FOLK MUSIC CHALLENGE ON MUSIC METER ESTIMATION: PREDICTION USING LOGISTIC REGRESSION MODELS ON *QUANTIZED LAG VECTORS*

Pierre Beauguitte
Dublin Institute of Technology
pierre.beauguitte@mydit.ie

Bryan Duggan
Dublin Institute of Technology
bryan.duggan@dit.ie

John D. Kelleher
Dublin Institute of Technology
john.d.kelleher@dit.ie

ABSTRACT

This report describes our entry in the Folk Music Challenge on Music Meter Estimation, where the goal is to automatically estimate the musical meter of a recording of Greek traditional music. Our approach is described in details in (Beauguitte et al., 2018).

1. METHOD

In (Beauguitte et al., 2018) we presented a new method, based on signal processing techniques and machine learning, to infer rhythmic information from recordings of Irish traditional music. Our task was related to, but distinct from, music meter estimation. In particular we wanted to be able to distinguish between two different rhythms (reels and hornpipes) that share the same $\frac{4}{4}$ time signature. The musical genre we considered did not include any assymetric meters. Nonetheless, it appears that our method can be applied with some success to Greek music.

Our method can be summarized as follows:

- an onset detection function is computed by spectral difference on a Bark spectrogram of the audio file;
- auto-correlation function (ACF) of the obtained curve is performed on a sliding window;
- the ACF curve is smoothed with a Gaussian kernal, and all its peaks (local maxima) are found;
- quaver duration q is obtained by applying the *fuzzy histogram* algorithm [1] to the set of peaks;
- the peaks are grouped in sets

  $$P_i = \{p \in P \text{ where } \mathit{round}(p_l/q) = i\}$$

  and the *quantized-lag vector* is defined by

  $$ql_i = \begin{cases} \left(\sum_{p \in P_i} p_v\right) / |P_i| & \text{if } P_i \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

- a multinomial logistic regression model is fitted to predict the musical meter from the *ql* vector.

[1] in the context of Irish music, for which this algorithm was primarily intended, the obtained duration does correspond to the quaver. This might not hold for Greek music, but for consistency we decide to keep the terminology of our original publication

2. DATA

Machine learning algorithms typically need large amount of training examples to work well. The public dataset for the challenge is fairly small; in particular, it contains only one instance of the $\frac{11}{8}$ class. Hence we decided to collect a larger training set. All of the added audio tracks are available on YouTube, and the list is provided alongside the code. Because none of the authors of the present paper are familiar with Greek music, the list of songs given in (Fouloulis et al., 2013) was a precious resource to collect these supplementary data.

As in our study on Irish music, the models are trained in a balanced manner: all instances are given a weight inversely proportional to the relative frequency of their class, resulting in each class having an equal importance in the loss function. In addition to this, a second weight is added so that the public dataset has the same importance as the one we curated ourselves. The former is much smaller than the latter, hence less diverse, but also certainly of better quality as it was designed by experts of this musical genre.

3. FINE-TUNING

Some parameters of our method were fine-tuned to obtain the best possible predictions.

- The sliding window for computing the ACF function is 10 seconds long, instead of 5.
- The "fuzz" parameter of the fuzzy histogram, allowing to match durations up to a certain ratio, is 0.25 instead of $1/3$.
- The *ql*-vectors are of size 64 instead of 16, and the P_i sets are now defined as

  $$P_i = \{p \in P \text{ where } \mathit{round}(2 * p_l/q) = i\}$$

  The 2 factor in the numerator allows for rounding at "semiquaver" positions.

The first modification is a consequence of the sampling rate used in the challenge (22.05kHz), twice lower than the one used in our study. The other two were obtained by performing a grid search, using the public dataset as a validation set.

4. REFERENCES

Beauguitte, P., Duggan, B., & Kelleher, J. D. (2018). Rhythm Inference from Audio Recordings of Irish Traditional Music. In *Proceedings of the 8th International Workshop on Folk Music Analysis*.

Fouloulis, T., Pikrakis, A., & Cambouropoulos, E. (2013). Traditional asymmetric rhythms: a refined model of meter induction based on asymmetric meter templates. In *Proceedings of the Third International Workshop on Folk Music Analysis*, (pp. 28–32).