

# **How Many Engineers Does it Take to Transcribe a Cretan Dance Tune?**

## **Exploring the Potential of Machine Learning for Automatic Transcription**

**Andre Holzapfel, KTH Royal Institute of Technology, Sweden**

**Emmanouil Sifakis, University of Crete**

Engineering research within the field of Music Information Retrieval (MIR) has for several decades developed software that aims at automatically obtaining music transcriptions, with strong focus on Western staff notation and piano roll representations. Transcription work in ethnomusicology, on the other hand, makes increasing use of computers in the transcription process for tasks such as formatting, but largely excludes the application of automatic music transcription models. In the present paper, we will report results from a recently published study (Holzapfel et al., 2021), in which we collected manual transcriptions of traditional dance tunes from Crete, Greece, and compared them with automatic transcriptions by state-of-the-art systems. We will provide an overview of the kinds of problems that are encountered when using automatic transcription, and discuss the most important points of divergence between human and machine transcription.

In this presentation, we go beyond those published results by investigating the influence of the most recent developments in machine learning on the quality of transcriptions. The goal in this process is to take advantage of the possibility to adapt machine learning models to the instrumental timbres in the music to be transcribed, which are Cretan dance tunes in our context. To this end, we adapted source separation models to focus on the lead instrument of Cretan music, the lyra, in order to obtain an audio signal that attenuates the volume of all other instruments. We then used this lyra-focused audio signal in several state-of-art machine learning based transcription models, and will discuss whether the approached adaptation to the instrumental timbre in a recent machine learning model brings quantitative and qualitative improvement of transcription compared to our previous study.

In all the involved evaluation, strong emphasis will be given on the provision of examples of transcription outputs, as well as audio examples of the intermediate source separation stage.

## Reference

Holzapfel, Andre, Emmanouil Benetos, Andrew Killick, and Richard Widdess. 2021.

[“Humanities and Engineering Perspectives on Music Transcription.”](#) *Digital Scholarship in the Humanities* 37(3): 747-764.