A posthumous improvisation by Toots Thielemans

Marc Chemillier (EHESS), Ke Chen (UCSD), Mikhail Malt (IRCAM), Shlomo Dubnov (USCD)

In jazz, the presence of the musician is of such central importance that the recordings considered to be the best were often live recordings, preferably made in small intimate clubs. The evolution of live performance technologies tends to blur this notion, first with the gigantism of concerts, and more recently, with the appearance of simulacra in rap and pop music in the form of holograms (Tupac), virtual stars (Hatsune Miku) and video game avatars (Travis Scott). Toots Thielemans was one of those jazzmen whose presence left an indelible mark on a concert or a recording session, so strong was his musical personality and his harmonica playing immediately recognizable. We propose here to try an experiment consisting in creating a musical avatar of Toots Thielemans.

We use a music improvisation software developed at IRCAM in collaboration with the CAMS (EHESS) in Paris [1]. This machine learning system allows to capture the phrases played by an instrumentalist and to extend them by a virtual improvisation system restoring the sound of the musician, his phrasing and his accents, but playing something different. The system is able to synchronize with the pulse of a human orchestra and can also follow a given chord progression.

The overall software architecture is based on three main modules. A first one built in Max, the body, containing the listening machine, the sound memory unity, the tempo following functions and allowing to interface external MIDI controllers. The mind, an external set of algorithms, made in common lisp and embedded in the OpenMusic [2] computed aided composition software, in charge of calculating new improvisation sequences (according to harmonic and other constraints) sent to the main interface. And the third piece, the Antescofo [3, 4] Max object, that is the clock and the bridge allowing to synchronize the Body and Mind of our software, by messages. Currently, this last module is also able to infer tempo, adapting past sound recorded slices to actual improvisations sequences calculated by our OpenMusic Mind taking into consideration swing and tempo changes.

To realize the experiment and to make the machine learn Toots' playing, it was necessary to have Toots' solos in separate tracks. We used a system of separation of the audio sources. From the field of machine learning, we adapted a zero-shot audio source separator [5], which comprises of a three-component pipeline including a sound event detector, an embedding processor, and a query-based source separator, to extract both the leading instrument and the left accompaniment of the jazz music for the machine’s improvisation. The advantage of this model is that it only requires a small audio clip of the source to extract it from the mixture audio without being pre-trained for solely separating this source. We used the only solo part, a two-sec harmonica clip around 5:56-5:58 in the jazz song "Body & Soul" (from the album *Affinity* with Toots and Bill Evans, 1979), to extract the whole harmonica piece of this song. And we used an existing accompaniment part, about an 80-sec clip around 2:44-4:09, to extract the whole accompaniment. The separation results are correct enough to be used for the following machine improvisation.

The interest of this experiment lies in the problem of acceptability raised by the improvisations produced by the machine [6]. Will an amateur who knows well the playing of Toots Thielmans be taken in by listening to this fake harmonica player? Moreover, it is possible to set the artificial improviser to produce solos more or less similar to the model. At what threshold will the Toots fan consider that these improvisations are not real Toots? Beyond these questions, the device would also allow us to explore aspects of Toots' style. For example, we can distinguish several periods in his stylistic evolution, and the computer model would allow us to hybridize the old Toots and the young Toots by having them meet and play together.

Fundings: This research is supported by the European Research Council (ERC) REACH project, under Horizon 2020 programme (Grant 883313) and by Agence Nationale de la Recherche (ANR)

project MERCI (Grant ANR-19-CE33-0010).

[1] Gérard Assayag, Georges Bloch, Marc Chemillier, Arshia Cont, Shlomo Dubnov, OMax brothers: a dynamic topology of agents for improvization learning, *AMCMM '06: Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, October 2006, pp. 125-132.

[2] Gérard Assayag, Camilo Rueda, Mikael Laurson, et al., Computer-assisted composition at IRCAM: From PatchWork to OpenMusic. *Computer music journal*, 1999, vol. 23, no 3, p. 59-72.

[3] Arshia Cont, A coupled duration-focused architecture for realtime music to score alignment. *IEEE Trans Pattern Anal Mach Intell* 32(6):974–987, 2010.

[4] José Echeveste, Arshia Cont, Jean-Louis Giavitto, Florent Jaquemard. Operational semantics of a domain specific language for real time musician–computer interaction. *J. of Discrete Event Dynamic Systems*, vol. 23, n°4, pp 343-383, 2013.

[5] Ke Chen, Xingjian Du, Bilei Zhu, Zejun Ma, Taylor Berg-Kirkpatrick, Shlomo Dubnov, Zeros-Shot Audio Source Separation through Query-based Learning from Weakly Labeled Data, in *Thirty-Sixth Association for the Advancement of Artificial Intelligence Conference*, AAAI 2022,

[6] Marc Chemillier, Jérôme Nika, « Étrangement musical » : les jugements de goűt de Bernard Lubat à propos du logiciel d'improvisation ImproteK, *Cahiers d'ethnomusicologie*, n° 28, 2016, pp. 61-80.

Biographies

Marc Chemillier
Musician, computer scientist and anthropologist, Marc Chemillier studied jazz piano (Schola Cantorum, CIM). He entered the ENS de Fontenay-aux-roses in mathematics in 1981 and studied harmony-counterpoint at the CNSM in Paris. He made a PhD thesis in collaboration with IRCAM. In ethnomusicology, he worked on the harp of the Nzakara of the Central African Republic (CD *Musiques des anciennes cours Bandia* in 1995), then on the zither of Madagascar. In 2000, he created with the OMax Brothers (Gérard Assayag, Marc Chemillier, Shlomo Dubnov, Georges Bloch) the OMax improvisation software. Director of studies at the EHESS in Paris, he published *Les Mathématiques naturelles* in 2008 (Odile Jacob) and continues his research on computer-assisted improvisation and its anthropological and social issues. In 2021, he published the book-CD *Artisticiel* with Bernard Lubat and Gérard Assayag.

Ke Chen
Audio and Music Researcher. Ke Chen now is a Ph.D student in Computer Music at UC San Diego, California, USA. He is co-advised by Prof. Shlomo Dubnov and Prof. Taylor Berg-Kirkpatrick. His research interest lies in the interdisciplinary between music and computer science, focusing on: Music Generative System and Music Information Retrieval (including singing melody extraction, music source separation, and music recommender system). He proposed a music generative framework, Music SketchNet, to allow the user to specify their own ideas in automatic music generation. He also proposed a zero-shot audio separator to allow a more efficient separation pipeline to separate all possible sources. Ke Chen obtained his Bachelor's degree in Software Engineering at Fudan University, China. He is also the website maintainer of China Conference on Sound and Music Technology (CSMT) and New Interfaces for Musical Expression (NIME).

Mikhail Malt, having a twofold training, scientific and musical (Engineer, composer and musical conductor) started out his musical career in Brazil as both flutist and orchestral conductor. He has a

PHD grade with a thesis at the “Ecole des Hautes Etudes en Sciences Sociales” dedicated to the use of Mathematical models in Computer Assisted Composition, and a HDR (Habilitation à Diriger des recherches) degree with the dissertation : “Representation in Computer Aided Composition and Computational Musicology”. Nowadays he is researcher in the Music Representations Team - UMR 9912 STMS (IRCAM, CNRS, UPMC), associated researcher at IReMus – Sorbonne Universités and Computer Music Designer Teacher in the Educational Department at Ircam, Paris-France. He is currently pursuing his research and composition activities in the fields of musical modeling and discovery, computer assisted performance, and musical representation epistemology.

Shlomo Dubnov is a computer music researcher and composer. He is a Professor in the Music Department and Affiliate Professor in Computer Science and Engineering and a founding faculty of the Halıcıoğlu Data Science Institute in the University of California, San Diego, where he has been since 2003. He is the Director of the Center for Research in Entertainment and Learning (CREL) at UC San Diego's Qualcomm Institute. He is best known for his research on Machine improvisation in Computational creativity and Stylometry of music. He is also known for his contributions to the field of Computer Audition by inventing the method of information dynamics and use of Spectral flatness.