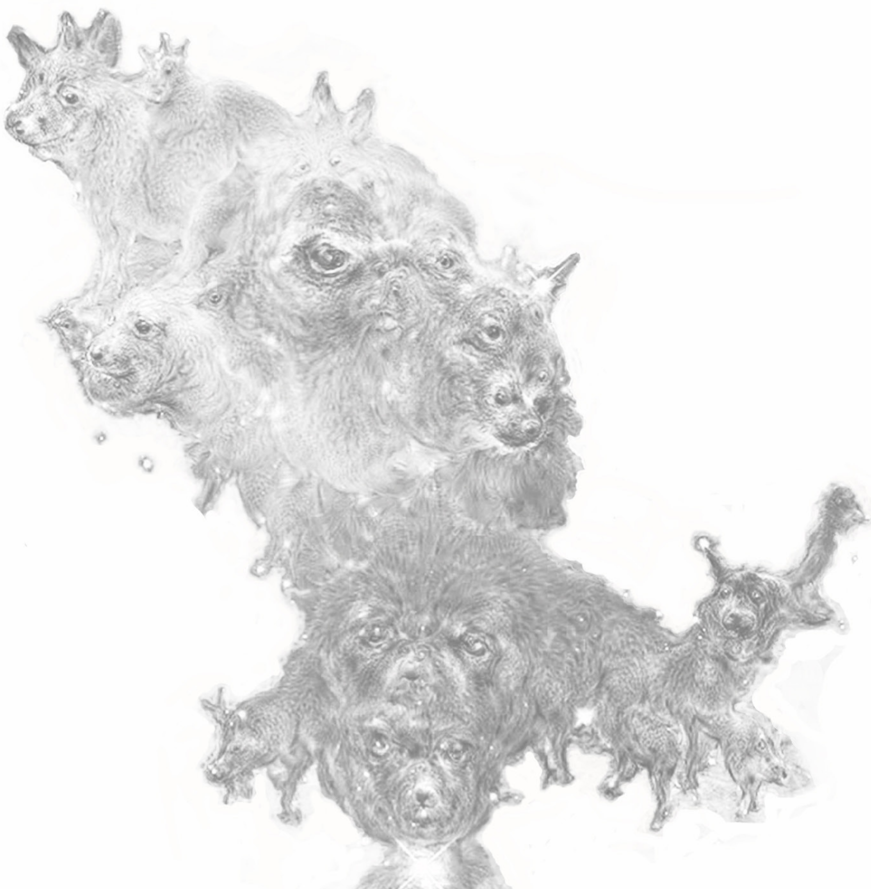
I N S Δ M

JOURNAL OF CONTEMPORARY MUSIC, ART AND TECHNOLOGY



Book review: Jorge Solis and Kia Ng (Eds.), *Musical Robots and Interactive Multimodal Systems* (Springer Tracts in Advanced Robotics 74), Berlin Heidelberg: Springer Verlag 2011, ISBN 978-3-642-22291-7

Jelena Čupić

INSAM Journal of Contemporary Music, Art and Technology
No. 2, Vol. I, July 2019, pp. 144–148.

Jelena Čupić*
University Archive of University of Music and Performing Arts
Graz, Austria

Jorge Solis and Kia Ng (Eds.), *Musical Robots and Interactive Multimodal Systems* (Springer Tracts in Advanced Robotics 74), Berlin Heidelberg: Springer Verlag 2011, ISBN 978-3-642-22291-7

In the era of technology and artificial intelligence, it is not a question of if, but, rather, when will robots become part of our everyday life. Between dystopian and utopian ways of seeing this robotic future, it is definitely more appealing to choose to side with the utopian perspective. Bruno Siciliano, the editor of the *Springer Tracts in Advanced Robotics (STAR)* stresses in his Foreword that *STAR* is devoted to bringing the most recent advances in the robotics field to the entire research community. The collection of scientific papers published in this volume, *Musical Robots and Interactive Multimodal Systems,* edited by Jorge Solis and Kia Ng are the first in the series to cover the subject of musical robotics, a new emerging field of human-robot interaction. The volume consists of fifteen chapters divided into three parts: one introductory chapter[1] and two sections with seven chapters each.

The preface states that the fourteen carefully selected scientific contributions should "[…] highlight cutting edge research related to […] exploring musical activities, interactive multimodal systems and their interactions with robots to further enhance musical understanding, interpretation, performance, education and enjoyment." The chapters are thematically organized, with the first section being *Understanding Elements of Musical Performance and Expression*.[2] This section

*Author's contact information: jelenacupic@gmx.at

1 Jorge Solis and Kia Ng, "Musical Robots and Interactive Multimodal Systems: An Introduction", pp. 1–12.

2 The studies in this section are: Rolf Inge Godøy, "Sound-Action Chunks in Music", pp. 13–26, Fabrizio Agenti, Paolo Nesi and Gianni Pantaleo, "Automatic Music Transcription: From Monophonic to Polyphonic", pp. 27–46, Antonio Camurri and Gualtiero Volpe, "Multimodal Analysis of Expressive Gesture in Music Performance", pp. 47–66, Joseph Malloch, Stephen Sinclair, Avrum Hollinger and Marcelo M. Wanderley, "Input Devices and Music Interaction", pp. 67–84, Diana S. Young, "Capturing

concentrates on the development of multimodal systems to provide untaught and efficient human-machine interaction with the goal of conceiving more advanced methods for the analysis, modeling, and understanding of musical performance and innovative interfaces for musical expression. The second section, entitled *Musical Robots and Automated Instruments*,[3] focuses on the advancement of automated instruments and anthropomorphic robots designed to study human motor control from an engineering point of view and it aims to propose new ways of musical expression. Although the singular chapters are, as stated above, connected on a general thematic level, every one of them focuses on one specific subject and/or experiment and by doing so, they aim to cover all of the relevant aspects that make up the "bigger" section theme, in as much detail as possible, at least as far as the current technology allows.

The seven chapters in the first section are all about musical performance, what makes it unique, why it is important, how it can be better understood and improved, and what the current technology can offer to musical performance. So, the exact focus in this section varies. In his study, "Sound-Action Chunks in Music," Rolf Inge Godøy explores the connections between sound features and action features, i.e. musical features and body movement, and brings those features together, focusing on the micro- and meso-levels of the sound-action links, which, according to his proposal, manifest at the timescale of a chunk, meaning in excerpts in approximately the 0.5 to 5 seconds range, forming one sound-action chunk. The chapter "Automatic Music Transcription: From Monophonic to Polyphonic" focuses on audio analysis and transcription. Fabrizio Agenti, Paolo Nesi, and Gianni Pantaleo explain automatic music transcription as the process of analyzing a musically recorded signal, or a musical performance, and converting it into a symbolic notation or any equivalent representation. Their aim is to investigate the evolution of algorithms that help us understand music, as well as the evolution of the two models – from monophonic to polyphonic, based on the most used techniques in the recent literature, such as: Nonnegative Matrix Factorisation, Hidden Markov Models, Bayesian models, generative harmonic models, and the use of jointed frequency and time information. The authors credited MIREX (Music Information Retrieval Evaluation Exchange) as one remarkable achievement in unifying the existing approaches. The rising

Bowing Gesture: Interpreting Individual Technique", pp. 85–104, Kia Ng, "Interactive Multimedia for Technology-Enhanced Learning with Multimodal Feedback", pp. 105–126, Frédéric Bevilacqua, Norbert Schnell, Nicolas Rosamimanana, Bruno Zamborlin and Fabrice Guédy, "Online Gesture Analysis and Control of Audio Processing", pp. 127–142.

3 The studies in this section are: Eiji Hayashi, "Automated Piano: Techniques for Accurate Expression of Piano Playing", pp. 143–164, Roger B. Dannenberg, H. Ben Brown and Ron Lupish, "McBlare: A Robotic Bagpipe Player", pp. 165–178, Koji Shibuya, "Violin Playing Robot and *Kansei*", pp. 179–194, Jorge Solis and Atsuo Takanishi, "Wind Instrument Playing Humanoid Robots", pp. 195 – 214, Ajay Kapur, "Multimodal Techniques for Human/Robot Interaction", pp. 215–232, Guy Hoffman and Gil Weinberg, "Interactive Improvisation with a Robotic Marimba Player", pp. 233–252, Jorge Solis, Klaus Petersen and Atsuo Takanishi, "Interactive Musical System for Multimodal Musician-Humanoid Interaction", pp. 253–268. The studies are followed by an Author Index.

significance of the expressive gesture for music performance is the focus of the fourth chapter: "Multimodal Analysis of Expressive Gesture in Music Performance." In this paper, the research of Antonio Camurri and Gualtiero Volpe joins findings from numerous disciplines: psychology, biomechanics, computer and social sciences, and performing arts, the result is an automatic system that can classify gestures according to basic emotion categories and simple dimensional approaches, which is exceptionally important for group playing. The topics included are interaction between the performers, between the performers and conductor, as well as the interaction between performers and audience. In the joint paper, "Impute Devices and Music Interaction," by Joseph Malloch, Stephen Sinclair, Avrum Hollinger, and Marcelo M. Wanderley the understanding of playing an instrument is discussed. The authors state, that impute devices are extremely relevant to our ability to understand the playing of an instrument. The focus of the study is the design and conceptualization of the digital musical instruments (DMIs) and approaches to instrument design are presented in three different contexts: application to new music performance, use within specialized medical imaging environments, and interaction with virtual instruments. Chapter six: "Capturing Bowing Gesture: Interpreting Individual Technique" is dedicated to the string players, and more specifically, to the myriad of ways in which they control their bow. Diana S. Young focuses on the significance of the bowing parameters and introduces a measurement system for violin to accurately capture bowing techniques during realistic playing conditions. The implication being, that by capturing the individual bowing techniques, the understanding of the physical elements of performance will be improved. Right after this paper understanding the individual techniques for string players, Kia Ng introduces in his paper an interactive multimedia system for music education. The study "Interactive Multimedia for Technology-Enhanced Learning with Multimodal Feedback" focuses on the i-Maestro 3D Augmented Mirror (AMIR) which can offer offline and online feedback for technology-enhanced learning of strings. Since playing an instrument is a physical activity, the idea behind this project is to capture a performance in detail and be able to use the video capture in order to improve posture. This kind of musical robot can assist both teachers and students. The last chapter of the first section "Online Gesture Analysis and Control of Audio Processing" presents the notion of temporal mapping. In this paper, Frédéric Bevilacqua, Norbert Schnell, Nicolas Rosamimanana, Bruno Zamborlin, and Fabrice Guédy show the importance of the temporal aspects of the relationship between gesture, sound, and musical structures.

The second part of the volume concentrates on the development of anthropomorphic robots and automated musical instruments that allow us to study human motor control, facilitate the human-robot interaction from a musical point of view, and propose innovative ways of musical expression. Every chapter of this section is concentrated on a specific difficulty that instrument players have, whether it is a technical or even a biological issue. Combining these studies and observing them together, one can get a fairly clear picture of how the human body functions. As for the technical difficulties, such as those involved in "touch", a superior automatic piano is designed to produces soft tones. The development of this piano and the analysis of its actions are presented by Eiji Hayashi in the paper "Automated Piano: Techniques for Accurate Expression of Piano Playing." The next two papers, "McBlare: A robotic Bagpipe Player" by Roger B. Dannenberg, H. Ben Brown and Ron Lupish and "Wind Instrument Playing Humanoid Robots" by Jorge Solis and Atsuo Takanishi focus on human anatomy. McBlare aims to exceed the physical limits of a human bagpipe player and give composers new possibilities, and Humanoid Robots – such as flute- and saxophone-playing robots – serve to study human motor control. For these purposes, some human organs (lungs, lips, tongue, arms and fingers) were mechanically reproduced. On an emotional level, Koji Shibuya introduces a violin-playing robot and the idea of *kansei* (Japanese word similar to "feeling" or "mood"), the purpose being to develop a musical robot that can perform expressive musical sounds. However, expressive playing is not the only goal of this chapter. It is rather important, according to the author, to develop a robot that understands and can express human *kansei* to facilitate smooth human-robot communication. With the introduced topic of human-robot communication it is almost 'natural' for the volume to culminate with research on human-robot interactions. In his paper "Multimodal Techniques for Human-Robot Interaction¸" Ajay Kapur concentrates on fusing together musical gesture extraction, musical robotics, and machine musicianship, stating that by using multimodal systems for machine perception of human interaction, robots can be trained to generate a mechanical response. The studies are implemented in India to promote Indian culture and show how technology can be useful in making new music. The last two chapters of this section: "Interactive Improvisation with a Robotic Marimba Player" by Guy Hoffman and Gil Weinberg, and "Interactive Musical System for Multimodal Musician-Humanoid" Interaction by Jorge Solis, Klaus Petersen and Atsuo Takanishi focus on human-robot communication. In their paper, Guy Hoffman and Gil Weinberg describe "Shimon," an interactive robotic improvisation system for marimba-playing. All the techniques used in order to achieve the requirements of a performing robotic musician are described. Shimon is a pioneer project and the first of its kind, which uses anticipatory gesture-based methods to music viewing. The results emerged from a number of human-subject studies, testing the effect of robotic presence on the synchronization of musicians, as well as the audience's perception of the duo.

Rounding up all the separate subjects covered in the volume, the last chapter brings into focus the concept and implementation of an interactive musical system for multimodal musician-humanoid interaction. It is explained that, in order for the machines to communicate with humans, they must be able to emulate two of human's most important perceptual organs: eyes and ears. When it comes to musical interaction, a great part of the performance is still based on improvisation. The future prospects for this project are nonetheless very encouraging. The researchers are working on implementing a feature that would allow musical robots to improve their own performance by listening and analyzing the sound they produce. Still, the biggest challenge lies in the research regarding the ability of musical robots to emulate emotion and recognition through musical sounds, but the technologies are promising and the researchers are optimistic.

Musical Robots and Interactive Multimodal Systems is a very intriguing collection of papers, which shows how the limits of the human body can be easily exceeded thanks to robotics and at the same time underlines that the limits of robotics are further than we could ever imagine. The fourteen chapters give but a glimpse into the world of technology and one should stay eager to see the future developments from the respective authors, or to see what kind of new developments their work will inspire by other researchers in their respective fields.

Article received: May 30, 2019
Article accepted: June 1, 2019