

# Interaction guidée par machine learning pour la composition musicale gestuelle d'objets sonores non-conventionnels

Romina Romay<sup>1</sup>, Jean-François Trubert<sup>1</sup>, Diane Lingrand<sup>2</sup>, Lucile Sassatelli<sup>2</sup>

<sup>1</sup>Université Côte d'Azur, CTCL

<sup>2</sup>Université Côte d'Azur, CNRS, I3S

*Mots-clés* : objets sonores électro-acoustiques, geste musical, musique visible, composition, réseaux de neurones artificiels, auto-encodeurs variationnels

L'analyse de la création musicale des XXe et XXIe siècles a largement mis en lumière que la notation musicale ne doit pas avoir pour but la reproduction à l'identique de sons fixés. C'est ainsi qu'est progressivement apparu un graphisme musical, visant à provoquer des processus d'interprétation à formes multiples, l'image, stimulation énigmatique pour des sons inconnus, n'ayant quasiment plus de relation avec les sons produits. C'est en particulier John Cage qui en a été le précurseur avec les graphiques de *Cartridge Music* (1960) [1] ou des *Variations I* (1958) [2]. La lecture musicale de cette musique visible dépasse ainsi la réception du texte – certes souvent difficile – pour en venir à un comportement actif et globalisant : le graphisme musical exige des musiciens une interprétation totale basée sur l'énergie. Les symboles souvent polysémiques peuvent produire des musiques différentes selon la façon dont on se laisse influencer par eux, plus activement, en changeant le type de reproduction – tantôt en s'en tenant le plus précisément possible au texte, tantôt en le sur-interprétant, voire en s'opposant à lui (ainsi dans *visible music I*). On peut aussi changer l'interprétation elle-même en comprenant autrement le texte. Ainsi, Schnebel a proposé de favoriser les « représentations imaginaires » de la musique « au moyen d'expositions » de graphismes musicaux. Ceci a fait l'objet d'un accrochage lors de l'exposition *L'œil écoute*, de mai 2017 à avril 2018 au Centre Pompidou. C'est de ce point de départ que le projet ANR GEMME (Geste Musical, Modèles et Expériences) entre l'IRCAM et l'Université Nice Sophia Antipolis, a mis en lumière et élaboré, de 2012 à 2016, les manières de composer la musique basées sur le geste des instrumentistes, et plus largement sur la notion de « geste musical ».

C'est ainsi que le geste peut être envisagé comme un intermédiaire à part entière pour la composition musicale, permettant d'interagir d'abord avec des instruments réels, mais également avec des objets sonores électro-acoustiques, mais aussi avec non plus des instruments, mais ce que nous appellerons des « dispositifs ». C'est cette catégorie d'objets – qui sous le terme « dispositif » regroupent toute une chaîne d'action et de production sonore – que nous considérons ici comme « Drôles d'objets » et pour lesquels nous présenterons dans la démonstration une manière possible d'interagir à travers le

geste. L'interaction est construite à partir d'un modèle génératif de geste obtenu avec un réseau profond de neurones artificiels.

La première étape est de sélectionner un corpus sonore à partir duquel une synthèse musicale peut être contrôlée. Pour cela, nous utilisons l'outil CataRT (*concatenative real-time*) de l'IRCAM comme système de synthèse de son [4]. Il permet, à partir d'un large corpus de sons segmentés et pour lesquels des descripteurs sont extraits, de réaliser une synthèse granulaire selon la proximité du curseur à une position cible dans l'espace (3D ou plus) des descripteurs. Cet espace de descripteurs du corpus est aussi appelé matrice sémantique des objets sonores du corpus. Nous sélectionnons donc un corpus, et une personne musicienne va générer une base de gestes correspondant à une composition musicale sur cette matrice sémantique.

La deuxième étape consiste à entraîner un réseau de neurones profond de type auto-encodeur variationnel, pour encoder la base de gestes. Lors de l'interaction d'un utilisateur non-expert avec le système, ses gestes seront utilisés pour parcourir l'espace latent des distributions. Chaque échantillon obtenu à partir de l'échantillonnage de ce parcours des distributions latentes, sera décodé (par la partie décodeur de l'auto-encodeur profond), et générera ainsi une suite de sons.

Présenter lors de la conférence ces nouvelles techniques d'interaction avec ces nouveaux objets sonores non-conventionnels sera une excellente opportunité de discuter leurs enjeux, tant d'un point de vue artistique que du point de vue de l'interaction humain-machine.

## Références

- [1] John Cage, « Cartridge Music », Peters, New York, n° EP 6703, 1960.
- [2] John Cage, « Variations I », Peters, New York, n° EP 6767, 1958.
- [3] Dieter Schnebel, « Visible Music I », 1960.
- [4] IRCAM, « CataRT », disponible à : <http://imtr.ircam.fr/imtr/CataRT>