EXPLOITING ONLINE RESOURCES TO IMPROVE CHORD RECOGNITION ACCURACY

Matt McVicar
University of Bristol

Tijl De Bie
University of Bristol

Introduction. In this late-breaking paper we demonstrate the potential for using online chord annotations maintained by guitar enthusiasts to enhance chord recognition performances. A more comprehensive explanation of our method can be found [1]. Methods for chord estimation have in recent years focused on chromagram feature vectors and either Hidden Markov Models (HMMs), or more general Dynamic Bayesian Networks [2], [3]. The website e-chords.com features chord annotations written above the lyrics of songs. The times of onset or duration of chords is not present, so we refer to such chord sequences as *Untimed Chord Sequences* (UCS's). We scraped the chord transcriptions for 154 Beatles songs from e-chords and noticed that several of these songs had multiple UCS's. These may have been uploaded by a different user, or be transposed into a more guitar-friendly key. We shall refer to these multiplicities as *redundancies*. We found 26 Beatles songs with 1 or more redundancy, so took these as our test set and trained an HMM on the remaining 154 songs in The Beatles' catalogue. The ground truth chord annotations were kindly provided by Chris Harte and the alphabet was restricted to 24 major/minor chords and a 'no-chord' symbol.

Methodology. We used the UCS's in three different ways to see how they performed over standard Viterbi decoding. First, we only allowed the Viterbi algorithm to output chords from the UCS (Alphabet Constrained Viterbi, ACV). Secondly, we only allowed chord transitions which occurred in the UCS to be predicted (Alphabet and Transition Constrained Viterbi, ATCV). Lastly, we restricted the Viterbi algorithm further by only allowing the chords from the UCS to be predicted, in the same order as they appear in the UCS. This method is analagous to aligning the UCS to the chromagram, so we refer to it as Untimed Chord Sequence Alignment (UCSA).

Evaluation methods. As we mentioned already, all the test set songs had one or more redundancy, and some may be in the incorrect key. To counteract this we transposed the UCS's into each key and chose the best redundancy and key using two methods. Firstly, we chose the (redundancy, key) pair which maximised performance. Secondly, we chose the pair which had the best likelihood score in the Viterbi decoding (note that this can be done without knowledge of the ground truth). We refer to these methods as Best Accuracy (Acc) and Best Likelihood (Lik) respectively. We also took the ground truth stripped of all repetitions as a UCS, which represents the situation where e-chords is entirely noise-free and we label as True UCS.

Results. Our results can be seen in the table below. Row 1 shows the impressive potential of using these additional data sources. A more realistic setting is seen in the rows 2-3, where we see improvements on standard Viterbi for the ACV and ATCV methods. Best Accuracy UCS selection only slightly outperforms the best Likelihood UCS, suggesting that best Likelihood is a suitable proxy for UCS choice. We noticed that aligning e-chords UCS's to audio did not offer an improvement over Viterbi. Upon investigation this was because although the UCS shared many chords with the ground truths, text comments from the website such as *'Play verse chords twice'* were not understood by our scraper. We are currently working on a solution to this problem by relaxing the alignment process in such a way that groups of chords can be repeated with small probability.

	Viterbi	ACV	ATCV	UCSA
Using the true UCS from the ground truth				
True UCS	77.03%	81.18%	84.53%	88.00%
Using two or three redundancies				
Acc UCS	77.03%	80.41%	81.96%	71.60%
Lik UCS	77.03%	79.65%	81.63%	71.60%

1. REFERENCES

[1] M. McVicar, T. De Bie. Enhancing chord recognition accuracy using web resources. *3rd International Workshop on Machine Learning and Music*. Firenze, Italy, 2010.

[2] Y. Ueda, Y. Uchiyama, T. Nishimoto, N. Ono, and S. Sagayama. HMM-based approach for automatic chord detection using refined acoustic features. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP*, 2010.

[3] M. Mauch, S. Dixon. Approximate note transcription for the improved identification of difficult chords. *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, 2010.