

Cadence Detection in Western Traditional Stanzaic Songs using Melodic and Textual Features

Peter van Kranenburg and Folgert Karsdorp - Meertens Institute, Amsterdam

How to automatically detect melodic cadences?

Data

4,120 digitally encoded Dutch folk songs. Phrase ends (cadences) are annotated.

Trigrams

Three musical staves are shown. The first staff has lyrics 'Jan Al-berts die uit het rij-en wou gaan.' The second and third staves have lyrics 'Die zag daar van ver een mooi meis-je al staan.' Brackets above the notes indicate trigrams. Below the notes, binary sequences represent pitch and note trigrams.

We use both pitch trigrams and note trigrams.

Groups of Features

- Pitch Features
- Contour Features
- Rhythmic Features
- Narmour Closure Features
- Contextual Features
- Textual Features

Single labels and label-trigrams

A musical staff with a single note highlighted. Below it is a binary sequence: 0 0 0 0 1 0 0 0 0. Below that is a sequence of 9 groups of three binary digits: 000 000 000 001 010 100 000 000 000.

Predicting label-trigrams decreases the probability of incorrectly prediction two cadences in a row.

Classification

Random Forest Classifier

Evaluation

10-fold cross validation. The folds are at the level of songs rather than individual trigrams.

Class	pr	rec	F_1	σ_{F_1}	support
<i>note-trigrams</i>					
cadence	0.84	0.72	0.78	0.01	23,925
nocadence	0.96	0.98	0.97	0.01	183,780
<i>pitch-trigrams</i>					
cadence	0.85	0.69	0.76	0.01	23,838
nocadence	0.95	0.98	0.96	0.00	130,992

Table 1. Results for single labels.

Class	pr	rec	F_1	σ_{F_1}	support
<i>note-trigrams</i>					
cadence	0.89	0.72	0.80	0.01	23,925
nocadence	0.96	0.99	0.98	0.00	183,780
<i>pitch-trigrams</i>					
cadence	0.89	0.71	0.79	0.01	23,838
nocadence	0.95	0.98	0.97	0.01	130,992

Table 2. Results for classification with label trigrams.

Ablation Study

We successively remove one group of features and we only use the group of features.

Subset	pr	rec	F_1	σ_{F_1}
F_{all}	0.89	0.72	0.80	0.01
$F_{all} \setminus F_{pitch}$	0.88	0.72	0.79	0.01
F_{pitch}	0.84	0.04	0.08	0.01
$F_{all} \setminus F_{contour}$	0.88	0.73	0.80	0.01
$F_{contour}$	0.00	0.00	0.00	0.00
$F_{all} \setminus F_{rhythmic}$	0.79	0.49	0.60	0.01
$F_{rhythmic}$	0.90	0.35	0.50	0.01
$F_{all} \setminus F_{textual}$	0.85	0.58	0.69	0.02
$F_{textual}$	0.70	0.40	0.51	0.01
$F_{all} \setminus F_{narmour}$	0.83	0.55	0.66	0.01
$F_{narmour}$	0.95	0.30	0.45	0.01
$F_{all} \setminus F_{contextual}$	0.87	0.67	0.76	0.01
$F_{contextual}$	0.71	0.45	0.56	0.01
$F_{all} \setminus F_{rest}$	0.87	0.67	0.76	0.01
F_{rest}	0.97	0.27	0.43	0.02

Table 3. Results for various feature subsets for class 'cadence'.

Conclusions

Both melodic and textual features contribute to establishing a melodic cadence.

All groups of features, except for the contour-features, contribute to the classification, while none of the groups is crucial for classification.

Employing label-trigrams increases performance substantially.

A musical staff with lyrics and binary sequences. Three blue boxes highlight 'Cadence' points. Arrows point to 'Lyrics', 'Phonetic Representation', and 'Detected Ryme'.