

# A dataset of TV-Show Opening Theme Song for MIR analysis

Carla Kim Gaieski

Music Information Retrieval

carlakg@yonsei.ac.kr

## ABSTRACT

There are multiple audio datasets available for research purposes. Most contain data on different genres of songs, information according to streaming platforms, or data on instrumental audio, which provides the possibility of the creation of other datasets with different audio and features. In the entertainment industry, sound and music play a crucial role when paired with the images shown on screen. And in TV series, songs are an important aspect of the show. The opening theme song characterizes the show, it becomes its identification tag, which people can recognize it from. Therefore, it can be explored for research purposes. This study is the preliminary construction of a dataset of TV show opening theme songs and their features. By using multiple wav files of different songs, their audio features of zero-crossing rate, RMS, spectral centroid, spectral roll-off, spectral flatness, chroma, beat, tempo, length, and 19 MFCCs values were extracted. Hence, this study requires a continuation of the creation of the dataset, for the possibility of conducting further MIR and qualitative analysis of TV theme songs.

## 1. INTRODUCTION

TV shows have been a means of entertainment for decades. Since its early days of TV, TV shows have had a theme song, a song at the beginning that shows short clips of the show while also displaying the production team and the main actors' names. This song not only characterizes the show but also becomes something that everyone who knows the show can recognize.

Additionally, when watching different genres of TV shows, there also seems to be a correlation between the theme song and the genre of the show. This study explored the relationship between TV shows' opening theme songs and their musical characteristics (symbolic and audio representation), to understand how TV shows set the tone to their genre.

## 2. LITERATURE REVIEW

MIR projects require a large quantity of audio file data. There are multiple papers that use different instruments, such as piano or drums, and different genres of music,

such as blues and pop, and music related to emotions, such as happiness or sadness, that are available for the MIR community to use. There are also different projects in online communities, such as Kaggle, where users work on the GTZAN dataset, or with the resources of streaming apps, such as Spotify. All of these, generally, consist of multiple audio files and their features, beat, tempo, pitch, and many others. This makes it possible for the application of MIR: music search, music recommendation, music classification, and music creation, for example.

However, there is a gap in the audio datasets. As most of them focus on lyrics, instruments, voice, or genres, there is no data available regarding audio from the entertainment industry, more specifically, from TV shows.

In A computational Lens Into How Music Characterizes Genre in Film, Benjamin Ma, Timothy Greer, Dillon Knox, Shrikanth Narayanan explored the soundtracks of different Hollywood movies, to conduct a MIR study of genre classification. The team was successfully able to collect soundtracks from different movies and implement a neural network of CNN to classify movies based on their genre. Also making it available for the public to use their dataset.

Therefore, there is a research utility for musical data present in the entertainment industry with the soundtracks, main title theme songs, and ending credit songs. And the existing audio datasets were important to determine which features are the most important to be extracted.

## 3. METHODOLOGY

### 3.1 Data Collection and Cleaning

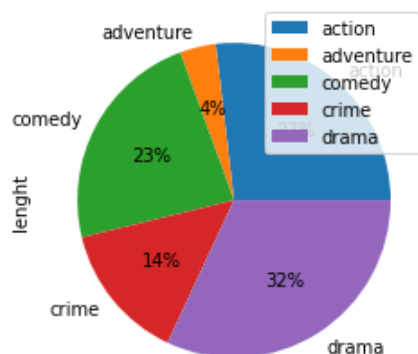
The TV shows were selected according to IMDb ranking of best comedy, best drama, best action, best thriller, best adventure, best sci-fi, and best crime TV series. IMDb is an online platform with information, being a trustworthy source for detailed information regarding TV shows and their ranking and popularity, and their genre classification. The audio files were obtained via YouTube, and converted into .wav files for easier manipulation when using librosa package. In total, 109 audio files were collected from TV series dating from the late 90s until 2022.

Usually most tv series are multi-labeled, that is, they belong into more than one genre. For example, a show can be drama, crime, and thriller (“Breaking Bad”) or comedy and romance (“How I Met Your Mother”), In order to generate a single labeled dataset, only the first genre of each series was considered, as it is the genre that better classify the show as a whole, having greater importance.

Once each series was classified, only five main categories remained: action, adventure, comedy, crime, and romance. However, given the change in the original label by IMDb, this led to a decrease in the number of TV shows for each genre.

Genre	Number of TV Shows
Action	26
Adventure	5
Comedy	33
Crime	12
Drama	33

**Table 1.** The number of TV shows according to their genre that composes the dataset.



**Figure 1.** Percentage of each genre present in the dataset

The majority of the dataset, over 50%, is composed of drama and action TV series opening theme songs.

### 3.2 Feature Extraction

The python librosa package was used to extract the audio features to create the dataset. The features were chosen according to the features that are more commonly used when analyzing audio files with MIR, as well as the ones considered more important to define an audio file, as it

indicates the uniqueness characteristics of the specific file: the domain and frequency-domain features. The symbolic representation of the audio files was also considered important, as it represents how the music is perceived by humans, being an easier factor to be analyzed during future dataset usage.

Audio Features	Features
Time-Domain	Zero Crossing Rate, RMS
Frequency-Domain	Spectral Statistics: centroid, flatness, and roll off; Chroma; MFCC
Symbolic Music Representation	Beat, Tempo, Length

**Table 2.** Features extracted from each audio file.

For the time-domain features, the zero crossing rate indicates how much the signal changes within a frame, being used for analyzing the timber, and the RMS returns the loudness of the sound. For the frequency-domain features, the spectral centroid captures the center of the energy of the audio, its brightness; the spectral flatness returns the noisiness, and the spectral roll off returns the roll-off frequency, The MFCC was divided into its different coefficients, as having them individually can give more precise audio information about the genres. Beat is the basic unit of rhythm, which is the steady pulse of the song, and the tempo is the speed of the beat, which is easy for humans to detect and perceive the difference between songs.

## 4. RESULTS AND IMPLICATIONS

After applying the feature extraction algorithm to all audio files, it was possible to generate a dataset containing 109 TV show opening theme songs with their zero-crossing rate, RMS, spectral centroid, spectral roll off, spectral flatness, chroma, beat, tempo, length, and 19 MFCCs values. The full dataset is available at [MIR/TVOpenS.csv at main · carlakim/MIR \(github.com\)](https://github.com/carlakim/MIR-TVOpenS.csv). It became possible to see some characteristics of each genre, according to the features extracted.

Feature	Genre ranking (highest to lowest)
Chroma	Action, drama, comedy, adventure, and crime

Spectral centroid	Comedy, action, crime, adventure, and drama
Zero crossing rate	Comedy, adventure, action, crime, and drama
Tempo	Crime, drama, adventure, action, and comedy

**Table 3.** Separation of the genres in a ranking according to a feature

However, given that the dataset is still small to generalize, it is the starting point for TV show's theme songs studies. It is necessary to expand the dataset, in order to be able to have a significant amount of data to utilize it for MIR research.

## 5. APPLICATIONS

A TV show theme song dataset can be used in qualitative research analysis of TV series theme songs, as well as different MIR research fields.

### 5.1 Understanding of Emmy Awards classification

The annual Emmy Awards is a TV show award ceremony. In it, there are different categories that highlight different aspects of a TV show, one of which being the opening theme song. There is an award for "outstanding original main title theme music" (*Nominees / Winners*, n.d.), which is given for the best theme song of the year. However, it is not disclosed what characteristics or features that are evaluated in the song, for it to win the title. So, it would be possible to gather the features of the nominees, and the winners to discover patterns and possibly find a pattern in the songs, which ultimately led them to win the award.

### 5.2 Musical analysis over different decades

There has been constant change in music. If enough audio data from TV series of different eras are added to the dataset, it would be possible to use it in an analysis of how the opening title signs also changed over the years.

### 5.3 MIR Research

#### 5.3.1 TV show genre classification

Having a TV show dataset with the audio files genres and its features can be used for a genre classification. With it, it would be possible to apply classification models, such as SVM or neural networks, to analyze the possibility of classifying TV shows solely based on their opening theme songs. Not only, but it would also allow to generate a better classification system, given the results, as there could be mislabeling problems with the current IMDb classification.

## 6. CONCLUSION

It is possible to do audio feature extraction with TV series opening theme songs. This dataset provides the start for future studies to be made under this field. However, given the difference in number of genres in the dataset, and overall low quantity of data, it is necessary to first expand the dataset, and gather more opening theme songs and their features, to create an even and balanced dataset with the same number of audio files for each genre. This is needed to be able to make any further analysis and research using the dataset.

## 7. REFERENCES

- [1] Ma, B., Greer, T., Knox, D., & Narayanan, S: "A computational Lens Into How Music Characterizes Genre in Film" *PLOS ONE*, 2021.
- [2] *Nominees / Winners*. (n.d.). Television Academy. <https://www.emmys.com/awards/nominees-winners>
- [3] *Browse Movies and TV by Genre*. (n.d.). IMDb. <https://www.imdb.com/feature/genre/>