Clustering Analysis of Middle Eastern and East African Music:

Mapping the Sounds of the Swahili coast and the Arab Mashriq

Oscar Gomez, Konstantinos Trochidis, Carlos Guedes

Abstract

This project focuses on creating interactive visualizations of the East African and Middle Eastern music collections from the NYU Abu Dhabi library. In our study we explore the cross-cultural similarities, interactions and patterns of the music excerpts from the different regions and understand these similarities by employing visualization and dimensionality reduction techniques to the data.

The project was implemented in several stages. First, musical features such as Mel Frequency Cepstral Coefficients (MFCCs) and the logarithmic short-time Fourier transform (STFT) were extracted from clips of each song in the collection using the librosa Python library. These features were then collected and reduced to two-dimensional representations using both deep autoencoders and the t-SNE algorithm. Finally, they were clustered using the standard k-means algorithm and plotted using the bokeh Python library, in a dashboard web application that displays the projections of the collections and lets users explore clusters of similar songs and artists. This project constitutes a first step in exploratory analysis with the broader goal of developing new browsing strategies and recommendation systems for these types of music.

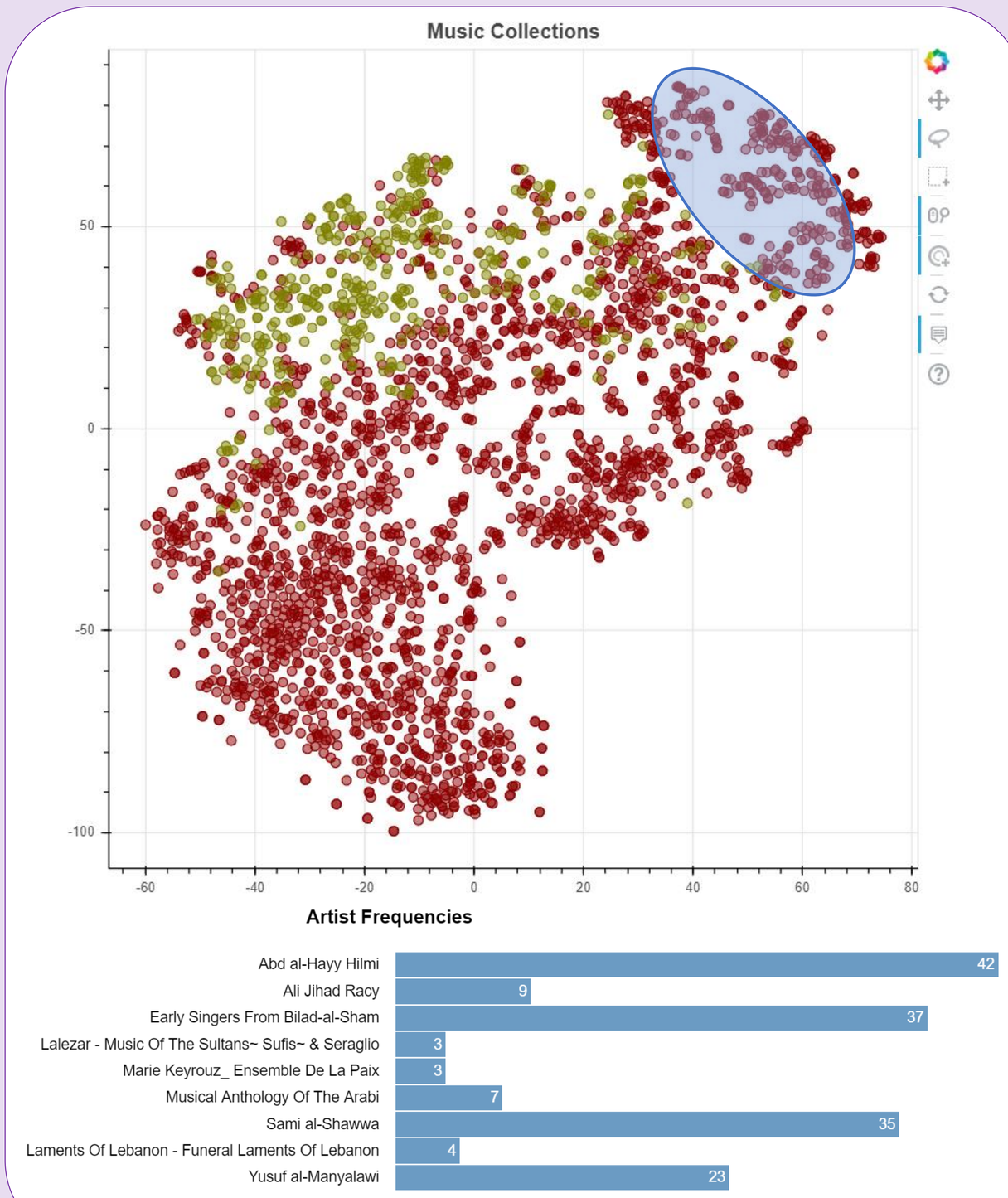

Figure 1: Mapping MFCCs for the two collections using t-SNE.

Music Collections

The corpora for analysis consist of two collections: the Eisenberg Collection and the Music Compendium from the Arab Mashriq. The Eisenberg Collection of East African Commercial Sound Recordings contains 500 sound files and associated metadata of commercial recordings produced for East African Swahili coast audiences between the late 1920s and the fist decade of the twenty-first century. The second corpus of this work consists of a digital compendium of 2827 recordings collected from the Library's collection of Arab audio on compact disc. The ethnic group and region of the digital compendium comes from Jordan, Kurdistan, Turkey, Lebanon, Morocco, Egypt, UAE, Bahrain, Yemen, Afghanistan, Beirut, and Azerbaijan.

Computational Analysis

As a baseline model we extracted standard features (MFCCs) which are spectral representations and are best used to describe the instrumentation and genre/style of the recordings. We tested our baseline model against an unsupervised analysis of the raw representation of the spectrogram (STFT) that is fed to a deep autoencoder and investigate if its able to learn more complicated relationships and patterns of attributes of the music structure. In our deep learning model we are using a series of hidden layers that encode and decode the spectrogram to learn a compressed representation of the important features of the spectrogram (Figure 3). The bottleneck of this autoencoder layer is our final feature representation for the music excerpts. The feature extraction for both MFCCs and STFT were calculated for a duration of 5 seconds taken from the middle of each excerpt so that we have a representative sample of each. We used the t-SNE algorithm to reduce both representations to two dimensions and visualize them in the dashboard application.

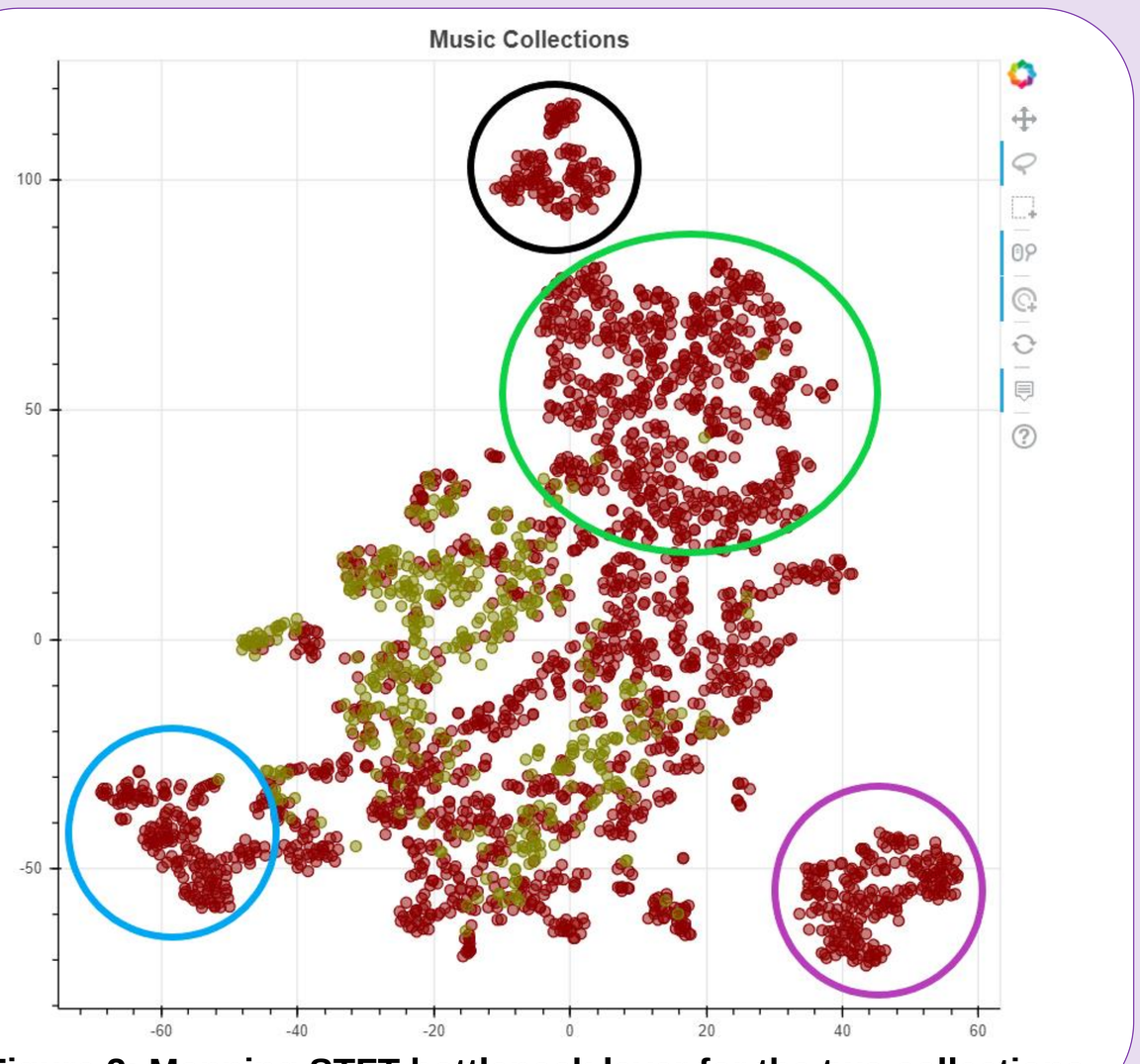

Figure 2: Mapping STFT bottleneck layer for the two collections.

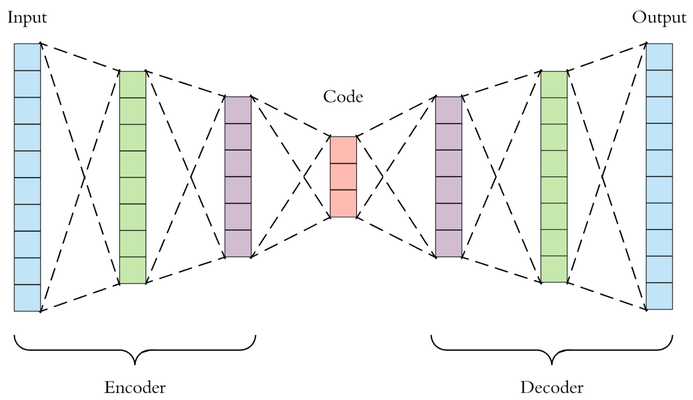

Figure 3: Architecture of the autoencoder model.

Results

Figures 1 and 2 present the mapping of the MFCC and STFT features using t-SNE into a 2D embedding. Red points represent the corpus of the Music of the Arab Mashriq while the green points represent the corpus of the Eisenberg Collection. For the MFCC + t-SNE embedding there is a separation of the two corpus, but other than this separation there is no other clear cluster. On the other hand, the autoencoders + t-SNE embedding of STFT shows more interesting clusters between the two corpus. The first cluster circled in purple includes Persian instrumental music of Tar and Sitar with artists such as Dariush Talai. Another cluster circled in blue includes traditional vocal music from Syria, Lebanon, and Palestine of the fist 3 decades of 1900's with artists such as Ahmad al-Sheikh, Ahmad al-Mir, and Antoine al Shawwa. A third cluster circled in black includes modern electronic and pop Arab music including artists such as Ahlam, Mesaed Al-Belushi, and Abdullah Al Rowaished. Finally, there is another cluster circled in green that includes modern contemporary and classical Arab excerpts including artists such as Sabreen. Folk music artists from both the Music Compendium of the Arab Mashriq and Eisenberg collection are clustered together at the center of the map. Figure 1 shows another feature of this dashboard application where statistics and frequency distributions of the artists for the different areas of the mapping are presented. This is particularly helpful in our exploratory analysis and to get insights regarding the similarities of the artists within the clusters and different areas.