
Titre : Apprentissage de problèmes inverses en acoustique musicale

Mots clés : Problème inverse, timbre musicale, modèles physique, apprentissage

Résumé :

La production sonore dans le monde physique implique un résonateur et un exciteur, tels que le corps d'un instrument et un musicien. Décoder et démêler les effets de ces éléments permet une analyse plus précise des performances musicales expressives et ouvre la voie à la conception d'instruments de musique numériques plus ergonomiques. En général, le décodage du timbre en informations physiques nécessite soit la présence d'un modèle physique, soit des preuves d'interaction entre le musicien et l'instrument acquises dans le monde réel. Le premier, un modèle idéal, ne tient pas compte des écarts par rapport à la réalité, tandis que le second exige un effort considérable pour l'acquisition et l'annotation des données. Cette thèse combine ces deux approches et examine dans quelle mesure la « physicalité » est implicite dans un son instrumental. Trois contributions sont présentées, chacune utilisant des données à différents niveaux de réalisme : des simulations basées sur un modèle physique, des performances artificielles enregistrées dans des environnements contrôlés, et des performances humaines.

La première contribution vise à estimer les paramètres physiques liés à la fabrication d'instruments à partir de sons synthétiques de percussions. Elle propose une méthode innovante pour accélérer une fonction de perte perceptuellement pertinente dans l'entraînement de réseaux neuronaux. La deuxième contribution se concentre sur l'extraction des paramètres physiques de sons de cordes pincées enregistrés par un joueur artificiel en laboratoire. Elle aborde le problème du transfert de connaissances obtenues par simulation vers le monde réel dans un contexte de données limitées. La troisième et dernière contribution commence par l'enregistrement d'un ensemble de données audiovisuelles de techniques isolées de jeu du guqin. Elle présente ensuite une étude de cas visant à identifier computationnellement les corrélations entre les gestes et les sons. Ensemble, ces trois travaux mettent en évidence les défis distincts rencontrés et les approches sur mesure développées pour aborder le problème de l'extraction de la physicalité des sons.

Title : Unearthing the physicality of instrumental timbre

Keywords : inverse problem, musical timbre, physical modeling, machine learning

Abstract : Sound production in the physical world involves a resonator and an exciter, such as an instrument body and a musician. Decoding and disentangling the effects of these causes enable a more accurate analysis of expressive musical performances, and paves ways to the making of more ergonomic digital musical instrument. In general, decoding timbre into physical information requires either the presence of a physical model or evidence of player–instrument interaction acquired in the real world. The former is an idealized model and does not account for deviations from reality, while the latter requires a tremendous effort of data acquisition and annotations. This thesis adopts both of the above approaches and investigates the extent to which “physicality” is implicit in an instrumental sound. Three contributions are presented, each utilizing data at different levels of realism: simulations from a physical model, artificial performances recorded in controlled environments, and human performances. The first contribution learns to regress physical parameters relevant to instrument making from synthetic, percussive drum sounds. It introduces a novel method to accelerate a perceptually relevant loss function for neural network training. The second contribution learns to extract physical parameters of plucked string sounds recorded from an artificial player in lab environment. It addresses the problem of transferring knowledge obtained from simulation to reality under data scarcity. The third and last work begins with recording an audiovisual dataset of isolated guqin playing techniques. It then presents a case study for computationally identifying the correlations between gestures and sounds. Together, the three works demonstrate the distinct challenges encountered and the tailored approaches employed in tackling the problem of extracting physicality from sounds.

