Master thesis proposal

By  
Arthur Parmentier

Soundpainting language recognition for sound synthesis

# Context

* 1. Music through movement

As shown in previous proposal, several tools to create or influence sound synthesis through movements are and have been developed recently.

One example that I started my proposal with:

The HEM (Haute Ecole de Musique) of Geneva is developing with IRCAM (Institute for Research and Coordination in Acoustics/Music) a tool named [GeKiPe](https://pdfs.semanticscholar.org/4824/2fb9b50b5fc45a5683c24fa64ca939607b75.pdf) (Geste Kinect et Percussion), a gesture-based interface for audiovisual performance by Philippe Spiesser, Thomas Penanguer, José-Miguel Fernandez and Alexander Vert.  
Watch it [in action](https://www.youtube.com/watch?v=mjHkCf6NRAs) or in this more [explanatory (&old) video](https://www.youtube.com/watch?v=yH6uUcxFyFE).  
The research project led by Philippe Spiesser in Geneva will probably continue and cover other topics, for instance questions of gesture notation; a project is to be submitted early 2020 to the FNS linking the DCML & HEM.

Another interesting example of technical development for music through movement is [The MiMu gloves](https://mimugloves.com/). I have learned from their developers that they were originally made thinking about soundpainting applications. They created a tool based on the same principles, without a grammar, but rather with simple 1 dimensional controls of each synthesis parameter.

* 1. Soundpainting (SP) & semantic gesture recognition

Explained by its creator, « Soundpainting is the universal multidisciplinary live composing sign language for musicians, actors, dancers, and visual Artists. Presently (2019) the language comprises more than 1500 gestures that are signed by the Soundpainter (composer) to indicate the type of material desired of the performers. The creation of the composition is realized, by the Soundpainter, through the parameters of each set of signed gestures. The Soundpainting language was created by Walter Thompson in Woodstock, New York in 1974. »

Recent research[[1]](#footnote-0) in soundpainting and performance has shown the potential of using soundpainting not only for a frontal performance linking a composer and performers but as a communication and synthesis language [between the performers themselves](https://vimeo.com/331394981?utm_source=email&utm_medium=vimeo-cliptranscode-201504&utm_campaign=28749&fbclid=IwAR0TKa0rMbnou6Bv94mRRJG9-vL2xEmMKpE1MkIIk7E-m_ECgtRefbvYnCA) or [between the performers and machines](https://hal.archives-ouvertes.fr/hal-01973067).

A few documented attempts at using sounpainting gestures semantic recognition have been documented recently ([for instance at ESTIA](https://hal.archives-ouvertes.fr/hal-02162875/document) or [Université de Toulouse; IRIT](https://dl.acm.org/citation.cfm?doid=2636879.2636899)) but lack efficiency, grammatical processing and more advanced synthesis possibilities.

* 1. EPFL resources

At EPFL, a small artistic collective (co-founded by myself) is born around the practice and research with the soundpainting language as part of the Musical, a commission of the AGEPoly. We organized several workshops in and outside EPFL and are taking the idea further by preparing with Constance Frei (musicologist and lecturer at EPFL) and the CdH-culture the next workshop under the ARTLab in february 2020: Soundpainting & Cinéma, inviting Walter Thompson himself. The workshop will be a good start to answer some theoretical questions with Walter and advance on the definition of a simplified soundpainting grammar.

Then, several actors can play their role on this project:

* Sarah Kenderdine from the EM+ will supervise the project and finance the necessary equipment, although the lab will be of no technical support, but rather artistic/conceptual support.
* The IIG group at EPFL led by R. Boulic could be a relevant co-supervisor and give very important help/insight into the critical technical points.
* Constance Frei (EPFL-Unil) is an important external advisor since the beginning of the project.
* Robert Lieck from the DCML at EPFL will be able to help on the topic of soundpainting grammar implementation.   
  A conversation with J.C. Chappelier (expert on the topic of computational semantic/grammar processing) is also planned, but cannot be considered a regular helper.
* Sabine Süsstrunk will eventually also give a help on the topics of ML and computer vision, if necessary and relevant. As of today, it seems rather unlikely that it will be needed.
* The soundpainting group will be the main test and training group for the soundpainting gestures.

# Project description (master thesis)

* 1. Concept

The project, conducted over 17 weeks at EPFL, starting mid-february, would consist in creating a sound synthesis tool that respond to simple soundpainting queries and that can later be trained to recognize additional soundpainting signs and concepts.

It divides into the following steps:

* A motion tracking system allowing for precise finger/hand gestures tracking and a broader full body tracking, in a zone of approximately 10m² corresponding to the zone of expression of the soundpainter
* A machine learning sign classifier
* A soundpainting parsing module (syntaxic analysis + sentence forming) based on the temporal serie of signs expressed by the soundpainter
* A simple sound synthesis module reacting to the parsing module.

The resulting tool can be thought as a kind of “virtual instrument” in which settings or requests are not being formed with a physical interface in the form of an object, nor potentiometers, keys etc… but rather a sign language.

* 1. Hardware and tools

The motion tracking system will be based on a glove finger tracking system (Ntoitom Hi5 gloves OR VMG 8 gloves - will be decided soon), and a kinect (V2) for the full body tracking. One of the challenges will be to combine the output of the two sensors inside unity, although they can at first be considered two independent inputs.

The ML classifier will be based in the “[InteractML](http://interactml.com/)” tool. Other tools have been already reviewed in case of failure with InteractML (unlikely).

The parsing module will probably be implemented in Max/MSP with python, or directly in Max/MSP (not tested yet).

The sound synthesis module will be implemented with Max/MSP and optionally Ableton Live/Reaper.

* 1. Goals

At the end of my project, the soundpainter will be able to send request to a computer that creates sounds with certain parameters. In a way, the computer will react specifically to the soundpainter while as in most of soundpainting performances, several degrees of freedom are left to the improvisers; here, to the synthesis module.

More specifically, the soundpainting should be at least able to:

* Specify 2 or 3 types of instrument he would like to play with based on the WHO gestures
* Choose between 2 or 3 different contents (eg. patterns, long tones…)
* Modify dynamically the volume, spatialization and temporality of the sounds
* Define when (right now and slow entrance)
* Make use of both the symbolic and mechanical gestures interpretations offered by the soundpainting language

The final presentation would contain a short commented performance using the tool that show its capabilities.

* 1. Ideal calendar

|  |  |
| --- | --- |
| Weeks | Description |
| 0 (10/02/2020) | Soundpainting workshop @ EPFL. Work on soundpainting grammar |
| 1+ (17/02/2020) | First weeks of the master thesis. Gloves and kinect acquisition and first tests with InteractML, Wekinator, Max/MSP to build a simplistic synthesis module. Getting to have a working, simple classifier with 2 hand gestures and 2 different sound outputs. Ideally, having the kinect working with the gloves already (both inputs). |
| 3+ | Starting the ML training of soundpainting gestures. Performance assessment and improvement of the ML module. Kinect+glove combination. |
| 5+ | Starting the implementation of a simple soundpainting parser based on the classified signs. Playing with simple requests and tweaking the synthesis module to play more interesting contents/sounds than simple notes, according to the possible requests. Building a simple video feedback. |
| 12+ | At this point, ideally, we should already be able to recognize simple requests and have an artistically interesting tool, that can take sound inputs in different ways and switch instruments. Starting overall performance assessment, preparation of a small demonstrative performance. Starting implementing complex soundpainting rules (modes). |
| 15+ | Preparing final report & documentation. Optimization of overall performance and result quality. (Optional) Building a more advanced video feedback. |
| 17+ | End of the master thesis at EPFL. Starting the [Cargo Bike Band](https://www.epfl.ch/about/sustainability/fr/communaute/act4changelab/les-projets/cargobikeband/) 1 month trip through CH and FR, using the new tool! Augmenting the sound synthesis possibilities for band/live applications and continuous optimisation of the setup during the CBB tour. |
| September | Project defense |

# Further extensions

The development of this tool would make possible several analyses that are interesting for future research.

* By capturing soundpainting sequences and their interpretation by the program, one will be able to analyze the linguistic content and grammars of the gestures that are used, should they have semantic content or not. From this data, a statistic interpretation of style and grammars in soundpainting could be inferred, moreover explicitly showing its use of non-semantic content as part of the collaborative creation process.
* Throughout the use of immersive visualization contexts, one could build a tool that makes use of panoramas and domes as a visual interface between the artists and the synthesis tools. Inside a panorama, one could instead navigate into a map of artistic contents (sounds, visuals etc) and trigger dynamically the desired content; having access to a possibly huge palette of sounds, instruments and other artistic contents.
* The tool could be the basis for the study of the soundpainting generative properties and explore how this system combines and makes use implicitly of both
  + Semantic elements (as the American Sign Language)
  + Implicit cultural representations (mapping low-high frequency/volume to coordinates, temporal spacing to visual spacing, the « style » of a gesture, strong movement for a loud stroke… making use of implicit synthesis references and processes)

# Conclusion remarks

The project has been highly technically revisited from its first version to the new one. The second version is now based on more efficient tools that are independent from external structures. It has been raised that this project is ambitious and require extensive technical effort.  
I realize that I am not expert in any of the technical fields used in my project, but I think that y broad background on all of these fields will allow me to solve on my own many challenges and advance autonomously.  
Being a percussionist, soundpainter and EPFL student myself, this project is to me at the meeting point of several of my interests and passions. The key ingredient for me to be confident in being able to achieve what I am proposing here is the huge amount of motivation I have for building this tool and using it in my own artistic field, as soon as this summer with my “Cargo Bike Band” project.  
I would like to thank Constance Frei and Lorenzo Cantelli for their important help in designing and reshaping this project, Loni Mahé for the horizon he has opened with our soundpainting group and all the passionate discussion we had on the topic.

1. For instance, take a look at [Conducting with the body](https://www.facebook.com/ConductingWithTheBody/) or [similar duets](https://www.facebook.com/groups/4383252451/permalink/10156376454242452/), [Audrey Vallarino and the Tours soundpainting orchestra](https://www.facebook.com/ToursSoundpaintingOrchestra/videos/vb.401085769989676/1217561198395241/?type=2&theater), [Col·lectiu Free't. Soundpainting](https://www.facebook.com/pg/colectiufreet/photos/?tab=album&album_id=467792383556650) [↑](#footnote-ref-0)