherche de l'Ircam et des membres sont mis à disposition pour la création musicale, le spectacle vivant, les arts numériques et la production sonore. Le Forum Ircam accompagne cette diffusion par une politique éditoriale accompagnant la documentation et la promotion des nouveautés technologiques.

Alors que le Forum Ircam fête en 2023 ses 30 années d'existence, sa communauté connaît depuis quelques années une forte croissance et regroupe près de **30 000 membres dans le monde**, principalement professionnels et étudiants.

La plateforme collaborative en ligne forum.ircam.fr

Une nouvelle version totalement réécrite de la plateforme du Forum a été publiée en 2019. En sus des fonctions déjà existantes de diffusion de logiciels Ircam et de communication, elle permet aux membres de publier leur propres travaux et offre une nouvelle éditorialisation du contenu et des échanges sous forme de *projets* associant différents briques technologiques, *d'articles*, *d'événements* et de *salons de discussion* en regroupant l'ensemble de ces informations en *collections* thématiques. Une refonte complète du design de la plateforme offrant une meilleure expérience utilisateur est prévue en 2023.

Le Forum Ircam facilite ainsi le développement collectif de nouvelles idées, favorise la transmission des pratiques, l'archivage et l'indexation des connaissances, et offre une vue d'ensemble sur les outils et innovations disponibles dans les domaines de la musique et du son.

Les Ateliers du Forum Ircam

Organisés une fois par an à l'Ircam en présence de l'ensemble des équipes de recherche, les Ateliers Forum constituent le principal rendez-vous annuel de la communauté Forum : trois jours de conférences, workshops, démonstrations et de partage autour de la musique et du son, dans un esprit pluridisciplinaire. Ses éditions depuis 2021 font l'objet d'une nouvelle organisation thématique avec une dimension en ligne renforcée pour répondre aux contraintes sanitaires et pour étendre la diffusion internationale des rencontres.

Les Ateliers Forum Hors les Murs

Organisés en collaboration avec un réseau international d'universités, centres de recherche et institutions culturelles, ces événements accueillis une fois par an par nos partenaires à l'étranger regroupent conférences, workshops et performances artistiques (Séoul 2014, São Paulo et Buenos Aires 2015, Taipei 2016, Santiago du Chili 2017, Shanghai 2019, Montréal 2021, New York 2022).

L'offre Forum Ircam

L'inscription au Forum Ircam est gratuite pour les usages expérimentaux d'un grand nombre de logiciels. Les offres payantes *Premium individuel* et *Premium institutionnel* proposent sous forme d'abonnement annuel un accès exclusif à un ensemble de logiciels et à des réductions sur des formations et sur des collections de logiciels commerciaux ainsi qu'un droit d'usage plus large, notamment pour l'enseignement et la formation dans le cadre de Premium Institutionnel.

Contact : +33 (0)1 44 78 49 62 | ircam-forum@ircam.fr

Les équipes au 2 novembre 2022

■ UMR 9912 – Sciences et technologies de la musique et du son

stms-lab.fr

Directrice

Brigitte d'Andréa-Novel (SU)

Directeur adjoint

Nicolas Misdariis (Ircam)

Secrétaire générale

Anne-Marie Vaudeville (CNRS)

Responsables d'équipe

Gérard Assayag [Représentations musicales], Frédéric Bevilacqua [Interactions son musique mouvement], Clément Canonne (CNRS) [Analyse des pratiques musicales], Thomas Hélie (CNRS) [Systèmes et signaux sonores : audio/acoustique instruments], Nicolas Misdariis [Perception et Design sonores], Axel Roebel [Analyse et synthèse des sons], Olivier Warusfel [Espaces acoustiques et cognitifs]

Administration de l'UMR

Brigitte Cruz-Barney, Viktoriya Uscumlic, Anne-Marie Vaudeville (CNRS), Vasiliki Zachari

Documentaliste

Aurore Baudin

Chercheurs et ingénieurs

Carlos Agon (SU), Benoît Alary, Baptiste Bacot (CNRS), Joachim Borg, Riccardo Borghesi, Frederik Bous, Henri Boutin (SU), Antoine Caillon, Thibaut Carpentier (CNRS), René Caussé (émérite), Elaine Chew (CNRS), Alice Cohen-Hadria (SU), Frédéric Cornu, Guillaume Doras, Philippe Esling (SU), Laurent Feneyrou (CNRS), François-Xavier Féron (CNRS), Emmanuel Fléty (CNRS), Emma Frid (KTH Royal Institute of Technology), Lawrence Fyfe (CNRS), Jean-Louis Giavitto (CNRS), Karim Haddad, Olivier Houix (ESAD TALM Le Mans, Coralie Joucla, Marco Liuni,

Mikhail Malt, Benjamin Matuszewski, Rémi Mignot, Jérôme Nika, Markus Noisternig, Nicolas Obin (SU), Charles Picasso (CNRS), Emmanuel Ponsot (CNRS), Robert Piéchaud, David Roze (CNRS), Pierre Saint-Germier (CNRS), Diemo Schwarz, Patrick Susini, Isabelle Viaud-Delmon (CNRS), Coralie Vincent (CNRS), Marc Wijnand (SU), Daniel Wolff

Chercheurs associés et collaborateurs extérieurs

Moreno Andreatta (université de Strasbourg), Pablo Arias Sarah (Université de Lund, Suède), Mondher Ayari (Université de Strasbourg), Georges Bloch (Université de Strasbourg), Alain Bonardi (université Paris 8), Arthur Faraco (Université de São Paulo), Maria Larrosa (Universidad Politécnica de Madrid).

Artistes invités, artistes résidents en recherche artistique :

Baptiste Bacot (CNRS), Per Bloland (Résidence), Sasha Blondeau (Résidence), Richard Causton (Résidence), Andrea Cera, Basile Chassignain (Résidence), Judith Deschamps (Résidence), Aaron Einbond, Rémi Fox (Résidence), Emily Graber (CNRS), Aki Ito (Résidence), Joëlle Léandre, Claudy Malherbe, Roque Rivas, Nadine Schütz, Hervé Sellin, Marco Stroppa, Jean-Philippe Lambert (Résidence), Maxime Mantovani (Résidence).

Technicien

Arnaud Recher (CNRS)

Doctorants

Daniel Bedoya, Giovanni Bindi, Baptiste Bouvier, Daniel Cabanzo, Antoine Caillon, Nils Demerlé, Ninon Devis, Constance Douwes, Matthieu Duroyon, Loïc Forma, Martin Fouilleul, Valérian Fraisse, Marta Gospodarek (New York University), David Genova, Aliénor Golvet, Nadia Guerouaou, Paul Lascabettes, Antoine Lavault, Clément Le Moine Veillon (CDD Ircam), Vincent Martos, Victor Paredes, Iseline Peyre, Lenny Renault, Claire Richards, Thomas Risse, Gonzalo Romero, Léane Salais, Tom Souaille, Yann Teytaut, Alexis Thibault, Marion Voillot, Colette Voisembert.

■ Département Innovation et Moyens de la recherche

Directeur

Hugues Vinet

Administration, coordination

Brigitte Cruz-Barney, Louise Enjalbert, Ghislaine Montagne

Forum Ircam

Greg Beller, Karim Haddad, Stéphanie Leroy, Paola Palumbo, Tom Debrito

Valorisation industrielle

Pierre Guillot

Service Web

Responsable : Clément Gérard
Laura Frémy

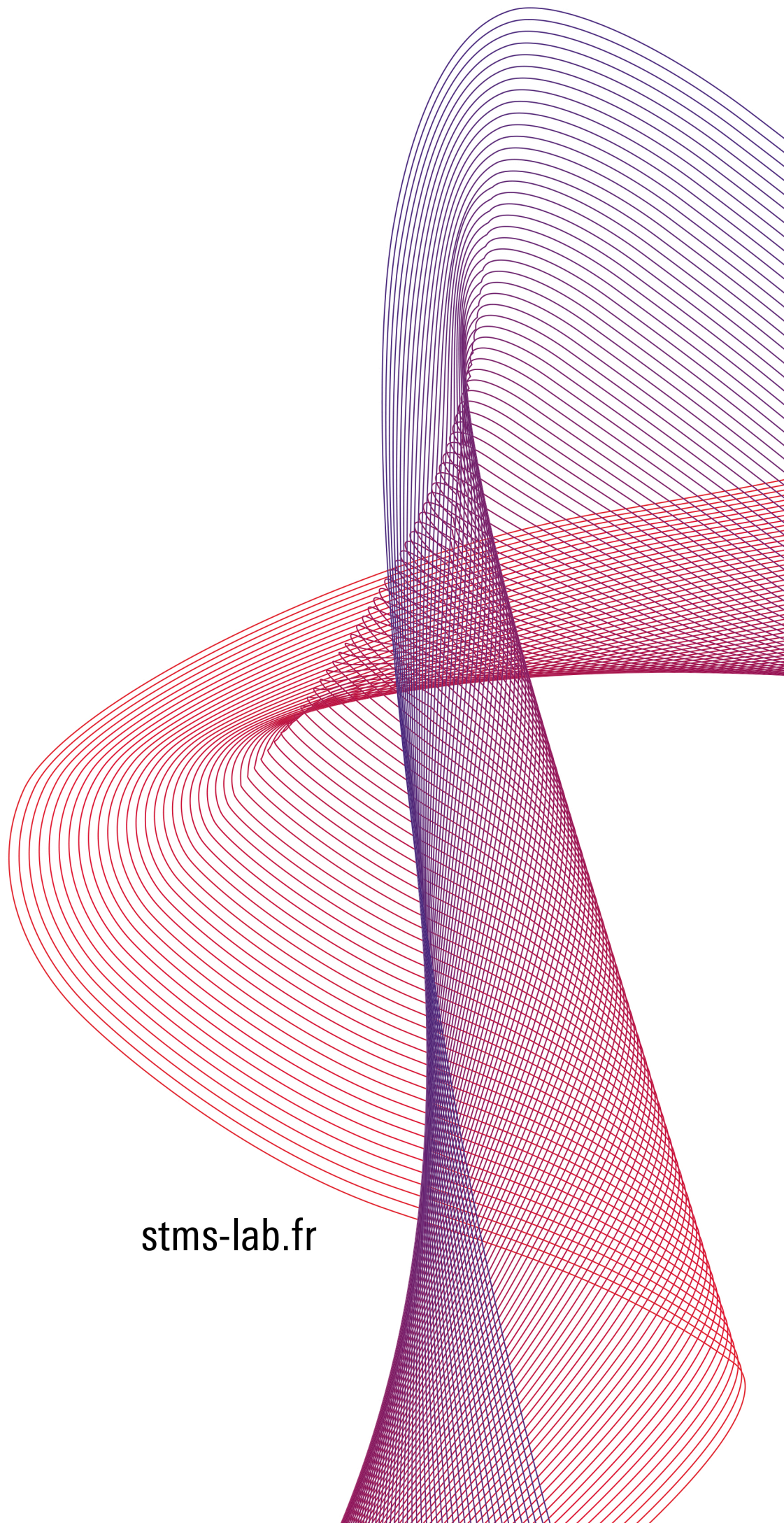
Equipe Web Audio Musique

Responsable : Guillaume Pellerin
Guillaume Piccarreta

Service informatique

Responsable : Rémi Desmonet
Benjamin Boudoir, Olivier Labat, Antoine Lorreytte

Nous écrire : prenom.nom@ircam.fr



stms-lab.fr