Music Information Retrieval

Jocelyn Kavanagh

Georgia Tech

Lesson Outline

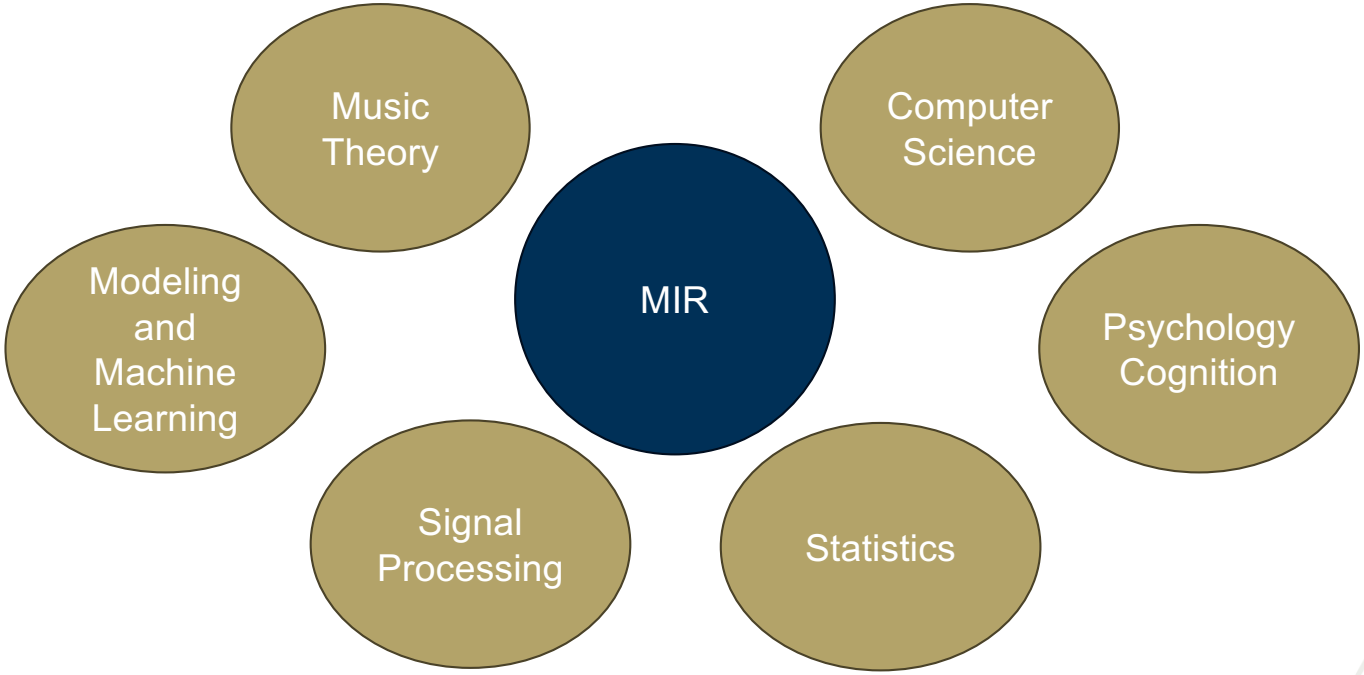
- MIR
- Audio Content Analysis
- Applications
- Up Next…

Georgia Tech

What is MIR?

Music Information Retrieval

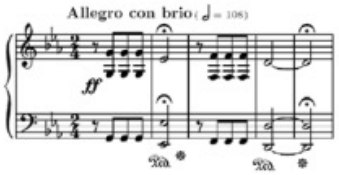
- Taking ”raw” musical data and turning it into something understandable/useful
- Interdisciplinary field



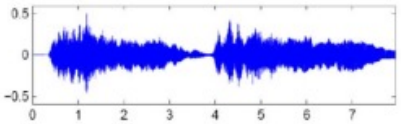
Georgia Tech

What forms does music come in?

Sheet Music (Image)



CD / MP3 (Audio)



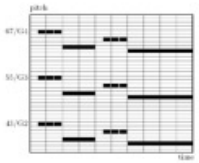
MusicXML (Text)

```
<note>
  <pitch>
    <step>E</step>
    <alter>-1</alter>
    <octave>4</octave>
  </pitch>
  <duration>2</duration>
  <type>half</type>
</note>
```

Dance / Motion (Mocap)

Music

MIDI



Singing / Voice (Audio)

Music Film (Video)

Music Literature (Text)

Georgia Tech

Music Analysis Through Different Lenses

- Content-Based
  - Analyzing the audio or score
- Context-Based
  - Analyzing the metadata – streams, shares, lyrics, artists, tags
- Human-Centered
  - Analyzing how humans experience the music

Georgia Tech

Why are we trying to analyze this data?

- We can process that into more data!
  - Leading to segmentation of the data
  - Converting into different forms
- This can help us
  - Study musical structures and patterns
  - Make predictions about other music
  - Generate new music
  - Build tools for musical analysis, performance, creation, production, education…

Georgia Tech

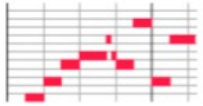
Common MIR Tasks

Georgia Tech

Melodic Similarity and Pattern Matching

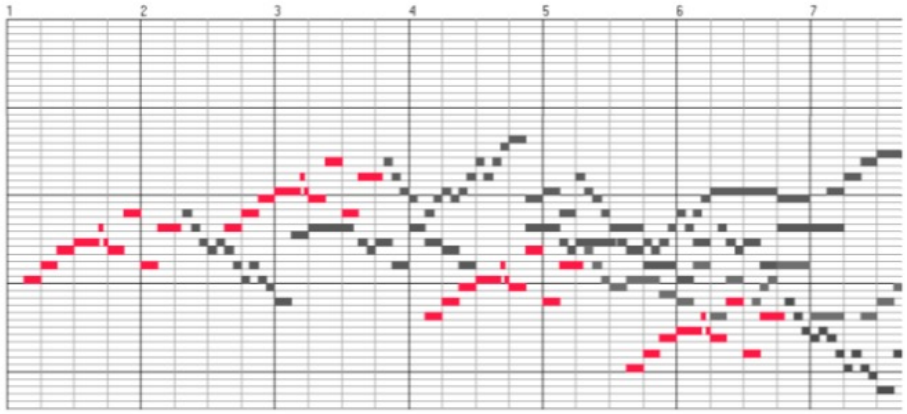
Piano Roll Representation (MIDI)

Query:



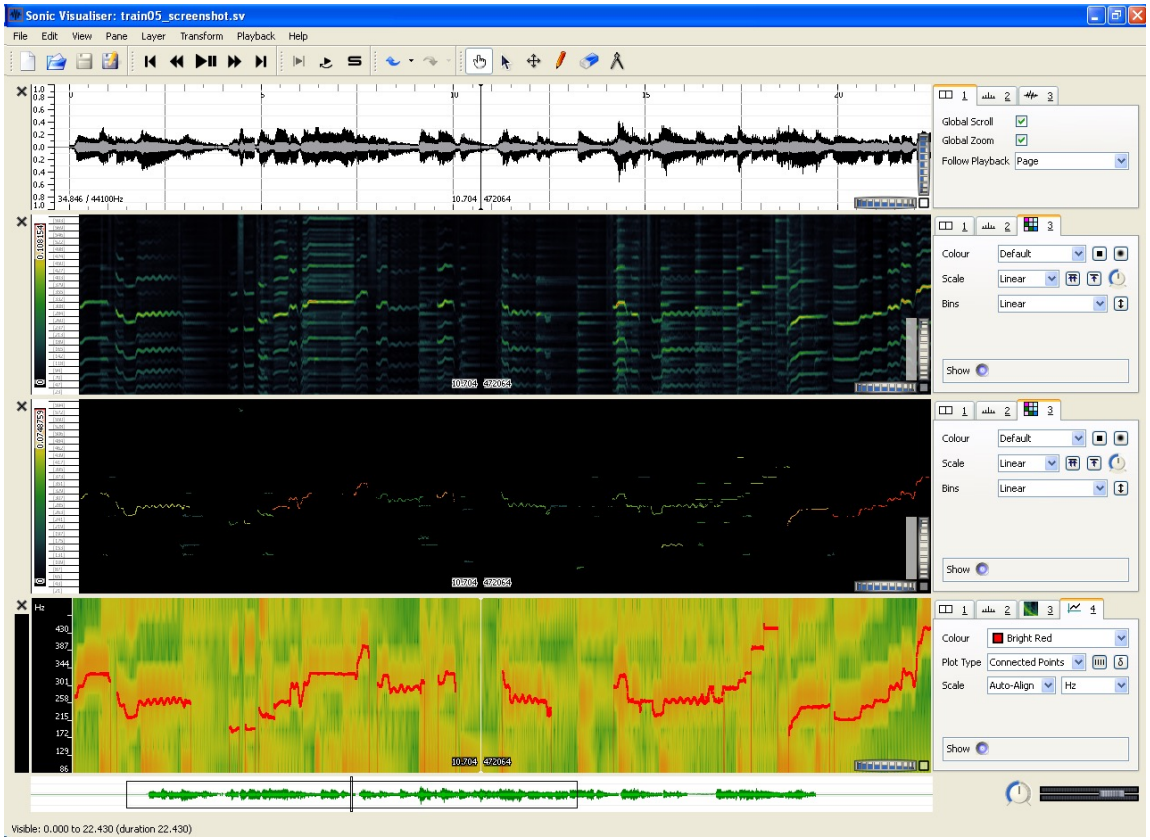
Goal: Find all occurrences of the query

Matches:

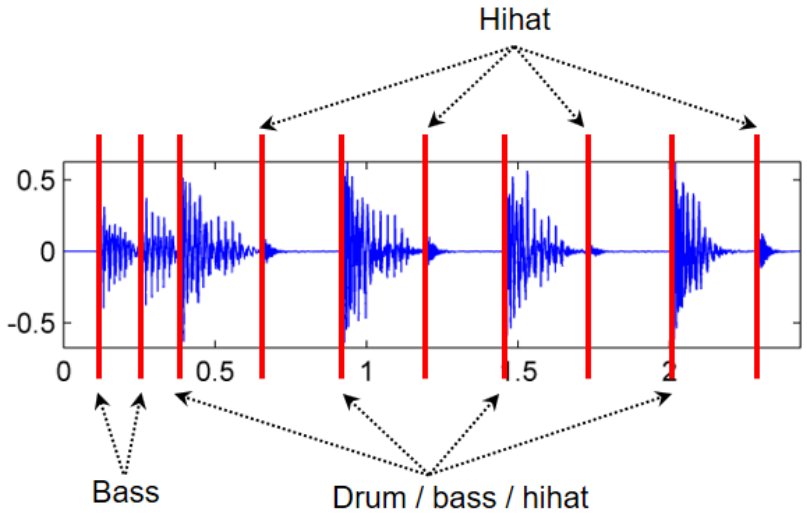


Georgia Tech

Melody Extraction and Source Separation



Tempo Estimation and Beat Tracking



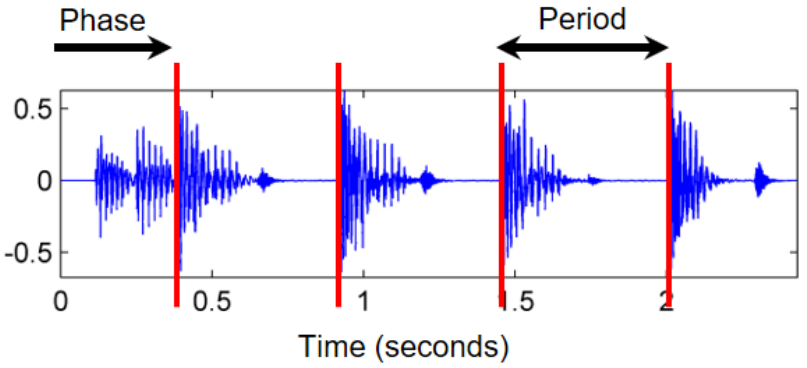
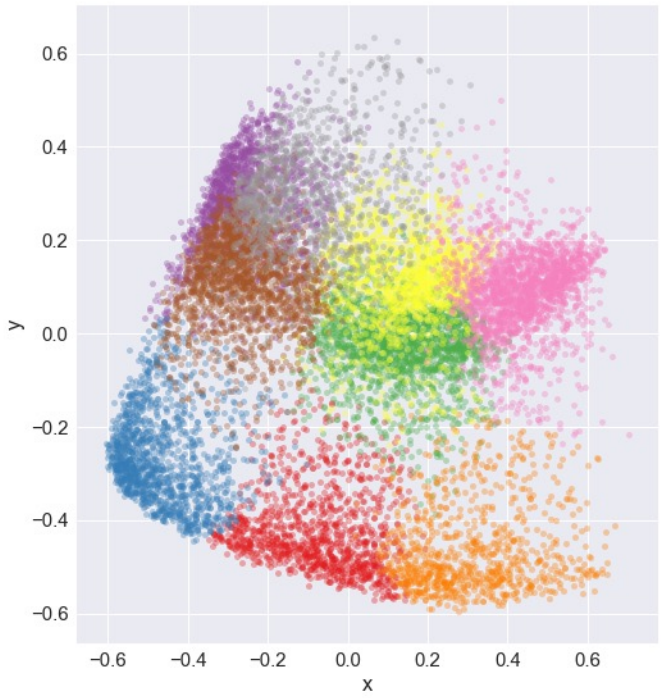
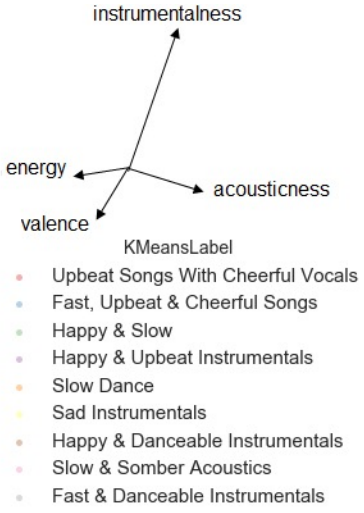


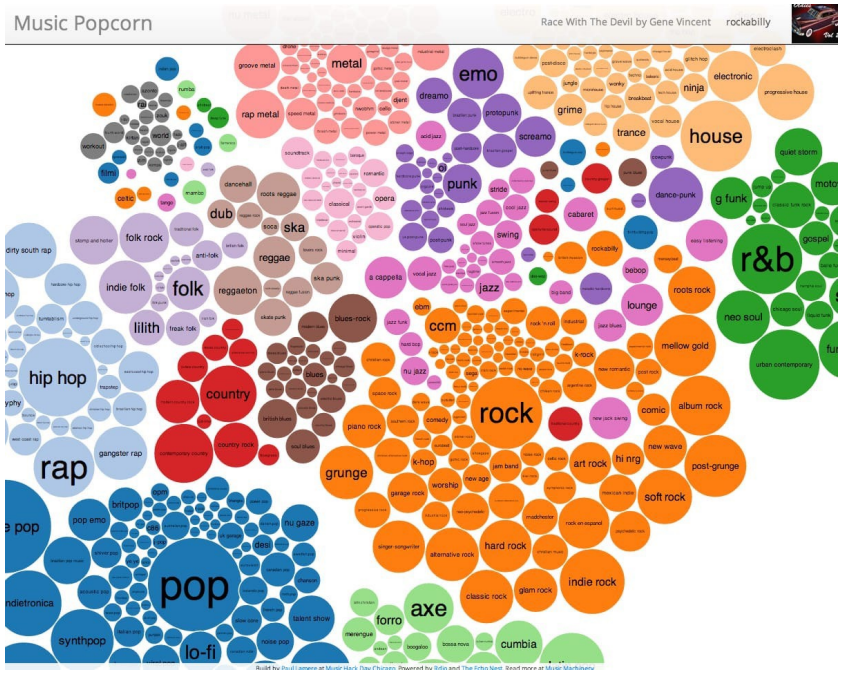
Figure 6.1 from [Müller, FMP, Springer 2015]

Georgia Tech

Categorization and Classification







Georgia Tech

How does this all work?

Georgia Tech

Feature Extraction

- We can gather “features” related to auditory models, spectrum, or other mathematical properties in a signal/stream
  - Organized content extracted from the data in a meaningful way
- We collect the data into frames and extract the features
  - Features don’t necessarily tell us everything
  - We look for correlation between metadata, perceptual data, other features
- There are high-level and low-level features

Georgia Tech

Low Level Features

- Zero-crossing rate
- Root mean square amplitude
- Max power
- Spectral Centroid
- Spectral Flux
- Spectral Fall-off
- MFCCs

High Level Features

- Onset Detection
- Pitch detection
- Melody extraction
- Key and chord recognition
- Beat tracking
- Instrument recognition

Georgia Tech

Semantic Gap

- Difference between the extracted information and the interpretation of information
- Mismatch between low-level data and human perception/cognition

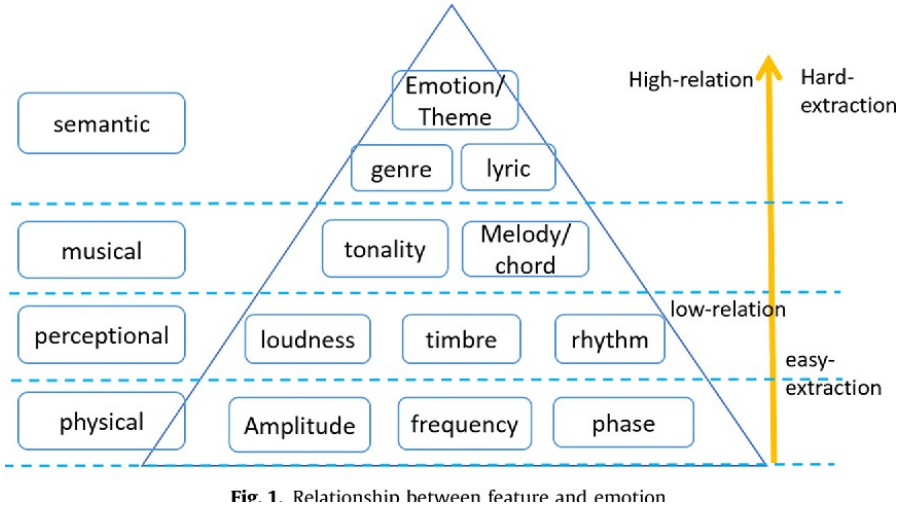
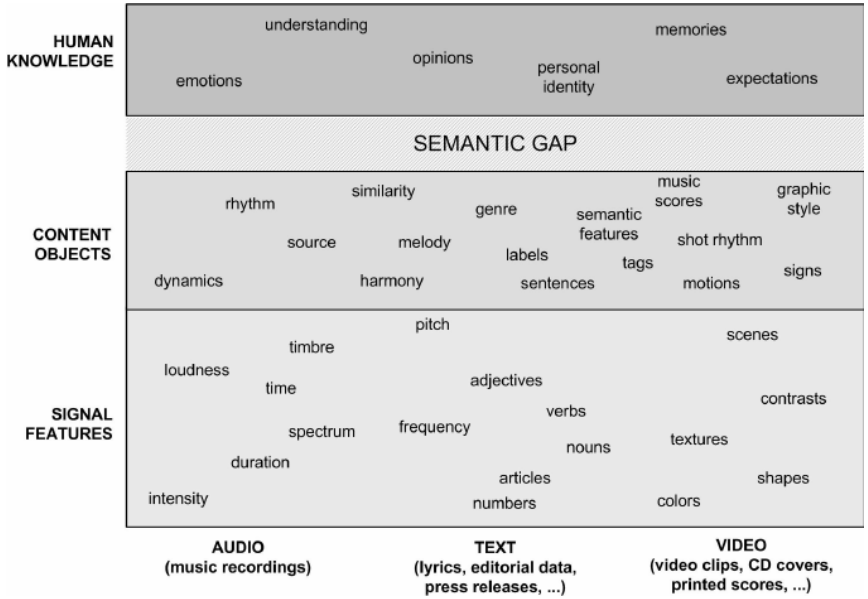


Fig. 1. Relationship between feature and emotion

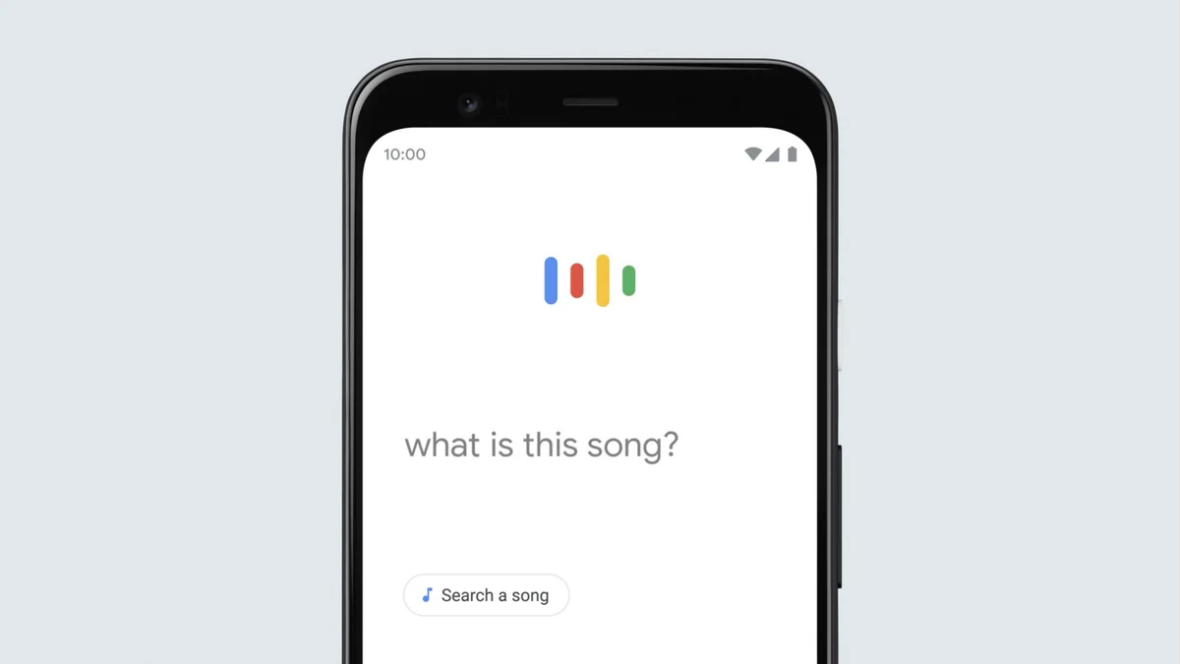


Georgia Tech

Applications

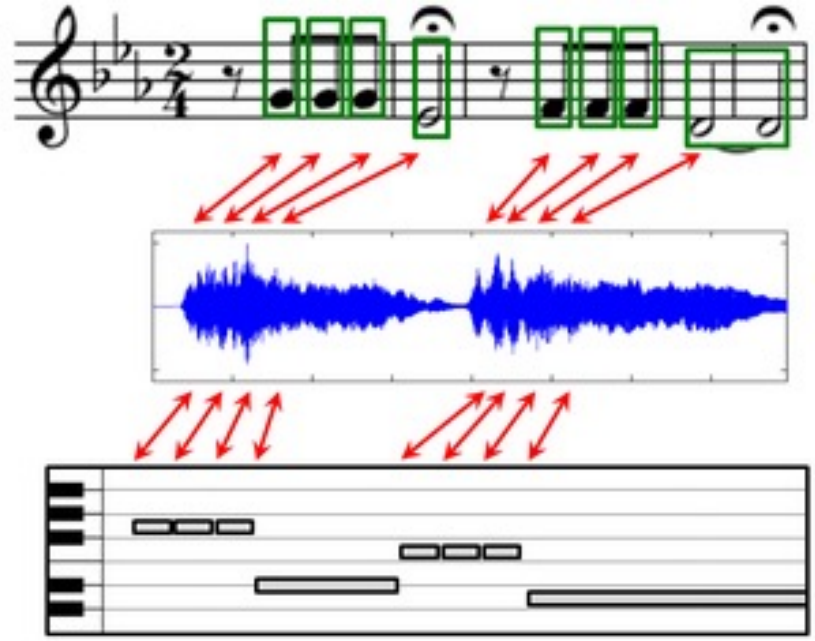
Audio Identification and Classification

- Query by humming
- Audio fingerprinting



Audio and Score Alignment

- Score navigation
- Automatic mixing



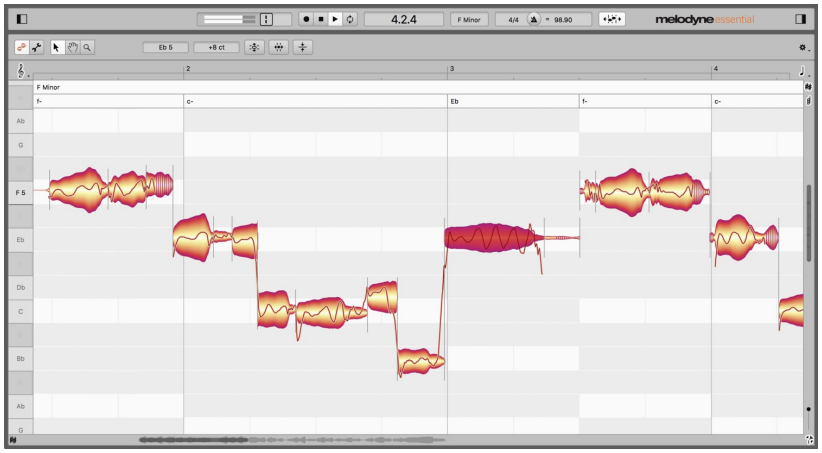
Music Analysis and New Music Generation

- Chord detection, key detection, structural analysis, classification
- Automatic tagging

Georgia Tech

Automatic Mixing and Plugins

- Melodyne
  - Pitch correction tool
- Izotope’s Nectar and Neutron
  - Mixing assistant plugins



Georgia Tech

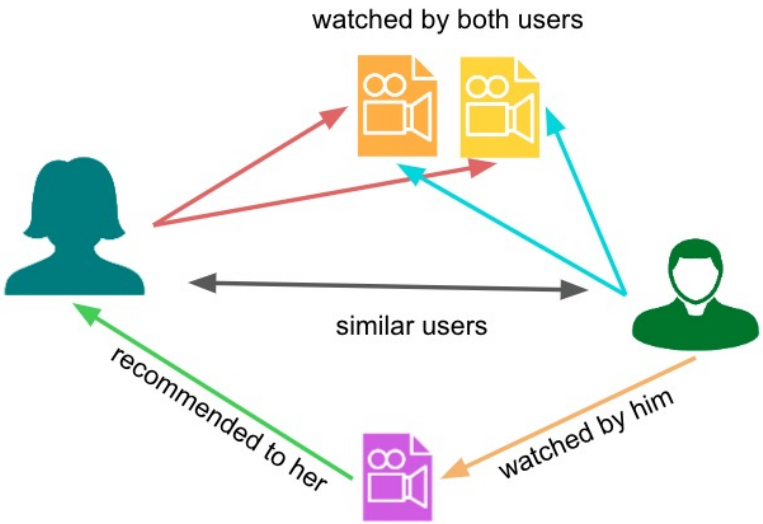
Music Recommender Systems

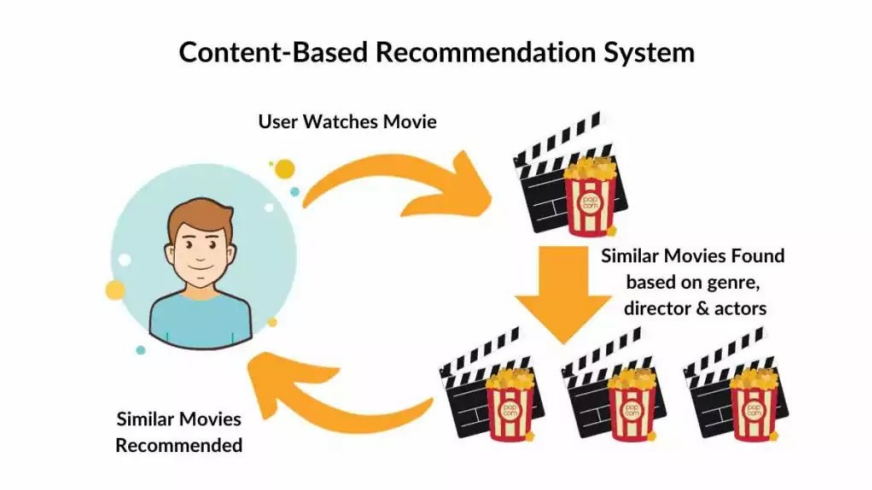
- Content-based vs Context-based

Georgia Tech

Different Ways to Recommend Songs or Build Playlists

- Manual
- Collaborative Filtering
- Content-Based Filtering
- Context-Based Filtering





Georgia Tech