Implicit and explicit judgements on the melodic similarity of cases of plagiarism and the role of computational models

Anna Wolf,[*1] Daniel Müllensiefen[#2]

[*]*Hanover Music Lab, Hanover University of Music, Drama and Media, Germany*
[#]*Department of Psychology, Goldsmiths College, University of London, United Kingdom*
[1]`anna.wolf@hmtm-hannover.de`, [2]`d.mullensiefen@gold.ac.uk`

ABSTRACT

Background

Computational similarity measures have proven to be invaluable in the classification, retrieval and comparison of melodies (e.g. Eerola & Bregman, 2007). A commercially very relevant application is their use in cases of musical plagiarism (Müllensiefen & Pendzich, 2009; Cason & Müllensiefen, 2012). However, apart from a few notable exceptions (e.g. Müllensiefen & Frieler, 2004) there is surprisingly little psychological evidence to validate the cognitive adequacy of the proposed algorithms.

Aims

The aims of this study are to model human similarity perception and to assess how closely human perception and computational measures match actual court decisions.

Method

In an implicit memory paradigm participants were exposed to 20 melodies while performing cover tasks. In a subsequent test phase participants listened to 30 melodies (15 similar to melodies from the initial phase, 10 neutral, 5 identical) to identify which ones they had heard before. The dependent variable was the confusability between similar melodies. For the explicit ranking task participants had to rank four melodies compared to a target.

For both tasks we used melodies from court cases from the US and the Commonwealth. Participants were 36 adults mainly from a student population with a wide range of musical expertise.

The similarity measures have been selected from Müllensiefen & Pendzich (2009) and complemented with the Earth Mover's distance measure (Typke, Wiering & Veltkamp, 2007).

Results

In the implicit task, participants' judgments agreed fairly well with the courts' decision on copyright infringement (AUC of .70). Many of the computational measures of similarity correlate highly with the participants' data, such as a Tversky (1977) feature-based measure (r = .59) and a duration-weighted Edit Distance (r = .51).

For the explicit task the average inter-participant correlation was .30 (as measured by Kendall's τ, sd = .19), which was compared to the mean correlation of one measure with all participants. Again, the weighted Edit Distance ($\tau = .23$, sd = .17) and the Tversky measure ($\tau = .22$, sd = .19) performed most similarly to the mean participants' correlation.

The court decisions are best classified by an EMD measure (AUC of .84) and the Tversky measure (AUC of .69).

Conclusions

Participants are able to distinguish between those melodies classified or rejected as plagiarism to a good degree. However, it has to be noted that, aside from melodic similarity, factors such as knowledge of either song, lyrics or the title can also significantly influence the court's decision.

Keywords

Melodic similarity, Music plagiarism, Similarity measures

REFERENCES

Cason, R. J. S., & Müllensiefen, D. (2012). Singing from the same sheet: computational melodic similarity measurement and copyright law. *International Review of Law, Computers & Technology, 26*(1), 25-36.

Eerola, T., & Bregman, M. (2007). Melodic and contextual similarity of folk song phrases. *Musicae scientiae, Discussion Forum 4A,* 211-233.

Müllensiefen, D., & Frieler, K. (2004). Cognitive Adequacy in the Measurement of Melodic Similarity: Algorithmic vs. Human Judgments. *Computing in Musicology, 13,* 147-176.

Müllensiefen, D., & Pendzich, M. (2009). Court decisions on music plagiarism and the predictive value of similarity algorithms. *Musicae Scientiae, Discussion Forum 4B,* 257-295.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327–352.

Typke, R., Wiering, F., & Veltkamp, R.C. (2007). Transportation distances and human perception of melodic similarity. *Musicae scientiae, Discussion Forum 4A,* 153-181.