Towards incorporating the notion of feature shape in music and text retrieval

Yi-Yun Cheng[1], David M. Weigl[2], J. Stephen Downie[1], Kevin R. Page[2]

[1] School of Information Sciences, University of Illinois at Urbana-Champaign, USA
[2] Oxford e-Research Centre, University of Oxford, UK
yiyunyc2@illinois.edu

Abstract. Extracted feature data augment information resources with concrete characterizations of their content, but only approximate to the meaningful high-level descriptions typically expected by digital musicology scholars (domain experts with some technological affinity, but with no expertise in signal processing or feature data). Feature shapes provide abstract aggregations of feature types which share common characteristics when applied in extraction workflows. We explore the feasibility of feature shape-based filtering and querying within a large audio dataset of live music performances, employing operation sequences as specified by the Audio Feature Ontology and Vocabulary. We further implement analogous semantic structures for the HathiTrust Extracted Feature Dataset to demonstrate the general applicability of feature shapes in music and text retrieval.

Keywords: Feature data, feature shape, Music Information Retrieval, HathiTrust.

1 Introduction

In Music Information Retrieval (MIR), extracted feature data are algorithmic quantifications and categorizations of specific aspects of symbolic musical structures or of audio signals. MIR processes operate on feature data to provide mathematical approximations of musical or musicological concepts [1]. These features may not be immediately accessible to end users in terms directly applicable to their studies, e.g. musicology. This situation is not limited to MIR; other areas of information retrieval, including text, also encounter the gap between the mathematical approximations provided by feature data, and more meaningful domain-specific descriptions expected by users.

Feature shapes are higher-level abstractions of the characteristics shared between different subsets of features, intended to better reflect user expectations. For instance, a musicologist wishing to conduct a harmonic analysis could be guided toward features sharing a *harmonic shape* (operating in the spectral domain), without requiring extensive signal processing background knowledge. By implementing the notion of feature shapes, we make the feature data more accessible and retrievable.

The Audio Feature Ontology and Vocabulary (AFO/AFV) [2] provides a generic, implementation-independent semantic description of audio features informed by a survey of existing MIR feature taxonomies. They incorporate process descriptions specifying the operation sequence of each feature, comprising a series of discrete steps in the feature extraction process. E.g., the chromagram feature's operation sequence comprises Windowing, Discrete Fourier Transform, Logarithm, and Sum. These granular process descriptions afford the definition of feature shapes as aggregations of multiple feature types according to shared operation sequence subsets (Figure 1).

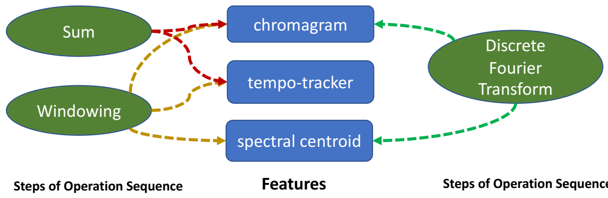

Fig. 1. Multiple features sharing different steps of their operation sequence

2 Feature shapes in the Live Music Archive

Here, we explore the feasibility of feature-shape based filtering and querying within a large collection of extracted audio features and associated metadata describing recordings in the Internet Archive's Live Music Archive [3], provided as Linked Data. To evaluate the feasibility of incorporating operation sequences in audio feature retrieval, we provided a mapping scheme to align the AFO/AFV to RDF descriptions of Vamp [4] feature extraction plugins employed within the live music dataset[1].

We customized the SPARQL queries provided in [2] to demonstrate i) the steps of operation sequences tied to a given Vamp feature extractor; and ii) how the Vamp features that share specific steps of their operation sequence are retrievable (Listing 1).

Listing 1. SPARQL queries and results sets.*Left*: Retrieve the operation sequence for the qm-chromagram Vamp plugin.*Right*: Find all Vamp audio extractors to perform a 'Windowing' step.

```
SELECT distinct ?optype WHERE {
  BIND(plugbase:qm-chromagram as ?vamp).
  ?opid a ?optype.
  ?fopid afo:next_operation* ?opid .
  ?seqid afo:first_operation ?fopid .
  ?model afo:sequence ?seqid .
  ?feature afo:model ?model .
  ?vamp vamp:computes_event_type ?feature .
  FILTER (?optype != afo:LastOperation).
    }
```

```
SELECT distinct ?vamp WHERE {
  BIND(afv:Windowing as ?optype).
  ?opid a ?optype .
  ?fopid afo:next_operation* ?opid .
  ?seqid afo:first_operation ?fopid .
  ?model afo:sequence ?seqid .
  ?feature afo:model ?model .
  ?vamp vamp:computes_event_type ?feature
    }
```

optype
afv:Windowing
afv:DiscreteFourierTransform
afv:Logarithm
afv:Sum

optype
plugbase:qm-tempotracker
plugbase:qm-chromagram
plugbase:spectral_centroid
plugbase:zcr

[1] https://github.com/yiyunyc2/OIDLPP/blob/master/Mapping_vamp_afv.n3

3 Feature shapes in textual information retrieval

The HathiTrust Extracted Feature Dataset (HTEFD) is a collection of textual features derived from the content within the HathiTrust Digital Library. Like the extracted audio features described above, these textual features are also generated by feature extractors employing specific sequences of operation. For instance, when parsing text, the process incorporates a sequence of sentence segmentation, tokenization, and part of speech tagging. We have reviewed the developer's manuals for the Apache OpenNLP[2], Natural Language Toolkit[3] (NLTK), and the Stanford CoreNLP[4], considering the commonalities and divergences between the operation sequences defined by AFO/AFV and analogous processes within HTEFD to gain a more generic understanding of feature shape-based explorations in information retrieval.

We have created an RDF vocabulary analogous to the AFO/AFV that describes the operation sequences of a subset of the text features published by the HTEFD (Figure 2). By cross-application of SPARQL queries from the audio feature domain, we demonstrate applicability of a feature shape approach to textual retrieval (Listing 2).

```
opennlp:LDA a owl:Class ;
    rdfs:label "LDA"@en ;
    rdfs:subClassOf htrc:TextFeature ;
    afo:model htrc:LDAModel.

htrc:LDAModel a afo:Model ;
    rdfs:label "LDA Model"@en ;
    rdfs:comment "Topic Modeling technique that
    afo:sequence htrc:LDA_operation_sequence.

htrc:LDA_operation_1 a htrc:Tokenization ;
    rdfs:label "LDA_operation_1"@en ;
    rdfs:comment "LDA sequence step 1";
    afo:next_operation htrc:LDA_operation_2 .

htrc:LDA_operation_2 a htrc:POStagMethod ;
    rdfs:label "LDA_operation_2"@en ;
    rdfs:comment "LDA sequence step 2" ;
    afo:next_operation htrc:LDA_operation_3 .

htrc:LDA_operation_3 a htrc:ProbsMethod ;
    rdfs:label "LDA_operation_3"@en ;
    rdfs:comment "LDA sequence step 3" ;
    afo:next_operation htrc:LDA_operation_4 .
```

```
htrc:LDA_operation_4 a htrc:TopKSequencesMethod ;
    rdfs:label "LDA_operation_4"@en ;
    rdfs:comment "LDA sequence step 4" ;
    afo:next_operation htrc:LDA_operation_5 .

htrc:LDA_operation_5 a htrc:Lemmatization ;
    rdfs:label "LDA_operation_5"@en ;
    rdfs:comment "LDA sequence step 5" ;
    afo:next_operation htrc:LDA_operation_6 .

htrc:LDA_operation_6 a htrc:CaseFolding;
    rdfs:label "LDA_operation_6"@en ;
    rdfs:comment "LDA sequence step 6" ;
    afo:next_operation htrc:LDA_operation_7 .

htrc:LDA_operation_7 a afo:LastOperation,
    htrc:StopListing;
    rdfs:label "LDA_operation_7"@en ;
    rdfs:comment "LDA sequence step 7".

htrc:LDA_operation_sequence a afo:OperationSequence ;
    afo:first_operation htrc:LDA_operation_1.
```

Fig. 2. Example of the LDA feature and its operation sequence

Listing 2. SPARQL queries and result sets. *Left*: Retrieve the operation sequence for OpenNLP LDA. *Right*: Find all HTEFD features that perform a 'Tokenization' step.

```
SELECT distinct ?optype WHERE {
  BIND(opennlp:LDA as ?feature).
  ?opid a ?optype .
  ?fopid afo:next_operation* ?opid .
  ?seqid afo:first_operation ?fopid .
  ?model afo:sequence ?seqid .
  ?feature afo:model ?model .
  FILTER (?optype != afo:LastOperation)
    }
```

```
SELECT distinct ?feature WHERE {
  BIND(htrc:Tokenization as ?optype).
  ?opid a ?optype .
  ?fopid afo:next_operation* ?opid .
  ?seqid afo:first_operation ?fopid .
  ?model afo:sequence ?seqid .
  ?feature afo:model ?model .
    }
```

---

[2] https://opennlp.apache.org/docs/1.8.1/manual/opennlp.html
[3] http://www.nltk.org/
[4] https://stanfordnlp.github.io/CoreNLP/

optype	optype
htrc:Tokenization	htrc:Chunker
htrc:POStagMethod	htrc:LDA
htrc:ProbsMethod	htrc:NER
htrc:TopKSequencesMethod	htrc:POStagger
htrc:Lemmatization	htrc:Tokenizer
htrc:CaseFolding	
htrc:StopListing	

4 Conclusion

We have investigated the application of *operation sequences* to inform the notion of *feature shape* in feature-based information retrieval. Our SPARQL queries demonstrate the feasibility of our approach to both audio and textual retrieval. We have applied this approach to augment the Computational Analysis of the Live Music Archive dataset with an additional semantic layer mapping audio features to AFO/AFV concepts[5]. We will build upon this conceptualization of feature shapes to inform ongoing work on information systems providing domain-agnostic, usable access to feature data.

Acknowledgements

This work was supported by the Fusing Audio and Semantic Technologies for Intelligent Music Production and Consumption project, funded by the EPSRC (EP/L019981/1), the Mellon Funded Workset Creation for Scholarly Analysis and Data Capsules project and the HathiTrust Research Center. We additionally thank colleagues for their support and insight, particularly David Lewis, Graham Klyne, and Ryan Dubnicek.

References

1. Page, K., Fields, B., De Roure, D., Crawford, T., Downie, J. S.: Reuse, remix, repeat: the workflows of MIR. ISMIR, 409–414 (2012).
2. Allik, A., Fazekas, G., Sandler, M.: An ontology for audio features. ISMIR, 73–79 (2016).
3. Page, K., Bechhofer, S., Fazekas, G., Weigl, D., Wilmering, T.: Realising a layered digital library: exploration and analysis of the Live Music Archive through linked data. JCDL, 1–10 (2017).
4. Cannam, C., Sandler, M., Jewell, M. O., Rhodes, C., & d'Inverno, M. (2010). Linked data and you: Bringing music research software into the semantic web. Journal of New Music Research, 39(4), 313-325.

[5] https://github.com/yiyunyc2/OIDLPP/blob/master/analyses-subset.ttl.zip