INNOVATION AT IRCAM: RECENT SUCCESS STORIES

IRCAM'S SUCCESS IN THE EUROPEAN H2020 PROGRAM
Over the past two years of the H2020 program (2014-15), IRCAM, despite its small size, is among the top 20 French beneficiaries in the domain of information and communication technologies, and the top beneficiary in creative industry fields.

DECIBEL D'OR PRIZE
The Conseil national du son (French National Council for Noise) awarded the Decibel d'Or 2014 prize to IRCAM's Sound Design and Perception team for its unique research in sound design that combines scientific work in auditory perception and cognition on environmental sounds with applications in the domains of sound quality and conception.

CREAM: THE FIRST ERC PROJECT AT IRCAM
Obtained at the end of 2013 by Jean-Julien Aubert, a newly recruited CNRS researcher in the SIMS laboratory, the CREAM project (Cracking the Emotional Code of Music) was selected by the prestigious European Research Council and aims to characterize the emotional mechanisms induced by listening to music.

THE OPENMUSIC SOFTWARE
reached 100,000 downloads and sets the standard for music research and computer-assisted composition.

THÉORIES DE LA COMPOSITION MUSICALE AU XXE SIÈCLE
This monumental work made up of 1860 pages, 63 chapters, and 2 volumes, published by Editions Symétrix, edited by Nicolas Durante and Laurent Penney, brings together contributions by today's top specialists, providing an overview of compositional theories of the last century. Published at the end of 2013, this work is the fruit of an unprecedented editorial project at IRCAM carried out over more than a decade.

INTERLUDE: WINNER OF THE 2013 ANR AWARD FOR SOCIETAL IMPACT AND THE 2011 GUTHMANN PRIZE
This project, supported by the Agence nationale de la recherche and coordinated by IRCAM in collaboration with Grane, NoDesign, DaPace, Voxler, and Atelier des Présentiments, examined expressive gestural interaction with musical contents.

ANTESCOF: WINNER OF THE 2013 AWARD FOR INDUSTRIAL INNOVATION AND THE 2011 SPECIAL AWARD FROM THE FRENCH MAGAZINE LA RECHERCHE
This real-time software simulates a musician's intelligence, capable of listening to other musicians and accompanying them according to their performance. It is the result of Arnaud Cent's PhD, winner of the Gillels Raun award in 2009.

POLYPHONIC TRANSCRIPTION AUDIOZNOTE INTEGRATED IN LIVE 9
Version 9 of the software Live, launched in March 2013, includes the pioneering technology AudiozNote by IRCAM, the result of a collaborative project with Ableton in Berlin, making automatic transcription in MIDI of polyphonic audio recordings possible.

NEW MASSIVE MULTICHANNEL AUDIO SYSTEM IN THE ESPACE DE PROJECTION
Inaugurated in November 2012, this unique system is comprised of 365 loudspeakers installed in the Espace de projection - a room with modular acoustics - and offers high-resolution spatialized sound reproduction through the combination of the Wavefield Synthesis and High-Order Ambisonics systems.

SUPERVP ACCLAIMED BY THE AUDIO INDUSTRY
Software for high-quality sound processing, SuperVP is the international leader in its category and has been licensed to numerous software publishers and audio services. SuperVP is the engine for the Audiosculpt program.

SPAT: WINNER OF THE RESOLUTION MAGAZINE AWARD FOR THE BEST AUDIO PLUG-INS
Offering a novel approach to the simulation of sound spaces based on perceptual criteria. Poten and included in the MPEG’s audiovisual standard, Spat is sold in the IRCAM Tools collection in collaboration with the company Flux...

SOUND DESIGN OF CONCEPT AND PRODUCTION CARS BY RENAULT
IRCAM has worked closely with car manufacturers, including Renault, for the perception of the sounds heard in the passenger compartment as well as the sounds made by electric cars (e.g., the ZOE Renault concept car).

KEY FIGURES IN IRCAM'S R&D

7 research teams work in various scientific fields related to music acoustics, signal processing, computer science, psychology, musicology.

110 contributors (researchers, engineers, technicians, coordinators, & 40 interns) and guest researchers per year.

30% of staff are of foreign nationality.

200 publications per years, 50 of which appear in scientific journals or as chapters in books.

90 annual partnerships with 60 labs and public organizations and 30 with private companies.

10 software environments developed and distributed to 6,000 professional users registered to the IRCAM-User (software user group).
FOREWORD

This five-part brochure, written for fellow researchers and professionals, presents an up to date and exhaustive view of IRCAM’s research and development activities.

The research activities hosted by IRCAM are a part of the STMS (Science and Technology of Music and Sound) joint research lab that brings together the CNRS, the Université Pierre et Marie Curie, and the French Ministry of Culture and Communication. (Director: Gérard Assayag, five-year period 2014-2018).

The major events during this period include:

- In January 2017, a new team S3AM (Systèmes et signaux sonores – Audio/acoustique et instruments) directed by Thomas Hélie was created. This team brings together the Instrumental Acoustics and activities of the Sound Analysis & Synthesis teams on the theme of modeling multi physic systems and sound signals.

- 2017 also marks the important process of drawing up the UMR STMS’ research project for the five-year period 2019-2023 in view of an assessment by HCERES in articulation with a reorganization throughout IRCAM, renewing the themes and organization on musical research.

- March 2017 is the launch of a new annual event at the Centre Pompidou called Mutations/ Créations co-organized by IRCAM featuring exhibitions, symposia, and artistic events with the goal of exposing and debating the evolution of artistic practices within today’s scientific and technological ecosystem. The theme for the inaugural edition in March 2017, in connection with the exhibits “Imprimer le monde” and “Ross Lovegrove” is “Simulated Space or Digital Forms”.

- IRCAM drafted the selected European Vertigo project, as project coordinator, for the call H2020 ICT36-2016. The project is a part of the STARTS initiative (Science, Technology and the ARTS) created by the European Commission’s program DG Connect and aims to coordinate, on a European level, collaboration between artists and R&D projects in innovative technology fields via residences, the development of an online collaborative platform, and an annual public presentation of work carried out by members of the community during the Mutations/ Créations event at the Centre Pompidou.

- The selection of Vertigo, together with iMuSciCA dedicated to educational applications for sound synthesis through physical modeling, supports the recent success of IRCAM in the European Commission’s H2020 calls and its position among the top French actors involved in the H2020 ICT program.

- 2016-2017 saw the launch of new, innovative products connected to IRCAM software: The Snail-Analyser (smart tuner) and its extensions for new platforms, a new, massively multichannel version of Spat Revolution IRCAM Tools which obtained the NAMM Tech Award in January 2017, the version 2 of IrcamMax for Ableton’s Live software marketed by Ableton, and a new, extended version of Ircam TS Pro in the Ircam Lab collection.

Hugues Vinet, January 2017
1

RESEARCH AT IRCAM

Science and Technology of Music and Sound
STMS Lab
(IRCAM, CNRS, UPMC)
MISSION STATEMENT

The fundamental principle of IRCAM is to encourage productive interaction among scientific research, technological developments, and contemporary music production. Since its establishment in 1977, this initiative has provided the foundation for the institute's activities. One of the major issues is the importance of contributing to the renewal of musical expression through science and technology. Conversely, specific problems related to contemporary composition have led to innovative, theoretical, methodological, and applied advances in the sciences with ramifications far beyond the world of music. Focused on artistic production, in all its particularity and sensitivity, this highly original research dynamic addresses modes of representation and of extended manipulation of sound and music, thus eliciting scientific and technological applications that touch an increasingly wide-ranging public including sound and music professionals, industrial concerns, academics, music devotees, etc.

Central to the interaction between research and musical production is the development of software tools for musicians, composers, performers, and musicologists using the models and prototypes created by the research teams in various music-related domains including information technology (e.g. languages, human-computer interaction, real-time, and databases), signal processing, acoustics, auditory perception and cognitive psychology, and musicology. The work carried out in the STMS (Science and Technology of Music and Sound) joint research lab supported by the CNRS and the Université Pierre-et-Marie-Curie (UPMC-Sorbonnes Universités).

Research, the source of innovation, seeks to expand our knowledge of music-related issues. Because of its essentially cross-disciplinary character, a great deal of research is carried out in collaboration with laboratories in France and abroad, academic institutions, institutional partners (particularly the CNRS, the Université Pierre et Marie Curie – Sorbonne Universités, and INRIA), and private partners. The integration of graduate and post-graduate students enables research training. The skills developed at IRCAM can be applied to domains beyond the musical arena and are frequently included in projects carried out with industrial partners or in the context of European, French, or multinational research projects.
The development pole seeks to adapt the knowledge, models, and prototypes produced by the research teams into software environments. The principle applications are used to create tools that assist musical creation by putting in place open and programmable environments, adaptable to a variety of aesthetic approaches and that allow for the continued integration of models and their updates. This configurable design makes it easier to develop simplified versions of the software programs intended for the general public: Ircam Tools (for audio professionals), IrcamMax (for live electronic music), IrcamLab (for a broad range of users). Central to software development is the design and evaluation of specific human-computer interaction interfaces and the permanent integration of the rapidly evolving technology from the computer industry.

The IRCAM Forum—an IRCAM software user group—encourages the distribution of the institute’s software programs within an international community of music and sound professionals (e.g. composers, multimedia and visual artists, sound engineers and designers, researchers, teachers, etc.) with an estimated 5,000 users since its beginnings. The IRCAM Forum also provides companies with licenses for either internal use or distribution purposes. These licenses can apply to either complete applications (e.g. the Max program, used by tens of thousands worldwide, is licensed to the company Cycling ’74 in California) or to specific functional models (e.g. audio analysis and processing, indexing, spatialization, synthesis) that are included in commercial environments.
INSTITUTIONAL LINKS

IRCAM, a non-profit making association with recognized public utility, is associated with the Centre Pompidou, under the aegis of the French Ministry of Culture and Communication.

To reinforce its presence in the national and international research scenes, the development of strong ties with large research institutions and universities is one of IRCAM’s priorities. In addition to the numerous collaborations with laboratories that specialize in several different fields, its major associations are:

- The STMS Joint Research lab (Science and Technology of Music and Sound – UMR 9912) brings together researchers and professors from the CNRS, IRCAM, and the UMPC (Université Paris 6) to work on a multidisciplinary scientific project. The UMR (unité mixte de recherche) is associated with the CNRS Institutes for Information Sciences and Technologies (INS2I), for Engineering and Systems Sciences (INSIS), for Humanities and Social Sciences (INSHS) and of Biological Sciences (INSB). In the case of the UPMC, the lab is part of the Faculty of Engineering and in addition a member of the Sorbonne Universités’ Collegium Musicae. IRCAM is also a supervising institution, and the STMS lab is a member of the Labex SMART that focuses on man-machine interfaces, coordinated by the UPMC. A collaboration has also begun with the INRIA. This collaboration began in 2013 with the creation of a common team project called MuTant that focuses on the elaboration of synchronous languages for music and the analysis of audio streams.

- The ATIAM Master’s program (Acoustics, Signal Processing, and Computer Science Applied to Music) hosted and organized by IRCAM (coordinator: Moreno Andretta) is a multidisciplinary scientific training program applied to music. This unique program is organized as a part of the Master’s in Science and Technology offered by the UPMC in collaboration with the École Télécom ParisTech. IRCAM is also a partner in the Sound Design Master’s Degree organized with the EPCC - École supérieure des Beaux-Arts Tours Angers Le Mans, Université du Maine and ENSCI-Les Ateliers.

- The STMS Joint Research lab is an accredited host lab within its specialized fields for the UPMC doctoral programs, notably EDITE (a Paris post-graduate school for information technology, telecommunications, and electronics), EDMMAER (mechanical sciences, acoustics, electronics), and ED3C (brain, cognition, behavior). In the framework of the Sorbonne University group and the Idex SUPER, the Université Paris 4 – Sorbonne (ED Concepts et languages), the Institute for Doctoral Education at the UPMC, and IRCAM established a new doctoral program for composers called Doctorat de musique - Recherche en composition that began with the 2014-2015 academic year.

On both the national and the European level, IRCAM’s active participation in a variety of collaborative research and development projects (e.g. the Cap Digital competitiveness cluster, national research networks such as ANR, PIA, FUI, BPI, the European Commission’s programs (FP7, H2020), also provides the opportunity to maintain existing and establish new relations with numerous public and private partners.
ORGANIZATION AND FIELDS OF STUDY

IRCAM provides a unique framework to welcome and coordinate various scientific points of view on the musical phenomena, including physics (acoustics, mechanics), signal processing, information technology, cognitive psychology, and musicology. Activities within the UMR-STMS Joint Research lab are structured by themes, entrusted to specialized teams. Each team assumes responsibility for all the work carried out within its domain: research, software development, contracts, collaborative projects, and diffusion. The people working within IRCAM’s research and development department and the UMR-STMS Joint Research lab (110 research workers, engineers, doctoral students and post-doctoral researchers, technicians, and administrative staff members) are divided among the following teams in the Research and development department directed by Hugues Vinet and the SMTS joint research unit directed by Gérard Assayag.

ACOUSTIC AND COGNITIVE SPACES
The research and development activity carried out by the Acoustic and Cognitive Spaces team led by Olivier Warusfel, concentrates on the notion of space in its manifestation and its sonorous reproduction as well as in its cognitive reality. The scientific domains concerned include signal processing for reproducing a sound field and computer science applied to the design of spatialization control interfaces. Spatial perception and cognition are also a part of the reflection on the spatial dimension when listening to music or to our environment in association with other sensomotor modes of perception. The artistic applications concern the musical arena for concert or post-production situations as well as the domain of dance via issues surrounding sound/space/body interaction. The integration of sound spatialization functions in virtual reality environments opens the possibility of scientific applications for neuroscience research, therapeutic systems, or simulators for transportation.

SOUND PERCEPTION AND DESIGN
Established in 2005 by Patrick Susini who has headed the team since its creation, the Sound Perception and Design team brings together researchers in psychoacoustics, sound signal processing/synthesis, experimental psychology, and, more recently, cognitive neurosciences. This multidisciplinary team has created a program that combines fundamental research on the perception of sounds with acts of creation and education in the field of sound design. The team’s work centers on the perception of sounds concerning: the perceptive characterization of sounds that range from signal descriptors to associated auditory percepts and concepts; the cognitive mechanisms involved in the identification of a sound source and in the emotional processing of sounds; interactions between auditory perception and motor processes. The applications of sound design are connected to the study and design of our everyday sound environment: sound-signaling in public environments; browsing through human-computer interfaces; sound design for electric vehicles; sound interaction, quality, and identity of industrial products.

SOUND ANALYSIS & SYNTHESIS
Headed by Axel Roebel, the Sound Analysis & Synthesis team studies the procedures of analysis, synthesis, and transformation of sounds using signal processing methods or from the acoustic modeling of sound sources. Control systems are then designed, based on dedicated user interfaces and rule-based systems. Much of the team’s current research is focused on the analysis and automatic indexing of musical content. This research has led to the development of audio synthesis and processing tools for composition (SuperVP, AudioSculpt, and Diphone Studio). Many applications are found in fields such as speech processing, sound simulation, indexing, and multimedia.

SOUND SYSTEMS AND SIGNALS: AUDIO/ACOUSTIC, INSTRUMENTS
The team Sound Systems and Signals: Audio/Acoustic, InsTruMenTs—S3AM—led by Thomas Hélie, creates theoretical, technological, and experimental tools based on multiphysical systems and the sound signals they produce. The global approach, “systems and signals” creates a synergy in the conception of these tools, combining several disciplines and scientific fields: physics, theory of systems and control, differential geometry, numerical analysis, signal processing, computer science, electronics, mechatronics, and robotics.

MUSICAL REPRESENTATIONS
The Musical Representations team, headed by Gérard Assayag, investigates the computer formalization of musical structures. Its work is focused on the development of the OpenMusic environment that enables the computation of musical structures using graphical programming languages. The team’s recent work reflects a broader interest in automatic analysis of musical structures from symbolic representations, synthesis control, computer-assisted composition or orchestration, style modeling, computer-generated improvisation, the application of computer modeling on ethno-musicology, and in computer-assisted music education.
ANALYSIS OF MUSICAL PRACTICES
The Analysis of Musical Practices team, headed by Nicolas Donin, focuses its research on musical practices from the 20th and 21st centuries. The team’s work centers on musicology and the humanities and branches out in four directions: empirical research on contemporary composition and interpretation, musicological works that lead to the creation tools for listening and musical analysis, historical and epistemological investigations into scholarly musical practices of the 20th century, and the publication of scientific works and composers’ writings.

SOUND MUSIC MOVEMENT INTERACTION
The Sound Music Movement Interaction team, previously called the Real-Time Musical Interactions team, headed by Frédéric Bevilacqua carries out research and development on interactive systems for music and live performances. These systems implement the team’s work on multimodal (gesture and sound) sensing and analysis and interactive sound synthesis. Current projects focus on collective interaction and sensori-motor learning processes. Applications can be found in artistic creation (music, dance, interactive installations), education, new digital musical instruments, musical games, and rehabilitation.

The research and development department also encompasses logistical teams that work in collaboration with all IRCAM’s departments.

The IRCAM Resource Center headed by Nicolas Donin ensures the management and access to both physical documents in the multimedia library and multimedia documents via the website ressources.IRCAM.fr.

The Information technology service, directed Raphaël Sallé de Chou, manages all IRCAM’s computer systems and networks—approximately 400 computers, including 50 physical servers that house 250 virtual servers connected via a dual 10 Gbits/s connection through the Renater network.

The mechanics workshop, coordinated by the S3AM team, designs and produces mechanical prototypes according to the specifications given by each team.
CURRENT ISSUES
AND ORIENTATIONS

The evolution of IRCAM’s research and development projects relies, to a large extent, on the development of scientific and technological knowledge in relation to artistic activity and the repercussions of such scientific expertise on more widespread societal issues. The current scientific objectives can be expressed in terms of four themes that run through the research and development department. These themes that focus on the production of sound material, writing, interaction, and the contribution of other scientific disciplines provide structure for the complementary knowledge found in the different teams and their areas of expertise of musical research.

1. MODELLING, SIMULATION AND ANALYSIS

This field applies to all sound synthesis and processing procedures and is based on the findings of IRCAM’s various research teams—and in the refinement and renewal of these findings. This field applies particularly to the following aspects:

- Signal Models for Sound Analysis, Synthesis, and Processing
  The subject of this research has found numerous applications for studio work and live performances using real-time technology. Beyond perfecting existing models—particularly the phase vocoder via the SuperVP engine—current research issues focus on methods for automatic analysis of polyphonic audio streams—temporal segmentation and source separation—and on the elaboration of models of specific signals of different sound source classes such as instruments (source-filter dynamic filters) and sound textures (environmental sounds like rain, wind, rubbing).

- Voice Synthesis and Processing
  While general methods for sound processing apply to certain manipulations of vocal sound signals (e.g. transformation of a speaker’s gender or identity), specific methods, involving an understanding of different levels of representation of linguistic contents, phonology, and semantics, are carried out on syntheses from text as well as on analyses and models of prosody and expressivity. New research directions are currently being launched on sung voice synthesis. These works are particularly interesting for musical creation, as they provide a novel form of manipulation of either spoken or sung vocal material and its confrontation with instrumental and electronic sources.

- Music Information Retrieval
  Musical indexing consists of automatically extracting data (descriptors) from recordings. The research in this domain is carried out in response to important societal issues associated with music distribution based on new modes of accessing music that rely on databases of recordings. This work leads to the elaboration of methods for automatic description, analysis, and classification of pieces of music that, beyond their applications for cultural industries, are of interest for musical creation, in that they provide an extension of the traditional musical vocabulary (pitch, intensities, temporal structures) introduced via descriptors of new sound categories.

- Virtual and Augmented Instrument-Making
  This theme is connected with the study of phenomena associated with the physical production of sound, especially musical instruments. Research focuses on modeling and simulating new interactions, non-linear behavior (excitation, propagation) and sound radiation, as well as the study of active control applied to instruments. These studies lead to applications for creation such as virtual, augmented or hybrid instrument-making and also to simulation and virtual reality.

- Sound Spatialization
  The use of sound spatialization techniques has become common in all styles of musical creation and is also widely used in applications for audiovisual productions, virtual reality, and multimedia simulations. Current research focuses on advanced systems of audio diffusion (Wave Field Synthesis, Ambisonic, binaural and transaural) and the perfection of new hybrid models that combine perceptual and physical approaches, extending the concepts of Spat. Inaugurating an important step for artistic and scientific experimentation in sound spatialization, an unparalleled 450-loudspeaker system that combines a peripheral WFS system and a high order Ambisonic system was installed in IRCAM’s Espace de projection in the beginning of 2012.
2. REPRESENTATIONS AND LANGUAGES
This set of themes, set primarily in the domain of computer science, concerns modes of representation and manipulation of musical contents. As an integral part of the problems surrounding writing for composition, this set of themes focuses on the elaboration of computer languages and environments that formalize the abstract structures describing musical processes and modes of programming and calculating these structures in connection with techniques for sound analysis and synthesis found in the preceding section, Modeling for Sound Synthesis and Processing. These studies involve the implementation of relationships between the different levels of representation of sound and music information. This subject has been a topic of study for several years at IRCAM under the appellation signal/symbol. The studies carried out include the following research themes:

• Computer-Assisted Composition
In the continuation of the research and development carried out on the OpenMusic environment, the work here focuses on the extension of existing functions such as representations and operations on temporal structures and rhythm, new man-machine interfaces intended for composers (augmented paper). OpenMusic is evolving into a reactive environment that combines high-level symbolic representations and the functions of interactive environments.

• Writing Sound and Spatialization
Establishing the relationships between structures intended to assist instrumental writing on one hand and techniques for analysis and synthesis created during the first theme on the other hand in addition to the formalization of electronic processes that is currently the object of important advances, for both synthesis control and spatialization.

• Computer-Aided Orchestration
A unique and recent case of computer-assistance for composition, this project addresses the question of orchestration through automatic search of instrument combinations in function of different criteria of acoustic similarity, from a target defined by the composer. Current research focuses on this paradigm of dynamic orchestration, following the targets with sound characteristics that vary throughout time.

• Writing Time and Interaction
The objective of this research is to develop concepts, tools, and languages that go beyond the notions used in today’s environments, notably Max (combination of reactive asynchronous processing and synchronous processes for signal processing) to provide composers with new modes of formalization and execution of interactive musical processes. Continuing work on computer-performer synchronization carried out on score following, an important objective is to apply and extend the formalizations of synchronous languages to the case of musical processes.

3. MUSICAL AND SOUND INTERACTIONS
This third theme concerns the study and development of interactions that associate human subjects, particularly in performance situations, with musical or sound processes, and possibly in the larger context of multimedia systems. The applications concern the performing arts (music, dance, theater, etc.), installations, interactive sound design, and more generally, situations of interaction where the coupling of performance and sound play an important part. The research directions are the following:

• Gesture Analysis
The work on the analysis of gesture includes two complementary aspects. Work focuses firstly on the design and production of gesture sensor systems for specific situations (such as the wireless system MO) and secondly, on the study and modeling of gestural signals that come from different sensing systems.

• Gesture/Sound Interaction
This research and development theme brings together several fundamental studies on the relationships between gesture and sound in different situations and on the creation of systems for interaction, calling upon real-time sound synthesis techniques and mapping methods, continual form recognition and following in relation to the synthesis parameters with the gestural information.

• Symbolic Interaction
This third part of this theme concerns the development of musical interaction situations by computer, particularly in the context of improvisation. The use of symbol sequence models (Oracle of factors, OMax system) is a component of the memory of past performances, enabling the generation of new sequences that compare the performer with his computer-double that reacts to its own performance. Another application is the extension of notation of electronic instruments for the generation, depending on the performance, of sequences from analysis of predefined musical corpora.
4. COGNITIVE, THEORETICAL, AND MUSICOLOGICAL FOUNDATIONS

This final cross-disciplinary theme groups together a body of research that nourishes the three aforementioned themes prior to their execution by bringing knowledge and methodologies from a range of disciplines and finding applications from different aspects of IRCAM’s activities.

• Psychology of Hearing and Sound Design
The objective these studies is twofold: to contribute to the advancement of knowledge in the domain of experimental psychology, in hearing perception and cognition (especially environmental sounds), and also to apply the results and methods used to sound design and to the design of sounds that play a part in our daily surroundings.

• Neurosciences for Music
In connection with research in psychoacoustics, these studies use cerebral imagery to understand cognitive activity and emotional response while listening to music with the aim of creating new computational models for musical description and interaction.

• Intermodal Spatial Cognition
In connection with research carried out on sound spatialization, these works study the role and relationships between different sensorial modalities in the creation of mental representations of space, in particular in situations of multimodal interaction and virtual reality.

• Contemporary Musicology and Analysis of Musical Processes
This research theme focuses on contemporary works, expanding the methods of analysis practiced in musicology to the study of the process of the production of these works. Structuring these findings is the subject of exclusive hypermedia developments and provides the contents for the presentation of musical works.

• Music and Mathematics
This research field is dedicated to the study of mathematic formalisms describing certain musical processes encountered in contemporary works.

• Musical Knowledge Engineering and Preservation of Works
This research theme is centered on clarifying and formalizing knowledge found in the technical realization of interactive works using digital supports, particularly for the purpose of long-term documentation and preservation.
RESEARCH DISSEMINATION AND TRANSFERS

Central to the societal and economic issues found in both culture and information technologies, the research conducted at IRCAM is present in the international research community as a point of reference for multidisciplinary projects in the sciences and technologies of sound and music constantly in touch with society’s new demands and uses. This mediation takes on a variety of forms depending on the targets and domains concerned.

SCIENTIFIC ANIMATION AND MULTIDISCIPLINARY DEBATES

Every year, IRCAM is home to several international specialized conferences (e.g. NIME’06, IHM’07, ICAD’08, MCM’11, DAFx’11, ELS’14, MoCo’14, UMSM’14, WAC’15, and TCPM’15) and also organizes multidisciplinary meetings on subjects relevant to current research themes surrounding musical research and societal issues in culture and technology: Writing Sound and Interaction (2006), Melody and Melodic Function as an Analysis Object (2006), Politics of Musical Analysis (2007), New Paradigms for Computer Music (2007), Expressivity in Music and Speech (2008), Complexity in the Arts (2009), How to Analyze Improvisation (2010), Body/Space/Sound (2010), Model/Prototype/Work (2010), Producing Time (2012), Digital Dark Age (2013), Composing (with) Gesture (2014), New Technologies for Music Notation and Representation (2015), Emotionnal Archetypes (2016).

SCIENTIFIC VALORIZATION AND TECHNOLOGY TRANSFERS

IRCAM brings together a wide-ranging set of scientific and technological competences for application in multiple domains focused on the production and diffusion of music, sound, digital arts, and cultural industries as well as in domains that incorporate a sound facet: sound design (automotive, transportation, urban environment), multimodal interfaces, virtual and augmented reality, simulation, etc. Armed with several decades of experience in the management of innovation and collective creativity, IRCAM’s advice is often solicited on this subject by large industrial groups in a variety of fields such as innovation, marketing, human resources, and communication.

IRCAM is recognized by the French Ministry of Research and Education in respect of the research tax credit; services carried out by the research teams for private partners benefit from tax advantages. Connections with industry can be found in numerous collaborations: national and European R&D projects (20 currently underway, a third of which are coordinated by the laboratory), research services, co-directing CIFRE theses, accompanying young start-ups (Phonotonic, Mogees, and Niland are the most recent examples), and licensing agreements. IRCAM is particularly active in this final domain; several dozen licensing agreements for technological bricks are in place with both small and large businesses primarily in Europe and North America concerning technological fields in which IRCAM is today’s leader (e.g. sound synthesis and processing, musical indexing, sound spatialization, sound/performance interaction).

The culture of professional technological development—a peculiarity of IRCAM—included in its research teams leads to the production of finalized, optimized modules that are directly transferrable to the industry. More generally, the activity of development can be found in various forms of production adapted to different target categories. More than a dozen software environments are developed within the teams (see section 4), regularly perfected in sync with the latest advances in research, used by other IRCAM departments for their artistic production or educational projects (composers, computer-music designers, sound engineers, musicians, dancers, etc.), and commercialized by the IRCAM Forum—an international community of over 5,000 professional users. Simplified applications derived from software products can be found in a range of commercial packs.

1. forumnet.ircam.fr
SOFTWARE COLLECTIONS
Derived from IRCAM software environments, these collections and sound libraries, developed and commercialized in collaboration with well-known partners, ensure the diffusion of IRCAM’s technologies and knowledge with a broad range of users.

• Ircam Tools
Developed in collaboration with and sold exclusively by Flux; Ircam Tools offers high quality audio processing modules for professional musical and audiovisual production (www.ircamtools.com). The Spat module from this collection received the award for best plugin from the magazine Resolution in 2011.

• Ircam Max
Entirely developed at IRCAM, these devices compatible exclusively with the Live sequencer from the German company Ableton, offer experimental modules for interactive electronic music that include technologies from IRCAM. Ircam Max is sold by Ableton.

• Ircam Solo Instruments
Produced in collaboration with the company Univers Sons, this collection offers sample banks of acoustic instruments including contemporary playing modes (IRCAM Solo Instruments and IRCAM Prepared Piano). This collection is sold by Univers Sons.

• Ircam Lab
Developed completely at IRCAM and sold by the company Don’t Crack (www.dontcrack.com), Ircam Lab offers a suite of independent applications for Mac and PC. The first item in this collection, IRCAM T.S., was released at the end of 2014. These high quality modules are designed for simplified use and address more specifically the sound and video industry, digital education, as well as amateur and professional practices.

T.S. enables real-time audio manipulation (transposition and stretching) and remixing of sound components. These high-level modules have been designed to be easy to use and are intended for use by professionals in the audio/video industry as well as digital education. The latest addition to this collection, The Snail-Absolute Tuning, released in January 2016, offers an advanced tuner and a precise means to visualize music and sounds using a patented algorithm, the result of a collaboration between the CNRS and IRCAM in the STMS joint research lab.
EXAMPLES OF RECENT OR CURRENT PARTNERSHIPS
(LICENSES, RESEARCH SERVICES, AND CIFRE THESES)
2

RESEARCH TEAMS
ACOUSTIC AND COGNITIVE SPACES
Head Researcher: Olivier Warusfel

ACTIVITIES
- The research carried out by the Acoustic and Cognitive Spaces team centers on the notion of space, in both its sonorous manifestation and reproduction and its cognitive reality. The two major research axes focus on techniques for spatialization and spatial auditory cognition in the context of multi-sensory interaction. More concretely, the team concentrates on developing methods to capture, reproduce, or synthesize the spatial dimensions of a sound scene, locating the sound sources in space, and the acoustic signature of the environment. In parallel to this, the team studies the mechanisms used by the central nervous system to interpret acoustic spatial information and integrate it, if need be with information that comes from other sensory modes (i.e. vision or proprioception). From this double approach, we expect mutual enrichment, knowing that the study of spatial cognition is neuroscience requires the intervention of perfected techniques and these techniques could be optimized considering the perceptive capacities of the listener.

The work carried out concerning spatialization techniques are concentrated on synthesis procedures based on a physic formalism of the sound field (binaural, Wave Field Synthesis, and Ambisonics). The two major issues are: the mastery of a 3D sound field that enables the synthesis of virtual sound sources and the effect of a specific space in the entire auditory sphere and high-resolution synthesis of the directivity of virtual sound sources. While this axis of research focuses primarily on sound-signal processing, the objective is also to provide tools for high-level control, using descriptors taken from musical composition and employing them in compositional environments such as those developed by the Musical Representations team.

These 3D audio spatialization techniques associated with a system that captures the movements of a performer or a member of the audience, constitute an organologic base essential for addressing questions on ‘musical, sound, and multimedia interaction’. They offer the opportunity to reflect on the ‘cognitive foundation’ related to the feeling of space, in particular on coordination necessary among various sensory methods for the perception and cognition of space. More specifically, we wish to highlight the importance of the processes of integration between the idiothetic indices (related to our motor actions) and the acoustic indices (localization, distance, reverberation, etc.) used by the central nervous system to create a spatial representation of the perceived environment.

On the musical level, our ambition is to provide models and tools that enable composers to include sounds in the space throughout the compositional process: from writing to concert. This contributes to making spatialization a parameter of musical writing. In the arts, this research also applies to post-production, to interactive sound installations, and to dance via the issues surrounding sound/ space/body interaction. The incorporation of sound spatialization functions in virtual reality environments creates the opportunity for scientific applications to be used in neuroscience research, therapeutic systems, or transportation simulators.

MAJOR THEMES AND ASSOCIATED PROJECTS
- Reproduction of sound fields (Wave Field Synthesis, High Order Ambisonic, binaural and transaural technologies, BILI project, p. 40, 38)
- Analysis, transformation of sound scenes, hybrid reverberation (ORPHEUS, SOR2, 3DR3 projects, p. 29)
- Analysis and synthesis of sound scenes (directivity controlled sources)
- Spatialization control (spatialization control and spatialized sound synthesis, (EFFICAC(e), COSIMA projects, p. 45, 51)
- Spatial auditory perception and cognition (Verve, HC2 projects, p. 42)
- Spatialisateur, ToscA, and Panoramix software (p. 80, 81, 82)

SPECIALIST AREAS
- Signal processing, physical acoustics, spatial perception/ cognition, virtual and augmented reality

COLLABORATIONS
- Bayerischer Rundfunk, BBC, Biocom, Ben Gurion Univ. (Tel Aviv, ISR), Cnsmdp (FR), CNES (FR), Elephant Candy, Flux:: (FR), France Télévisions (FR), Fraunhofer ISS, Hôpital de la Salpêtrière (FR), HEGP (FR), Hôpital Universitaire de Zürich (CH), Inria (Sophia Antipolis, FR), IRBA, IRT, Joanneum Research (Graz, AU), ITA RWTH, Limsi (FR), Orange-Labs (FR), RWTH (D), Radio France, Trinity College Dublin (IRL)

TEAM
SOUND PERCEPTION AND DESIGN
Head Researcher: Patrick Susini
Deputy Head Researcher: Nicolas Misdariis

ACTIVITIES
- The team’s research focuses on the perception and cognition of sounds, bringing together knowledge from psychoacoustics, sound signal processing/synthesis, psychology, and cognitive neurosciences. More generally, the team’s research revolves around creation and educational activities in the field of sound design.

Research focuses primarily on environmental sounds. This rarely studied subject provides access to different levels of cognitive representation in connection with the surface characteristics of sounds and the sources that produce them. The team’s work therefore concerns the perceptive characterization of these sounds and of the cognitive mechanisms used to identify them.

Environmental sounds are often produced in a context of interaction combined with movement when manipulating an object. One of the major themes in the team is to continue to study environmental sounds in a process that combines perception and action. This is a vital and unique gateway for research in auditory cognition and for interactive sound design applications.

The team’s research topics have been extended to vocal imitations as well as musical sounds and to the voice in connection with environmental sounds. This creates new perspectives for better understanding the cognitive processes involved by the identification of a sound source and by the emotional processing of sounds.

Applications for the team’s work are found primarily in sound design: we create the rules, the tools, and the methods—derived from our fundamental work—with the aim of facilitating the creation of functional and pleasant sounds for our everyday lives. The societal impact of our work was awarded in 2014 with the Décibel d’or Recherche prize granted by the French Comité National du Bruit.

Creation in sound design is carried out combining the “knowledge” with “know-how” of the team’s researchers and associated composers (Andrea Cera, Sébastien Gaxie, Alexander Sigman, Hiroshi Kawakami, Roland Cahen) in different projects. There are three types of creative projects carried out by the team: industrial projects (telecommunications, the automotive and the luxury industry); research projects for the creation of sound signaling; and educational projects for a variety of workshops on subjects that range from sports equipment to mass transit.

Finally, the team has associated disparate educational actions:
- Organization of applicative workshops in partnership with industrial companies and collectivities as a part of the DNSEP degree program in Sound Design at the École Supérieure des Beaux Arts TALM (Le Mans) in collaboration with the design school ENSCI - Les Ateliers (http://lemans.esba-talm.fr/etudes/option-design-sonore).
- Organization of professional training courses that combine a project with classwork in connection with IRCAM’s department of Education and Cultural Outreach (http://www ircam.fr/formations.html).

MAJOR THEMES AND CURRENT PROJECTS
- Loudness of non-stationary and multi-source sounds (LoudNat, p. 67 and MoUVie, p. 66)
- Analysis and perceptive characterization of complex sound scenes and environments (Houle, p. 37 and MoUVie, p. 66)
- Sound, imitation, and identification (SkAT-VG project, p. 65)
- Sound and interactions (Legos project, p. 47)
- Sound and emotions (CREAM project, p. 64)
- Sound signaling and HNI (RoadSense project, p. 67)
- Tools and methods for sound design (CIFRE SNCF thesis, SkAT-VG project, p. 65)

SPECIALIST AREAS
- Psychoacoustics, experimental psychology, cognitive neurosciences, acoustics, sound design, data analysis

COLLABORATIONS
- TU-Berlin (Berlin, Germany), J. Gutenberg Universität (Mainz, Germany), McGill University (Montréal, Canada), University of Canberra (Canberra, Australia), IUAV (Venice, Italy), Politecnico di Torino (Turin, Italy), Carnegie Mellon University (Pittsburgh, United States), KTH (Stockholm, Sweden), Lund Universitet (Lund, Sweden), ZHdK (Zürich, Switzerland), LVA- INSA (Lyon), LMA-CNRS (Marseille), CR-ICM (Paris), LAM (Paris), LAPPS (Paris), Ebam TALM (Le Mans), ENSCI (Paris), Université de Strasbourg, Centre d’Études Techniques de l’Équipement (Paris), IFITTAR (Marne-la-Vallée), Renault, PSA, SNCF, EDF, Klaxon, LAPS-Design, Genesis, Axium, Univers Sons, Région Île de France

TEAM
- Researchers: J.-J. Aucouturier (CNRS), O. Houix, N. Misdariis, P. Susini
- Researchers Associated with Projects: L. Goupil, M. Liuni (CREAM), M. Vannier (MoUVie), E. Ponsot (CREAM)
- Doctoral Students: R. Leiba (MoUVie), L. Rachman (CREAM)
**SOUND ANALYSIS & SYNTHESIS**
Head Researcher: Axel Roebel

**ACTIVITIES**

The Sound Analysis & Synthesis team carries out research and development activity in sound analysis, transformation, and synthesis of sound signals. The analysis of sounds includes methods enabling the permanent extraction or automatic structuring of diverse sorts of information given off by the signal, such as the fundamental frequency or the spectral evolution determining the pitch and timbre of a perceived sound.

Information outside what is strictly musical is also taken into consideration, notably concerning industrial acoustics, sound design and multimedia as well as the automatic indexing of recorded sounds. The methods used are based on signal processing, statistical analysis, information theory, machine learning and recognition techniques, but also on knowledge of auditory perception.

Techniques of transformation and synthesis are initially created to respond to the needs of musicians in the production of new sounds and new music. A typical example of this is the computer synthesis of a virtual choir for an opera. However, the work is frequently applied to the fields of mobile telecommunications, video games, and virtual reality.

Analysis and synthesis is based on the conception of both signal models (the modeling of sound effects produced via their signals) and physical models (the acoustical modeling of the source of sound production). These models are integrated into pieces of computer software for PC or Macintosh (Mac OS X, Windows, or Linux) equipped with graphical interfaces specifically designed for musicians, but adaptable for sound engineers, acousticians and amateurs.

**SPECIALIST AREAS**

Signal processing, statistics, information theory, learning techniques, form recognition.

**MAJOR THEMES**

- Signal Models (additive sound analysis and synthesis, phase vocoder project, page 28, Sample Orchestration projects 1 & 2, p. 55, 29, Physic, p. 29) A. Roebel
- Sound Characterization (automatic indexing, projects ABC-DJ, Bee Music, SkAT-VG, MiréS, Quaero, 3DTV, SemanticHIFI, Ecoute, p. 33, 34, 35, 61) G. Peeters
- Analysis, Transformation, and Synthesis of Voice and Speech (Chantier, Voice4Games, Rhapsodie, Angel Studio, Affective Avatars, Vivos, p. 31, 32, 74) A. Roebel, N. Obin
- Analysis of Sound Scenes (Automatic Transcription, 3DTV, Houle, p. 35, 36) A. Roebel
- Robotized experimental platforms (p. 36)
- Software: Diphone Studio, SuperVP (p. 74), AudioSculpt (p. 75), IRCAMLab TS (p. 76), SnailAnalyser-Tuning (p. 76), A. Roebel, F. Cornu C. Picasso

**COLLABORATIONS**

- Aalto University (Finland), Ableton (Germany), Acapela Group, Acoustic Research Institute (Austria), AudioGaming, Arte, Artipolis (Luxembourg), Dualo, École des Mines de Paris, Exalead, ExeQuo, France Télécom R&D / Orange, Flying Eye (Germany), Game Audio Factory, Genesis SA, HHI Berlin (Germany), Idol, INESC, ISAE-Sup’Aéro, ISIR/UPMC, IUAV (University of Venice, Italy), Kantar Media, KTH (University of Stockholm, Sweden), LIMSI, LMA Marseille, LIA (Université d’Avignon), LPL (Laboratoire Parole et Langage) - Hôpital La Conception, NuHag (University Of Vienna, Austria), OFAI, ParisTech, OMUL, Queen Mary University (Great Britain), SCREAM National Cheng Kung University (Taiwan), SmartLog, Smartsound, Stupeflix, Ubisoft,UPF/MTG (Spain), University of Alberta (Canada), University of Huddersfield (Great Britain), Univ. Paris 8, Univers Sons (UVI), Viddiga, Vizion’R, Voxygen, Xtranormal (Canada), Yacast

**TEAM**

SOUND SYSTEMS AND SIGNALS: AUDIO/ACOUSTICS, INSTRUMENTS (S3AM)

Head Researcher: Thomas Hélie

ACTIVITIES

The team Sound Systems and Signals: Audio/Acoustics, Instruments creates theoretical, technological, and experimental tools based on multiphysic systems and the sound signals they produce. The team is interested in exploring, understanding, reproducing, and inventing sound objects in audio, acoustics, focusing on musical instruments and the production of the human voice.

More precisely, the objectives are to model, simulate, identify, and optimize these systems (voice, musician/instrument, loudspeakers, audio electronic effects, etc.) using physics to reveal and benefit from their intrinsic structures. The team creates methods for analysis, transformation, control, and simulation of sounds as well as tools to assist conception in virtual, real, or hybrid paradigms.

This global approach, “systems and signals” creates a synergy in the conception of these tools, connecting—not merely juxtaposing—several disciplines and scientific fields: physics, theory of systems and control, differential geometry, numerical analysis, signal processing, computer science, electronics, mechatronics, and robotics.

Target applications concern scientific, artistic, educational, and healthcare domains.

INTEREST AREAS AND ASSOCIATED PROJECTS

- Physical modeling of musical instruments and voice
- Sound synthesis based on physical modeling
- Experimental robotic platforms: robotic artificial mouth to play brass instruments, robotic vocal apparatus at a 1:1 scale
- Identification of non-linear systems
- Control of non-linear systems
- iMuSciCA project (p. 26)
- UMN project (Unfold Mechanics Network)
- INFIDHEM project
- Instrumental acoustics (projects: Cagima, PAFI, Augmented Instruments, Imarev, Smart Instruments, p. 26, 27)
- Software: Modalys, p. 77; The Snail, p. 76

SPECIALIST AREAS

- Acoustics, mechanics, non-linear systems, control, signal theory, differential geometry, digital analysis, experimentation, mechatronics, real and virtual instrument-making, sound synthesis

COLLABORATIONS

- Athena-RIC (Greece), Cabrilog SAS (France), Cambridge University (UK), C2RMF-Louvre (Paris), Centre Bernoulli-EPFL (Switzerland), EPCC-Edinburgh Univ. (UK), GIPSA-lab (Grenoble), IJLRA-UPMC (Paris), IMJPRG-UPMC (Paris), Imperial College London (UK), ISAE-Sup’Aéro (Toulouse), LAGEP-université Lyon-1, LaSiE-université de la La Rochelle, LEOPOLY (Hungary), LMA-CNRS (Marseille), LMD-ENS (Paris), Mines ParisTech (Paris), NTNU (Trondheim, Norway), Thalès Group (France)

TEAM

- J. Bensoam, D. Bouvier (doctoral student), R. Caussé, D. Chalabi (apprenticeship contract, CNRS), T. Hélie (CNRS), M. Jossic (doctoral student), T. Lebrun (doctoral student), R. Muller (doctoral student), R. Piéchaud, D. Roze (CNRS).
MUSICAL REPRESENTATIONS
Head Researcher: Gérard Assayag

ACTIVITIES

The Musical Representations team works on the formal structures of music and creative environments for composition and musical interaction. This work finds application in computer-assisted composition (CAC), performance, improvisation, performance and computational musicology. Reflection on the high-level representation of musical concepts and structures supported by original computer languages developed by the team, leads to the conception of models, which can be used for both musical production and analysis.

Since its widespread distribution to a large community of musicians, this software has provoked new ways of thinking, which exploit the computer possibilities for simultaneously representing and executing the final score, its formal underlying structures, and their algorithmic generators. It is important to note that a work elaborated in such a way contains, to a great extent, its own structural analysis.

For the past few years, the team has worked on symbolic interaction and artificial creativity in its work via projects on artificial listening, synchronization of musical signals and score following (in the INRIA team-project MuTant, for example), as well as in the engineering of intelligent agents capable of listening, learning, and interacting musically in improvisation situations.

On the musicology side, the representation and modeling tools enable a truly experimental approach, injecting considerable dynamism to this discipline. The team has a long history of collaborations with composers and musicians both from IRCAM and elsewhere. Archives of this work can be found in three volumes of the OM Composer’s Book, guaranteeing its international dissemination and continuity.

INTEREST AREAS

- Computer-assisted composition (assisted composition, orchestration assistance, p. 55, 56)
- Control of synthesis and spatialization, creative systems to write for time, sound, space, and interaction (p. 57)
- Mathematics and music, p. 60
- Computer languages for music, p. 55
- Modeling style, dynamics of improvised interaction (improvised musical interactions, DYC12, p. 51)
- New interfaces for composers and teachers
- Musicology and computational analysis, p. 52, 53
- Efficient search of temporal series (p. 36)
- Writing timed interactions ans musical synchronizations (p. 52)
- EfficacE, Inedit (p. 53)
- Computer-assisted composition (p. 54)
- Writing for sound and space (p. 57)
- Augmented paper for CAC (p. 58)

SPECIALIST AREAS

- Computer-assisted composition and analysis, computer musicology, cognitive musicology, artificial intelligence, computer languages, algebraic and geometric methods, symbolic interactions, languages for synchronous time

COLLABORATIONS

- Bergen Center for Electronic Arts (Norway), CIRMMT/McGill University (Canada), City University London, CNSMDP, Columbia New York, CNMAT/UC Berkeley, Electronic Music Foundation, Gmem, Grame Lyon, École normale supérieure Paris, EsMuC Barcelona, Harvard University, Inria, IREMus - Sorbonne Paris 4, Jylvaskyla University, University of Bologne, USC Los Angeles, Université Marc Bloch Strasbourg, Pontificad Javeriana Cali, Université Paris-Sud Orsay, University of Pisa, UPMC Paris, UCSD San Diego, Yale, University of Minnesota, University of Washington

TEAM


Guest Researchers: M. Ayari (Université de Strasbourg), G. Bloch, C. Rueda
ANALYSIS OF MUSICAL PRACTICES

Head Researcher: Nicolas Donin

ACTIVITIES

The Analysis of Musical Practices team carries out research on the works and scholarly practices that form IRCAM’s heart: composition and performance. To this end, the team is developing an empirical and technological musicology, whose goal is to describe (or even to modify) knowledge outside of the realm of music analysis, through novel methods of data collection. The goal is to characterize the phenomena under examination without reducing them a priori to their apparent structure (formalizable procedures, canonical texts), but in addressing them in all their dynamic, creative, contextualized, social, and cultural complexities. This can be applied to the study of contemporary activities and situations (e.g., a performer’s work at home, in rehearsal, in concert, a composition using computer-music technology, the production of a multimedia musical analysis for instructive purposes, etc.) or to the study of past practices. The various projects carried out offer a wide range of technological and social interdisciplinary configurations, depending on the study’s focus. Thus, a study of the compositional activity carried out by the composer Philippe Leroux from 2001 to 2006 demanded a connection be made between the artist’s manuscripts and cognitive anthropology: it coincided and interfered with the elaboration of a new work by the composer. Finally, it takes place in relationship with the development of a hypermedia document that traces the creative process and the way in which the composer listens to his own piece. Or again: the study of musical analysis in France (end of the 19th century, beginning of the 20th century) required the creation of a relationship between a micro-historic investigation into writing conditions as well as the reception of these texts, with a critical appraisal of the segmentation and categorization procedures of the 19th and 20th centuries and led to the creation of a computer tool for musical analysis. These activities are disseminated through publications and symposiums within the scientific communities concerned (primarily historic musicology, musical analysis, cognitive ergonomics, sociology, genetic criticism, and music technology); they are also made public in other ways: the creation of computer tools and hypermedia documents, short documentary films, teaching, and conferences.

INTEREST AREAS

- Composition Analysis (MuTeC project, p. 62, Gemme p. 60)
- Performance Analysis
- Contemporary Listening Practices
- Listening Practices and Musical Analysis: A Historical Approach
- Musicology and Humanities / synthesis work (Theories of Musical Composition in the 20th Century, p. 60)
- Multimedia Publishing Tools for Musicology, p. 61, Wave project, p. 62
- Projet Wasabi, p. 33

COLLABORATIONS

- Centre de Recherches sur les Arts et le Langage (EHESS, CNRS), CEAC/Université de Lille-3, Conservatoire national supérieur de musique et de danse de Paris, Haute École de Musique/Conservatoire de Genève, Institut des textes et manuscrits modernes (CNRS-ENS Ulm), OICCM/Université de Montréal

TEAM

- B. Bacot (doctoral student), O. Baudry (doctoral student), C. Cannone (CNRS), N. Donin, L. Feneyrou (CNRS), G. Pellerin
SOUND MUSIC MOVEMENT INTERACTION

Head Researcher: Frédéric Bevilacqua

ACTIVITIES

The Sound Music Movement Interaction team (previously known as the Real-Time Musical Interactions team) carries out research and development on interactive systems dedicated to music and performances. Our work relates to all aspects of the interactive process, including the capture and multimodal analysis of the gestures and sounds created by musicians, tools for the synchronization and management of interaction, as well as techniques for real-time synthesis and sound processing. These research projects and their associated computer developments are generally carried out within the framework of interdisciplinary projects that include scientists, artists, teachers, and designers and find applications in creative projects, music education, or in digital audio industrial fields.

INTEREST AREAS

- Modeling and Analysis of Sounds and Gestures
  This theme covers the theoretical developments concerning the analysis of the sound and gesture flow, or more generally, multi-modal temporal morphologies. This research concerns diverse techniques for audio analysis, the study of the gestures of performing musicians or dancers (see gesture analysis and recognition, p. 44).

- Technologies for Multimodal Interaction
  This theme concerns our tools for analysis and multimodal recognition of movements and sound; tools for synchronization (gesture following, for example) and visualization.

- Interactive Sound Synthesis and Processing
  This focuses essentially on synthesis and sound processing methods based on recorded sounds or large sound bodies (concatenative synthesis by corpus, p. 28, SDIF Format, p. 28).

- Systems for Gesture Capture and Augmented Instruments
  This theme focuses on the developments the team has made in terms of gestural interfaces and augmented instruments for music and performances (augmented instruments, p. 45).

KEY PROJECTS

- Legos (p. 47)
- COSIMA (p. 45)
- SkAT-VG (p. 65)
- Rapid-Mix (p. 45)
- MusicBricks (p. 46)
- Interlude (p. 48)
- Topophonie (p. 30)
- Same (p. 49)
- Sample Orchestrator 1 and 2 (p. 55 & 29)
- Wave (p. 62)
- ISMES (Labex SMART)
- SNESE (Labex SMART)

SPECIALIST AREAS

Interactivity, real-time computer science, human-computer interaction, signal processing, motion capture, modeling sound and gesture, statistical modeling and automatic learning, real-time sound analysis and synthesis

COLLABORATIONS

- Goldsmiths University of London (United Kingdom), Phonotonic (Paris), McGill University (Canada), CNMAT Berkeley (USA), Genoa University (Italy), ENSCI (Paris), LIMSI (Université Paris-Sud), ENSAD (Paris), GRAME (Lyon), iCK Amsterdam, Atelier des Feuillantines (Paris), Association Puce Musers (Rungis), Univers Sons/Ultimate Sound Bank (Paris), Cycling’74 (USA), Da Fact (Paris), Voxier (Paris), No Design (Paris), HKU (Hilversum), UserStudio (Paris), BEK (Norway), IEM (Austria), STEIM (Amsterdam), PSA (France), University Pompe Fabra (Spain), Nokia Research Center (Finland), TKK (Finland), KTH (Sweden), UCL (United Kingdom), UCSD (USA)

TEAM

- P. Antoniadis (doctoral student), F. Bevilacqua, R. Borghesi, B. Caramieux, M. A. Suarez-Cifuentes (doctoral student), B. Matuszewski, J. Larralde, K. Sanlaville (doctoral student), D. Schwarz, N. Schnell, H. Scurto (doctoral student), Michele Tarabella, Y. Zeitoun (doctoral student).
3

MAJOR RESEARCH PROJECTS
SOUND SYNTHESIS AND PROCESSING

PHYSICAL MODELING

IMUSICA

- Team Involved: S3AM
Funding: European Commission Call Ha2020 ICT-22-2016
Calendar: January 2017 — June 2019
Partners: Athena Research Center (coordinator, Greece), UC Limburg (Belgium), Ellinogermaniki Agogi (Greece), 3D for All Kft (Hungary), Cabrilog (France), Mathematics for More (Spain), University of Fribourg (Switzerland).

iMuSciCA is an educational project that intends to improve the acquisition of scientific skills in secondary-level education. iMuSciCA offers new learning methods using innovative “discovery” technologies—both personalized and collaborative—in order to make certain scientific subjects, in particular math and physics, more appealing. IRCAM brings its skills in virtual instrument making via Modalys, giving life to 3D instruments created by students. Some of these instruments feature parameters that can be tested and modified (geometry, material, etc.) in real-time. The iMuSciCA environment, implemented through a collaboration among different industrial and educational partners, will be tested and assessed during a pilot program with a large number of students and teachers in three European countries: Belgium, France, and Greece.

AUGMENTED INSTRUMENT-MAKING

- Team Involved: S3AM

This project is at the heart of the research on musical instruments. Prototypes of musical instruments, or components of musical instruments, that can be used in concert have been produced in this project. Work carried out includes a broadening of the range of mutes for brass instruments, studying a system for the automatic tuning of a kettledrum, the creation of a clarinet mouthpiece with a variable volume, and the creation of a bow and mouthpiece instrumented via several sensors. In addition, IRCAM has designed and owns an instrumentarium of SmartInstruments (guitars, string quartet, bass clarinet, brass mutes) equipped with sensors and actuators with acoustic properties that a musician or composer can modify, and can diffuse sounds from a variety of sources without an external loudspeaker. This led to the realization of COALA, an embedded low-latency active control system for SmartInstruments.

CAGIMA

Global Acoustic Conception of Reed Musical Instruments, In Tune and Homogeneous
http://cagima.ircam.fr

- Team Involved: S3AM, Sound Analysis & Synthesis
Funding: ANR - programme blanc
Calendar: December 2011 — November 2015
Partners: Laboratoire de mécanique et d’acoustique de Marseille, Laboratoire des sciences, des procédés et des matériaux de l’Université Paris 13, and Buffet Crampon
Coordinator: LMA Marseille-CNRS

The Cagima project focused on flaws found in reed instruments in tuning, homogeneity of emitted sounds, and timbre from both the perspective of the musician and also that of the instrument-maker who endeavors to satisfy the specific demands of each musician. The project’s ambition was to replace the historical incremental approach used by instrument-makers with a rational and global approach that aimed to design new “logical” instruments that minimized the identified flaws. This was a large step forward for the methodology and technology of instrument making. To begin, the constraints of sound production that are imposed by an instrument on a musician were evaluated via the measurement of a specific gesture (e.g. pressure in the mouth, pressing on the reed with a lip, configuration of the vocal tract) and were correlated to the acoustic flaws of the instruments in order to suggest novel methodologies for the general design for the bore and the lateral holes of reed instruments. For this, the study of global criterion was at the heart of this project; their understanding led to the fabrication of playable prototypes of musical instruments.
IMAREV
Active Musical Instruments with Virtual Tuning

- Team Involved: S3AM
- Funding: ANR - programme retour post-doc
- Calendar: October 2011 — September 2014
- Partners: Institut d’Alembert – Université Pierre et Marie Curie, Open University (United Kingdom)
- Coordinator: IRCAM

During the past 50 years, the sciences and technologies of music have made it possible to create and control new sounds from synthesizers. However, hundreds of millions of acoustic instruments are still used worldwide. The interaction with digital keyboards and hearing sound via loudspeakers is poor compared to the subtleties produced by acoustic musical instruments. A family of innovative musical instruments, called hybrid instruments, has recently been developed. Their principle is based on the use of acoustic instruments controlled by feedback with processing of the sound synthesis, with the goal of extending the sound possibilities of instruments. The final sound is therefore hybrid; it comes from the overlapping of the acoustic or mechanic vibration and its digital processing. Hybrid instruments have an important advantage over synthesizers: the interface with the musician remains the acoustic instrument while including the possibilities of sound synthesis. However, hybrid instruments are not as widely sold as synthesizers; today there is no unified methodology or tools for the design and creation of active instruments.

The focus of this project was to develop the models, algorithms, and tools necessary for the creation of optimized and unified hybrid instruments. The models include virtual tuning, the parameters for which are deduced from the knowledge and tuning of instrument makers. This facilitates an intuitive control by musicians as well as simplify the complexity of the design of active instruments. In addition to the applications for musical creation, this project made it possible to create “tunable” instruments, making use of a digital process after construction. This project included an aspect of fundamental research with applications for musical creation, construction of musical instruments, and the domain of active control.

PAFI
Modular Platform for Assisted Instrument Construction

- Teams Involved: S3AM, Mechanics Workshop
- Funding: ANR - programme Contenus et interactions
- Calendar: December 2008—May 2013
- Partners: Laboratoire d’acoustique de l’Université du Maine (LAUM, Le Mans), Télécom Paris, ESEO Angers, ITEMM Le Mans
- Coordinator: LAUM Le Mans

The PAFI project instigated an ambitious and completely new dynamic among research laboratories, a national pole of innovation for the music industry, and a group of instrument makers acting on behalf of professional associations of instrument makers, representative of the French small business culture.

The goal of this project was to overcome the difficulties associated with the reproduction and optimization of high-quality musical instrument design, distinctive of French instrument making. PAFI aimed to implement tools for characterization as well as mechanical and acoustic prediction for the analysis and creation of prototypes of virtual instruments.

During this project, the IRCAM team implemented an experimental characterization system for the accuracy of instruments via a system that measures the input impedance, the development of a tool that predicts accuracy, and as part of Pauline Eveno’s doctoral thesis, the installation of a methodology for control and innovation with instrument-makers using tools based on a model developed by the other project partners (tests such as artificial mouths and simulations via physical models of instruments being played).
**Signal Models**

**Processing by Phase Vocoder**

Team Involved: Sound Analysis & Synthesis

The phase vocoder, one of the most effective techniques for the analysis and transformation of sounds, represents the foundation of the SuperVP software program. With the phase vocoder, it is possible to transpose, stretch, or shorten sounds; it is possible to apply a practically limitless number of filters to sounds. By the same token, the level of sound quality of the transformed signals is extremely high when applied to speech. Numerous improvements and extensions have been introduced, for example:

- Reassigned spectrum
- Estimation of the spectral envelope via ‘true envelope’ transposition with the preservation of the spectral envelope transposition with the ‘shape invariant’ model
- Generalized cross synthesis enabling the synthesis of hybrid sounds
- Several methods for estimating the fundamental frequency (pitch) of a signal
- Classification by nature of the spectral, sinusoidal (voiced) or non-sinusoidal (non-voiced sounds or noises) peaks segmentation of the time/frequency zones into transitory and non-transitory regions and the increase or decrease of transitory sections
- Processing the sinusoidal, non-sinusoidal, and transitory time/frequency zones
- The LF model of a glottal source, making it possible to transform a voice, etc.

These different modules of analysis, synthesis, and processing are used in several software programs on the market today.

**Corpus-Based Concatenative Synthesis**

Team Involved: Sound Music Movement Interaction

Corpus-based concatenative synthesis uses a database of recorded sounds and a unit selection algorithm that chooses the segments from the database that best suit the musical sequence that we would like to synthesize by concatenation. The selection is based on the characteristics of the recording obtained through signal analysis and match, for example, the pitch, energy, or spectrum. The habitual methods for musical synthesis are based on a model of a sound signal, but it is very difficult to establish a model that conserves the entirety of the details and delicacy of the sound. However, concatenative synthesis—which uses real recordings—preserves these details.

Putting the new approach for concatenative synthesis by corpus in real-time in place enables an interactive exploration of a sound database and a granular composition that targets specific sound characteristics. It also makes it possible for composers and musicians to reach new sounds. This principle is carried out in the CataRT system. This system makes it possible to display a 2D projection of the descriptor space that can be browsed using a mouse or external controllers. Grains are then selected in the original recording and performed by geometric proximity, metronome, in loops, or continuously. It is also possible to define a perimeter around one’s present position that selects a sub-group of grains that are then played randomly. CataRT is used for musical composition, performance, and in various sound installations.

As this field of research is fairly young, several interesting research questions have been raised (or will be raised in the future) concerning the analysis and exploitation of the information found in the data of a corpus, the visualization, and real-time interaction.

![System of visualization used for sound synthesis by corpus](image-url)
**PHYSIS**

Physically informed and Semantically controllable Interactive Sound Synthesis

Teams Involved: Sound Analysis & Synthesis, Sound Music Movement Interaction
Funding: ANR — programme Contenus et interactions
Calendar: April 2012—April 2015
Partners: Audiogaming, LMA, Game Audio Factory
Coordinator: Audiogaming

Physis was an industrial research project centered on modeling, transforming, and a synthesizing diegetical sounds for interactive virtual worlds (video games, simulators, serious games) and for augmented reality. By diegetical sounds, we mean sounds created by identifiable objects in a virtual scene such as sounds made by weapons, liquids, fire, water, or fabrics, and their possible interactions including physical impacts, rubbing, sliding, and rolling. Physis considered the production of interactive sound contents in its entirety and in a systematic manner. It covered all steps necessary to meet the challenge: from fundamental research essential for the proper modeling of sound phenomena, to the production of portable code for game consoles or specific material. Physis also anticipated the creation of tools for sound designers, for the analysis and transformation of sound files, the creation of high-level semantic and physical controls as well as their implementation in an interactive situation.

The major findings of this project led to:

- Significant advances in modeling acoustic properties of sounds in a target corpus
- Synthesis models and new strategies intended for the creation and transformation of sounds interactively with semantic and/or physical controls
- Technological demonstrators that showcase these innovations.

The recent surfacing of complex video games and virtual universes like “Second Life” made apparent the limits of existing sound engines that use pre-recorded sounds with an image that is computed in real-time. Moreover, while the use of physical models enables a more realistic, more complex, and more varied graphical behavior, they have only a slight impact on the sound behaviors based on pre-recorded files. Improvements in computer materials have enabled a precise simulation of the audio and acoustic properties of everyday noises based on physical parameters.

In addition, new interfaces now included in smartphones, tablets, and video game consoles are changing the way we access digital information. These systems are underused from a sound point of view for the reason that reading pre-recorded sounds prevents correct interaction with this type of interface. Audio synthesis in real-time is perfectly adapted to these new interfaces and to the uses they imply. In conclusion, PHYSIS made it possible to create a solid scientific and technological foundation that replied to these new needs.

**SAMPLE ORCHESTRATOR 2**

Funding: ANR — programme Contenus et interactions
Calendar: November 2010 — October 2013
Partners: Univers sons, Conservatoire national supérieur de musique et de danse de Paris
Coordinator: IRCAM

This follow-up to the Sample Orchestrator project aimed to create innovative functions intended for a new generation of software samplers, in accordance with three complementary aspects:

- Hybrid techniques for real-time sound synthesis, intermediaries between models of signals designed for all types of sounds and samples, and taking advantage of variations made possible by the definition of parameters of the former and the effectiveness of the latter. The goal is to have specific models for each family of instruments.
- Hybrid techniques for real-time sound spatialization, intermediaries between the parametric models such as IRCAM’s Spat and the convolution of signals via measured concert hall impulsion responses. Here again, the issue is to create models that represent the best compromise in terms of variability and quality on one hand, and the efficiency of calculations on the other.
- Real-time methods for voicing and orchestration, providing an extension of the musical performance in function of orchestration methods modeled and/or learned in different musical corpora.

The project’s objectives lay beyond the then-current state of research and the project, via the synergy it created among different teams at IRCAM and with external partners, made important advances the domains of research and musical technologies including signal models of instruments, new techniques for spatialization and analysis/synthesis of sound fields, generative synthesis from a musical corpus.
Topophonies are virtual navigable sound spaces, composed of sounding or audio-graphic objects. Graphic and sounding shapes or objects are audio-graphic when visual and audio modalities are synchronized.

In virtual reality and videogames, we know how to make scenes composed of point-shaped elements: graphic and sound (i.e. a spot representing an object). However, there is no tool enabling navigation to make scenes consisting of a very great number of interactive visual and sound elements, nor of dispersed elements such as a crowd, a flow of traffic, foliage, or rain.

The research project Topophonie proposed lines of research and innovative developments for sound and visual navigation in spaces composed of multiple and disseminated sound and visual elements. By working in a multidisciplinary scientific group (digital audio, visualization, sound design) with companies specialized in the domain of interactive multimedia activities, the project Topophonie conceived and developed models, interfaces and audio-graphic renderings of groups of granular animated and spatialized objects. The project team was composed of researchers specialized in granular sound renderings and in advanced interactive graphical renderings, as well as digital designers and enterprises specialized in the relevant fields of application. The completed works produced interfaces that define and control multimedia scenes and tools for real-time rendering on synchronized audio and visual channels. The first production was an augmented audio reality installation, Topophonie Mobile, during the Futur en Seine festival. This installation augmented a position-determined walk in a public park with an acoustic ambiance on the theme of water via the Navidium application for interactive audio-graphical digital maps. Topophonie Mobile, won the Grand Prix de l’Innovation from the City of Paris in 2013 (mention special design). The other strong points of the research carried out in this project were the conception of a new method of synthesizing sound textures, and the dissemination of knowledge via two international workshops on audio-graphical modeling organized as a part of the project in 2011.
The objective of the Voice4Games project was to provide innovative solutions in the domain of voice processing for the production and post-production of audio contents for video games.

The Voice4Games consortium was created to reply to the demands for the analysis, recognition, and transformation of speaking and singing voices for the video game industry.

The innovative solutions that were developed during the project expanded the use of voice processing techniques (voice recognition and transformation) to the production and post-production of audio content for video games.

The project’s main innovations:

- Homogenizing the audio quality of multi-speakers and multi-language voices
- Voice similarity for multi-language voice casting
- Integration of innovative audio/video technologies (Kinect) for real-time voice interactions (vocal control, singing voice)

The role of IRCAM in the Voice4Games project was to manage the research on the analysis, recognition, and transformation of voices with the partners ExeQuo, Voxler, and Cyanide for the extension of voice processing applications for video game production and post-production.
ANGEL STUDIO
Generator of Personalized Avatars

Team Involved: Sound Analysis & Synthesis
Funding: Feder-drire, Cap Digital, Île-de-France
Calendar: March 2009 — June 2013
Partners: As An Angel, SoBuzzy, Télécom SudParis

The development of an avatar generator that uses photos and voice samples was the central concept in the Angel Studio Project. The photos and voice samples for a specific user are used to create a realistic 3D avatar face, with a synthesized voice with parametric possibilities. The scientific objective of the project was to be able to reproduce the facial and vocal movements of a user with a level of simulation acceptable for users. The commercial objective of the project was to include this generator in solutions developed by two small businesses involved in the project: As An Angel and SoBuzzy. The development of the conversational agent market, as well as that of Web3D, has come up against the challenge of what end users accept. Numerous projects and studies, especially those carried out by As An Angel (e.g. Project Agent Conversationnel Expressif), the Laboratoire Informatique de Grenoble (ex CLIPS), and the HUMAINE network have demonstrated this level of acceptance that can vary from one culture to another but remains relatively high. User expectations—modeled by writing, cinema, and science fiction—correspond to a plausible human imitation, both in terms of the capacity to converse and to the representation and animation of an artificial face and body. The Angel Studio project therefore aimed at creating conversational agents or avatars that meet users’ standards for two means of communication:

• Expressive facial animation of “photo-realistic” faces
• Plausible simulation of an expressive human voice

IRCAM developed and provided the conversion algorithms for the identity of the voice, the transformation algorithms for the expressivity of the voice, sentences synthesized from text (TTS) by a system that is on the market, or by a propriety IRCAM system (IRCAM TTS).

RHAPSODIE
A Prosodic Reference Corpus of Spoken French

Team Involved: Sound Analysis & Synthesis
Funding: ANR - SHS
Calendar: January 2008 — June 2012
Partners: Université Paris X UMR 7114, Lattice UMR 8094, CLLE Université Toulouse 2 UMR 5263, and LPL Université de Provence UMR 6057 Laboratories

The intention of the Rhapsodie project was to create a prosodic reference corpus of spoken French, sampled from a broad array of discursive styles of speech. In addition to prosodic annotations, the corpus contains syntactic and informative annotations that can be used to analyze the relationships between syntax and the status of prosody in the speech (e.g. placing of the communicative structure, conducting speaking in turns). The issue of standardization of annotation is therefore central to the project. The complementary objectives of the project’s partners can be found in 7 domains:

• Perfecting formants for the annotation and reading of intonosyntactic data to promote interoperability and exchanges.
• Placing resources of spoken French (30 hours of speech) at the disposal of the scientific community. Twenty percent of these resources have been annotated in terms of prosody (following the Text Encoding Initiative transcription guidelines, formant xML).
• Distributing the tools online to process and analyze these resources with comprehensive users’ manuals.
• Developing reasoned and explicit linguistic methods for the interpretation and generation of structures
• Enriching the intonosyntactical models of spoken French.
• Contributing to the improvement of prosody in speech synthesis.
• Organizing an international colloquium at the end of the project on prosody-speech interfaces to present the project’s results and compare them with work carried out by internationally known research teams.
INDEXING AND SEARCH ENGINES

AUTOMATIC MUSIC INDEXING

Team Involved: Sound Analysis & Synthesis

During projects presented hereinafter, the following subjects were addressed:

• Methods for the automatic extraction of musical descriptors for a piece of music such as the tempo, location of beats, metrical, tonality, or a temporal grouping for a chord. These descriptors facilitate the automatic classification of a piece and can be used for content-based searches in sound databases.

• Musical excerpt recognition methods, designed to automatically identify excerpts from pieces of music using reference databases. These methods are based on a compact sound signature (fingerprint) encoding the essential information. These algorithms compare each fragment of sound under investigation, with those in the database.

• Methods for the estimation of the temporal structure of a piece of music in terms of the repetition of a section being listened to and enabling browsing within the temporal structure of the given musical piece.

• Methods for the automatic creation of audio summaries making it possible to quickly pre-listen to the contents of a given musical piece via its key points.

WASABI

Web Audio Semantic Aggregated in the Browser for Indexation

Teams Involved: Analysis of Musical Practices, Sound Analysis & Synthesis, IRCAM Resource Center

Funding: Agence nationale de la recherche, programme générique

Calendar: October 2016 — March 2021

Partners: 135 (coordinator), Parisson, Deezer, Radio France

The goal of the WASABI project is to define an optimized methodology for indexing music for the Web for large databases by linking metadata from audio analysis, Semantic Web techniques, and the analysis of textual data such as song lyrics using natural language analysis and to compare this methodology with case scenarios to develop unique services and applications using Web Audio technologies.

The project entails using algorithms to extract musical information and Semantic Web techniques to produce more consistent musical knowledge bases for streaming services and music databases. Services using Web Semantic data like LastFM, MusicBrainz, or DBPedia use the extraction of structural data, connecting works to metadata such as the producer, the recording studio name, the composer, the release year, or the subjects in the lyrics for example. The data in free text like the lyrics are also analyzed to determine the musical context of the piece. Web Audio technologies make it possible to explore these musical spaces improved with analyses such as high-level musical indexation: detecting emotion and plagiarism, detecting and characterizing the singing voice, detecting the structure and separating the different sources.

Open source software bricks and “open data” online services will be proposed at the end of the project for:

• The visualization of audio metadata and listening to unmixed tracks in a browser as well as using the latest Web Audio API technologies (mixing in real-time, audio effects)

• Automatic processing of lyrics, recognition and merging named entities, collaborative annotation and correction

• Access to a Web service with an API offering an environment in which to study musical similarities from audio and semantic analyses

These software bricks will be used in the development of formalized demonstrators with our partners and collaborators (journalists and composers), using the new Web Audio API standard making it possible to develop musical applications accessible to the general public via a Web browser.
ABC_DJ  
Artist-to-Business-to-Business-to-Consumer  
Audio Branding System  
abcdj.eu

Team Involved: Sound Analysis & Synthesis  
Funding: H2020 ICT19  
Calendar: January 2016 – December 2018  
Coordinator: Technische Universität Berlin

The ABC_DJ project seeks to provide advanced technology to European creative agencies in the field of audio branding—the selection and diffusion of recorded music in relationship with the characteristics of products and brands (boutiques, communication, etc.). The primary goal is to enable these agencies to produce high-level contents so they can successfully compete with major actors in the market. The second goal is to open the audio branding value chain to European artists and independent recording labels by offering them new means of monetizing their works, notably through in-store music. The range of music that can be used by client agencies and brands will be considerably extended.

The scientific and technical approach employed in this project brings together research in engineering knowledge, sound signal processing, user studies, and market research to assist the creative process and the automatic broadcasting of music.

In this project, IRCAM is particularly involved in the creation of new algorithms for the analysis of audio musical contents (auto tagging by genre, emotion, instrumentation, estimation of the tonality and tempo) as well as new tools for automatic mixing (measuring the audio quality, segmentation, a complete hierarchic analysis of the structure, intelligent audio summaries, separation of audio sources).

BEE MUSIC

Team Involved: Sound Analysis & Synthesis  
Funding: FNS – National Fund for Digital Society  
Calendar: January 2013—December 2015  
Partners: Kantar Media, LTU, SmartJog, Idol, Sinequa  
Coordinator: Kantar Media

BIPP stands for the Base de données Interprofessionnelle des Producteurs Phonographiques (lit. Inter-professional Phonographic Producers’ Database). Owned by the SNEP (Syndicat National de l’Édition Phonographique which represents the major producers) in partnership with the UPFI (Union des Producteurs phonographiques Français Indépendants, gathering independent producers), its management was entrusted to Kantar Media in 2008. Its vocation is to reference all active digital and physical catalogues active in the French market. BIPP has currently referenced 2.5 million titles. The objective of this project that associates Kantar Media (under the aegis of the SNEP and the UPFI) with a consortium of academic and private partners is to provide the BIPP database with all the qualities necessary for it to be validated by the music industry:

- Exhaustiveness and homogeneity
- Univocal identification of the contents
- Wealth and automation of indexing
- A variety of search and recommendation methods, with an emphasis on musicological, visual, and semantic criteria in addition to the common contextual criteria
- Simple and in-depth interfaces that enable personalized exploitation of the data by targeted professional users

The only public laboratory associated with the project, IRCAM, is carrying out research in the domain of musical information retrieval, focusing on the automatic extraction of expected characteristics in pieces of music from the analysis of their recordings: selection of a representative excerpt, genre and mood, unique identifiers, etc.

Bee Music has a structuring nature; for the first time, research carried out over the past ten years by consortium members can be scaled up to fit within a truly operational database containing millions of tracks. In addition to the technical aspect, providing a rich musical database for educational purposes is a central idea of this project. Beyond the target of music industry professionals, the general public can also take advantage of the work carried out during Bee Music thanks to the wealth of information available online by the professionals that will create B2C services using BIPP.
**Quaero**

Team Involved: Sound Analysis & Synthesis  
Funding: Agence pour l’innovation industrielle (OSEO)  
Calendar: January 2008—December 2013  
Partners: Technicolor (coordinator), France Télécom, INA, Exalead, Yacast, Bertin, ITU, Synapse and Vesys, CNRS (INIST, LIMSI, IMMI), DGA, BNF, INRIA, Institut Télécom, IRIT, LIPN, MCG-INRIA, Université Joseph Fourier, University of Karlsruhe, University RWTH of Aachen

Quaero, a 200-million euro project, has been one of the largest industrial R&D projects in France. Quaero focuses on the design and production of new technological solutions enabling the extraction, analysis, classification, and use of information in digital multimedia and multi-lingual contents, concentrating on the automatic processing of speech, language, music, images, videos, and printed digitalized documents.

The Quaero consortium was created in a world where it is becoming easier and easier to have access to digital information via PCs, television, or handheld devices. The Quaero consortium aims to reply to new demands to analyze multimedia contents from the general public and professionals. The consortium makes it possible for the most up-to-date technologies developed in public and private laboratories in France and Germany to be used to their fullest potential by key industrial interests in this domain be they small, growing businesses or large industrial groups.

The developments made by the Quaero consortium contribute to the expansion of the services offered by portals, search engines, applications for interactive television, professional environments for the production or post-production of multimedia contents, and will facilitate uploading digital media online for libraries (e.g. books, films, television programs, etc.).

The Quaero project concentrates on five major application domains:

- Multimedia online searches
- Improvement of services that provide access to audiovisual contents via portals
- Personalized selection and display of video contents management of professional audiovisual resources digitalization and improvement of library contents, audiovisual
- Archives, and scientific publications

IRCAM will coordinate the research carried out by partner laboratories on musical indexing and act as a liaison with industrial partners such as Exalead and France Telecom in developing applications for multimedia indexing and in improving services to access audiovisual content.

**3DTVs**  
Study of techniques for description of 3D audiovisual contents

Team Involved: Sound Analysis & Synthesis  
Funding: European Commission, ICT, FP7  
Calendar: November 2011—January 2015  
Partners: The Aristotle University of Thessaloniki (coordinator), Fraunhofer HHI, Flying Eye, Velti, Sky-Deutschland, Arte

The 3DTVs project was dedicated to the study of description techniques for the contents, automatic indexing, searches, and browsing in three-dimensional audiovisual contents available on both mobile and stationary platforms. The objective of this project was to improve the performance of the analyses carried out on the three-dimensional formation of audio and video signals. The project also aims to expand audio/video indexing through the multimodality of the approaches. The role of IRCAM in this project focuses on the analysis methods for spatialized audio signals (multi-channel audio streams) in order to improve indexing and provide the localization of sound sources in three-dimensional space.

**MIReS**  
Roadmap for Music Information Research

Team Involved: Sound Analysis & Synthesis  
Funding: European Commission, ICT, FP7  
Calendar: October 2011—March 2013  
Partners: MTG, Stromatolite, OFAI, INCEC Porto, GMUL, BMAT

MIReS project is a Coordination and support action funded by the 7th Framework Programme of the European Commission. MIReS project aims to create a research roadmap for Music Information Research. The project studies the challenges associated with multimodal, multicultural, and multidisciplinary information.
EFFICIENT SEARCH OF TEMPORAL SERIES

Team Involved: Musical Representations

Searching for sounds can be a painful and tedious task when dealing with large-scale databases. Even when meta-information is available, query results are often far from the mental image imagined by the user. Today, there is no system that transforms the intuitive projection of a sound idea into an effective search; sound samples do not let users extract high-level information such as melody or lyrics from songs.

Beginning with this observation, we have developed a code that makes it possible to carry out effective searches on temporal forms and take into account the multidimensional nature of sound perception. This makes it possible to carry out searches based on the temporal form of the descriptors rather than on mean values. These descriptors are modeled to obtain their average, standard deviation as well as the form of their temporal evolution via a symbolic representation enabling both compact storage and an effective search.

However, it was essential that the comparison of temporal series that make it possible to obtain a similarity based on perceptive criteria for objects that could possibly be very different mathematically. Using an approach derived from Dynamic Time Warping (DTW), we have developed a robust measure of similarity following non-linear distortions such as range, noise sound, and unique values. Thanks to a new algorithm for indexing, it is possible to obtain the best element from a database containing several million sound samples almost immediately.

Our study then opened to the implementation of higher-level interactions. We studied the possibility of a query that is pertinent to several temporal curves simultaneously, going beyond the framework of the simple consideration of often less relevant criteria. Thanks to new heuristics, we have carried out the first precise multi-objective search algorithm for temporal series.

These techniques apply to all fields of scientific research due to the ubiquity of the temporal information. Multi-objective searches of temporal series are open to numerous applications in fields ranging from medical analysis to robotics.

This also enables the installation of a system of request by vocal imitation based on multiples of spectral descriptors. These advances have been implemented in an interface using iPad multi-touch technology.

ROUTE

Robot à l’écoute

Team Involved: Sound Analysis & Synthesis

Funding: Convergence de Sorbonne Universités – Emergence Programme 2014

Calendar: September 2015-September 2016

Partners: ISIR (Institut des systèmes intelligents et de robotique)

This project focuse on a common issue for two apparently distinct research domains:

- The automatic analysis of a sound scene, from a learning process using a dictionary. In audio signal processing, we find this aspect in several major applications for example computational auditory scene analysis (CASA), automatic indexing, source separation, detection and localization of sound objects.

- Artificial hearing, a recent field of study in robotics, for which the analysis of a sound stage gradually becomes a prerequisite for any modern application (e.g. monitoring the elderly, or studying human-robot interaction).

The central objective of the project is the design and development of a new method for the detection and localization of the main speaker in a sound scene. The method is intended to enable a robot to identify a vocal signal in the presence of noise and locate the main speaker’s position, in the case where there are several speakers. The problem is closely related to the current field of computational auditory scene analysis (CASA).

The objective of CASA is to design automatic systems with a perception that mimics human hearing, considering its physical and psycho-acoustical aspects. This project takes a different approach, while the audio processing tools are comparable (machine learning, source separation), hearing is treated from the robot’s point of view and the interest lies in the analysis of the audio scene.
MAJOR RESEARCH PROJECTS

INDEXING AND SEARCH ENGINES

HOULE
Hierarchical Object based Unsupervised Learning
houle.ircam.fr

Teams Involved: Sound Analysis & Synthesis, Sound Perception and Design, Musical Representations
Funding: ANR – programme Jeunes Chercheurs
Calendar: September 2011—January 2015

In the context of the HOULE project, IRCAM suggests innovative unsupervised learning techniques to solve the problem of the structuring of audio scenes (Computational Auditory Scene Analysis: CASA); a problem that consists of identifying and describing sound objects that make up a scene. CASA is a very active domain of sound processing with numerous applications.

The points that block the current methods are:

• Without being able to draw up a hypothesis on the nature of the processed objects, we can not model them
• The objects combined in the scene can not be observed isolated from each other
• The structure of the objects is governed by numerous relationships that is difficult to prioritize

The characteristics that we will use for our approach are the hierarchal organization of audio scenes (atoms brought together in objects that are examples of classes such as “Piano A4”, itself an example of “Piano note”) and the redundancy present at all levels of this hierarchy. This redundancy enables us to identify reoccurring motifs on which a rich and robust representation can be based.

This new method of representing audio scenes has led to the creation of an unsupervised learning algorithm designed expressly to process this data. The system features two components: the MLG (multi-level grouping) that structures the data and the Supervisor (a reflexive adaptation module) embodies the learning aspect by optimizing the MLG function on-the-fly in reference to the stocked memories of past executions.

The originality of our proposition lies in its distance from traditional CASA approaches, beginning with the paradigm of representing scenes and objects. Innovation is primarily present in the unsupervised learning methods that we will develop, including applications that go well beyond the CASA framework.

SEMANTIC HIFI
Browsing, Listening, Interacting, Performing, Sharing on Future Hifi Systems

Program: Information Society Technology (IST), FP6, European Commission
Calendar: December 2003—November 2006
Partners: Fraunhofer Institute (IDMT, Erlangen), Pompeu Fabra University (Barcelona), Sony Computer Science Labs- Paris (Sony CSL), Ben Gurion University-Israel, Native Instruments, Sony EUTEC
Coordinator: IRCAM

In the context of large-scale digital music distribution, the goal of the project was to develop a new generation of Hi-Fi systems, offering new functionality for browsing, interacting, rendering, personalizing and editing musical material.

This next generation of hard-disk based Hi-Fi systems will drastically change the home users’ relationship to music and multimedia content. They were able to interact with music, blurring the traditional limits between playing, performing and remixing. These Hi-Fi systems were as much open instruments as listening stations.

Main Functions Proposed:

• Personalized classification and content-based management of music pieces; query by humming; automated playlist
• Generation specified by global and content-based criteria
• High-quality home sound spatialization with automatic compensation of the acoustic listening context; 3D audio real-time navigation in the sound scene with assisted mixing, browsing within musical pieces,
• Personalized editing and composition tools, DJ application
• Instrumental and vocal tools and automatic accompaniment,
• Sharing of the indexing, composition, and performance work through P2P network
SOUND SPATIALIZATION

ANALYSIS & SYNTHESIS OF SOUND FIELDS

ANALYSIS OF A SOUND SCENE THROUGH SPHERICAL NETWORKS OF TRANSUCERS

Team Involved: Acoustic and Cognitive Spaces
Partners: Ben Gurion University, Supelec

The spatial properties of a sound scene are determinant for the perception of sounds in a room. These spatial properties are usually studied through separate investigations of either the network of microphones, or the network of loudspeakers. Measures that take into account both the microphone and the loudspeaker networks, called MIMO (multiple-input multiple-output), make it possible to augment the description of the sound scene. By using MIMO systems, the impulse responses of the room can be represented using a matrix and the analysis of the sound scene’s spatial properties can call upon linear algebra tools. For example, the rank of the matrix and its kernel reveal interesting information such as the number of significant reflections, the direction of their incidences on the microphone, and the direction of their emission from the loudspeakers. This project is the object of theoretical work of the formalization of the problem and the optimization of networks of spherical transducers and also includes an experimental aspect dedicated to the acoustic analysis of rooms.

WFS AND AMBISONIC SYSTEMS IN THE ESPACE DE PROJECTION

Team Involved: Acoustic and Cognitive Spaces
Funding: Programme Sesame (Conseil général d’Île-de-France), CNRS, IRCAM, ANR, UPMC

The Wave Field Synthesis (WFS) technique is type of holographic reproduction process that enables, by analogy with visual holograms, to capture or synthesize a sound scene while conserving the spatial characteristics of distance and direction. This approach, initiated by the University of Delft, goes well beyond the limits of conventional systems in terms of reproduction fidelity in a wide listening zone. While traditional stereophonic techniques (e.g. stereo, 5.1) are a kind of trompe-l’œil and can only be truly appreciated when one is positioned in the center of the system, the goal of holophonic technology is to reproduce a sound field in which listeners can move freely while maintaining a coherent perception of the localization of the sound sources. IRCAM has gained experience in this type of reproduction through its participation in the European CARROUSO Project in collaboration with the Swiss company SonicEmotion. These studies have made it possible to carry out productions in partnership with the Centre Pompidou such as the production of an installation at the end of 2005 for the DADA exhibit, an interactive installation in June 2006 (in collaboration with N+N Corsino), and an installation for the Samuel Beckett exhibit in 2007. From 2008–2011, IRCAM was able to acquire and install a WFS system in the Espace de Projection with the support of the Île-de-France Région, the CNRS, and the UPMC. This system is made up of a horizontal crown of 264 loudspeakers placed at regular intervals around the stage and the audience for WFS sound diffusion and of a dome of 75 loudspeakers for three-dimensional Ambisonic sound diffusion. This equipment is used to experiment with new methods of spatialization for musical creation and for scientific experiments with virtual reality and spatial cognition. The WFS component of this installation, inaugurated in 2008, was used in the musical creations by P. Schoeller in June 2009 (Operspective Händel), M. Jarrell in June 2010 (Le Père), M. Garcia-Vitoria in October 2011 (Mimesis), and in theatrical productions at the festival of Avignon (Shakespeare’s Richard II, directed by J. B. Sastre, July 2010). The final system was inaugurated on November 28th, 2012.

Acoustic measures at Grosses Festspielhaus in Salzburg combining a spherical microphone (64 capsules) developed at IRCAM and a spherical network of loudspeakers developed at RWTH (Aachen).
3DR3
Room Impulse Response Renderer
—
Team Involved: Acoustic and Cognitive Spaces, Creation-Diffusion Department

3DR3 brings together the Acoustic and Cognitive Spaces research team with the Creation-Diffusion department’s sound team to work on the development of a generic environment for multichannel mixing (see panoramix, p. 82), using a unique model of hybrid reverberation recently developed by the Acoustic and Cognitive Spaces team. This innovative model takes advantage of convolution reverberators that accurately reproduce the “natural” sound of the reverberation measured in a room compared to feedback delay network-based reverberators. This extends the principle of spatialized responses (see Analysis of a Sound Scene through Spherical Networks and Transducers) while associating the flexibility of Spat~ and its control via perceptive factors. The spatial disposition in 3D of the different sound sources exploits the room impulse responses of the rooms recorded in the same space as the concert, assuring the natural integration of the primary sound recording. The emblematic application of this development is the capture, mixing, and 3D diffusion of spatialized works.

ORPHEUS
Object Based Broadcasting
orpheus-audio.eu
—
Team Involved: Acoustic and Cognitive Spaces
Funding: H2020-ICT-2015
Funding Calendar: December 2015—May 2018
Partners: FhG IIS, BBC, Bayerischer Rundfunk, IRT, B-COM, Magix, Trinnov-Audio, ElephantCandy, Eurescom

Object-based media is a promising approach for creating and deploying interactive, personalized, scalable and immersive content, by representing it as a set of individual assets (video and/or audio) together with metadata describing their temporal and spatial relationships and associations. This allows media objects to be assembled in groundbreaking ways to create new user experiences, offering various levels of interaction: innovative ways for playing contents, spatial navigation, adaptation to listening contexts (enhancement of dialogues, mobility, etc.).

The ORPHEUS project aims to develop and validate an object-based audio production chain. IRCAM’s role in this project focuses on aspects of characterization as well as the synthesis and transmission of reverberation effects.
**BINAURAL REPRODUCTION TECHNOLOGY**

Team Involved: Acoustic and Cognitive Spaces

The realistic reproduction of sounds through headphones requires the use of binaural spatialization. This technique is based on a dynamic filtering of the sound source using transfer functions (HRTF - Head Related Transfer Functions) measured on the head of a listener or model. Despite the limited nature of its current application to music (except in virtual reality or interactive sound installations), binaural technology is a very important method of reproduction in the lab. Associated with systems that follow your position, this is currently the only technique that allows the complexity of a sound scene to be reproduced without artifact. This technique is particularly useful in perceptive validation work. At IRCAM it is used in the study of spatial auditory cognition including listener navigation and the perception-action process brought about by the motion control of a localized sound source.

**BILI**

bili-project.org

Team Involved: Acoustic and Cognitive Spaces

Funding: FUI

Calendar: January 2013—June 2016


Coordinator: France Télévisions

Personal devices (e.g. smart phones, tablets) now have several functions and represent one of the main means of accessing music, radio and television programs. As the public increases their use of these devices, binaural listening with headphones becomes more frequent and symbolizes the idea of a personal experience, providing access to a reproduction of 3D immersive sound scenes. However, the individual dependence head-related transfer function (HRTF) that is the basis of binaural technology has limited its availability to the general public. The BiLi project endeavors to provide solutions for individualized binaural listening via the support a network of research labs, small companies, and broadcasting companies (France Télévisions, Radio-France, Orange). The main focuses of the project are the assessment of the quality of the experience made possible by binaural listening, the research and development of solutions for individualizing the listening while avoiding tedious measurements in an anechoic chamber, and the definition of a format for sharing binaural data in anticipation of an international standard.

Virtual reality experiment in which the sound contents of the scene, recreated on headphones using the binaural technique that is dependant on the orientation of participant’s head and gestures. © CNRS Photothèque / F. Vrignaud
In the field of sound reproduction and communication, future audio technology will attempt to shift emphasis towards sensations of immersion and presence. Such notions are intimately linked to the spatial dimensions of a multimedia scene, and particularly in the field of sound, and are intensified in situations involving the participation of the listener. This participation may involve navigation within a scene or the gestural interaction of objects within it.

Under these conditions, made possible by binaural and holophonic technology, the congruence and real-time updating of auditory spatial clues in accordance with the listener’s movement or actions, have a major impact on the sensation of presence. Such a context led to the development of a set of experiments focusing on auditory spatial cognition, notably via the study of multi-sensorial integration processes, focusing on auditory and idiothetical modalities (clues induced by the subject’s movements including balance and proprioception).

Experimental methods call on behavioral experiences based on the observation of a subject’s performance, in terms of localization and navigation, when submitted to different exploratory contexts. We are also interested in the relationships between multisensory integration and emotions. We may study, for example, the effects of spatial conflicts between sound and vision on the emotional reaction of subjects, or assess the perception of numerosity (e.g. the quantification of a crowd) based on a sensory mode and its connection to emotion.

Our perception is based on the impression that our “self” is located in the same position as our bodies; in other words, the feeling of embodiment, of living in a body, of being located within the physical confines of a body. This feeling is not governed by an isolated sensory organ, but depends on multisensory afferences. We study this feeling by looking at the contribution of audio and vestibular interactions in the perception of our own body, using zero gravity conditions created during parabolic flights. In zero gravity vestibular indices are naturally troubled and this may modify the feeling of unity between the feeling of “self” and one’s own body. In other words, this may modify the perception of your body as a spatial reference.
VERVE

Team Involved: Acoustic and Cognitive Spaces
Funding: FP7 ICT Call 7 ICT-2011.5.5 ICT for Smart and Personalized Inclusion, European Commission
Calendar: October 2011—September 2014
Partners: Trinity College Dublin (Ireland), Centre hospitalier universitaire de Nice, INRIA Sophia-Antipolis, CNRS Télécom ParisTech (France), Testaluna (Italy), Kainos (United Kingdom), University of Zaragoza (Spain), and Deutsches Forschungszentrum für Künstliche Intelligenz (Germany)
Coordinator: Trinity College Dublin

The VERVE project aimed to improve the quality of life of people at risk of social exclusion due to the fear and apathy associated with ageing and neurological troubles. The consortium works on the specification, development, and testing personalized and populated virtual environments that can be used on different types of platforms from immersive rooms (CAVE) to smart phones. VERVE’s efforts focused on three situations: fear of falling and Parkinson’s disease; apathy related to cognitive decline and behavioral disturbances, in particular due to Alzheimer’s Disease; and other emotional disturbances linked to anxiety. Although focusing on these areas initially, it is expected that the results of the research will be applicable to a much wider range of situations.

3D sound is used in immersive situations to study the connections between visual-audio integration and emotion. The understanding of the emotional impact of visual and audio rendering of a virtual environment, from an emotional perception point of view, must enable the implementation of new therapeutic strategies for troubles linked to anxiety.

HC2

www.hcsquared.eu

Team Involved: Acoustic and Cognitive Spaces
Funding: European Commission, FP7, Programme Coordinating Communities, Plans and Actions in FET Proactive Initiatives
Calendar: October 2010—September 2013
Partners: Starlab, Catholic University of Milan, Johannes Kepler Universität Linz
Coordinator: Starlab

Human-computer confluence is an ambitious research program studying how the emerging symbiotic relation between humans and computing devices can enable radically new forms of sensing, perception, interaction, and understanding. This is a multidisciplinary research field, covering fields as diverse as cognitive sciences, fundamental computer science, and electronics. The HC2 project coordinated the promotion of identity building and defining future research and educational programs. HC2’s goals were reached by collecting updated information and by consolidating research agendas using a work methodology based on the active participation of the community and the creation of events (2 summer schools, 2 workshops, an encounter for the establishment of an ERA NET, an industry day).

REVA

Virtual Reality and Tinnitus

Team Involved: Acoustic and Cognitive Spaces
Funding: Tinnitus Research Initiative (TRI), Amplifon
Calendar: January 2008—June 2009
Partners: Université Paris Descartes, Hôpital Européen Georges Pompidou

This project centered on the design and testing of a form of therapy based on virtual reality to treat tinnitus – phantom auditory sensations that are felt in the absence of any external acoustic stimulation. The principle of the study was to place patients suffering from tinnitus in a 3D virtual reality situation (visual and audio) to favor brain plasticity and reduce the discomfort caused by the subjective tinnitus. In practice, a sound and visual avatar of the tinnitus was created and gave the patient the means to controlling its location; both its direction and distance.
GESTURE AND INTERACTION

GESTURAL INTERACTION

GESTURE ANALYSIS AND RECOGNITION

Teams Involved: Sound Music Movement Interaction, Creation and Production department

This research project replies to an increased interest for interactive musical systems based on gestural control. The applications concern not only music, but also performances such as dance or theater. The research carried out in the framework of this project is multi-disciplinary and includes engineering, physiology and biomechanics, cognitive sciences, and artistic domains. This work is carried out in synergy with the team’s developments on gestural interfaces.

The first style of research focuses on the study of instrumental gesture and its relationship with both musical writing and the characteristics of the sound signal. Measurements were taken on violinists and trumpet players during a collaboration with McGill University. In the case of string instruments, the movements of the bow, the violin, and the musician’s entire body can be measured using optical 3D measurement techniques.

The team has also developed systems for capture, compatible with use on stage or in educational settings. The ensemble of these methods makes it possible to measure and model musicians’ gestures. Diverse issues are also addressed in this study: motor control, learning in the case of the gesture of an expert, characterization of playing styles taking into account sound and gesture parameters, and the modeling of the phenomena connected to the gestural co-articulation similar to those of speech. A second axis concerns the development of systems for gesture analysis and recognition.

The team is favorable to general approaches that make it possible to apply their results to music, dance, or to interactive installations. Such systems could be included in the context of a live performance for the synchronization and interaction between the performers’ movements and a broad range of sound or visual processes. Recognition is based on a variety of temporal characteristics of movement, captured either by video, either by sensors attached to the performers’ bodies.

The tools developed during this research, such as “gesture follower”, make it possible to accurately recognize and characterize diverse high-level gestural elements defined by the artists.
In the past decade, we have seen a huge change in media and their relationship with bodies and space with tangible interfaces, augmented reality, and ambient Internet. The COSIMA project will explore the relationship between body, media, and space via new interfaces and tools for collaborative creation. The aim of COSIMA is to implement a platform for publishing and disseminating media situated in space and time, combining several sensory modalities. These contents partner proprioception and perception, involving the entire body in their creation and reception phases. The principle of augmented reality combines a digital space with tangible space. The use of this kind of space makes it possible to edit collaborative situated media. The COSIMA platform’s applications include artistic projects, innovative public services, events, and communication.

- Implementation of an Accessible Platform
  The principle objective of COSIMA is to democratize this new type of media via an accessible, generic, open, and interoperable platform. The project aims at the implementation of a platform that uses standard and open formats, enabling relationships with a variety of mobile technologies and existing systems. The platform will make it easier to carry out this type of experience at a lesser cost and in a more accessible manner for the general public (cultural institutions, public services, artists, etc.). Putting in place an environment of multimodal expression in tangible space requires a system made up of different components: client applications installed on mobile terminals connected to a network, a web service that offers an open API, a database server, and an administration interface for the contents.

- Create new Participative Experiences
  The creation of a platform also requires studying different uses and validation. For this, prototypes are developed at each stage of the project to test different scenarios and to collect and analyze the results to improve the platform. The COSIMA project therefore defines a context for experimentation according to different themes: the creation of visual and sound media associated with information about movement and movement to create maps or paths in augmented reality; situated media is shared in a social space for staging pervasive games or zones of public expression; dissemination of situated media in real-time to a community of users makes it possible to create mobs, performances, or large group experiences.

- Supporting New Economic Models and the Development of Communities for the Creation and Dissemination of Digital Media
  The combined development of Internet technology and mobile platforms defines new paradigms for the dissemination of media. The COSIMA platform favors the surfacing and structuring of communities around these new media. The assessment and selection of an adapted economic model will be a part of its development.

Rapid-Mix’s goal is to develop new modes of musical interaction by including multimodal data connected to movement and the body. From a methodology point of view, the project is based on user-centric evaluations and on the quick development of prototypes. Advanced techniques for signal analysis and automatic learning are also used. The transfer of these technologies to small business taking part in the project is an important aspect of the project. In this project, IRCAM is responsible for the development of application prototypes (agile prototyping) in addition to its implication in the other aspects of the project.
MUSICBRICKS
Musical Building Blocks for Digital Makers and Content Creators

Team Involved: Sound Music Movement Interaction
Funding: Program H2020-ICT18-2014-a
Calendar: January 2015—June 2016
Partners: Sigma-Orionis, Stromatolite Innovation Lab, Universitat Pompeu Fabra, Fraunhofer Institute IDMT, Technische Universität Wien
Coordinator: Sigma-Orionis

The goal of the European project Musicbricks is to facilitate the transfer of new musical technologies from major European research centers specialized in the domain to small digital creation companies. The project involves the creation of programming interfaces, of graphic and tangible user interfaces, the development of an ecosystem based on events during Music Tech Fest and in selected technology incubators to prepare access to the market. In this project, IRCAM is in charge of the diffusion of technologies connected to tangible and graphic user interfaces (TUI and GUI). A particular focus of the project is to develop analysis modules of gestural data and mapping between movements and sounds.

MIM
Enhancing Motion Interaction through Music Performance

Team Involved: Sound Music Movement Interaction
Funding: Program H2020-Marie Sklodowska-Curie Actions
Calendar: January 2016—December 2018
Partners: McGill University, Montreal Canada

This project focuses on Human-Computer interactions based on movement, leveraging a multidisciplinary approach between experimental psychology, music technology, and computational modeling. Initially, the project will look at sensorimotor learning mechanisms and expressive control in human movement. Computational models of these mechanisms will be developed based on experimental data gathered from the performers’ movements. Then, the models developed will be applied to the domain of Digital Musical Instruments (DMI), creating new types of instruments based on sensorimotor learning mechanisms. The project contributes to two fairly uncharted research areas. Firstly, it contributes to the fundamental understanding of sensorimotor learning processes by considering complex human motion such as the movements of musicians. Secondly, it represents the development and assessment of unique interactive musical systems using computational models of expressive musical gestures.
LEGOS
Sensorimotor learning in gesture-based interactive sound systems
http://legos.ircam.fr

Funding: ANR – programme blanc
Calendar: November 2010—March 2015
Partners: Laboratoire de Neurophysique et Physiologie (CNRS UMR 8199), Université Paris Descartes
Coordinator: IRCAM

The central idea of the LEGOS project was to foster interdisciplinary expertise in sonic gesture control technologies with neurosciences, especially in regards to sensori-motor learning. We believe that these aspects are not sufficiently taken into account in the development of interactive sound systems. A better understanding of the sensori-motor learning mechanisms of the gesture / sound coupling is necessary to provide efficient methodologies for their evaluation and optimization. Such advances would significantly expand the usability of today’s gesture-based interactive sound systems, often developed empirically.

The objective of the LEGOS project is to study systematically the coupling quality in gesture-sound systems using gestural interfaces. For this, we will evaluate the sensori-motor learning, and particularly its evolution over time, in various interactive devices. The aims are therefore to develop, evaluate, and compare interactive systems, with the mid-term goal of offering renewed paradigms for gestural interfaces control of digital media (including video games and sound design), as well as prototypes for medical applications such as rehabilitation.

The project will use extensively an experimental approach, considering these three perspectives:

• Sound Control: The first point corresponds to a case of sensori-motor learning where the goal is to produce a given sound through the manipulation of a gestural interface, as in the case of digital musical instruments. The sensori-motor learning is assessed in terms of the quality of the sound production.

• Learning Gesture with Audio Feedback: This second point corresponds to the sensori-motor learning where the goal is to make a gesture guided by an audio feedback. The sensori-motor learning in this case is assessed in terms of the gesture repeatability.

• Interactive Sound Design: This third point corresponds to sensori-motor learning in the case of tangible interfaces, where the goal is the proper handling of an object. The sensori-motor learning in this case is assessed through the quality of the objet manipulation for a given use.
INTERLUDE
New Digital Paradigms for Exploration and Interaction of Expressive Movement with Music
http://interlude.ircam.fr

Team Involved: Sound Music Movement Interaction
Funding: ANR - programme Contenus et interactions
Calendar: December 2008—July 2010

This project touched upon three major research topics in the Real-Time Musical Interaction team: paradigms for interactivity, multimodal modeling, and collaborative movement interfaces.

Recent developments in computer music have made it possible to analyze, process, and manipulate music and sound; both the audio signal and sound's symbolic representation (i.e. the score). These processing techniques and manipulations are generally carried out using traditional interfaces such as a keyboard or a mouse, perhaps using interfaces for mixing with potentiometers. Yet, the possibilities of interaction between gestures and larger body movements, as is the case with acoustic instruments with digital sound elements, still remains largely untapped.

The goal of the INTERLUDE project was to explore new means of musical expression by combining modular motion capture systems, innovative software programs for interactive sound synthesis, and dynamic systems for visualization. The expected findings touch artistic creation, musical education, and the world of musical games. This project focuses on a growing community of users that includes members of the general public who are interested in gestural and expressive use of new digital tools.

The new musical MO (Modular Musical Objects) interfaces that were developed in this project won first prize in the international Guthman competition for new musical instruments in 2011. These interfaces were featured in several international exhibitions such as “Talk to Me” at the MoMa in New York, the Biennale du Design St-Etienne, Objet(s) du numérique - Design d’un nouveau monde Industriel in Paris, and the Lift Experience in Geneva.

The MO - Modular Musical Objects interfaces being used in different situations (photos: nodesign.net)

The MO interfaces and associated software (MuBu, gesture follower) have led to concrete applications, in particular for musical education. These applications have been used successfully in the Atelier des Feuillantines music school. New forms of musical games have also been created such as Urban Musical Game (a musical ball game), presented during the Futur en Seine festival in 2011. This project has also led to the creation of Phonotonic, a start-up that develops a certain number of these advances. Certain project partners are also carrying out specific industrial developments.

This project has already been the object of over 17 scientific communications, including scientific journals, the proceedings from national and international conferences, numerous mentions in the press (printed and online), as well as several public presentations in Europe, in the United States, and in Asia.

This project was the winner of the 2013 ANR Digital award in the category “societal impact”.

The MO - Modular Musical Objects interfaces being used in different situations (photos: nodesign.net)
AUGMENTED INSTRUMENTS

Teams Involved: Sound Music Movement Interaction, S3AM

Augmented instruments are acoustic instruments that have been fitted with sensors so that information concerning gestural parameters can be transmitted in real-time. These instruments are specifically adapted for mixed acoustic and electronic music. This work is carried out in close collaboration with composers and performers, the goal being the integration of this technology in their works. These developments are also a part of IRCAM’s research into instrumental gesture.

This project began with the development of the augmented violin, which then took off due to the interest of various composers. We are now working on all the string quartet instruments as well as certain percussion instruments. Diverse computer techniques have also been developed to analyze, recognize, and follow instrumental gestures. For example, a list of musical elements such as playing styles or musical phrases could be defined by the composer and could be used as a foundation for interaction with sound processes including synthesis or spatialization. These tools make it possible to develop a new generation of score following. Finally, it is important to note that these augmented instruments are used in musical creations, but also in the educational arena.

SAME

Sound And Music for Everyone Everyday Everywhere Every Way

Teams Involved: Sound Music Movement Interaction, Acoustic and Cognitive Spaces

Funding: ICT 2007 (networked Media), FP6, European Commission
Calendar: January 2008—December 2010
Partners: University of Genoa (Ugdist, Italy), Nokia Research Center (Finland), KTH (Sweden), Pompeu Fabra University (Spain), Helsinki University of Technology (Finland)
Coordinator: University of Genoa (Ugdist)

Music making and listening are clear examples of a human activity that is above all interactive and social, two of the major issues facing new communication devices and applications. However, to date, music making and listening has been, in general, a passive, non–interactive, and non-context sensitive experience. The current electronic technologies, with their potential for interactivity and communication, have not yet been able to support and promote this essential aspect of music making and listening. This can be considered a significant degradation of the traditional listening and music making experience, in which the public was (and still is) able to interact in many ways with performers to modify the expressive features of a music piece.

The main objective of the SAME project was to create new end-to-end systems for mobile active, experience-centric, and context-aware active music listening. The project tried to answer to questions like “What device will correspond to today’s iPod in 5 years?” or “What potential new markets would such new devices open up?”

The objectives of the SAME project were threefold:

• To define and develop an innovative end-to-end research platform for novel mobile music applications for participative, experience-centric, context-aware, social/shared active listening of music for a broad target of non-expert as well as expert users.
• To investigate and implement new communication and interaction paradigms for mobile music applications based on high-level, expressive, non-verbal multimodal interfaces, enabling the user to influence, interact, and mold and shape the listened content, by intervening actively and physically into the experience.
• To develop new mobile context-aware music applications, starting from the active listening paradigm, which will bring back the social and interactive aspects of music to our information technology age.
IMPROVISED MUSICAL INTERACTIONS - OMAX & CO

Team Involved: Musical Representations

This project focuses on the development of improvised man-machine musical interactions. A new paradigm for interaction was invented at IRCAM and has been made available to the general public via the OMax software program.

Using machine learning techniques and formal languages, OMax learns in an unsupervised way from either a MIDI or an audio stream produced by a musician. The underlying process behind this interaction could be called “stylistic reinjection”. The musician is continually kept informed by several sources providing complex feedback. He hears himself play, he listens to others while memorizing sound images that flow from the present towards the past.

Using medium-term and long-term memory, these motifs, combined with even older images taken from the repertoire or musical culture, for example, can return to the present after undergoing several transformations, including one of the most common transformations in improvisation: formal recombination. OMax models this memory-based process and makes it possible to “reify” it, to make it heard on stage. It then re-injects musical figures taken from its short-term and long-term memory and reconstructs them in a manner that is both similar and innovative, providing the musician with stimuli that are familiar and stimulating.

This project has led to two new research projects: SoMax that explores the immediate reactivity of an artificial agent to its sound environment and Improtek (in collaboration with the EHESS) that explores the notion of guided improvisation in the framework of a specific scenario (e.g. a chord chart). The vocation of the ANR-funded project DyCII is to study the synthesis of these two different approaches with innovative artificial listening techniques.
DYCI2
Creative Dynamics for Improvised Interaction

Team Involved: Musical Representations
Funding: ANR - programme générique
Calendar: October 2014—September 2018
Partners: Inria Centre Nancy-Grand Est, Université de Bretagne Occidentale

The project on creative dynamics for improvised interaction focuses on the creation, adaptation, and implementation of effective models for artificial listening, learning, and interaction as well as the automatic generation of musical contents enabling the creation of digital musical avatars that are creative and either capable of being incorporated in an interactive and artistically convincing manner within a range of human systems such as live performances, (post) production, or education or to contribute perceptive and communicative skills to cyber-physical systems.

The project highlights improvised interaction, both as an anthropological and cognitive model of action and decision, as a schema of discovery and unsupervised learning and as a discursive tool for human–digital artifact interaction with the aim of modeling style and interaction.

The objective is to produce creative agents that become autonomous through direct learning resulting from contact with live performances by human improvising musicians, creating a loop of stylistic retroaction via the simultaneous exposition of humans to the productions of digital artifacts that improvised themselves. This creates a situation of human-artifact communication that evolves in a complex dynamic. Off-line learning with archives can also be anticipated to systematically “color” the digital individuality of the agents or to situate the experience within different genres (jazz, classic, pop, etc.). The live performance situation could be extended to novel applications such as interaction with users that have a range of skills, with audio-visual archives dynamically resuscitated in artistic or educations scenarios of co-improvisation, as well as in the general situation of new narrative forms of the work in interactive/generative digital media and virtual reality.

The goal is also to constitute procedural knowledge of music through this interaction and to produce a rich instantaneous human-digital experience, likely to provide aesthetic satisfaction for the user, to enrich his sound and musical production, to implement a dialog with him, to imitate or contradict him, and in general, to stimulate and revitalize the experience of collective performance. This human-artifact interaction will be extended to artifact-artifact interaction in diverse configurations (several humans, a network of several digital artifacts). During the experiment, an autonomous musical individuality will form an individuality capable of intervening in a plausible fashion in situations of collective interaction.

A creative entity in an audio-musical context thus subsumes a collection of concurrent, contributing, and competitive agents capable of interactive learning, of taking charge of artificial learning tasks, of discovering short and long-term temporal structures, of modeling style, of generating symbolic sequences, of real-time audio rendering, but also of visualization and human-machine interfaces.
IMPROTECH
Improvised Technologies and Music

Team Involved: Musical Representations
Funding: ANR - Sciences, technologies et savoirs en sociétés
Calendar: January 2010 — December 2012
Partners: CAMS (Centre d’analyse et de mathématique sociales, UMR 8557 CNRS-EHESS), LAHIC (Laboratoire d’anthropologie et d’histoire de l’institution de la culture, UMR 8177 CNRS-EHESS), CREM (Centre de recherche en ethnomusicologie, UMR 7186, formerly UMR 7173), ENS-LSH de Lyon (Junior Laboratory "penser l’improvisation")

This project aimed to study the impact of new technologies on musical creation considering the knowledge put to use. The scientific advances expected at the end of this study include a better understanding of the current evolution of the popular music in the context of the disorder caused by the generalization of digital technologies now found in all stages of production. As seen above, all music is related to knowledge, but the emphasis here is laid on that which, within the popular music, relies on specific musical capacities, such as the capacity to improvise, and given this, put in practice complex knowledge. This is the case of jazz, in particular.

It is therefore necessary to question the nature of this type of musical knowledge. Beyond the technical aspect, music has a semantic aspect. This means that during an improvisation, the sound support “doubles” as something else via the sharing of common elements at the heart of a community made up of people who play, listen to, and promote this music (even if the word “community” creates more problems than in traditional ethnography connected to a given territory). At the heart of this community, musical phrases are shared during a complex game of covers and borrowing. One of the goals of this project is to follow the paths of this sharing and demonstrate how individual creations connect among themselves, leaving space for innovation.

The unique aspect of this project lies in the comparison of two types of exogenic knowledge, musical knowledge on one hand, techno-scientific knowledge on the other, although this dualism must be moderated. The idea is to see if a sort of “transplant” works in the world of improvised music because of the increasing influence of digital tools. The criterion for success for this type of project is the clear vision that it will provide to understand the place of technology in improvised music today. It has been said that this music is full of creative energy, but that it is difficult today to draw up an overall picture of the situation. The implementation of this music is a choice made by a certain number of musicians with whom we work, and their presence is an essential element of the project’s success.
WRI TING TIMED INTERA CTIONS AN D MUSICAL SYNCHRONIZATIONS

Team Involved: Musical Representations

In a context of the performance of contemporary works with instrumental and computer parts, score following is a technique often used for the synchronization of the electronic accompaniment with a solo performer. An augmented score is recorded in the computer using a specific format that contains both the essential elements of the instrumental score a specification for the computer part written in a dedicated musical language. During the performance, the analysis in real-time of the performer’s sound and/or of the captured movement is aligned with the symbolic score. At every instant throughout the performance, the score follower’s algorithm determines the location in the score and synchronizes the processes programmed in the electronic part of the work.

The score follower (Antescofo) now includes a real-time reactive, synchronous and timed programming language, enabling the definition of complex interactions between performers and live electronics. The use of this technology makes it possible to ensure coherent writing between the electronic and instrumental parts of the composition, and a synchronous and polyphonic performance of electronic modules with musicians in real-time. Its development will include an adaptation for different families of musical instruments and recognition of musical forms as well as the extension of dedicated expressive structures to address various paradigms for the electronic parts.

This system includes two important issues in computer music: the recognition and extraction of musical data in real-time (artificial listening) from an audio signal and reactive synchronous programming for writing time and interaction.
**EFFICAC(E)**

Extended Frameworks For ‘In-Time’ Computer-Aided Composition

- Funding: ANR - ICJC Program
- Calendar: October 2013—March 2017
- Partner: CNMAT / UC Berkeley
- Coordinator: IRCAM

The objective of this project is the development of tools for computer-assisted composition (CAC) exploring new relations between computation, time, and interactions in the OpenMusic environment and in other technologies developed at IRCAM and CNMAT. This project endeavors to take CAC processes beyond the traditional “offline” domain to include them in a structured interaction in their context be it performance, execution, or composition (in the processes that lead to the creation of musical material). During this project, we will focus on a number of specific directions, such as:

- Reactive processes for computer-aided composition: communication and propagation of events and changes in compositional processes
- Management of temporal flows: Writing, dynamic scheduling, relations between musical execution time and offline computation processes; representation of temporal constructs
- Control, visualization and execution of sound synthesis and spatialization processes
- Gestural interactions in compositional processes

Several critical antagonisms in the domain of computer music are addressed in this project (signal versus symbolic approaches, offline versus real-time). By bridging high-level computer-aided composition systems with other disciplines and frameworks such as sound processing, spatialization and gestural integration, it includes control and interactions in abstract and expressive compositional models.

**INEDIT**

Interactivity in Writing of Interaction and Time

- Teams Involved: Musical Representations
- Funding: ANR - Contenus et interactions
- Calendar: September 2012—August 2015
- Partner: Labri, Grame
- Coordinator: IRCAM

This project aimed to enable the interoperability of sonic and musical creation, opening the path to new creative dimensions that couple writing for time and writing for interaction. Our motivations are reflected in the international efforts to bridge the current gap between compositional and performance aspects of existing tools for sonic creation. This gap is apparent in the categorization of tools for computer-assisted composition and real-time. For a composer, performance-oriented tools (e.g. sound processing and synthesis, mixing, audio post-production) are often limited to the metaphors of production and sound-processing sources (notably oscillators, filters, and physical models). However, these metaphors should be included in a symbolic language like the final score, enabling real musical writing and favoring more powerful methods of representation, conceptualization, and communication.

Inversely, composition-oriented tools still struggle to take into account a new problem, that of interaction: a sound system corresponds more or less to a work that is closed and entirely specified in advance, but must interact in real-time with an environment. This is the case, for example, for interactive musical works, video games, sound decors, or for multimedia systems and installations.

Our approach is based on a language vision: an interactive work is seen like a performer (in the computer sense) who must articulate locally synchronous low-level temporal streams (the audio signal) in a globally asynchronous event time (pertinent events for the composition process). The INEDIT project promotes a hybrid approach that enables cooperation among different tools and stylistic approaches at the heart of an environment that supports all the phases of the musical workflow, from composition to performance. New technologies must be developed to reach these goals: dedicated scheduler, mechanisms for cooperative coupling, on-the-go compilation, as well as innovative visualizations and tangible interfaces that enable the specification and control in real-time of musical processes in a natural process for the creator and designer.
COMPOSITION, DESIGN, MEDIATION

COMPOSITION TOOLS

COMPUTER-ASSISTED COMPOSITION (OPENMUSIC)

The purpose of research in computer-assisted composition (CAC) is to study and design models and computer techniques adapted to the creative process, incorporating paradigms for calculations as well as musical interactions and representations. This approach favors symbolic orientation using programming languages for artistic creation and processing harmonic, temporal, and rhythmic data in addition to other aspects that come into play in the compositional process. Our work in this domain is articulated primarily around the OpenMusic environment, a visual programming language based on Common Lisp and dedicated to musical composition. Contemporary music composers have used this environment for the past 15 years. Today, is regarded as one of the principle references in computer-assisted composition and has been downloaded by several thousand users from around the globe.

OpenMusic (OM) is a visual programming environment for composition or musical analysis assisted by computer. OM offers users a range of interconnected modules associated with specific functions, making up patches that enable the creation or transformation of musical data structures. OM also offers several editors to manipulate these data in addition to libraries in specialized sound analysis and synthesis, mathematical models, the resolution of constraint problems, etc. Unique interfaces like the maquette editor let users construct structures that include functional and temporal relationships among musical objects. OpenMusic is used by a large number of composers and musicologists; it is taught in all major computer-music centers and several universities worldwide.

Recently, a new calculation and programming paradigm was suggested for the OpenMusic environment, combining the existing demand-driven approach with a reactive approach inspired by event-driven, interactive real-time systems. The activation of reactive channels in visual programs increases the possibilities for interaction in a CAC environment: an event—a change or an action made by a user—made in a program or in the data it is made from, produce a series of reactions leading to an update (reevaluation). An event can also come from an outside source (typically, a MIDI port or an open UDP and attached to an element of the visual program). Two-way communication can be established between the visual programs and exterior applications or systems. The CAC environment finds itself inserted in the temporality of a larger system, potentially governed by events and interactions produced by or in this system. This temporality could be that of the compositional process, or that of the performance. The expressive wealth of CAC environments is largely found in the combinations it enables between temporal and functional relationships within musical structures. In traditional “off line” work, the processes of calculating and the execution of musical structures take place in distinct phases and temporalities. Here, we study interactive situations where the sequencing of musical events may depend on, or interact with generative processes put in place in a specific compositional or performance situation. A new architecture was defined with this objective in mind, connecting musical structures produced in a CAC environment with the sequencing structure. This architecture employs a hierarchical representation of musical objects that interact with their conversion by sequencing events.

Different prototypes for the sequence kernel were created for this architecture. These prototypes made it possible to integrate the idea of a dynamic sequencer in the OpenMusic environment, as well as the modification and generation of musical data “on the fly” during the performance of a score. The different models and optimizations, specified for each prototype make them more or less adapted for different cases and musical situations: static scores, static material with dynamic sequencing, and entirely dynamic material. The creation of a unifying model is one of the medium-term goals in this project.
COMPUTER-ASSISTED COMPOSITION (ORCHIDS)

Team Involved: Musical Representations

A unique and recent case of computer-assisted composition, this project addresses the question of orchestration via an automatic search of instrumentation and layering instruments approaching, depending on different acoustic similarity criteria, a target defined by the composer. Current research endeavors to make this dynamic orchestration paradigm heard, according to the targets with sonorous characteristics that vary with time.

Realized after the Orchidée software suite, Orchids is the first complete system for temporal computer-assisted orchestration and the optimization of timbre combinations. It provides an ensemble of algorithms making it possible to recreate any sound target that changes over time through a combination of instruments or samples, according to the psycho-acoustic criteria. This can help composers obtain unique timbre colors providing a multitude of effective solutions to best recreate the sound target. Through a large selection of functions, Orchids can also recreate the evolutions and abstract forms of spectral movements. Its results provide multiple orchestra scores that can be organized intuitively in order to quickly construct orchestral and musical ideas. This system provides several approximation algorithms that make it possible to conjointly optimize several timbre features. The advantages of the Orchids system lies in the fact that this approximation can be carried out separately on temporal forms, values, mean values or standard deviations (or any combination of the three) of each psycho-acoustic descriptor. In addition, users can also manually define a temporal deformation and carry out a multi-target search within several sound segments, making it possible to create full orchestral works in just a few seconds.

ORCHESTRATION

Team Involved: Musical Representations
Funding: Conseil de recherches en sciences humaines du Canada (SSHR/C)
Calendar: April 2015—March 2018
Partners: McGill University (Canada, coordinator), Haute École de musique de Genève (Switzerland)

One of the most complex and mysterious aspects of musical composition, still scarcely studied in the scientific domain, concerns writing timbre, in particular orchestration techniques. The project is based on a comparison of the state of the art in musicology, in psychology of perception, and in computer music to create new tools to address problems connected to orchestration, its perception, and its instruction. The objective is to develop models that can be generalized, that facilitate the instruction and practice of orchestration, assisted by new technologies. The long-term goal is to create an interactive treaty on orchestration comprising knowledge on orchestration practices, the perception of orchestral effects, and digital tools to help resolve orchestration problems including writing electronic sections of mixed-music works. The project relies on a large quantity of annotations made on the corpus of classical music using a range of pertinent categories, intended to supply both perceptive experiences and educational materials as well as to establish practice data for automatic learning algorithms.

IRCAM’s role in this project focuses on computer music applications and artificial intelligence for computer-assisted orchestration tools based on the software program Orchids and new research on deep learning techniques.
SAMPLE ORCHESTRATOR

Funding: ANR - Riam program
Calendar: December 2006—June 2009
Partners: Univers Sons
Coordinator: IRCAM

The Sample Orchestrator project aimed to produce new applications based on content management and manipulation of sample (isolated pre-recorded sounds) banks. Its production corresponds to the conjunction of two elements: the widespread availability of large sample banks on different supports (e.g. CD and DVD, online databases) but that present a limited number of applications (e.g. synthesizer samples) and recent scientific and technologic developments in indexation methods and management systems for audio databases, making it possible to imagine new musical functions, calling on global methods of content management and manipulation of the entire corpus of available samples.

• Content management of the sample bank: automatic classification, search by resemblance, concatenative synthesis based on high-level descriptions
• Audio processing by contents (based on a previous analysis of characteristics of the processed signals): transposition, dilation, hybridization, etc. These processing functions can be divided in pre-processing functions that address the preparation of samples in the studio, and post-processing functions that intervene in real-time when the sound banks are used via a new generation instrumental sampler.
• Orchestration tool, integrated in the OpenMusic environment, which finds, based on models of musical knowledge (generally empirical for composers), the best combinations of sounds with a specific musical effect in mind (fusion/separation, reorchestration from existing sounds, etc.)

Research activities on the identified major scientific and technological issues were planned as a part of the project: description, indexation, automatic classification of musical and sound contents, sound processing by content, engineering musical knowledge and its application for orchestration, software environments for audio processing in real-time. This project was the subject of numerous scientific publications (2 doctoral theses, 10 articles in journals and books, 18 conferences), of an industrial valorization (the technologies developed in the project led to the creation of the MachFive 3 software program by Univers Sons, marketed by the American company Mark of the Unicorn), and several musical creations at IRCAM. The project was selected for the label “major scientific advance” by the ANR in its 2006 report.
The technologies used for sound analysis, synthesis and signal processing have made it possible to imagine significant, even new, possibilities for composition. In this prospect, new functions have been developed in the OpenMusic environment to include these technologies in the context of the compositional process. Several specialized libraries make it possible to connect programs created in a computer-assisted composition environment with processes for sound processing or synthesis (created by IRCAM tools such as SuperVP, pm2, CHANT, Modalys, Spat, but also tools such as Csound and Faust). Bringing the domain of sound synthesis and computer-assisted composition together is a new approach to the representation of sounds via high-level programs that put structures of abstract data and models of advanced temporal structures to work, permitting total control of the processes of sound processing and generation of synthesis data.

Developed in collaboration with Marco Stroppa, the OMChroma library makes it possible to control the processes of sound synthesis with the help of matrix data structures. OMChroma’s extension for spatialization, OMPrisma, enables to implement “spatialized sound synthesis” processes, calling on spatialization (positions and trajectories, but also room characteristics, sound source orientation and directivity) when sounds are produced. Controlled in OpenMusic via an ensemble of graphical editors and operators, these tools propose a wealth of possibilities in the conjoint specification of synthesized sounds and spatialized spaces.

The OM-Chant project also made it possible to bring back the technology of FOF (Formant Wave Functions) synthesis at the heart of the process of computer-assisted composition, and to create at the core of CAC processes synthesized sounds inspired by the model of vocal production.

The theme of space, and the collaboration with the Acoustic and Cognitive Spaces team, has been reinforced with a project on the conception of interfaces to control sound spatialization processes. Using a participatory design approach, interviews were held with composers to better understand their needs with the goal of offering tools to facilitate the control and writing of spatialization. These interviews revealed the necessity of interactive tools to enter, visualize, and manipulate control data for spatialization, and led to the design of different prototypes.

OpenMusic objects and tools for 2D/3D spatial conception are based on a model of delayed data, enabling a homogenous management of temporal specifications and transformations. Connected to a sequencing motor and high-level temporal structures, these data make it possible to represent sound scenes and their evolution over time. An interactive object was developed using models of visualization, interaction, and spatialized rendering from the Spat library.
AUGMENTED PAPER FOR CAC

Teams Involved: Musical Representations, Analysis of Musical Practices
Partners: LRI Orsay, INRIA

Through the technology present in augmented paper that contains a motif visible only to the small camera inserted in a special pen, the computer can detect what the composer is writing. This ‘before’ connection between paper and computer is completed by an ‘after’ connection from the computer to paper by printing on augmented paper, therefore creating an entirely new production cycle.

The goal of this project, extended upon request by the ANR, is to work in collaboration with artists and research workers to identify the possible musical applications that can come from the link between the computer support and augmented paper. Rather than replace paper, with its recognized properties of flexibility, comfort, and free association our approach consists of developing applications that combine the advantages of paper and electronic documents.

In the long run, we would like to develop prototypes of applications in three core fields:

- Notation
  In the field of musical composition, we are interested in both notation and annotation as both are complementary in their reflective nature for musical creation.

- Writing Synthesis
  Due to its flexibility of use and the subtlety of handwritten interaction compared to interaction via a mouse or even with a graphical tablet, paper remains the preferred support for the informal expression of artistic gestural ideas, and especially for writing down ideas for synthesis. In collaboration with the OpenMusic software program, this technology facilitates a natural evolution of various parameters for synthesis.

- Teaching
  The technology present in augmented paper is exceptionally interesting for musicologists and music teachers. By connecting a simple support, the printed score, to the MusiqueLab 2 software program, it is possible to work collaboratively and create new forms of interaction.

© Inria / Photo H. Raguet

Interfaced used for the premiere of Quid sit musicus? by Philippe Leroux in June 2014. This interface associates electronic musical processes with the calligraphy of Guillaume de Machaut’s score.
The Vertigo project was written and selected in response to the first call during the European Commission’s H2020 ICT program (Information and Communication Technologies) connected to the STARTS initiative (Science, Technology and the Arts). STARTS encourages synergies between artists and R&D projects in the field of ICT supporting innovation and Vertigo is committed to coordinate a range of actions carried out in this regard throughout Europe focusing on three areas:

- The organization of artistic residencies assessed by an international and multidisciplinary jury. A budget of 900 € is dedicated to supporting 3 annual cycles for a total of 45 artistic residency projects that focus on creating, via the production of unique artistic works, prototypes for innovative products using the technologies developed.
- The public presentation of different actions carried out during the new, annual event at the Centre Pompidou: Mutations/Créations. This event features exhibitions, symposia, and artistic events and aims to expose and debate evolutions in artistic practices within the contemporary scientific and technological ecosystem. The theme of the inaugural edition in March 2017, presented in connection with the exhibitions “Imprimer le monde (Printing the World)” and “Ross Lovegrove”, is “Simulated Space or Digital Forms”.
- The development of a communal online platform enabling different parties (scientists, artists, manufacturers, business incubators, investors) to contact each other and offer support for their actions (communication, organization of specific artistic residency programs, etc.). In addition to the consortium partners, the project oversees an international network of 20 correspondents in the cultural field, including Zentrum für Kunst und Media de Karlsruhe, The Victoria and Albert Museum in London, the Venice Biennale, the festival Ars Electronica in Linz, etc.

The WAVE project aims to formalize new means of editing, visualization, and interaction with temporal audiovisual objects online. This project will lead to the design and development of software bricks concerning user interfaces and interactions, digital audio interfaces, client/server interfaces, and their data exchanges. These software bricks will come from the results of the analysis of expert musical practices compared to their ordinary use in web applications and standards, with the goal of offering innovative uses for them. The project’s objective is to integrate these software bricks in existing software to develop new services and improve existing ones. New means of purchasing and sharing musical and audiovisual documents online, these services will be put in place during the project and made available to users by project partners using the W3C standards (particularly HTML5 and the new Web platform), completing them if necessary, and also proposing interaction via new terminals in order to offer coherent interfaces, accessible and innovative, and new user experiences adapted for consultation, interaction, annotation, transformation, and sharing of temporal objects. In a fluctuating web technology situation, the formalization and instanciation of these multiple interfaces in different technologies supported by WC3 standards and recommendations will contribute to the surfacing of a truly hypermedia and interactive Internet. In the context of the cultural industry’s crisis, and particularly those connected to music, companies and institutions can make use of these cases and explore new markets via innovative systems for listening and consultation that are interactive and enhanced with temporal streaming from the Internet.

This project led to the organization of the first Web Audio Conference (WAC) in collaboration with Mozilla, Google, and the W3C at IRCAM on January 26, 27, and 28, 2015.
MATHEMATICS AND MUSIC

Team Involved: Musical Representations

The project “Modèles algébriques, topologiques et catégoriels en musiciologie computationale (Algebraic Models, Topologies, and Categories in Computational Musicology)” was selected by the CNRS upon the creation of a director’s position in the Musical Representations team. This project follows the MISA project (Computer Modeling of Algebraic Structures in Music and Musicology), broadening the spectrum of mathematical tools for computational musicology (tools from algebra and also from topology and category theory). At the same time, this project attacks the theoretical and practical articulations between classical and popular music (rock, pop, jazz, song). Four institutions support the project: the Society for Mathematics and Computation in Music (an international society for which the MISA project was a federating force), the GDR ESARS (Esthétique, Art et Science), a collaboration with Sorbonne Universités’ IReMus (during international encounters held at the Centre de Recherche on popular music), and a partnership with université de Strasbourg (in particular with IRMA and Labex GREAM). This project also supports the Mamuphi seminar on mathematics, music, and philosophy organized in collaboration with the école normale supérieure and the book collections: Musique/Sciences (Delatour France) and Computational Music Sciences (Springer).

The spatial program aims to model problems such as movements in a specific space or the transformation of spatial structures. It provides computer tools that enable the development of analyses in line with Set Theory. This work has already made it possible to explore the pertinence of topological tools for the representation and classification of musical objects such as calculating all-interval series, Neo-Reimannian harmonic theory, and the geometric representation of chord progressions.

The result is an experimental environment to assist analysis of musical sequences called HexaChord. HexaChord is an environment that makes it possible to build spatial representations associated with a group of chords and analyzes them via several topological notions. The spatial representations proposed include divers Tonnetz and simplicial complexes corresponding to groups of chords’ pitches. The software offers 2D and 3D visualizations of the representations produced.

GEMME
Musical Gesture: Models and Experiments

Team Involved: Analysis of Musical Practices, Sound Music Movement Interaction

Funding: ANR — programme blanc shs
Calendar: November 2012 — June 2016
Partners: Université de Nice-Sophia Antipolis (CTEL: Centre transdisciplinaire d’épistémologie de la littérature)
Coordinator: IRCAM

In the beginning of the 21st century, gesture became a significant means of interaction with technology for the general public, and also within the musician community. While musical writing has put a strong emphasis on deconstructing and reshaping the acting body of the musician-performer for nearly 50 years, it is interesting to note that for the past 10 years there has been a strong multidisciplinary convergence on this research/creation topic, drawing the attention of composers, performers, and computer scientists as well as the domains of engineering, psychology, physiology, biomechanics, and cognitive sciences. This concept of gesture, commonly used in numerous domains, notably in the performing arts such as theater and dance, has only been the subject of embryonic research in the domain of musicology. The GEMME project offers an analysis of theoretical texts and musical works, and also carries out investigations before and after the premiere of a score: what theoretical and technical possibilities of the formalization of gestures are available to composers? What gestural procedures can they test on paper and during the performance of a work? What means of transmission of the gestural information are created not only during the collaboration between composer and performer, but also when the performance of the work is taught? This project endeavors to answer these questions via four main themes:

1. Tacit Theories of Gesture: genealogy of the compositional notion of gesture, its categorizations and periodization, the current state of the art
2. Gesture and Stage: study of a paradigmatic method—that of Kagel—where the musical idea is connected to its staged expression in the framework of musical and instrumental theater
3. Gesture and Instrument: study of a contrasting paradigmatic method—that of Lachenmann—where the composition calls upon a breakdown of the organological possibilities of sound production in relationship with a political and social criticism of expressive conventions
4. Gesture and Technology: a series of musical analyses of a group of seminal scores, from Ferneyhough’s Time and Motion Study II to Luna Park by Aperghis, that offer a variety of technical and computing paradigms that formalize and/or accompany the instrumental gesture.
THEORIES OF MUSICAL COMPOSITION OF THE 20TH CENTURY

Team Involved: Analysis of Musical Practices
Funding: CNRS, Réseau international d’études des écrits de compositeurs (Université de Montréal), Éditions Symétrie (lyon)

This editorial project began in 2003 and aims to publish a substantial panorama of the main compositional theories of the 20th century. While certain theoretical corpora have been broadly published and commented (e.g. Schoenberg, Xenakis, Boulez, Stockhausen), others, just as important for history and the musical aesthetic, remain relatively unknown due to a lack of available introductory texts. Today, neither musicians nor musicologists have a comprehensive reference source on the fertile relationship between composition and theory during the last century. Even so, the fertility of this relationship was a determining factor in the development of contemporary music, and particularly of computer-music.

Crossing historical musicology, discourse analysis and musical analysis, the book edited by N. Donin and L. Feneyrou will fill the current void with a collective reference publication on the compositional theories of past century—a century characterized by a profusion of writings by artists, aesthetic manifests defining musical techniques and technologies, and research carried out on compositional systems. Sixty international specialists have come together to treat, on one hand, compositional theories specific to an individual or a school—from Schoenberg to Hindemith to Carter to Rihm—and on the other hand, the categories and problems that surfaced throughout the century such as algorithmic music, musical theater, live electronics, or spectralism. Theoretical corpora unknown in French, sometimes unpublished, are presented alongside the most widely explored themes by musicologists during the past 40 years. The contributions offer a synthetic presentation, both technical and historical, of these notions and/ or doctrines, based on the most current knowledge about the articles, manifests, etc. on the composers studied. The level of diversity of the participants reflects different aspects of 20th century musicology, from Philippe Albéra to Elena Ungeheuer, from Angelo Orcalli to Richard Toop. Éditions Symétrie published this two 900-page volume work in 2013.

MUTEC
Musicology and Techniques of Contemporary Composition

Team Involved: Analysis of Musical Practices
Funding: ANR - SHS
Calendar: December 2008 — December 2011
Partners: IEM, Conservatoire de Genève, CEAC, Université de Lille-3, Technico, Université de Technologie de Troyes
Coordinator: IRCAM

This project aimed to document and question the specificity of a set of modern compositional techniques, via a thorough and coordinated study of several creative processes representative of the music of 20th and 21st centuries, documented through archives of the creative process (sketches, drafts, etc.) and/or in-depth interviews with the composers (for practicing composers). The method and the references of the program have roots in a range of complementary disciplines: musicology, genetic criticism (in relation to the genesis of a work), and cognitive sciences.

The project develops a pioneering methodology based on the concept of “reenactment of the activity of composition”. The different creative processes selected cover a period of about 60 years, from the mid-20th century to contemporary works: music for a film for Cartier-Bresson by Koechlin, Pli selon pli by Boulez, Requiem pour un jeune poète by Zimmermann, Les espaces acoustiques by Grisey, Traiettoria by Stroppa, Trei II by Jarrell, L’Esprit des Dunes by Murail, a recent sound garden project by J.L. Hervé, and work in progress, Gramigna, by Gervasoni. Each work uses a specific technology such as a montage of sound recordings, music mixed with synthesized sounds, or a sound installation with real-time treatments. The empirical study of interaction between an artistic project and its technical tools makes it possible: to understand differently a group of works that are references for contemporary music, to better understand the relationship between technical systems and creativity, to shed light on certain characteristic traits of musical cognition via an in-depth analysis of one of its forms that demands the highest level of expertise; a form that is seldom researched.

The program organized the circulation of the notions and themes of transversal analysis through consecutive case studies. This circulation was facilitated through the design and use of a shared computer platform (developed by the team, based on a preexisting prototype of “genetic navigation” within a musical work) that enabled the digitization of different genetic corpora and hypermedia browsing within these corpora, including their annotation and variations in their forms of representation.
MULTIMEDIA PUBLICATION TOOLS FOR MUSICOLOGY

Team Involved: Analysis of Musical Practices

This work aims to experiment with, and in some cases, standardize the original formats of multimedia publications on music. One can argue that any analytic discourse on music contains a potential ‘multimedia’ aspect, as it includes a text (the oral commentary) and indicators for musical information (e.g., citations of musical examples, references to movements or passages in a work, etc.). Even so, it is not until recently that musicologists became interested in the new expressive possibilities of multimedia publication, online and on other physical supports that were limited to educational and cultural CD-ROMs until recently. We have begun a two-step process in order to contribute to this process.

Firstly, the experimentation carried out consists of producing various musicological analyses in a multimedia environment, specifying as we go along the environment chosen according to the steps carried out in the analysis. This experimentation materializes itself through the development of analysis tools developed in collaboration with musicologists (e.g., Charting the Score: A Tool for “Segmented Listening” in collaboration with Jonathan Goldman) and also through the publication of multimedia documents, on DVD-ROM (such as those included in the Inouï, the IRCAM 2005 and 2006 journal) and in online magazines such as DEMeter (Université de Lille-3) or Musimédiane—an audiovisual and multimedia music analysis journal.

Secondly, using a process of comparison and abstraction we will carry out a formalization of the analytic operations brought into play. This enables standardization of publication formats using text, sound, and image—not only for the specific needs of musicologists, but also for disciplines that are close to music publication found within the institutional projects carried out at IRCAM (e.g., the revamping of the Brahms database, the creation of the IRCAM repertoire, software documentation). The issues are those of documenting: using a well-defined format (to facilitate the processing of the information) but one that is not overly restrictive for others; joining visual and sonorous musical excerpts together; making several types of publication possible (online, a CD-ROM, a DVD-ROM, paper).

In order to do this, and given the sheer number of commands involved by these projects, it is unfeasible to depend on a manual layout of the data provided by authors for each publication. A tool dedicated to these types of publications must be used to integrate the aforementioned constraints. This was the focus of our participation in ECoUTE (Authoring environment for instrumental music listening, sound archive management, and multi-support publication) supported by the RIAM program and the RNTL Scenariplatform project. The sum of these actions has already made it possible to establish a publication format for musical analysis. This format was first tested in a model of online classes for the European project MusicWeb in 2003 and in a scientific publication on the analysis of a performance in 2005. The format was finally formalized with the production of a series of CD-ROMs associated with the Musique Lab 2 project featuring the Variationen op. 27 by Webern and Voirex by Philippe Leroux (2006-2007) as a part of an agreement between IRCAM and the Provence Alpes Côte-d’Azur region. These CD-ROMs offered two different ways to address the work:

- The first introduces the work by letting the user read the score as they are listening. The score was annotated with a few contextual elements.
- The second offers several analytic points of view on the work. Each position leads, via a hypermedia link, to key passages in the score annotated by the author via the Musique Lab Annotation software program. The contents were produced entirely by the author using the Scenari-CHAIN authoring tool and the Musique Lab software package.
**GAMELAN**

An environment for the management and archiving of music and digital audio content

- Team Involved: Musical Knowledge Engineering
- Financing: ANR — Contenus et interactions
- Calendar: December 2009 — October 2013
- Partners: Université de Technologie de Compiègne, Institut National de l’Audiovisuel, EMI Music
- Coordinator: IRCAM

The existing environments for sound and musical production consist of several tools designed for making and putting together sounds. Be it classical, contemporary, electro-acoustic, electronic, or pop music, each production manages a large number of sound samples and uses a broad range of tools for the creation, modification, or embellishment of sounds.

In a production setting, it is not unusual to find a variety of tools: recording tools, tools to ‘clean’ the sound, sound banks, processing tools, tools for sound synthesis, sequencers, audio effect libraries, tools for synchronization and mixing, in addition to all the other tools that are specific to a given production such as the score.

This project had four main objectives:

- **Production Environments**
  Follow all your actions from the source material to the finished product (e.g. action ‘X’ carried out on file ‘Y’ created a new file) and organize the elements that intervene in a production (files, software) in structures predetermined by the user including them in the environment’s components. Formalize the knowledge created during the production process. The application objective is to be able to work with any production software on the market and work with the environment whenever any of these programs are used.

- **Preservation Strategies**
  Utilize the production environment like a platform for archiving, extracting structures and knowledge that simplify future access to the environment. Apply OAIS methodologies that let future users use the environment and its components.

- **Reusing Productions**
  Restructure production material with new objectives in mind, add other material and modify the connections found in the global structure. Use parts of the environment to generate new environments. Deconstruction and reconstruction of processes to analyze the intentions.

- **Copyright Management**
  Enable a traceability of the contents during a production to manage the copyrights.
CREAM
Cracking the Emotional Code of Music
—
Team Involved: Sound Perception and Design
Funding: ERC Starting Grant 335536
Calendar: October 2014—September 2019
Partner: Centre de Recherche de l’Institut du cerveau et de la moelle épinière (CRICM UMR7225)

The CREAM project’s objective is to produce technology and knowledge that make it possible to characterize which type of musical signal sets off which type of brain mechanism of emotional induction.
Until now, research in musical cognition has focused on correlations between relatively indistinct emotional reactions and loosely controlled musical stimuli: we know that music creates emotions, but we do not know how. The CREAM project suggests combining current neuroscience methods with a high technical level of signal processing to create, for the first time, musical stimuli capable of selectively causing or inhibiting certain cortical circuits involved in emotional processing so they may be studied in isolation. For example, we will suggest targeting centers of interpretation of prosody and speech by constructing musical sounds that “tremble” like an anxious voice or that “rejoice” like a happy voice.
These new experimental control techniques make it possible to expand our current understand of brain mechanisms of emotional induction, but also to conceive several clinical applications for therapy or diagnosis of depression or neurodegenerative illnesses. In other words, the CREAM project will turn music into a real clinical technology capable of stimulating specific neuronal circuits in a non-intrusive and non-pharmacological manner.
SKAT-VG
Sketching Audio Technologies using Vocalizations and Gestures

Teams Involved: Sound Perception and Design, Sound Music Movement Interaction, Sound Analysis & Synthesis
Calendar: January 2014—January 2017
Funding: Objective ICT-2011-C, FET-Open: Challenging current thinking FP7, European Commission
Partners: IUAV (Universita Iuav di Venezia), KTH (Kungliga Tekniska Hogskolan), GENESIS (Genesis Acoustics)
Coordinator: IUAV (Universita Iuav di Venezia)

Sketching is at the root of any design activity. In visual design, hand and pencil are still the primary tools used to produce a large variety of initial concepts in a very short time. However, in product and media design the sonic behavior of objects is also of primary importance, as sounds may afford seamless and aesthetically pleasing interactions. But how might one sketch the auditory aspects and sonic behavior of objects, in the early stages of the design process? Non-verbal sounds, more than speech, are naturally and spontaneously used in everyday life to describe and imitate sonic events, often accompanied by manual expressive gestures that complement, qualify, or emphasize them.

The SkAT-VG project aims at enabling designers to use their voice and hands, directly, to sketch the auditory aspects of an object, thereby making it easier to exploit the functional and aesthetic possibilities of sound. The core of this framework is a system able to interpret users’ intentions through gestures and vocalizations, to select appropriate sound synthesis modules, and to enable iterative refinement and sharing, as it is commonly done with drawn sketches in the early stages of the design process. To reach its goal, the SkAT-VG project is based on an original mixture of complementary expertise: voice production, gesture analysis, cognitive psychology, machine learning, interaction design, and audio application development. The project tasks include case studies of how people naturally use vocalizations and gestures to communicate sounds, evaluation of current practices of sound designers, basic studies of sound identification through vocalizations and gestural production, gesture analysis and machine learning, and development of the sketching tools.

The project has several scientific goals:
- Expand existing knowledge in the domain of perception and production of vocal imitations and expressive gestures
- Develop algorithms for automatic classification of vocal and gestural imitations, based on imitation itself, combining the analysis of a sound signal with the physiological mechanisms of vocal production
- Explore the effectiveness of sound and gestural sketches in the domain of interactive sound design, by taking advantage of automatic classification for selection and parameters of sound synthesis models
- Develop applications for intuitive sound creation and processing using voice and gesture

Sketch of the synopsis of the SkAT-VG project. (© O. Houix).
MOUVIE
Mobility and Quality of Life in Urban Areas

Team Involved: Sound Perception and Design
Funding: Chaire d’Excellence — Fondation UPMC
Calendar: April 2014—April 2019
Partners: OSU Ecce Terra, Laboratoire Atmosphère, Milieux, Observations Spatiales (LATMOS), Inserm, UPMC — Institut Jean le Rond d’Alembert (MPIA Team)

The Impact of Atmospheric and Noise Pollution on the Health of Urban Populations
The chair in mobility and quality of life in urban areas will measure the impacts of pollution by gases, particles and noise on the quality of life and health of the residents of a city. In response to this societal issue, the car manufacturers PSA Peugeot Citroën and Renault have realized the importance of supporting the sponsored research project that UPMC is running together with France’s Institute of Health and Medical Research (INSERM), Atmosphere, Media and Spatial Observation Laboratory (LATMOS), Ecce Terra, the Jean le Rond d’Alembert Institute and the Institute for Acoustics and Music Research and Coordination (IRCAM). “By taking the many factors involved in chemical air pollution into account, as well as unconventional sources of pollution such as noise, the chair can bring a new dimension to the concept of the urban environment, which may be promising for the future,” commented Laurence Eymard, Chair.

An Inter-Disciplinary Research Team of Experts
A chair headed by Laurence Eymard, coordinating an inter-disciplinary team of experts in acoustics, air quality and health. Laurence Eymard will coordinate a scientific community of more than 1,000 people specialized in earth system and environmental sciences, ranging from biodiversity to health, and from deep-earth science to planetary science, via climatology. Régis Marchiano, a professor at UPMC, works on linear and non-linear acoustic wave propagation in complex media at the UPMC/CNRS Jean le Rond d’Alembert Institute. Nicolas Misdariis is joint head of the Sound Perception and Design team at the Institute for Acoustics and Music Research and Coordination (IRCAM). Sébastien Payan, professor at UPMC and a research fellow at LATMOS, researches various issues on the boundary between molecular physics and atmospheric physics. In health, Isabella Annesi-Maesano, head of the INSERM/UPMC Epidemiology of Allergic and Respiratory Diseases team, runs research projects to understand the etiology of rhinitis, asthma, chronic obstructive bronchopneumopathy, and identify subjects at risk. Annick Clément, professor of pediatric pneumology at UPMC, runs the pediatric pneumology ward at Trousseau Hospital. She also coordinates a European research group on rare lung diseases in children and sits on numerous scientific boards in France and around the world.

Major Issues for Private-Sector Sponsors of the New Chair
Sylvain Allano, Scientific and Future Technology Director at PSA Peugeot Citroën: “We thought a sponsored chair was an appropriate way to support high-quality inter-disciplinary research in areas with a strong societal dimension, such as air quality and noise pollution. The research results should provide the kind of solid, recognized scientific basis that often is missing from societal and environmental debates.” Rémi Bastien, Head of Innovation Engineering at Renault: “Society expects mobility, which is both a necessity and a fundamental freedom, to contribute to a higher quality of life, especially in urban areas. Current forms of mobility have increasingly unacceptable impacts. As manufacturers, we are part of the problem, so we want to be part of the solution. The best way is to cooperate with high-level inter-disciplinary scientists. The new chair offers us a unique opportunity to pave the way for sustainable, reliable solutions backed by in-depth scientific expertise.” Claire Martin, Vice President of CSR at Renault and Managing Director of the Renault Foundation: “The better we understand impacts in their complexity, the more able the government and industry will be to develop appropriate responses. It is also legitimate and necessary to analyze the technical, economic, social, health and other aspects of phenomena in order to send clear, objective messages to all audiences.”
**LOUDNAT**

Team Involved: Sound Perception and Design  
Funding: ANR Programme blanc 2011  
Calendar: November 2011—October 2015  
Partners: IVA-INSA, Lyon and IMA-CNRS, Marseille  
Coordinator: IVA-INSA

Loudness is a basic dimension of sound perception related to sound intensity. Models exist to predict this feeling based on measurements carried out on the sound signal, but they are limited to specific cases of sounds that are unrelated to natural environmental sounds. These restrictions consequently limit the use of these models to measure the loudness of environmental sounds that vary in time such as a passing car. The general impression of loudness for a listener therefore depends on the type of variation of the sound’s characteristics that must be taken into account in a predictive model. Moreover, the localization of a sound source as well as the obstacle created by the listener’s very body in an everyday situation lead to modifications of a sound and inter-aural differences which affect how loudness is judged by a listener. These modifications must also be taken into account in a predictive model adapted for environmental sounds. This project’s goal was to expand the validity of loudness models to environmental sounds; the study consists of psychoacoustic experiments that will test several hypotheses concerning perceptive and cognitive mechanisms that must be taken into account to modify and adapt existing loudness models.

**ROADSENSE**

Team Involved: Sound Perception and Design  
Funding: ANR, Programme VIT  
Calendar: October 2010—November 2013  
Partners: AXIMUM, CETE, LEPSIS, COLAS  
Coordinator: AXIMUM

RoadSense is an industrial research project that aimed to design, implement, and validate via experiments a system of assistance for drivers. This system consists of a delineation (rumble strips) that creates an alarm—a sound and vibrations—that is set off when the wheels pass a specific limit. This system is designed to alert drivers whose trajectory does not correspond with that of the road in the countryside.

To reach this goal, the RoadSense project aims to:

- Propose a framework for the functional analysis of the security of roads using current knowledge, the identification of specific problems, and the circumstances surrounding accidents on rural roads.  
- Design a digital simulator for relevant sound signals, implement them using a driving simulator and via physical tests, and test their effectiveness and acceptability of a “vigilance rumble strip” that uses sound delimitation of the roadway.  
- Carry out and test a new, low-cost system of assistance for drivers. This system will be able to be applied quickly and will work with all existing vehicles on roads, from small country roads to highways, via a gradual validation process in situ (on closed tracks and then on roads).
CLOSED
Closing the Loop Of Sound Evaluation and Design
http://closed.ircam.fr

Team Involved: Sound Perception and Design
Program: NEST (New and Emerging Science and Technology), MTI (Measuring the Impossible) Program, FP6, European Commission
Calendar: July 2006—June 2009
Partners: HGKZ (Hochschule für Gestaltung und Kunst Zürich), Technischen Universität Berlin, Universita Degli Studi di Verona
Coordinator: IRCAM

The vocation of the Closed project is to offer a new approach to sound design for tools— for artistic, measurement, and assessment purposes—for the design of everyday objects. This project is part of the larger domain of industrial design and can be represented graphically as seen here:

The design process is seen as an iterative loop. The initial phase of analysis leads to the creation of specifications of the qualities of the sound to be produced. The creation phase uses these specifications that are then validated based on the initial specifications. If need be, the production is refined and passes again through the creation-evaluation loop.

Closed focuses on the creation of tools for measurement destined to be used by designers in their creative processes. These tools are associated with functional, aesthetic, and emotional aspects of sound; the designer can shape the sounds of daily objects while assessing the results of their creation. The evaluation tool will be made up of a group of easily understood indicators measuring functional, aesthetic, and emotional responses that will be connected to a specific usage for the product.

This project brings together four European laboratories and is structured around four principal domains:

- Sound synthesis based on a physical modeling approach (solid, liquid, etc.). Synthesis is controlled via physical and perceptive parameters oriented toward interactive design.
- Study of the perception of environmental sounds and their perceptive and cognitive organization. Integration of the aesthetic and functional aspects perceived in sound in the design process and assessment of their emotional impact.
- Study of new uses for sound design. Creation of principles of sound interaction of an object in relation to perception and cognition. Elaboration of scenarios via sound prototypes that promote user/object interaction.
- Techniques for learning and automatic classification of sounds based on physiological models for an automatic aesthetic, functionality, and emotional evaluation of the sound included in the sound design process.
MINET
Measuring the Impossible NETwork

Team Involved: Sound Perception and Design
Funding: European Commission, 6th Framework programme as a nest pathfinder coordination action project
Funding: Coordination action 043297, NEST, FP6
Calendar: February 2007—March 2010
Partners: 21 European (and Israeli) partners, both academic and industrial
Coordinator: Stockholms Universitet

The MINET project coordinated activities for all European projects that are a part of the ‘Measuring the Impossible’ (MTI) program; the CLOSED project was a part of this program, as are specialists in metrology. The common point of these projects is their search to measure the impossible, meaning phenomena connected to being human, conventionally thought of as being immeasurable such as emotions or perceptions. The global objective of MINET was therefore to reflect on this idea of measurement applied to humans. Via an organization that facilitated exchanges, brainstorming groups, seminars, study groups, or summer schools among MINET project partners, this objective was divided into several points: improve the productivity of MINET project partners, structure and broaden the European scientific community interested in measuring quantities connected with the human being, developing and standardizing the metrology concepts used to measure humans.

SID
Sonic Interaction Design

Team Involved: Sound Perception and Design
Funding: Cost, European Commission
Calendar: July 2007—June 2011
Partners: thirteen European countries and one non-European partner — McGill University, Canada
Coordinator: Department of Art and Industrial Design (IUAV, Venice)

Sonic Interaction Design is the exploitation of sound as one of the principal channels conveying information, meaning, and aesthetic/emotional qualities in interactive contexts. The Action pro-actively contributes to the creation and consolidation of new design theories, tools, and practices in this innovative and interdisciplinary domain. While being advanced through a few sparse projects, this field relies on the COST - SID Action to strengthen the links between scientists, artists, and designers in the European Research Area. The COST - SID platform stands on four legs:
• Perceptive, cognitive, and emotional approaches in sound interactions
• Design of the sound produced
• Musical and artistic interactions
• Sonification
4

IRCAM
SOFTWARE
THE FORUM

IRCAM SOFTWARE USERS’S
http://forumnet.ircam.fr

The new Forum is intended, above all, to create a space for lively discussions where members can share knowledge and skills about music and technologies from IRCAM’s R&D department. The Forum proposes software programs, activities, and new services. The Forum was created in 1993 to make the results of the research carried out at IRCAM accessible to universities. Over time, a community made up of the users of IRCAM’s software and technology has been created, and the range of professions, cultures, and methods that have come to be found within this community have considerably enriched the Forum’s original offer. The new Forumnet platform has added a social dimension to the community, connecting software users to IRCAM’s R&D teams, enabling direct exchanges and discussion groups, workshops and working sessions held in Paris and elsewhere.

MEMBERS
The Forum technologies are intended primarily for:
- **Sound professionals** (sound engineers, sound designers, computer-music designers, acousticians) for voice and sound processing (film, installations, concerts), room acoustics (concert halls, cinemas, laboratories), and processing techniques for artistic creation (electronics, patches).
- **Researchers-developers** (researchers, developers, musicologists, instructors, students) for research on sound, the voice, instruments, and performance.
- **Artists from all musical genres** (composers, performers, multimedia artists, sculptors, choreographers, stage producers) that use computer music for the conception and production of a work.

The software distributed by the Forum is also intended for amateur use by anyone who is interested in new technologies applied to sound and music.

TECHNOLOGIES
Software is accessible via:
- A Premium membership (institution or individual) provides access to all the software programs.
- The Forumnet boutique lets users purchase individual software programs. This is the perfect option for students or institutions who are interested in a specific technology.

SERVICES
The Forum offers:
- forumnet.ircam.fr: a space reserved for members that includes and agenda, events, max patches, documentation, and more.
- An annual 3-day workshop at IRCAM in March.
- An annual workshop held in a university or institute abroad.
- Online technical support.
- Preferential rates for training courses organized by IRCAM’s department of Education and Cultural, for events (concerts, the ManiFeste festival), and on IRCAM publications products (CDs, books, CD-ROMs, audio samples). Members are also eligible for discounts on several software programs sold in stores (Ircam Tools, Ircamax, Ircam Solo Instruments, Ircam Prepared Piano, etc.).
A FEW NUMBERS
HOW MANY MEMBERS ARE THERE?
In 2016, there were nearly 6,000 members on Forumnet 2.0.
There are 569 paying members, 280 of which have individual Premium memberships, and 41 of which purchased institutional Premium memberships. 248 members purchased individual software programs. Over 86,930 people visited our website.

WHERE ARE THEY?
France 20.42%, Europe (not France) 58.20%, Europe (outside the EU) 7.56%, USA-Canada 18.88%, Asia 11.19%, Latin America 3.08%, Africa 1.12%

MEMBERS INCLUDE
France
Cité de la musique (Paris), Cirm (Nice), CNSMD de Lyon, CNSMD de Paris, CRR de Strasbourg, ENSA de Dijon, Ensemble TM+ (Nanterre), ESBA Le Mans, IReMus - CNRS (Paris), Théâtre National de Strasbourg, Université Jean Monnet (Saint-Étienne), Université Rennes 2, Université Paris-Sorbonne, les Centres nationaux de création musicale

Europe
Centre Henri Pousseur (Belgium), Charleroi/Danses - La Raffinerie (Belgium), Conservatorio Statale di Musica Giuseppe Verdi (Italy), Conservatorio Superior de Música de Aragón (Spain), De Monfort University (United Kingdom), Estonian Academy of Music and Theatre (Estonia), Experimentalstudio des SWR (Germany), Goldsmiths-University of London (United Kingdom), Gotlands Tonsattarskola (Sweden), Haute école de musique de Genève (Switzerland), Hochschule für Musik und Theater, Felix Mendelssohn Bartholdy Leipzig (Germany), Karl-Franzens-Universität Graz (Austria), Keele University (United Kingdom), Kungliga Musikhögskolan (Sweden), NOTAM (Norway), Sonic Arts Research Centre - Queen’s University Belfast (United Kingdom), Staatliche Hochschule für Musik und Darstellende Kunst (Germany), Universität für Musik und darstellende Kunst Wien (Austria), University of Arts in Belgrade (Serbia), University of Birmingham (United Kingdom), University of Edinburgh (United Kingdom), University of Huddersfield (United Kingdom), University of Liverpool (United Kingdom), University of Manchester (United Kingdom), University of Sussex (United Kingdom), Utrecht School of the Arts (Netherlands), ZKM Center for Art and Media (Germany)

Beyond Europe
Arizona State University - School of Arts Media and Engineering (United States), Boston University (United States), Chapman University (United States), CNMAT (United States), Columbia University (United States), Cycling’74 (United States), Columbia College Chicago (United States), Conservatoire de musique de Montréal (Canada), McGill University (Canada), Nagoya University of Arts (Japan), Shobi University (Japan), Tokyo University of the Arts (Japan), University of Cape Town (South Africa), University of South Korea, School of the Art Institute of Chicago (United States), SIAL Sound Studios - RMIT University (Australia), University of Auckland (New Zealand), Yale University (United States), Yong Siew Toh Conservatory of Music (Singapore), Western Sydney University (Australia)
SUPERVP, TRAX, AND LIBRARIES

Design and Development: Sound Analysis & Synthesis Team

SuperVP is a signal-processing library that is based on an implementation of a sophisticated phase vocoder algorithm. The library can be used to perform a large number of signal transformations with outstanding sound quality (time stretching, transposition of pitch and spectral envelope, denoising, noise-sinusoids-transient remixing, amplitude scaling of the spectral envelope, generalized cross synthesis, source filter cross synthesis, etc.) and provides access to an extensive set of parameters that offer complete and fine grained control over the different algorithms. In addition to sound transformation algorithms, the library includes a large collection of signal analysis algorithms (F0, onsets, spectrogram, re-assigned spectrogram, spectral envelope, etc.). One of the by-products of SuperVP is the VoiceForger library that offers a large number of real-time SuperVP processing and provides high-level access (see the Sample Orchestrator, Vivos, and Affective Avatars projects).

The TRaX tools have been designed so that the majority of the library parameters have been condensed in a set of high-level user parameters that enable an intuitive control of the application.

Main Features

Signal Transformation:
• Time Stretching/Compression: with and without preservation of transients, special mode with waveform preservation for speech transformation
• Transposition: independent transposition of pitch and spectral envelope, sample rate transformation (re-sampling)
• Denoising: interpolation of multiple noise keys and independent processing of sinusoidal noise
• Noise-Sinusoids-Transient Remixing: detection of sinusoidal, noise and transient components with independent mixing controls
• Envelope Transformation: estimation of the spectral and time envelope with independent mixing controls
• Cross synthesis: phase vocoder or source filter based cross synthesis with various controls

Signal Analysis:
• Spectrogram: Short-time Fourier transform (STFT) with and without reassignment
• Spectral Envelope: LPC, True Envelope
• Fundamental Frequency (F0): precise analysis with very low CPU usage, specialized instrument presets for various instruments (i.e. violin, flute, xylophone, vibraphone)
• Attacks: versatile algorithm for the detection of note onsets and other transients
• Voiced/Non-Voiced: detection of the cut-off frequency boundary

Technical Characteristics and Configuration

The library can be controlled on different levels of abstraction. It supports file and memory based sound input/output as well as real-time control of many parameters. The library is used in a wide variety of environments. There is a command line application using file based sound input/output that gives access to all available sound transformations and analysis. It is used for example in the AudioSculpt and OpenMusic applications (see OM-SuperVP). In Max, objects enable sound transformation and analysis in real-time; the TRaX application is based on these objects. The library is programmed in C++ using advanced optimized methods, such as SIMD operations (SSE2, SSE, Altivec) found in today’s processors. The library runs on Mac OSX operating systems, Windows XP, and Linux on Intel and PowerPC. Its multithread implementation makes it possible to exploit the calculated capabilities of the multi-core and computer processors.

Technologies and Partners
MakeMusic (SmartMusic), Flux;, Univers Sons, Xtranormal, Voxler, MXP4, etc.
AudioSculpt uses a graphical approach to “sculpt” (to modify) a sound file. After a phase of analysis, the user modifies the analysis’ result directly, using only a mouse to apply the desired changes to the sound. Several types of analysis of the spectral content are shown and the user can modify the sound several different ways. The main types of sound modification include filtering, moving segments of the sonogram in time and frequency, or applying one of the numerous high-quality transformations available.

Applications
Composition, sound design, post-production, movies, multimedia, acoustics, education, musicology, ethnomusicology, etc.

Main Features
• Sound representation – Sound file at a glance, zoom from overall amplitude envelope down to sample level, sonogram, spectral envelope, processing sequencer
• Sound analysis – LPC, True envelope, fundamental frequency, partial following, virtual pitch, estimation of time positions, tools for listening to specific parts of spectrum and showing harmonic relationships
• Annotation: text, segment, and MIDI notes can be added on the sonogram
• Sound processing – a sequencer makes it possible to group several tracks, each with different processing, and to listen to their effects in real-time before finalizing the result
• Filtering – directly on the sonogram using a pencil or eraser or by specifying with the mouse the points forming a polygon. Manual source separation and “denoising” is possible. Compression/expansion – allows modification of the length of a sound without changing its pitch or timbre and preserving the quality of the attack and transient
• Transposition – using a specific editor the user can transpose a sound without loss of data or quality of sound
• Denoising – spectral subtraction with estimated noise interpolation
• Cross synthesis – application of spectral data of one sound to another sound in order to make a spectral transition from one sound to another or an hybrid sound
• Partial synthesis – creation of a new sound from partials found by one of the analysis
• Collage – sound zones defined in terms of their time-frequency can be copied and pasted
• Listening – transformations can be heard in real-time. The scrub mode enables navigation in a sound at any speed
• Bpf curves – users can draw curves directly on the sonogram or on the waveform to control processing

Technical Features and Configuration
AudioSculpt is based on a vocoder Sound Analysis & Synthesis engine called SuperVP, which provides the means for performing most analyses and sound modifications. Sinusoidal Sound Analysis & Synthesis is performed in AudioSculpt using the Pm2 tool. AudioSculpt is the intuitive interface that allows access to SuperVP. The user can load text files containing commands addressed directly to SuperVP.
AudioSculpt is compatible with high-quality multi-track sounds (up to 32-bit integer of floating/192kHz) and uses SDIF format for analyses, which can be exchanged with other software programs.

Platform
Macintosh on Mac OSx (10.5 or higher). Compatible with AIFF/AIFC, WAV, and Sound Designer II and more.
IRCAMLAB TS

Design and Development: Sound Analysis & Synthesis Team

TS is a standalone software application for sound transformation and time stretching that uses SuperVP and elements of the AudioSculpt 3 interface.

Applications
Composition, sound design, post-production, teaching

Main Features
• Visualization - display of the sonogram, waveform, and spectrum. Interactive zoom from the temporal signal to sample, linear and logarithmic sonogram. FFT, LPC, and True Envelope analysis
• Compression/Expansion - modification of the length of a sound without changing its pitch or timbre and preserving the quality of the attacks and transitions. Presets for different transposition modes with the possibility of adjusting the max FO and spectral envelope
• Remix - remix the transient, sinus, or noise components of the sound signal
• Control - audio playback with standard controls. Results can be recorded in real-time or offline by modifying the parameters. Possibility of piloting and automation via a MIDI controller

Platform
Macintosh on MacOSX (10.5 or higher). Compatible with AIFF/AIFC, WAV, and other formats via the libsndfile library

THE SNAIL-ABSOLUTE TUNING

Design and Development: Sound Analysis & Synthesis Team

The Snail-Absolute Tuning is a sound frequency analyzer that uses an inventive representation with chromatic alignment. Compared to standard analyzers, this representation makes it possible to easily visualize active zones with light (like on a spectrogram) and to organize frequencies by note on a tempered scale, in tune with a diapason.

Applications
Musical visualization of a sound signal, education, tuning

Main Features
• Visualization - display of The Snail-Absolute Tuning’s chromatic abacus simplified for a musical image of sound
• Interface - configurable analysis engine. Settings for different modes of visualization: phase, spectrum, or snail (tuner). Settings for the frequency region visible in MIDI notes. Diapason tuning

Platform
Macintosh on MacOSX (10.6 or higher), iOS (iPad). Compatible with all audio files (all formats supported by the native MacOSX library) or signal input via the audio input of the equipment used
Modalys is an environment that lets users create unheard of virtual instruments based on simple physical objects such as strings, plates, tubes, membranes, plectrum, bows, or hammers, and making them interact. It is possible to construct objects with complex forms using a three-dimensional mesh or resulting from measurements. Modalys brings these virtual instruments to life by calculating how they vibrate when played. By combining different physical, resonating objects and by specifying the type of interaction (striking, rubbing, mouthpiece, bow, etc.), the user creates a virtual instrument and then decides how it will be played. OpenMusic and Max communicate directly with the Modalys synthesis engine via intuitive graphical interfaces.

Applications
• Musical composition: contemporary, techno, electro, etc.
• Sound design: natural work on sound quality (metallic, woody, etc.)
• Education: modal theory in practice
• Cinema & video: sound design
• Scientific research & development: comparison of theoretical and measured modes

Main Features
Modalys' synthesis engine makes it possible to hear virtual instruments made from simple physical objects or 3D mesh. The user can make an instrument and then decide how it will be played. To do this, the user must define the following in four steps:
• Objects which simulate the physical structures (tubes, strings, membranes, plates, complex inexistent forms, plectrums, reeds, hammers) which are calculated according to the default values given by the system or can be modified by the user
• Access to the objects (points on the object on which an action will take place)
• Connections which make it possible to define playing modes: hitting, plucking, rubbing, blowing, etc.
• Controllers that will describe how the parameters of a connection will evolve over time

Use
Chosen by a large number of composers and sound designers, Modalys for Max is the incarnation of Modalys in the real-time environment Max, used to construct sophisticated instruments graphically using strings, plates, membranes, tubes with any type of materials (metal, wood, diamond, etc.), size (unlimited), and interactions: hitting, plucking, rubbing, blowing, etc. This interface can be controlled live from Modalys in real-time. Modalys for Max requires Max 5 or higher (including 7) and is available for Mac OS X (10.6 or higher) and Windows (7 or more recent). Modalys can also be used from ModaLisp, a textual programming interface that can be used to construct and play an instrument with a script. Using this approach, the user writes and executes Lisp programs to obtain the desired musical result. IRCAM's musical programming software, OpenMusic, also based on Lisp, enables a musical or graphical control of the models, for instance from a score and/or temporal envelopes. For a scientific approach to model synthesis, Modalys can easily be used with MatLab through dedicated libraries.

Platform
Mac OSX 10.5 or higher, Window 7 or higer (Modalys for Max only)
OpenMusic is a visual programming environment for creating computer aided composition applications. This environment offers a number of visual modules to the user, each of which is associated with a specific function and visual icon. With just the click of a mouse users can create a program, for generating or transforming musical material. Patches can be imbedded in each other to create programs and more elaborate structures.

Today, OpenMusic is used by a large number of composers and musicologists. It is taught in all the major computer-music centers and in several universities in Europe and the United States.

**Main Features**

- Graphical construction of functional programs (abstraction/application methods, higher-order functions, conditional structures, iterations, etc.) or object-oriented construction
- High-level functions for the creation and transformation of musical structures (e.g. pitch, rhythm, etc.)
- Algorithmic data manipulation in MIDI, audio (WAV, AIFF), and SDIF formats, communication with OSC
- Mathematical tools for analysis and composition
- Graphical editors: score editing functions for traditional and proportional music notation (with import and export to main software programs), 2D and 3D curves, and breakpoint functions
- Synthesis control tools (OM Chroma environment, interfaces with Csound, CHANT, and SuperVP)
- Mathematical tools for analysis and composition
- "Maquettes": a graphical editor that makes it possible to give a temporal and hierarchical form to patches and define temporal programs

**Specialized Libraries**

- LZ: statistical analysis and generation of improvised music
- Esquisse: spectral and frequency functions
- OMRC: programming rhythmic constraints
- OMCS: programming by generalized constraints
- Repmus: The Music Representation Team’s function collection
- Profile: manipulates harmonic profiles and contours
- Morphologie: number and symbol sequence analysis functions

**Technical Features**

Programming based on Common Lisp

**Platform**

Mac OS X, Windows, Linux

Open Source Distribution: LGPL license for the basic environment

**Technology and Partners**

LispWorks (http://www.lispworks.com/)

MidiShare and LibAudioStream © Grame (www.grame.fr)

Collaboration with BEK · Linux version (http://www.bek.no/)
**ORCHIDS**

Design and Development: Musical Representations Team

Orchids is the first complete system for temporal computer-assisted orchestration and timbral mixture optimization. It provides a set of algorithms and features to reconstruct any time-evolving target sound with a combination of acoustic instruments, given a set of psychoacoustic criteria. It can help composers to achieve unthinkable timbral colors by providing many efficient sets of solutions that best match a sound target. With our extended set of features, it can now also reproduce abstract spectral descriptors evolutions (bypassing the need for a sound file). Its results provide multiple orchestral scores that can be arranged in an intuitive way in order to perform a fast sketching of musical ideas. The system provides several algorithms for approximating jointly several timbre properties. The advantages of the Orchids system is that this approximation can be made separately over temporal shapes, mean values or standard deviations (or any combinations of the three) of each psychoacoustic descriptor. Furthermore, users can also define a manual temporal warping, and even perform multi-target searches inside parts of a sound, quickly providing full orchestral pieces in seconds. The new system provides an extensive out-of-the-box database of orchestral possibilities but can also be unlimitedly extended, even with synthesis sounds, by a simple drag-and-drop operation. Finally, the software provides an intelligent time series querying system that allows to easily search for temporal shapes inside the sound database.

**Applications**
Orchestral composition, re-orchestration, timbral writing, music education

**Main Features**
- Generation of new combinations and orchestral combinations by optimization of a target
- Possibility of timbral writing through orchestral mixtures and optimization of the temporal evolution
- Definition of abstract targets and temporal evolution of spectral descriptors
- Simple interface enabling the spatialization of an orchestra or efficient work of the model of an orchestral work
- Simple extension of a knowledge base and efficient temporal searches
- Interaction and exports towards several systems of musical notation

**Platform**
Orchids is a stand alone software program that works with all Mac OSX operating systems. Orchids is also available as Max patches and as a C++ server.
SPAT~

Design and Development: Acoustic and Cognitive Spaces Team

Spat~ IRCAM’s spatializer is software dedicated to sound spatialization in real-time. Originally designed as a library, it enables musicians and sound engineers to control the spatial sound processing for various sound broadcasting systems. Applications range from reproducing sound in a home setting (with headphones, a stereo system, or 5.1) to concert situations (loudspeaker distribution) to holophonic and interactive sound installations.

Applications

• Concerts and Real-Time Spatial Sound Processing
  Composers can map a particular room effect or sound position to each note or musical event in their score by using a score-following program. The Spat~ can be controlled via a sequencer, a score-following program, or by using a high-level control dedicated to composition (e.g. ListenSpace, OpenMusic, etc.).

• Post-Production
  Spat~ can be added to each channel of a mixing console or virtual mixing environment, enabling intuitive control of the position of each sound source and the effect of a virtual room. Simulation and Virtual Reality Spat~ is well suited for immersion through binaural rendering when wearing headphones. The effect is even more impressive if the system is associated with a head-tracking or position-following system. The sound sources maintain their localization according to the user’s navigation orders.

• Holophonic Reproduction
  Spat~ has been adapted to the holophonic reproduction system, facilitating the reconstruction of an acoustic sound field while preserving the consistency of spatial characteristics in an extended zone for the audience. For interactive installations, this reproduction mode enables natural spatial sensations to be retrieved during the exploration of the sound scene.

Technical Features

• Perceptual Controls: azimuth and elevation, distance, source presence, room presence, running and late reverberence, enveloping, frequency balances.

• Low-Level DSP Controls: equalization, Doppler effect, air absorption, multi-channel reverberation, source direction in function of the reproduction system (e.g. binaural, transaural, stereo, 3/2 stereo, 4 to 8 HP multi-track, VBAP for 3D loudspeaker configurations, or Wave Field Synthesis).

Platform

Spat~ is a library of objects for Max (IRCAM/Cycling ‘74) that runs on Mac OSx or Windows XP. The Spat~ is also available as a collection of software plug-ins developed in collaboration with the company Flux:: that can be found in software for music edition and in major brand-name synthesizers (e.g. Protools, LogicAudio, Digital Performer, etc.).

Technology and Partners

France Télécom, Cycling’74, Flux::
TOSCA

- Design and Development: Acoustic and Cognitive Spaces Team

Tosca is a plugin that makes it possible to send and receive the automation parameters of a digital audio workstation towards other applications using the OSC protocol. Its application typically concerns the production of object-oriented spatialized mixes independently of the host software constraints.

Main Features
Tosca can be inserted into a DAW (digital audio workstation) for each track you want to spatialize. Tosca does not affect the audio signals, which are simply bypassed. However, the plugin provides a certain number of parameters that can be automated. In order to ensure compatibility with a wide range of digital audio workstations, the plugin exposes a maximum of 32 parameters per track. During playback, active automation tracks are read by Tosca and the corresponding OSC messages are sent over UDP. When the automation tracks are armed for recording, Tosca accepts incoming OSC packets from remote applications and data can be written in the sequencer tracks.

Each Tosca instance has an ID that can be adjusted by the user (see illustration). This ID is the root pattern of the OSC messages sent and received by Tosca. Their syntax is as follows: “/ID/ParameterName ParameterValue” (e.g. “/3/azimuth 135.0”). For the sake of simplicity, the parameters in Tosca are all double precision floating-number points. Other types of data that can be included in an OSC message (such as integer numbers or text strings) are not currently supported.

Tosca is not connected to a particular spatialization or synthesis engine. In other words, the automation parameters are generic. By default, these 32 parameters are called param1, param2, and so on. The labels of the parameters can be modified. This is carried out via an XML mapping file that can be loaded into Tosca (see illustration).

The XML mapping file makes it possible to specify the automation scale for each parameter: inside the DAW environment, Tosca scales the tracks in a [0 – 1] floating point range towards the [min – max] domain selected by the user. This scaling affects data output and the inverse scaling is applied to input data.

Platform
Tosca was developed with the Juce framework. Juce offers a plugin wrapper that makes it possible, using a common base code, to generate plugins in different formats (VST, AudioUnits, etc). Therefore, Tosca is available for MacOSX and Windows, in VST, VST3, AU, and AAX formats and in either 32 or 64 bit configurations. Tosca is distributed by the IRCAM Forum.

[Image of Tosca in ProTools and an example of a file for automation parameter settings]

1. Plugin inserted in track 1
2. List of parameters exposed
3. Plugin window
4. Automation tracks

A view of Tosca in ProTools and an example of a file for automation parameter settings
PANORAMIX

Design and Development: Acoustic and Cognitive Spaces Team

Panoramix is a post-production workstation for 3D-audio contents. This tool offers a comprehensive environment for mixing, reverberating, and spatializing sound materials from different microphone systems: main tree, spot microphones, Higher Order Ambisonics capture. Several 3D spatialization techniques (VBAP, HOA, binaural) can be combined and mixed simultaneously in different formats. Panoramix also provides conventional features found in other mixing tools (equalizer, compressor/expander, grouping parameters, routing of input/output signals, etc.), and it can be controlled entirely via the Open Sound Control protocol.

Like a traditional mixing desk, the panoramix interface is designed as vertical strips (see Figure) which are organized in two main categories: input tracks and busses. Tracks are used to process signals from a recording session. Panoramix supports several input formats: mono (typically from a spot microphone or an electronic track), multi (a multi-track is basically a grouping of several mono tracks), SMA (used to process recordings from a spherical microphone array like the EigenMike), or tree (used to process a main microphone tree).

Three types of busses are available: a panning bus, a reverb bus, and the LFE (low frequency enhancement) bus. The purpose of the panning bus is threefold: 1) it allows for the summation of multiple input tracks, 2) it determines the spatialization technique (VBAP, HOA, or binaural) being rendered, 3) it provides various settings for the encoding and decoding of the signals (e.g., HOA decoder settings or HRTF attributes in the case of binaural reproduction). The reverberation busses create and control the late/diffuse sections of the artificial reverberation. The current version implements a feedback delay network.

Finally, all the output signals from all the busses are collected in the master strip and then sent to panoramix’s physical outputs or bounced to disk. While each session only has one master track, it is possible to create mixes in several formats simultaneously.

Main window of the Panoramix station 1) Input units 2) Panning and reverb bus 3) LFE bus 4) Master track 5) Session options (group management, preset imports/exports, etc.) 6) Insert module(s) (equalizer, compressor, etc.) 7) Geometric positioning interface
IRCAM SOFTWARE

MAX
The world standard for real-time sound interaction

Max (formerly Max/MSP) is a visual programming environment for building real-time musical and interactive applications. It is currently the world’s leading software program for interactive sound installations. Max is a bundling of Max (IRCAM/Cycling ’74) for real-time control of interactive musical and multimedia applications through MIDI, of MSP, a large set of objects for real-time analysis, synthesis, and processing of audio signals, and of Jitter, a set of video, matrix, and 3D graphical objects for the environment. Max was developed for musicians, sound designers, and researchers who want to develop interactive real-time programs. Cycling ’74 (United States) develops Max under an exclusive license from IRCAM.

Composition
Generation of musical structures using mathematical or random models.

Live Performance
Production of mixed works in which the electronic part reacts according to the instrumental performance using incoming instrumental capture, incoming MIDI signals, or direct modification of the instrumental sound as source material. Controlling the diffusion of sounds and audio files in time Max acts as a sampler. The dynamic control of complex programs that pilot musical signals, video, and lighting effects is also possible.

Audio Post-Production
Virtual control of several external devices such as the mixing console, direct-to-disk system, and synthesizers.

Real-Time Video Processing
A library of objects for video processing (Jitter) is included with the software. It provides users low-level control and the ability to program several applications.

Education
Max is widely used in universities, music schools, and conservatories as an educational tool. The documentation included in the software features numerous interactive tutorials that can be used to teach the fundamental aspects of MIDI systems, audio signals, or video.

Research
Max is used in research to create prototypes of synthesis or sound signal processing in real-time. New external objects can easily be written in JavaScript, Java, and C and dynamically linked to the program.

Technical Features
Several hundred objects for synthesis, control, processing, sampling, and analysis. The Max 6 sound engine offers multi-processor support, 64-bit signals, and several improvements in sound quality. Max 6 includes a system of on-the-fly code generation and compilation based on patches, enabling the quick creation of new objects directly in Max. Gen is can be used in different domains: audio (gen~), Jitter (jit.gen, jit.pix), and OpenGL (jit.gl.pix).
ANTESCOFO

Design and Development: Musical Representations and its MuTant Project Team

Antescofo is a modular score following system as well as a synchronous programming language for musical composition. The module allows for automatic recognition of the player’s position and tempo in a musical score from a real-time audio stream coming from a performer, thus making it possible to synchronize an instrumental performance with computer generated sounds. Antescofo brings together the description of instrumental and electronic parts of a work in a single score via a synchronous language designed for mixed music, with the goal of increasing the expressivity of the process of writing in real-time in a format adapted for the musical language. After changing the score, Antescofo is able to follow the position and tempo of the musicians in real-time and synchronizes the programmed actions for the computer-music portion of the score (setting off and controlling the electronic parts). Antescofo uses a polyphonic audio stream as input.

Applications
Interactive musical works, automatic accompaniment, musical education

Main Features
- Recognition in real-time of the position and tempo of the polyphonic audio stream in a score
- Writing the electronic components of a score in relative, absolute, and rational time
- Interactive sequencing and display via the NoteAbility Pro environment
- MIDI or MusicXML scores can be loaded with NoteAbility Pro

Platform
Antescofo is a library of objects for Max (IRCAM/Cycling’74) that can be used with Mac OSX or Windows XP. Antescofo is also available as a PureData object for Max OSX, Windows, and Linux.
GESTURE AND TEMPORAL SHAPE FOLLOWING

Design and Development: Sound Music Movement Interaction Team

Several Max objects are available to follow temporal morphologies from Markov models. The voicefollower~ object shows in real-time the temporal progression of a performance in comparison with an audio recording. This enables the synchronization of a number of sonorous or visual processes with a pre-recorded text. This object was tested in a theatrical context with spoken and singing voices.

The motionfollower~ object is similar to the voice follower~ (they both use a similar interface) but functions with gestural parameters, like those that come from sensors. This makes it possible to follow augmented instruments or a beat with a motion sensor.

The gf object (for gesture follower), included in the Mubu modules, enables the recognition and following of the temporal evolution of gestural and/or sound morphologies, comparing them with recorded examples.

Applications
Interactive musical and multimedia works, music, dance, gesture control of sound synthesis

Main Features
- Synchronization between gestures and digital media
- Continual gesture recognition from the beginning of the movement
- Similarity measurements between continual temporal morphologies

Platform
Max, MacOSX

MUBU FOR MAX

Design and Development: Sound Music Movement Interaction Team

MuBu (multi-buffer) is a set of Max externals for signal representation and processing around a container of temporal data. The MuBu container is optimized for the representation of temporal data memory, taking into account multiple representations such as signals, descriptors, segments, and annotations of the sound, gesture, and movement. Simplified musical representations and parameters for synthesis and spatialization control can also be integrated.

MuBu also includes PiPo (Plugin Interface for Processing Objects) modules for the extraction of information from sounds and data streams from motion capture as well as a set of Max modules that implement extractors of audio descriptors and filters using the same plugin format.

The development of the MuBu and PiPo modules is the result of the experience acquired during the development of the FTM & Co libraries as well as development carried out on software components during research projects such as the MuBu, ZsaZsa, and RTA libraries.

Applications
Musical works and interactive multimedia projects, interactive sound synthesis, music, dance

Main Features
- Multi-buffer with a graphical interface
- Recording data streams and sequences
- Processing data streams and data store in the multi-buffer
- K-NN unit selection
- Concatenative synthesis, granular synthesis, PSOLA
- Additive synthesis

Platform
Max, MacOSX
CATART

Design and Development: Sound Music Movement Interaction Team

Concatenative corpus-based synthesis makes use of a database of recorded sounds and an algorithm for the selection of units that makes it possible to choose the segments of the database in order to synthesize by concatenation a musical sequence. The selection is based on the characteristics of the recording that are obtained by an analysis of the signal and other features, such as the pitch, energy, or specter.

Using this new approach of sound synthesis by the CataRT system in real-time enables an interactive exploration of a sound base in real-time and a targeted granular composition with specific sound characteristics. This makes it possible for composers and musician to reach new sounds, and for sound designers to quickly explore a large sound corpus. CataRT exists as a stand alone application or as a modular system implemented in Max with the FTM, Gabor, and MnM extensions. Interaction relies on a simple interface made up of a 2D projection of the descriptors that can be navigated with a mouse. Grains can also be selected and played by geometric proximity.

Applications
Interactive musical and multimedia works, research in sound databases

Main applications
• Audio import with segmentation and automatic analysis
• Analysis, visualization, and sound interaction
• Analysis data export
• Composition via interactive navigation in a descriptor space
• Mixing and juxtaposition of massive amounts of sound samples
• Flexibility of granular synthesis
• MIDI compatibility
• Efficient interactive browsing of events or particularities in several or in long recordings, parasite noise separation

Platform
• Application Version: Mac OSX 10.5 or higher, Universal Binary. Compatible with MacPPC and Mac Intel
• Modular Version: Mac OSX 10.4 or higher, PC Windows with Max and FTM&Co

Audio File Formats
• Up to 32-bit/192 kHz integer and floating point: AIFF, AIFC, Wav, FLAC
• SDIF files for importing segmentation markers from AudioSculpt
• MIDI files for importing segmentation and pitch annotation
• Text files for importing segmentation and label annotation (Audacity, for example)
Omax is an environment for improvisation with a computer that analyzes, models, and re-improvises in real-time the performance of one or several musicians, in audio or in MIDI formats. OMax is based on a computer model called "Oracle Factors", a graph that connects all the motives, from smallest to biggest, and provides a map to the logic of the motives learned from the musician, leading to a large number of stylistically coherent variations. OMax bases its recognition on either notes (pitch following) or on timbres (spectral following). OMax 4 and higher are new versions rewritten in C and in Max.

Applications
Improvisation with a computer, interactive musical works, automatic accompaniment

Main Features:
- Artificial Listening: Extraction of the F0 descriptors, MFCC, MIDI input, automatic segmentation, adaptive learning of the symbolic alphabet of musical units
- Modeling: Incremental construction of a model of knowledge representing a full array of musical motives according to diverse similarity functions
- Real-time visualization of knowledge models, multiple selection functions, focalization on the visual representation
- Generation of sequences according to heuristic browsing through the model, taking into account its formal characteristics and musical criteria of continuity of motifs, rhythms, and energy
- Modular architecture that makes it possible to allocated improvisers, players, and to connect them as you like (e.g. several players connected to the same improviser can create a cannon with a rhythmic increase)

Platform
Max 5 or higher and Jitter, Max OSX 10.5 or higher, storage in the JSON format
THE RESEARCH AND DEVELOPMENT TEAM
(SEPTEMBER 2016 – AUGUST 2017)

Directors
Hugues Vinet, Director of the Research and Development Department
Gérard Assayaq, Director of the UMR STMS

Administration
Sylvie Benoit, Ghislaine Montagne, Carole Tan, Anne-Marie Vaudeville (CNRS), Vasiiliki Zachari (CNRS)

Head Researchers
Gérard Assayaq, Frédéric Bevilacqua, Nicolas Donin, Thomas Hélie, Axel Roebel, Raphaël Salé de Chou, Patrick Susini, Olivier Warusfel

Researchers and Engineers
Carlos Agon Amado (UPMC), Moreno Andreatta (CNRS), Jean-Julien Aucourturier (CNRS), Mondher Ayari (Univ. Strasbourg), Elie-Laurent Benaroya, Joël Bensoam, Mathieu Bernard, Jose-Luis Besada Portas, Louis Bigo, Georges Bloch (Univ. Strasbourg), Alain Bonardi (Univ. Paris 8), Riccardo Borghesi, Benjamin Boudoir, Jean Bresson, Clément Cannone (CNRS), Baptiste Caramiaux, Thibaut Carpentier (CNRS), Tatiana Catanzaro, (Univ. Campinas), Céline Chabot-Canet, Thomas Coffy, Arshia Cont, Frédéric Cornu, Gilles Degottex, Rémi Desmonet, David Doukhan, Gaël Dubus, José Echeveste, Philippe Eslin (UPMC), Laurent Feneyrou (CNRS), Emmanuel Fléty, Jules François, Jérémy Garcia, Matthias Geier, Jean-Louis Giavitto (CNRS), Samuel Goldszmidt, Louise Goupil, Karim Haddad, Thomas Hélie (CNRS), Olivier Houix, Florent Jacquieraud (Inria), Olivier Labat, Jean-Philippe Lambert, Joseph Larralde, Fabio Leao Figueiredo (Univ. Sao Paulo), Guillaume Lemaitre, Andreas Lind (CNRS), Marco Liuni (CNES), Mounira Maazaoui, Michelle-Agnès Magalhaes, Mikhail Malt, Gabriel Meseguer Brocal, Enrico Marchetto, Benjamin Matuszewski, Rémi Mignot, Nicolas Misdariis, Jérémy Nika, Markus Noisternig, Nicolas Obin (UPMC), Geoffroy Peeters, Kevin Perros, Charles Picasso, Robert Piéchaud, Emmanuel Ponsot, Sébastien Robaszkiewicz, Francis Rousseaux, David Roze (CNRS), Norbert Schnell, Diemo Schwarz, Marine Taffou (UPMC), Tsubasa Tanaka, Michael Vannier (UPMC), Isabelle Vlaud-Delmon (CNRS)

Technicians and Librarians
Djellal Chalabi, Camille Dianoux (UPMC), Roseline Drapeau, Sandra El-Fakhouri, Mikhaël Gautier, David Messina, Raphaël Raquin, Arnaud Recher (CNRS), Jean-Paul Rodrigues, Alain Terrier.

Doctoral Students

Guest Researchers and Doctoral Students
Anne-Sylvie Barthe-Calvet, Jonathan Cross, Edward Large, Pierre-André Taillard, Vincent Tiffon, Alex Chechile, Whitney Coyle, Olivier Migliore, Hai Morgenstern, Alessandro Olto, Nakai Tomoya, Wander Vieira Rodrigues, Oliver Olsen Wolf, Yohan Zeitoun

Interns
TO CONTACT US: firstname.lastname@ircam.fr

R&D DIRECTION
ircam.fr/recherche

Huques Vinet / Gérard Assayaq*
* UMR STAPS Director

Administration
Sylvie Benoit (IRCAM)
Anne-Marie Vaudeville (CNRS)

INDUSTRIAL RELATIONS
ircam.fr/innovations

Frederick Rousseau

SCIENTIFIC TEAMS

Acoustic and Cognitive Spaces
Olivier Warusfel

Sound Perception and Design
Patrick Susini/ Nicolas Misrahi

Sound Analysis/Synthesis
Axel Roebel

Systèmes et signaux sonores: audio/acoustique, instrumMents (S3AM)
Thomas Hélie

Musical Representations
Gérard Assayaq

Analysis of Musical Practices
Nicolas Donin

Sound Music Movement Interaction
Frédéric Bevilacqua

IT Service
Raphaël Salié de Chou

RESOURCE CENTER
ressources.ircam.fr

Manager
Nicolas Donin

Multimedia library
Roseline Drapeau

SCIENTIFIC AND MUSICAL INTERFACES
forumnet.ircam.fr

Manager
Grégory Beller

Administration
Stéphanie Leroy
Paola Palumbo

Forum - Software Support
Karim Haddad