Music Curation & Analytics

Individual Lab Assignment

Week 1: Musical Datasets

**Theme: Film Soundtracks (focussing on scores written specifically for films)**

I chose Film Soundtracks/Scores as the theme for this project as films are one of my greatest passions and interests in life. I absolutely love film music and admire how it is such a huge and significant component of the filmmaking process. Good music can undoubtedly lift the quality of a film and go a long way to help establish the tone or mood of a particular scene or cinematic sequence. I particularly enjoy how music can exist in a diegetic or non-diegetic way in cinema: either as part of the world created by the film (e.g. music playing on a radio that the characters can hear) or outside of the filmic world (orchestral score given only to the viewer).

The pieces I have chosen to focus on are six of my favourite examples of the film soundtrack genre: three written for marimba/xylophone and three written for piano. They were all originally written for their respective movies, with the exception of the first two, ‘You’re So Cool’ and ‘Gassenhauer’, which are variations of the same song (this will be explained in greater detail further down this page). All pieces also fall into the non-diegetic category, with the exception of ‘Cantina Band’ in Star Wars, which of course is played on-screen by the titular Cantina Band.

Three types of musical data:

1. Descriptive data
2. Notated data
3. Acoustic data

**Descriptive Data**

There are many online databases dedicated exclusively to film soundtracks and a couple of the most popular examples are detailed below. Additionally, the IMDb (Internet Movie Database) is a good place to find a wide variety of metadata relating to film soundtracks.

**Soundtrack Collector (**[**soundtrackcollector.com**](http://soundtrackcollector.com/)**)**

* Online soundtrack database of substantial size. The site claims to host over 53,000 movie titles, 74,000 labels, 11,000 composers, 21,000 movie posters, 49,000 cover images, 37,000 track listings and 11,000 links to reviews.
* Allows users to explore the site by movie titles (A-Z), composers (A-Z) or search the entire database by movie title, soundtrack, composer, label number or track.
* When exploring by composer, the list is arranged in alphabetical order of surname (e.g. Hans Zimmer is under ‘Z’). When a composer is selected from the list, the user is presented with the following information: full name, an image of the person, birth date, country of origin, a link to their website and a reasonably-sized biography detailing some key films and dates of the composer’s career. Beneath this is a link to view the filmo/discography of this specific composer.
* When viewing the discography of a specific composer, the list is presented in a chronological order, beginning with the composer’s earliest work. Six types of metadata are displayed:
  + Year
  + Movie Title
  + Label
  + Country
  + Star Rating (out of five)
  + Format (e.g. CD, DVD, LP etc)
* The user can view certain time periods by selecting options from the “Go to years” dropdown menu (e.g. 1996-2000).
* Once the user has selected a movie title, they are given information about the film including its year of release, country and genre. The user can then view all of the existing productions of the film soundtrack including features in compilation albums.
* Unfortunately, although the Soundtrack Collector database is well-organised and well-populated, there are no tools available to the user to analyse the data to any extent.

**Soundtrack.net (soundtrack.net)**

* Similar purpose and structure to that of the Soundtrack Collector database.
* The home page features a news section, new soundtrack releases, upcoming soundtrack releases, movies currently in cinemas and more.
* The user can browse the database by movie release date, title, composer or soundtrack (title, composer or manufacturer) and can also view news & reviews and trailer music.
* In terms of the way in which this site presents its data, there is a stand-out problem to note: when viewing the list of movies & tv shows, the list is sorted alphabetically. However, any title that begins with “The” is sorted under ‘T’. This is problematic because of the high volume of titles which fit this category, making it harder for the user to find their desired title. Most film databases avoid this issue by disregarding the “The” and instead sorting by the word that follows straight after. Unfortunately, that is not the case with Soundtrack.net.

**Notated Data**

* Sheet music for film scores and soundtracks can be found in both free and paid versions in online repositories. Some of the best sources include Musescore.com, Musicnotes.com and Noviscore.com. These scores generally take the form of downloadable pdf files. However, in the case of MuseScore, a signed-in user can download the file in many different formats: a MuseScore file (playable in the MuseScore desktop program), pdf, MusicXML, MIDI and also mp3. In my experience, MuseScore also has the best collection of online scores and the data on the website is well-structured and organised for easy search and retrieval.
* Another popular source of notated data is printed sheet music. Traditional bricks-and-mortar music shops will stock soundtrack sheet music, but a far greater selection will always be found in online stores such as Musicroom.com and Thomann.de and the usual e-commerce giants like Amazon and Ebay.
* Notated data for film scores and soundtracks can also take the form of guitar tablature: an alternative to traditional common Western music notation. Again, these can be found in printed books, but the best sources are online: notably Ultimate-guitar.com, Guitartabs.cc and Hooktheory.com.

**Acoustic Data**

* Official original soundtrack albums are generally released in conjunction with, or closely following, the cinematic release of their filmic companions. Different physical formats have been used over the years including vinyl, cassette tapes and CDs but the focus has now shifted to online music streaming platforms and digital downloads. The compressed mp3 audio file is the most common format in these modern cases, however there are high-fidelity streaming services such as Tidal that offer uncompressed audio with higher bit rates in the form of FLAC and ALAC files.
* In the cases of streaming services including Apple Music and Spotify, there are entire sections dedicated to film and tv soundtracks and users can also find curated playlists relating to screen music.
* Another source of acoustic data is live concert performances of film music. These have also been recorded and captured on video. DVDs and Blu-ray discs are a common way of accessing these performances, although digital versions are again available to buy online or stream on platforms such as Netflix and Amazon Prime Video.
* Finally, YouTube is always a reliable option for finding recordings of film soundtracks. There are user-created playlists and videos from official movie studio channels and with the sheer volume of content available on the online platform, users should always be able to locate what they are looking for.

Week 2: Notated Music

I have chosen to transcribe, compare and analyse six pieces of music from some of my favourite films: three written for marimba/xylophone and three written for piano.

**Gassenhauer**

**Gassenhauer nach Hans Neusiedler (1536)**

Short piece from Carl Orff's Schulwerk (a music education method developed by Carl Orff and Gunild Keetman in the 1920s). As the full title indicates, it is either an arrangement of, or inspired by, a much older work by the [lutenist](https://en.wikipedia.org/wiki/Lutenist) Hans Neusiedler from 1536. Credited to Orff's long-time collaborator, Gunild Keetman, on at least one recent release of the Schulwerk.

Different versions of this piece have been used in various films and television programmes.

Examples:

* A very true-to-original version of the piece was used in the soundtrack of *Badlands* (1973).
* Hans Zimmer wrote a re-worked version of the piece for the soundtrack of *True Romance* (1993).

**1st piece – Hans Zimmer – You’re So Cool (from *True Romance*, 1993)**

* Transcribed from MuseScore pdf file: <https://musescore.com/user/2674566/scores/5005336>
* Notes:
  + The pdf sheet music that I used as a reference for my transcription was written just for piano. I changed the instrument to Marimba (the instrument on which the original piece was intended to be played). This threw up a small issue that can be seen in my MuseScore file: some of the notes are highlighted in yellow and red signifying that they are outside of the pitch range of the instrument.
  + I also increased the tempo from the one detailed on the pdf sheet music so that it was a closer match to that of the original piece.
  + This is a shortened version of the full original piece. However, this was all I was interested in as it contained the main theme of the soundtrack that runs throughout the whole film.

**2nd piece – Car Orff – Gassenhauer (from *Badlands*, 1973)**

* Transcribed from sheet music by Schott Music: <https://en.schott-music.com/shop/gassenhauer-no334913.html>
* Notes:
  + The only reference used to transcribe this piece was the preview images given on the Schott Music website. A payment was required to fully download and access the sheet music file, but I was able to work from the preview images. This was the best I could find as there weren’t many examples of sheet music for this piece available online. This may be attributed to the age of the piece of music or to the fact that the origins are hard to trace.
  + The only changes I made in comparison with the Schott Music file were to set the tempo and set the instruments to marimba to match the original piece of music. I also removed the very final note as I didn’t think it sounded well, and it wasn’t consistent with the original piece either.

**3rd piece – Thomas Newman – Dead Already (from *American Beauty*, 1999)**

* Transcribed from pdf sheet music online: <https://filmscorearrangements.files.wordpress.com/2014/02/dead-already.pdf>
* Notes
  + No changes made by me with the exception of setting a tempo that was consistent with the original piece of music.
  + One of the unique challenges of transcribing this piece was having to add multiple voice in the latter part of the score.

**4th piece – John Williams – Cantina Band (from *Star Wars: Episode IV – A New Hope*, 1977)**

* Transcribed from MuseScore pdf file: <https://musescore.com/user/9359561/scores/2877071>
* Notes:
  + I only transcribed the first 32 bars of the piece as this was the only part I was interested in. The opening movement is the most well-known section of the piece and after this, it goes on for a long while and becomes very repetitive, and thus was not of much value to me.

**5th piece – Yann Tiersen – Comptine D’un Autre Été, L’après Midi (from Amélie, 2001)**

* Transcribed from MuseScore pdf file: <https://musescore.com/user/934881/scores/919886>
* Notes:
  + This time, I did transcribe the whole piece. This piece is a great deal longer than the other pieces, and although more challenging, there were certain things that made it manageable. For example, the piano left-hand (or bass clef) had the same repeating progression throughout the whole song.
  + This piece also gave me the opportunity to apply extra pieces of notation data such as repeating barlines, octave lines, prima & seconda voltas among other dynamics.
  + The only thing I changed when carrying out my transcription from the pdf sheet music was the tempo. I decreased it slightly as I believe the tempo detailed in the pdf file was a little too high in comparison to the original piece.

**6th piece – Michael Andrews – The Artifact and Living (from *Donnie Darko*, 2001)**

* Transcribed from MuseScore pdf file: <https://musescore.com/user/16279/scores/37355>
* Notes:
  + I had to make several changes from the pdf sheet music from which I was transcribing the piece. This mainly consisted of rectifying the note ties that weren’t quite correct in the pdf.

Week 3: Encoding Notated Music

**Notes on MusicXML:**

* At the top of the file, there is the title and composer of the piece, but not the subtitle. The subtitle can only be found inside the “credit” element later on in the file.
* Structured in a very standardised way: a MusicXML file will look extremely similar regardless of the piece. All descriptive and notated data is represented in nested elements. The drawback of this is that it is not as easily-legible by a human, however it is much better for cross-platform sharing and collaboration. Also uses far more lines of code (4 times that of a MEI file).
* Note pitch is represented by the ‘step’ (letters) and ‘octave’ (numbers) elements.

e.g. <pitch>

<step>F</step>

<octave>4</octave>

</pitch>

* Note length is represented by the ‘duration’ (number) element. For example, a duration value of 4 would signify a crotchet (1/4 note); 2 would signify a quaver (1/8 note) etc.
* Dotted notes are signified by a </dot> element. Note duration is consistent and intuitive. For example, a dotted crotchet would be displayed as <duration>6</duration> and a dotted quaver would be <duration>3</duration>.
* Other note-specific data given in MusicXML includes the following:
  + Position (given in x and y coordinates)
  + Voice (number)
  + Type (quarter, eighth etc)
  + Stem (up or down)
  + Staff (number)
  + Chord
  + Rest

**Notes on MEI:**

* Only the title of the piece is given. There is no inclusion of composer or subtitle anywhere in the MEI file.
* MEI is structured in a far less standardised manner in comparison to MusicXML. It can vary wildly from file to file, piece to piece, so is not favoured for cross-platform sharing and collaboration. However, as notation data is given in-line in the form of attributes, it is far easier for a human to read, and it doesn’t use as many lines of code as MusicXML.
* Note pitch is represented by the ‘pname’ (letters) and ‘oct’ (numbers) elements.

e.g. <note xml:id = “note-000…” oct = “4” pname = “f” />

* Note length is represented by the duration (number) attribute. For example, dur = “8” signifies a quaver (1/8 note), dur = “4” signifies a crotchet (1/4 note) and so on.
* If two or more notes are played at the same time, the ‘note’ elements are nested inside a ‘chord’ element. If this is the case, the note duration and stem direction attributes will be included inside the ‘chord’ element, avoiding the need to repeat these attributes for every note.
* MEI deals with dotted notes in a slightly different, and more counterintuitive way, in comparison to MusicXML. If a note is dotted, it gives a ‘dots’ attribute with an accompanying numerical value (e.g. dots = “1”). However, the note duration is not affected. Therefore, a dotted crotched will still have dur = “4” but will additionally include dots = “1”.
* Other notation data given in MEI includes the following:
  + Measure number (e.g. <measure n = ”1”>)
  + Staff number (e.g. <staff n = ”1”>)
  + Layer number (e.g. <layer n = ”1”>)
  + Beam
* All elements in MEI are also displayed with a unique ‘xml:id’ attribute.

e.g. <measure xml:id="measure-0000001884856637" n="1">

**Notes on creating different versions of MEI files:**

* I created one different version of the MEI file for the piece You’re So Cool. I decided to adjust every single note’s octave element (‘oct’) from the original file so that the whole piece was one octave higher. For example, notes with the attribute oct = “5” were change to oct = “6” and so on. I saved this new version separately from the original file and both MEI files are included in my dataset.

**Notes on MIDI:**

* MIDI is an acronym that stands for Musical Instrument Digital Interface. It is a technical audio standard that allows devices that make and control sound (such as synthesisers and samplers) to connect and communicate with one another.
* MIDI is far more useful for machine readability and connection/communication between electronic devices. It is not so easy for humans to use as the files do not contain crucial metadata such as title, subtitle, composer and notes among more when they are graphically represented.

**Notes on Verovio:**

* Verovio was used to convert MusicXML files into MEI files. This was mostly successful but there were a few problems with Verovio with regards to encoding and rendering the files. For example, the tempo of the piece would not be recognised so the playback speed was often wrong.
* On the plus side, Verovio did successfully recognised all of the notes and their respective pitches and durations etc. The instruments, bar lines, key signatures, time signatures, accidentals, dynamics and all other variables associated with the pieces of music were properly acknowledged and encoded by Verovio.
* In the specific case of comparing the two versions of the two ‘You’re So Cool’ MEI files (the original and higher octave version), Verovio recognised the change in the oct attribute and thus rendered and displayed the new file appropriately, with the notes sitting much higher on the stave in the octave above.