

T2M: Sonifying Text Using Audio Tags

Bruno Mesz^{1*}, Lucas Samaruga²

¹ MUNTREF Arte y Ciencia – Universidad Nacional de Tres de Febrero
Valentín Gómez 4752, Caseros, Buenos Aires, Argentina, (B1678ABH)

² EUDa, Universidad Nacional de Quilmes – CONICET
Roque Sáenz Peña 352, Bernal, Buenos Aires, Argentina (B1876BXD)

bruno.mesz@gmail.com, samaruga.lucas@gmail.com

Abstract

We present T2M, a generative text-based music system. T2M takes words from tweets produced by the audience or Twitter™ trending topics and composes music using Freesound™ sound-files labeled with those words as raw material.

1. Introduction

T2M explores intra-sonic and extra-sonic meaning of sounds given by tags of audio files in the Freesound database for generating soundscapes from textual input. The motivation for this installation was our work on music-taste crossmodal correspondences [1], which studied musical representations of the extra-musical semantic domain of taste words.

Previous systems mapping text to music are for example [2] and [3]. Semantics of music in connection with language is a vast topic, see for instance [4]. Language is considered an important factor in crossmodal associations [5].

2. T2M SuperCollider Quark

The base implementation for this installation [6] is made with SuperCollider [7] as an external library (quark in SC jargon). It consist in a set of classes that organizes information retrieval, both from Freesound [8] and Twitter [9] using external command-line utilities on Linux, processing and storage of that information, sequencing, sound synthesis and visual feedback of original messages and retrieved sounds.

An autonomous T2M execution loop is schematically as follows: tweets query, text parsing and storage, sound query based on selected words, sound-files pull and storage, sound processing and reproduction and visual presentation of information.

Freesound querying and sound-file download depends on the Freesound.sc quark [10] which is a higher level interface to communicate with the Freesound API from within slang (SC programming language). Twitter CLI [11] program is used for querying tweets using Twitter's API. A basic text parsing consisting in noun and verb selections is performed by a command-line utility provided by Apertium [12]. Because performance attendees were Spanish-speaking and due to that Freesound metadata is predominantly in English the translation platform Aперitum was introduced to locally translate information before querying.

2.1 System Description

Being an interactive performance, the latency of the involved systems with respect to the attendees input was an important factor to consider regarding sound and visual feedback as an aesthetic proposal. The latency of the system relies mainly on Twitter server's refresh rate and sound-file download time. Former latency is not possible to control but sound data transfer time was reduced by limiting sound-file length by search, downloading compressed (ogg) files and reusing duplicates.

These traits provide processing

* Supported by Programa de Investigación Sistemas Temporales y Síntesis Espacial en el Arte Sonoro, EUDa, UNQ.

constraints that affect possible outcomes. To work with these system the idea was to use the metaphor of a diffuse cloud of sounds of things happening within a dilative time span considered as the present (i.e., 30 seconds). A given input from attendees do not generate instant discrete sound events but contributes to a sound landscape which runs as far as there is new input data in which the present carries the near past. Because the installation uses flexible network APIs data search was possible either by defining specific tags, so input come from the present public, or external phenomena such as trending topics. The cloud metaphor applies consistently to both cases.

Sound synthesis is based on temporal granularity and spatial diffusion (originally stereo but it can be easily expanded for different setups). Due to some arbitrariness implicit in the search by (user provided) metadata and the unstructured nature of the Freesound database, retrieved sounds were often as different as complex sequences, e.g., field recordings, instrumental samples and loops or, in the most annoying cases, plane sinusoids with full amplitude. Thus the reproduction of these raw materials was transformed by random selection of different moments panned to different positions with different gain. In this way, the resulting mix of different sound-files creates a continuous diffuse texture which varies slowly with the input data evolution.

3. Performance at the MAR

We presented T2M at the MAR museum in Mar del Plata, Argentina. Four specially designed perfumes were successively spread in the auditorium, and the audience was required to tweet any associations evoked by the aromas. The tweets were displayed on a screen while T2M generated a sonic atmosphere.

4. Conclusions and Future Work

In this installation, different sets of information are set up in real time to reflect an ongoing human experience of the world.

After generating an olfactory atmosphere in the room, we ask the audience to exteriorize, via Twitter, free verbal associations produced by the time-varying smell. The installation then elicits a “music of concepts”, drawing on symbolic meaning generated by arbitrary user defined tags from Freesound that match the words in the tweets.

Future plans include generating soundscapes for a wine tasting and from Twitter trending topics. Additionally, we will implement the analysis of the tweets using natural language processing methods, to gather information on the role of semantics in crossmodal associations [5].

References

- [1] Mesz, B., Trevisan, M. A., & Sigman, M. *The taste of music. Perception*, 40(2), 209-219. 2011
- [2] http://www.erikbunger.com/html/let_them_sing.html
- [3] <http://melobytes.com/app/melobytes>
- [4] Zbikowski, L. M. *Conceptualizing music: Cognitive structure, theory, and analysis*. Oxford University Press. 2002.
- [5] Spence, C. *Crossmodal correspondences: A tutorial review. Attention, Perception, & Psychophysics*, 73(4), 971-995. 2011.
- [6] <https://github.com/smrg-lm/ttm>
- [7] McCartney, J. *Rethinking the Computer Music Language: SuperCollider*. *Computer Music Journal*, 26(4), 61-68, 2002.
- [8] <http://www.freesound.org>
- [9] <https://twitter.com>
- [10] <https://github.com/g-roma/Freesound.sc>
- [11] <https://github.com/sferik/t>
- [12] http://wiki.apertium.org/wiki/Main_Page

Art

