An information-dynamic model of melodic segmentation

Marcus T. Pearce, Daniel Müllensiefen and Geraint A. Wiggins

Centre for Cognition, Computation and Culture
Goldsmiths, University of London, UK
- melodic grouping (single level)
- e.g., Mozart Symphony 40 in G minor
  [Lerdahl and Jackendoff, 1983]

sequences of elements $e$ from an alphabet $\mathcal{E}$

model: $p(e_i|e_{i-1}^{i-1})$

Information content (unexpectedness):

$$h(e_i|e_{i-1}^{i-1}) = \log_2 \frac{1}{p(e_i|e_{i-1}^{i-1})}.$$

Entropy (uncertainty):

$$H(e_{i-1}^i) = \sum_{e \in \mathcal{E}} p(e_i|e_{i-1}^{i-1}) h(e_i|e_{i-1}^{i-1}).$$
- *n*-gram model
- combines models of different order
- uses a long- and short-term models
- estimates the probability of an event based on its pitch and onset (IOI and preceding silence).
- 10-fold cross-validation for training/testing
- [Pearce, 2005]
focus here on information content profile
we interpret this as a boundary strength profile $s_n$

pick peaks in the profile at locations where:

1. $S_n > S_{n-1}$
2. $S_n \geq S_{n+1}$

$$S_n > k \sqrt{\frac{\sum_{i=1}^{n-1} (w_i S_i - \overline{S}_{w,1\ldots n-1})^2}{\sum_{i=1}^{n-1} w_i}} + \frac{\sum_{i=1}^{n-1} w_i S_i}{\sum_{i=1}^{n-1} w_i}.$$
Experimental Psychology: infants/adults segment syllable/tone sequences on the basis of local statistics [Saffran et al., 1999]

Cognitive Linguistics: difficulty of word comprehension related to information content [Levy, 2008] and entropy [Hale, 2006];

Machine Learning: algorithms based on information content and entropy can identify word boundaries with some success [Brent, 1999, Cohen et al., 2007]
IDyOM: with $k = 2$

Grouper: [Temperley, 2001]

LBDM: [Cambouropoulos, 2001] with $k = 0.5$

GPR2a: [Lerdahl and Jackendoff, 1983] with $k = 0.5$

GPR2b: [Lerdahl and Jackendoff, 1983] with $k = 0.5$

GPR3a: [Lerdahl and Jackendoff, 1983] with $k = 0.5$

GPR3d: [Lerdahl and Jackendoff, 1983] with $k = 2.5$

Always: every note falls on a boundary

Never: no note falls on a boundary

$k$ optimised from set $\{0.25, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4\}$
Essen Folk Song Collection: Erk

- 1705 German folk melodies
- 78,995 sounding events
- average of 46 events per melody

annotated with phrase boundaries by musicologists

- 12% of notes fall on boundaries
<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hybrid</td>
<td>0.87</td>
<td>0.56</td>
<td>0.66*</td>
</tr>
<tr>
<td>Grouper</td>
<td>0.71</td>
<td>0.62</td>
<td>0.66*</td>
</tr>
<tr>
<td>LBDM</td>
<td>0.70</td>
<td>0.60</td>
<td>0.63*</td>
</tr>
<tr>
<td>GPR2a</td>
<td>0.99</td>
<td>0.45</td>
<td>0.58*</td>
</tr>
<tr>
<td>IDyOM</td>
<td>0.76</td>
<td>0.50</td>
<td>0.58*</td>
</tr>
<tr>
<td>GPR2b</td>
<td>0.47</td>
<td>0.42</td>
<td>0.39</td>
</tr>
<tr>
<td>GPR3a</td>
<td>0.29</td>
<td>0.46</td>
<td>0.35</td>
</tr>
<tr>
<td>GPR3d</td>
<td>0.66</td>
<td>0.22</td>
<td>0.31</td>
</tr>
<tr>
<td>Always</td>
<td>0.13</td>
<td>1.00</td>
<td>0.22</td>
</tr>
<tr>
<td>Never</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
GPR2a does well - importance of rests
LBDM/Grouper comparable to other studies
Hybrid model outperforms its component models
Information dynamic model performs surprisingly well
  - developed as a model of pitch prediction
  - not optimised for melodic grouping
focus on boundaries not indicated by rests
use boosting to create a hybrid model
  - optimise viewpoints for segmentation
  - other information dynamic measures
    - entropy
    - predictive information
explicit Bayesian models of phrase segmentation [Brent, 1999]
References I

An efficient, probabilistically sound algorithm for segmentation and word discovery.

The local boundary detection model (LBDM) and its application in the study of expressive timing.

Voting experts: An unsupervised algorithm for segmenting sequences.

Uncertainty about the rest of the sentence.

*A Generative Theory of Tonal Music.*
MIT Press, Cambridge, MA.

Expectation-based syntactic comprehension.
*Cognition*, 16(3):1126–1177.
References II

Meaning in music and information theory.

University of Chicago Press, Chicago.

*The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition.*
PhD thesis, Department of Computing, City University, London, UK.

Statistical learning of tone sequences by human infants and adults.

*The Cognition of Basic Musical Structures.*
MIT Press, Cambridge, MA.