**JOHNSON-ROBERSON — Susurrant**

<0 images>

**Susurrant: A Tool for Algorithmic Listening in Networked Soundscapes**

Johnson-Roberson, C.

Many methods for computational analysis of audio have been developed in the fields of music information retrieval and machine learning. Typically used for copyright protection and music recommendation, such methods can also be used to analyze audio artifacts embedded in a specific online community.

Susurrant, an open-source tool I am currently developing, enables exploratory data analysis on a corpus of digital audio accompanied by textual and social network metadata. (Such a corpus can be derived from a music-centered social media platform like SoundCloud.) This tool utilizes probabilistic topic modeling, a family of techniques for inferring latent variables (‘topics’) that could have generated the observed data (Blei, 2012).

Sonic features (both rhythmic and spectral content) and the text of user tags and comments are analyzed jointly in a single model. The result is a kind of auditory concordance for the corpus, linked directly to associated textual features. This builds upon work using topic models to identify latent sources in audio (Hoffman et al., 2009), model musical influence (Shalit et al., 2013), and analyze shared taste in online communities (Dietz, 2009).

Susurrant is meant to help researchers gain a new perspective on their audio corpora. It facilitates what Wendy Hsu terms ‘augmented empiricism’, a combination of ethnography and computational analysis that aims to ‘[document] social and cultural processes with empirical specificity and precision’ (Hsu, 2014). In this short paper, I will give a brief demo of the software (available at www.susurrant.org) and suggest ways that it might be useful for other scholars working with sound.

**Methodology**

Susurrant comprises several components: a database for storing audio metadata with textual data and social network graph, a script for semi-automated data collection, and a browser-based application that provides both visual and auditory display of analysis results. For the most part, these components make use of existing software packages such as the Neo4j graph database, the Essentia feature extraction library (Bogdanov et al., 2013), and the machine learning library MALLET (McCallum, 2002); thus, the visualization interface constitutes the bulk of this work’s contribution.

The browser-based visualization/sonification of the corpus lets one listen to a resynthesized version of the variables characterizing each audio ‘topic’ along with representative audio samples and user comments. In another mode, it displays the subsets of the social network most closely associated with each topic, enabling comparison of the distributions of different sonic features across the community.

I present a case study using an initial corpus of audio files and user commentary downloaded from SoundCloud. This corpus consists of music played in nightlife spaces that cater to queer and trans people of color, and is an integral part of my ethnographic work on queer of color nightlife and its online mediation. I will show how I am using this software to analyze sound in a specific social and cultural context.

**Theoretical Implications**

As well as offering a software package to facilitate multimodal analysis, this work can contribute to the theorization of ‘algorithmic listening’, a term for modes of computational analysis that have mostly been used in commercial applications but could readily be put to other uses.

We are already surrounded by algorithms that listen. For the most part, these algorithms act as censors (as in YouTube’s Content ID system) or as recommendation engines (like those of Pandora, Clear Channel’s iHeartRadio, or SoundCloud). They rely on massive data stores and computing resources that are inaccessible to the end user, stripping away context and rendering their operations opaque even as they come to shape more and more of what we hear.

Instead, Susurrant captures a cross-section of meaningfully related sounds as they exist within a specific interpretive community. It analyzes audio and textual features together and presents them in an integrated fashion to the researcher, ensuring that important contextual details are retained. Further, it helps us understand how an algorithm ‘hears’, when we listen to the variables learned by the model together with the original samples from which they were inferred.

**Conclusion**

The tool provides a novel means for exploratory data analysis of sound. By using sonic as well as textual data, this method calls attention to the aspects of sounds that have significance within an interpretive community and allows for exploration along different axes than those provided by extant platforms. I welcome dialogue about ways to improve this tool and extend it for other use cases.

**References**

**Blei, D. M.** (2012). Probabilistic Topic Models. *Communications of the ACM,* **55**(4), 77–84, http://dl.acm.org/citation.cfm?id=2133826.

**Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., et al.** (2013). ESSENTIA: An Audio Analysis Library for Music Information Retrieval. *International Society for Music Information Retrieval Conference (ISMIR’13)*, pp. 493–98.

**Dietz, L.** (2009). Modeling Shared Tastes in Online Communities. In *NIPS Workshop on Applications for Topic Models: Text and Beyond*. http://www.umiacs.umd.edu/~jbg/nips\_tm\_workshop/8.pdf.

**Hoffman, M., Blei, D. and Cook, P. R.** (2009). Finding Latent Sources in Recorded Music with a Shift-Invariant HDP. In *Proceedings of the 12th International Conference on Digital Audio Effects*, Como, Italy, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.154.845&rep=rep1&type=pdf.

**Hsu, W. F.** (2014). Digital Ethnography toward Augmented Empiricism: A New Methodological Framework. *Journal of Digital Humanities,* **3**(1),

http://journalofdigitalhumanities.org/3-1/digital-ethnography-toward-augmented-empiricism-by-wendy-hsu/.

**McCallum, A.** (2002). MALLET: A Machine Learning for Language Toolkit. http://mallet.cs.umass.edu.

**Shalit, U., Weinshall, D. and Chechik, G.** (2013). Modeling Musical Influence with Topic Models. In *Proceedings of the 30th International Conference on Machine Learning*, Atlanta, 16–21 June 2013, pp. 244–52.