Melody classification using patterns

Darrell Conklin

Department of Computing
City University London
United Kingdom
conklin@city.ac.uk

Abstract. A new method for symbolic music classification is proposed, based on a feature set pattern representation. An ensemble of sequential patterns is used to classify unseen pieces using a decision list method. On a small corpus of 195 folk song melodies this method achieves a good classification accuracy of 77%, though with only a 43% recall rate. A competing n-gram model achieves a slightly lower accuracy of 75% though with the advantange of 100% recall rate. The new proposed method may be applicable to a wide variety of music classification problems.

1 Introduction

Melody classification is an important task in music informatics. The ability to assign unseen melodies to existing classes — by composer, region, genre, period, etc. — is a powerful inference capability. Methods can be divided into two main types [1]. The *global feature* approach encapsulates information about a melody into a feature vector, where standard attribute-value data mining methods can be employed. The *event feature* approach views a piece as a sequence of events, each with its own features. The standard method to deal with such sequential data is the n-gram model and its extension, the multiple viewpoint model [2].

Bridging the global and event methods is the *sequential pattern* approach, which uses short patterns to classify an unseen melody. Patterns on one hand can be viewed as global boolean features of pieces, but on the other hand they retain some sequential structure of events.

A few preliminary studies exist on the use of short sequential patterns for music classification [3–6]. All of these methods use a pattern vocabulary comprising fixed-length n-grams, and most do some selection to reduce the exhaustive pattern set to those patterns that are predictive of the class. The problem with relying on fixed n is that short patterns may not be predictive of any class (may occur with relatively equal frequency in all classes), and long patterns have the danger of overfitting the training corpus, leading to limited matches of the pattern collection to unseen pieces.

The use of a collection of (non-sequential) patterns for classification is an ongoing area of research in *class association rule* mining [7]. Class association rules are simply association rules that are restricted to contain a class attribute in the right-hand side of the rule. Of general interest is how an exhaustive collection

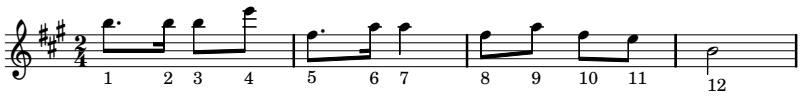
in MML 2009: International Workshop on Machine Learning and Music, Bled, Slovenia, 37-41, 2009.

of association rules can be pruned to an ensemble of rules that are together useful for future classification. Also of interest is the development of theory and methods for pruning, ranking and weighting rules to attain a final class prediction from the collection [8]. Ranking of rules is usually based on rule *confidence*: the probability of the class given the pattern.

This paper describes and evaluates a novel classification scheme for music based on feature set patterns. These are sequential patterns, of arbitrary length with arbitrary feature conjunctions within their components, thereby permitting a high degree of abstraction and flexibility. Extending the work presented in [9], a collection of mined feature set patterns is used to classify folk song melodies. For a small trial corpus of Swiss and Austrian folk songs, the proposed method attains 77% classification accuracy, though with only a 43% recall rate, as evaluated by 10-fold cross-validation.

2 Methods

A *pattern* is a sequence of feature sets. Patterns can be viewed as unary predicates: functions from pieces to boolean values. A piece w *instantiates* a pattern P, written $P(w)$, if the pattern occurs (possibly multiple times) in the sequence: if the components of the pattern are instantiated by successive events in the sequence. For example, the pattern $[\{\texttt{pcint:2}, \texttt{dur:3}\}, \{\texttt{pcint:3}\}]$ has two components and occurs at position 5 of the melodic fragment in Figure 1.



	1	2	3	4	5	6	7	8	9	10	11	12
contour(pitch)	⊥	=	=	+	−	+	=	−	+	−	−	−
contour(dur)	⊥	−	+	=	+	−	+	−	=	=	=	+
rest	⊥	0	0	0	0	0	0	0	0	0	0	0
pcint	⊥	0	0	5	2	3	0	9	3	9	10	7
diaintc	⊥	P1	P1	P4	M2	m3	P1	M6	m3	M6	m7	P5

Fig. 1. Example of a fragment of a Shanxi folk song (Essen folk song database code J0002), and a viewpoint matrix.

A pattern P *subsumes* (is more general than) a pattern Q if all instances of Q are also instances of P: if $\forall w\, Q(w) \Rightarrow P(w)$ is valid. This subsumption relation can be determined by searching for a contiguous, subsumption-preserving mapping from the components of P to the components of Q. For example, the pattern $[\{\texttt{pcint} : 3\}]$ subsumes the pattern $[\{\texttt{pcint} : 3, \texttt{dur} : 1\}]$, which in turn subsumes, for example, the pattern $[\{\texttt{pcint} : 2\}, \{\texttt{pcint} : 3, \texttt{dur} : 1\}, \{\}]$ (note the empty feature set in the last component, which is instantiated by any event).

The *piece count* of a pattern P in a class $\oplus$ is the number of pieces containing the pattern and in the class, notated by $c^{\oplus}(P)$. A *class association rule* is a structure $P \Rightarrow \oplus$ where P is a pattern and $\oplus$ is a class. The *confidence* of a rule is the quantity $p(\oplus|P)$, estimated from counts in the training corpus.

For classification tasks, we seek out patterns that have high confidence: $p(\oplus|P)$ greater than a specified threshold in the training corpus. The mining algorithm performs a general-to-specific search of a virtual subsumption taxonomy of patterns, terminating a branch with success when a confident rule is reached. The algorithm therefore finds a decision list of patterns that are both confident and maximally general (not subsumed by any other confident pattern).

To classify an unseen query piece w, it is probed against each rule in a decision list, looking for the first rule $P \Rightarrow \oplus$ such that $P(w)$, then predicting the class $\oplus$ for w.

To mine for a pattern decision list, the following method is used. For each class $\oplus$, the corpus is mined with the specified minimum confidence $p(\oplus|\cdot)$ and piece count $c^{\oplus}(\cdot)$, using all other classes as the anticorpus. The collection of all results for each class is used to build a final decision list, sorted from high to low confidence.

2.1 Folk song dataset

Two geographic regions represented in the Essen folk song database [10] were chosen to illustrate the method: Austrian (102); and Swiss (93). These were chosen as a pair due to their geographic proximity and presumed suitability as anticorpora for one another. The 5 viewpoints presented in Figure 1 were used in the experiment.

3 Results

To evaluate the classification method, a 10-fold cross-validation is performed. The order of pieces in the complete corpus is scrambled, then for each fold, 90% of the dataset is used to build a pattern decision list: for each class, split into a corpus and anticorpus and mined for patterns. The remaining 10% of the pieces are then classified using the resulting decision list.

In this initial evaluation of the method, a pattern P is retained for the decision list if $p(P|\oplus) \geq 0.1$ (P occurs in at least 10% of the corpus) and $p(\oplus|P) = 1$ (the rule is maximally confident).

The pattern classification method achieves 77% classification accuracy on the complete corpus, though 111 pieces are not classified by a rule, giving a recall rate of only 43%. Only short patterns were necessary to correctly classify instances: length 1 (2 patterns); 2 (34); 3 (10); 4 (16), and 5 (3). To put the results into some context, a zero-rule classifier (always predicting the most frequent class) achieves $102/195 = 52\%$ classification accuracy. A trigram model of linked interval/duration pairs achieves a slightly lower 75% classification accuracy in a 10-fold cross-validation on the complete corpus. A classifier using

99 global features [11] with the Weka [12] logistic regression implementation, a configuration performing well in a recent study of a larger folk song corpus [1], achieves 66% 10-fold cross-validation accuracy.

Corpus	Pattern
austrian	$\left[\left\{ \texttt{diaintc} : \texttt{M3} \right\} \left\{ \texttt{diaintc} : \texttt{P1} \right\} \left\{ \texttt{diaintc} : \texttt{P1} \right\} \left\{ \texttt{contour(dur)} := \right\} \left\{ \texttt{contour(dur)} := \right\}\right]$
swiss	$\left[\left\{ \begin{array}{c} \texttt{contour(dur)} : + \\ \texttt{restv} : 0 \\ \texttt{diaintc} : \texttt{m7} \end{array} \right\} \left\{ \texttt{contour(dur)} : + \right\}\right]$
austrian	$\left[\left\{ \begin{array}{c} \texttt{contour(dur)} := \\ \texttt{diaintc} : \texttt{M3} \end{array} \right\} \left\{ \texttt{diaintc} : \texttt{P1} \right\} \left\{ \texttt{diaintc} : \texttt{P1} \right\}\right]$
swiss	$\left[\left\{ \begin{array}{c} \texttt{contour(dur)} := \\ \texttt{restv} : 0 \\ \texttt{diaintc} : \texttt{M2} \end{array} \right\} \left\{ \texttt{contour(dur)} : - \right\} \left\{ \texttt{contour(pitch)} : - \right\} \left\{ \texttt{contour(pitch)} : - \right\}\right]$
austrian	$\left[\left\{ \begin{array}{c} \texttt{contour(dur)} := \\ \texttt{contour(pitch)} : - \\ \texttt{pcint} : 11 \\ \texttt{diaintc} : \texttt{M7} \end{array} \right\} \left\{ \begin{array}{c} \texttt{contour(dur)} := \\ \texttt{diaintc} : \texttt{P1} \end{array} \right\} \left\{ \texttt{contour(dur)} := \right\}\right]$
austrian	$\left[\left\{ \begin{array}{c} \texttt{contour(pitch)} : + \\ \texttt{pcint} : 4 \\ \texttt{diaintc} : \texttt{M3} \end{array} \right\} \left\{ \texttt{diaintc} : \texttt{P1} \right\} \left\{ \texttt{diaintc} : \texttt{P1} \right\} \left\{ \texttt{contour(dur)} := \right\}\right]$
austrian	$\left[\left\{ \texttt{pcint} : 6 \right\} \left\{ \texttt{diaintc} : \texttt{P1} \right\}\right]$
swiss	$\left[\left\{ \begin{array}{c} \texttt{contour(dur)} : + \\ \texttt{restv} : 0 \\ \texttt{contour(pitch)} : + \\ \texttt{diaintc} : \texttt{m2} \end{array} \right\} \left\{ \texttt{contour(dur)} := \right\} \left\{ \begin{array}{c} \texttt{contour(pitch)} : + \\ \texttt{pcint} : 2 \\ \texttt{diaintc} : \texttt{M2} \end{array} \right\}\right]$

Table 1. The patterns of the decision list trained on the complete corpus.

Table 1 shows the pattern of the decision list trained on the complete corpus.

4 Discussion

This paper has presented a method for melody classification based on feature set patterns. As reported earlier [9], patterns can have musicological interest in isolation, and this study showed that they can also be viewed as boolean global features for describing music for supervised learning. In this sense, pattern discovery in music can be viewed as a feature selection task.

Though initial results are encouraging, the low recall rate of the method does not yet provide a viable classifier. Further work is necessary to increase the recall rate of the method, and also to apply the pattern classification algorithm to several other and larger music corpora.

References

1. Hillewaere, R., Manderick, B., Conklin, D.: Global feature versus event models for folk song classification. In: ISMIR 2009: 10th International Society for Music Information Retrieval Conference, Kobe, Japan (2009)
2. Conklin, D., Witten, I.: Multiple viewpoint systems for music prediction. Journal of New Music Research **24**(1) (1995) 51–73
3. Pérez-Sancho, C., Rizo, D., Iñesta, J.M.: Genre classification using chords and stochastic language models. Connection Science **20**(2&3) (2009) 145–159
4. Westhead, M., Smaill, A.: Automatic characterisation of musical style. In: Music Education: An AI Approach. Springer-Verlag (1993) 157–170
5. Lin, C.R., Liu, N.H., Wu, Y.H., Chen, A.: Music classification using significant repeating patterns. In: Database Systems for Advanced Applications. Volume 2973. Springer-Verlag (2004) 506–518
6. Sawada, T., Satoh, K.: Composer classification based on patterns of short note sequences. In: Proc AAAI-2000 Workshop on AI and Music, Austin, Texas (2000) 24–27
7. Liu, B., Hsu, W., Ma, Y.: Integrating classification and association rule mining. In: KDD-98. (1998) 80–86
8. Sulzmann, J.N., Fürnkranz, J.: An empirical comparison of techniques for selecting and combining local patterns into a global model. Technical Report Technical Report TUD-KE-2008-03, Technische Universität Darmstadt, Knowledge Engineering Group (2008)
9. Conklin, D.: Discovery of distinctive patterns in music. In: MML08: Machine Learning and Music, Helsinki (2008) to appear in Intelligent Data Analysis.
10. Schaffrath, H.: The Essen Associative Code: A code for folksong analysis. In Selfridge-Field, E., ed.: Beyond MIDI: The Handbook of Musical Codes. The MIT Press (1997) 343–361
11. `http://jmir.sourceforge.net/jSymbolic.html`
12. `http://www.cs.waikato.ac.nz/ml/weka/`