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

ILLINOIS School of Information Sciences

1. Introduction

In Music Information Retrieval (MIR):

- extracted feature data quantify specific aspects of musical structures or of audio signals
- MIR processes operate on feature data to provide mathematical approximations of musical or musicological concepts

Problems:

- These features may not be immediately accessible to end users in terms directly applicable to their studies, e.g. musicology
- A constant gap between

1. the mathematical approximations provided by feature data
2. The more meaningful domain-specific descriptions expected by users

2. Feature Shape

Definition:

Higher-level abstractions of the characteristics shared between different subsets of features, intended to better reflect user expectations

Feature Shape as a solution:

By implementing the notion of feature shapes, we make the feature data more accessible and retrievable

Motivating example:

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

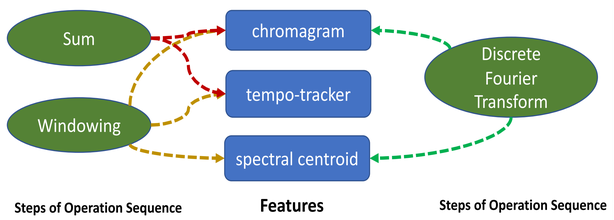

Figure 1. Multiple features sharing different steps of their operation sequence

Feature Shapes in the Live Music Archive

Live Music Archive:

A large collection of extracted audio features and associated metadata describing recordings in the Internet Archive's Live Music Archive, provided as Linked Data

Mapping scheme:

We provided a mapping scheme (on GitHub) to align the AFO/AFV to RDF descriptions of Vamp [4] feature extraction plugins employed within the live music dataset

SPARQL queries:

1. the steps of operation sequences tied to a given Vamp feature extractor
2. how the Vamp features that share specific steps of their operation sequence are retrievable

QUERY 1: Retrieve the operation sequence for the chromogram Vamp plugin

```
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).
}
```

RESULTS:

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

QUERY 2: : Find all Vamp audio extractors to perform a 'Windowing' step

```
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
}
```

RESULTS:

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

Feature Shapes in Text Information Retrieval

HathiTrust Extracted Feature Dataset (HTEFD) :

A collection of textual features derived from the content within the HathiTrust Digital Library. When parsing text, the process incorporates a sequence of sentence segmentation, tokenization, and part of speech tagging

Creating an RDF vocabulary for HTEFD analogous to AFO/AFV:

Based on our review of the features in Apache OpenNLP, Natural Language Toolkit (NLTK), and the Stanford CoreNLP, we created an RDF vocabulary to describe the operation sequences of a subset of the text features published by the HTEFD

QUERY 1: Retrieve the operation sequence for OpenNLP LDA

```
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)
}
```

optype
htrc:Tokenization
htrc:POStagMethod
htrc:TopKSequenceMethod
htrc:Lemmatization
htrc:StopListing

3. The Audio Feature Ontology and Vocabulary (AFO/AFV)

Description:

- A survey of existing MIR features
- Provides a generic semantic description of audio features
- Incorporates process descriptions specifying the **operation sequence** of each feature
- Comprises a series of discrete steps in the feature extraction process

Example of operation sequence in chromagram:

- **Step 1:** Windowing
- **Step 2:** Discrete Fourier Transform
- **Step 3:** Logarithm
- **Step 4:** Sum

4. Contributions

We have:

- Applied *operation sequences* to inform the notion of ***feature shape*** in feature-based information retrieval
- Used SPARQL queries to demonstrate the feasibility of our approach to both audio and textual retrieval
- Applied this approach to augment the Computational Analysis of the Live Music Archive dataset

We will:

- Build upon the conceptualization of feature shapes for ongoing work on information systems providing domain-agnostic, usable access to feature data

Acknowledgments

This work was supported by the FAST for Intelligent Music Production and Consumption project (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.

Reference & Contact Information

Email:
yiyunyc2@illinois.edu

GitHub Project Repository:
https://github.com/yiyunyc2/OIDLPP

Scan me!