Understanding Music Genre Similarity

Gabriela Groth
Stanford University
gngroth@stanford.edu

1 Introduction

A lot of work has been done on automatic genre classification for music. There are various sources of difficulties in this task, which have made human tagging of music tracks still far better than automated systems. Much of the work on music classification has focused on using supervised learning to classifying music into broad genres like "Hip Hop", "Rock", "Jazz", and "Classical." More recent work has focused on automatically tagging songs with descriptive tags like "cheerful" and "melancholic" for automatic playlist creation and song recommendation programs [1].

One big issue with studying music files is that many of the elements that humans use to construct and deconstruct music, such as pitch, timbre, melody, tempo, vocal qualities, and rhythm, are not easily extracted from music files. Thus there has been an evolution of the features used to classify songs. One of the most common low-level features to use are Mel-frequency cepstral coefficients (MFCCs) which are transformations of the sound frequencies in the audio files. Researchers have been successful to some extent in genre classification using these types of features and supervised machine learning.

The aim of this project is to investigate the substructures of genres within popular music (rock, hip hop, pop, etc.) First music subgenres (ex. blues-rock) will be created using unsupervised learning on audio features that capture timbre, rhythm, and tempo. Then these subgenres or clusters will be further explored by examining the types of songs present in each. Finally, the relationships between subgenres will be explored. Unlike other projects in genre classification, the purpose of this project is not to classify music into human curated genres, but instead to find new genres and relationships by using unsupervised machine learning.

2 Dataset

The data used for this project comes from the MusiClef 2012 Multimodal Music Data Set [1]. This is a collection of music features such as MFCCs, Block-Level Features, and PS09 features for 1355 popular songs by 218 artists. The dataset also includes expert curated tags for each song in the dataset. In order to evaluate and draw conclusions from the resulting clustering of songs, genre information was added to this data set for each artist from Wikipedia. For the purposes of this project, every song by the same artist is associated with the same set of genres. However, expert tags are associated directly with songs not artists.

3 Features and Preprocessing

I used three sets of features to obtain clusters from: block-level features (BLFs), PS09 features and MFCC features. To evaluate these clusters I used expert tags from the MusiClef dataset and genre tags scraped from Wikipedia.

The BLFs are a set of features found to perform well when using supervised learning to classify songs [2]. BLFs include Spectral Pattern (SP), Delta Spectral Pattern (DSP), Variance Delta Spectral Pattern (VDSP), Logarithmic Fluctuation Pattern (LFP), Correlation Pattern (CP) and Spectral Contrast Pattern (SCP). Respectively, these features reflect the timbre, onsets, strength of onsets, rhythmic structure, harmonic relations, and "tone-ness" [2]. The authors who came up with BLFs also created a similarity measure which can be calculated between two songs based on these component audio features. The weight of each component in this similarity measure was optimized to best classify a music dataset using kNN. In light of this, I chose to use the pairwise similarity measures between each song as features.

The PS09 features consist of: Fluctuation Patterns (FPs), Onset Patterns (OPs), and Onset Coefficients (OCs). These roughly measure patterns of loudness and tempo [3]. The authors who created this feature extraction method also created a similarity measure. This measure is a combination of the three novel feature distances and MFCC distances between two songs. The weight of each feature distance in the similarity measure was chosen to optimize classification of songs using kNN. I chose to use this pairwise distance measure which roughly models the difference in loudness patterns, tempo and timbre (because of the added MFCC distance component).

MFCC features were also used by finding pairwise Euclidian distances between MFCC representations of songs and using this distance matrix to perform k-means. In summary, I used similarity scores based on BLFs, PS09, and MFCC features to cluster songs.

4 Models

K-means clustering was used to divide the 1355 songs into clusters of various sizes. I applied k-means to the BLF pairwise similarity scores and the PS09 similarity scores separately and then together. I ran k-means with k ranging from 5 to 100. Since the aim of this project is to understand structure within larger genres, I did not choose a k that was too small. Choosing a k that was too large would create small clusters of very similar songs, as opposed to subgenres. Similar analysis was performed for various choices of k but the presented results will be from k means run with k = 16. Clusters from the BLF similarity scores and PS09 scores when k = 16 provide interesting genre insights without creating an unwieldy number of genres.

Once the songs were clustered, the clusters were examined by looking at the distribution of genres, artists, and tags within each cluster in conjunction with looking at the distances between clusters.

5 Results

Clustering was performed over MFCC, BLF, and PS09 similarity matrices. For every combination of k and similarity matrix, I created a heat map to visualize the clusters (Figure 1 and 2). The heat maps are made by ordering songs by assigned cluster along the x and y axes, and then plotting the pairwise similarity scores using colors to represent similarity. Red is used to signify songs that are similar and yellow represents songs that are very dissimilar. I used these visualizations to narrow down which cluster assignments to further investigate.

The most promising clusterings resulted from running k-means separately on the BLF similarity matrix and the PS09 similarity matrix for k = 16 (Clustering BLF16 and Clustering PS16, respectively). These heat maps are shown in Figures 1 and 2.

Figure 1. K means clusters with blfs, k = 16

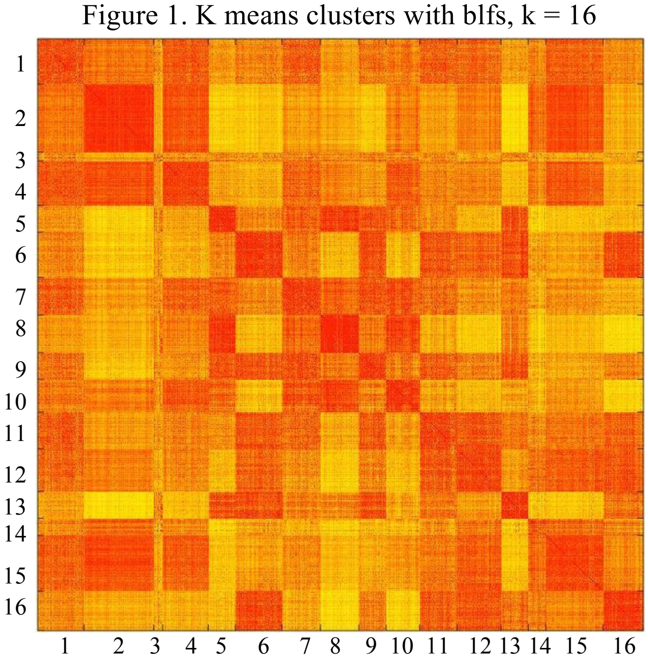
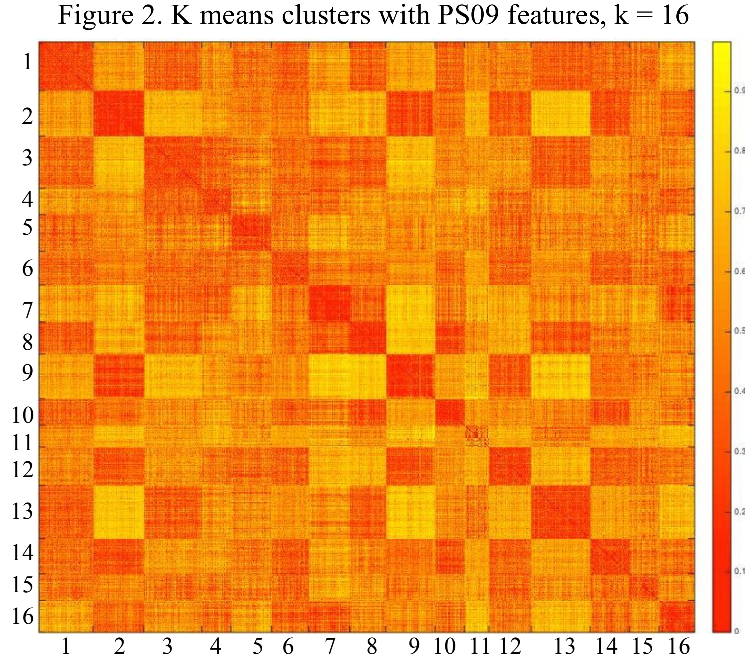


Figure 2. K means clusters with PS09 features, k = 16



After determining cluster assignments for each song, I used expert tags, genre tags, and artists to inspect these clusters. Table 3 in the appendix shows a summary of what types of songs were found in each cluster. I have listed the top five most common tags, genres, and artists found in each cluster. Each of the clusters the BLF16 clustering and PS16 clustering is given a name that corresponds to the subgenre that cluster represents. This name was created by combining the top two expert tags associated with each cluster and the top two Wikipedia genres associated with each cluster.

Next, I found the sum for each cluster of the squared distances between each point and its assigned centroid. I also determined the closest and farthest cluster (based on centroid location) for each cluster. All of this is summarized in Tables 1 and 2.

Table 1. Clusters resulting from BLF features and k = 16.

Clustering on block-level features with k = 16			
Cluster	Within Cluster Sum of Point to Centroid Distance	Clostest Cluster	Farthest Cluster
1 Catchy rock pop rhythm and blues	1636	4 Fast-hard rock	8 Repetitive-beat hip hop rhythm and blues
2 Fast-hard-alternative-punk rock	968	15 Energy pop hard rock	13 Soft romantic rhythm and blues soul
3 Repetitive rhythm and blues rock and roll	499	5 Repetitive-catchy rhythm and blues soul	2 Fast-hard-alternative-punk rock
4 Fast-hard rock	1158	1 Catchy rock pop rhythm and blues	13 Soft romantic rhythm and blues soul
5 Repetitive-catchy rhythm and blues soul	674	8 Repetitive-beat hip hop rhythm and blues	2 Fast-hard-alternative-punk rock
6 Soft-romantic rock rhythm and blues	1291	16 soft romantic pop rock	2 Fast-hard-alternative-punk rock
7 Catchy-movement rhythm and blues rock	1330	9 Catchy-soft rhythm and blues soul	16 soft romantic pop rock
8 Repetitive-beat hip hop rhythm and blues	655	5 Repetitive-catchy rhythm and blues soul	16 soft romantic pop rock
9 Catchy-soft rhythm and blues soul	899	13 Soft romantic rhythm and blues soul	2 Fast-hard-alternative-punk rock
10 Repetitive-fast hip hop rhythm and blues	940	8 Repetitive-beat hip hop rhythm and blues	16 soft romantic pop rock
11 Soft pop rock rhythm and blues	1373	12 Melancholic rock pop	8 Repetitive-beat hip hop rhythm and blues
12 Melancholic rock pop	1279	11 Soft pop rock rhythm and blues	8 Repetitive-beat hip hop rhythm and blues
13 Soft romantic rhythm and blues soul	736	9 Catchy-soft rhythm and blues soul	2 Fast-hard-alternative-punk rock
14 Atmosphere pop rock and roll	656	15 Energy pop hard rock	5 Repetitive-catchy rhythm and blues soul
15 Energy pop hard rock	1232	2 Fast-hard-alternative-punk rock	5 Repetitive-catchy rhythm and blues soul
16 soft romantic pop rock	1035	6 Soft-romantic rock rhythm and blues	8 Repetitive-beat hip hop rhythm and blues

Table 2. Clusters resulting from PS09 features and k =16.

Clustering on ps09 features with k = 16			
Cluster	Within Cluster Sum of Point to Centroid Distance	Clostest Cluster	Farthest Cluster
1 Soft-romantic pop rock	1819	3 Catchy rock pop rhythm and blues soul	9 Chaotic hard -punk-alternative rock
2 Hard-alternative pop rock	1114	9 Chaotic hard -punk-alternative rock	13 Soft romantic rhythm and blues soul
3 Catchy rock pop rhythm and blues soul	2028	13 Soft romantic rhythm and blues soul	9 Chaotic hard -punk-alternative rock
4 Repetitive movement rhythm and blues rock	1113	6 Catchy pop rock rhythm and blues	9 Chaotic hard -punk-alternative rock
5 Fast-repetitive pop rhythm and blues	1653	15 Catchy rock and roll rhythm and blues	7 Beat repetitive hip hop rhythm and blues
6 Catchy pop rock rhythm and blues	1346	15 Catchy rock and roll rhythm and blues	9 Chaotic hard -punk-alternative rock
7 Beat repetitive hip hop rhythm and blues	1033	4 Repetitive movement rhythm and blues rock	9 Chaotic hard -punk-alternative rock
8 Soft romantic pop rhythm and blues	922	3 Catchy rock pop rhythm and blues soul	9 Chaotic hard -punk-alternative rock
9 Chaotic hard -punk-alternative rock	1116	2 Hard-alternative pop rock	13 Soft romantic rhythm and blues soul
10 Melancholic pop rhythm and blues rock	950	14 Energy blues rock pop	9 Chaotic hard -punk-alternative rock
11 Atmosphere-acoustic rock and roll country	902	13 Soft romantic rhythm and blues soul	2 Hard-alternative pop rock
12 Fast hard rock	1376	14 Energy blues rock pop	13 Soft romantic rhythm and blues soul
13 Soft romantic rhythm and blues soul	1589	3 Catchy rock pop rhythm and blues soul	2 Hard-alternative pop rock
14 Energy blues rock pop	1283	2 Hard-alternative pop rock	13 Soft romantic rhythm and blues soul
15 Catchy rock and roll rhythm and blues	1245	6 Catchy pop rock rhythm and blues	7 Beat repetitive hip hop rhythm and blues
16 Repetitive beat hip hop soul	1195	4 Repetitive movement rhythm and blues rock	13 Soft romantic rhythm and blues soul

6 Discussion

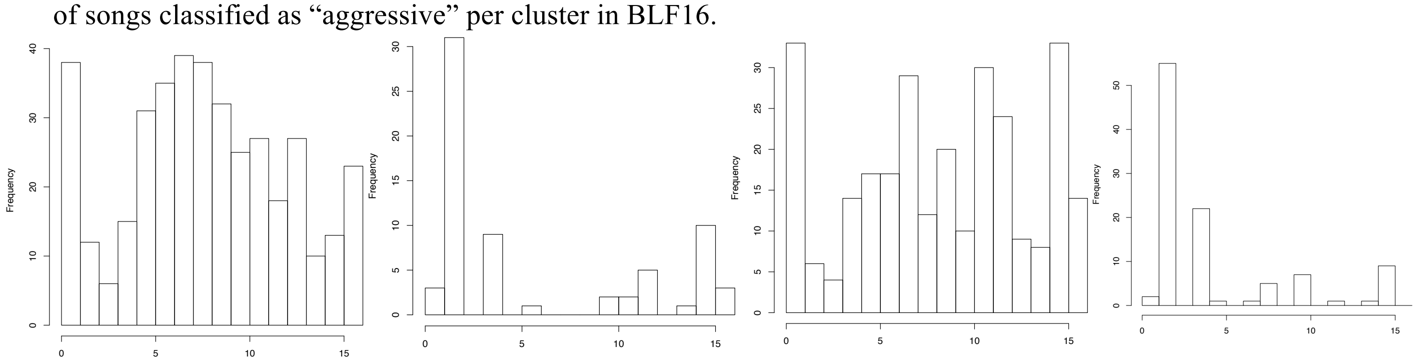
Using artists, genres, and tags to examine the clusters created using k-means, we can see that the audio features are capturing music qualities. First, Table 3 clearly shows that this algorithm combined with these features are able to correctly separate artists into clusters. More importantly, the clusters clearly have different distributions of tags and genres. Although, Figures 1 and 2 show that the genres aren't perfectly and clearly delineated, this is to be expected given that a songs musical influences are usually varied and numerous.

Still, we can see that songs in the same cluster are more similar than songs in other clusters. We can also see certain genres are extremely dissimilar. For example, in Figure 1, Cluster 2 "Fast-hard-alternative-punk rock" is very different from Cluster 12 "Melancholic rock pop", but very similar to Cluster 15 "Energy pop hard rock." Similarly in Figure 2, Cluster 9 "Chaotic hard -punk-alternative rock" is very different from Cluster 7 "Beat repetitive hip hop rhythm and blues" and Cluster 8 "Soft romantic pop rhythm and blues".

From table 2, we see that Cluster 13 "Soft romantic rhythm and blues soul" and Cluster 9 "Chaotic hard -punk-alternative rock" are the most prevalent clusters in the "Farthest Cluster" column. From this, we can conclude that perhaps these clusters are very different from many other musical genres.

Graphs 1 through 4 give further insight into the qualities of music in different genres. For example, Graph 1 shows that virtually every cluster has many songs associated with the "Rhythm and Blues" genre, which is to be expected given the profound influence of rhythm and blues on modern popular music. However, Graph 2 shows that "Heavy Metal" songs are overwhelmingly found in Cluster 2 because Heavy Metal has very distinct characteristics. Similarly, songs tagged as "catchy" can be found in every single cluster, but songs tagged as "aggressive" are also overwhelmingly found in Cluster 2. Thus, this type of analysis helps identify what genres are influential in many types of music versus specialized genres. We can also learn what types of tags are indicative of certain subgenres and what descriptions are vague and easily applicable to many genres.

Graphs 1-4. 1) Number of songs in "Rhythm and Blues" genre per cluster in BLF16. 2) Number of songs classified as "Heavy Metal" per cluster BLF16. 3) Number of songs tagged with "catchy" in BLF16. 4) Number of songs classified as "aggressive" per cluster in BLF16.



7 Conclusion and Future Work

One future direction of interest is to explore the evolution of genres throughout the 20th century. To do this, first I would like to augment this data set with more features and more songs. Specifically, I would like to add data about the year a song was written, the year a song was recorded, the artist's place of origin, and genre information by song as opposed to artist. In addition, an ideal data set would have all of these features with more songs for each year.

Given this augmented data set, similar analysis could be done to find relationships between artist birthplace and song year and subgenre. Further, I would like to create clusters similar to those created in this project but within time periods. Then, I would examine the closest clusters between time periods in order to model the evolution of genres and subgenres. This could also give insight into how past genres are combined to create new genres with multiple influences.

The question of discovering hidden relationships between music genres across time is not only a topic of interest, but it also has potential commercial applications. This exploration could potentially lead to using machine learning to determine artist influences, which is directly applicable to playlist creation and song recommendation.

Table 3. Summary of BLF16 clusters and PS16 clusters.

Clusters using Block-Level Features and k = 16								Clusters using PS09 Features and k = 16							
Cluster	Songs In Cluster	Top 5 Tags	# songs with tag	Top 5 Genres	# songs with genre	Top 5 artist	# songs by artist	Cluster	Songs In Cluster	Top 5 Tags	# songs with tag	Top 5 Genres	# songs with genre	Top 5 artist	# songs by artist
1	104	catchy	33	rhythm and blues	38	Allman Brothers Band	5	1	112	soft	66	pop	44	Bee Gees	5
		rock pop	32	pop	35	Carl Perkins	4			romantic	57	rock	40	Carpenters	5
		movement	30	rock	31	The Animals	3			melancholic	49	rhythm and blues	28	Righteous Brothe	5
		fast	27	rock and roll	28	Gene Vincent	3			love	44	folk rock	23	Neil Young	4
		energy	24	soul	28	Jackie Wilson	3			rock pop	44	country	21	Patti Smith	4
2	156	hard rock	91	alternative rock	39	Ramones	8	2	105	hard rock	43	alternative rock	36	Verve	6
		fast	65	punk rock	39	Sex Pistols	8			rock pop	40	rock	27	Weezer	5
		chaotic	63	hard rock	36	Guns N Roses	7			powerful	39	blues rock	25	Yeah Yeah Yeahs	5
		agitated	60	blues rock	32	Metallica	7			energy	38	hard rock	25	AC/DC	4
		powerful	56	rock	32	Black Sabbath	6			fast	36	punk rock	19	Aerosmith	4
3	20	repetitive	9	rock and roll	7	Chubby Checker	3	3	120	catchy	39	rhythm and blues	53	Bobby Darin	6
		rock & roll	8	rhythm and blues	6	Kanye West	3			rock pop	32	soul	44	Earth Wind & Fire	5
		movement	7	country	4	The Animals	1			movement	20	rock	40	Jackie Wilson	5
		ballroom	6	gospel	3	Bob Dylan	1			vintage	20	pop	33	Otis Redding	5
		carefree	5	hip hop	3	Booker T. & The MGs	1			rhythm and blues	26	rock and roll	28	The Temptations	5
4	102	fast	49	rock	27	The Clash	4	4	60	repetitive	27	rhythm and blues	28	Africk Bambaataa	5
		hard rock	42	hard rock	22	Franz Ferdinand	4			movement	21	rock	22	Prince	5
		energy	39	rock and roll	22	Gnarls Barkley	4			fast	16	soul	17	Little Richard	4
		powerful	36	blues rock	21	AC/DC	3			energy	15	disco	16	Beastie Boys	3
		repetitive	31	rhythm and blues	15	Billy Joel	3			rock pop	15	pop	14	Gloria Gaynor	3
5	61	repetitive	21	rhythm and blues	47	James Taylor	6	5	84	fast	26	pop	29	Four Tops	7
		catchy	17	soul	35	Solomon Burke	6			repetitive	23	rhythm and blues	27	Cream	6
		movement	16	pop	35	Aaron Neville	4			rock pop	23	rock	23	Stevie Wonder	4
		vintage	15	rock	34	Jackson Browne	4			movement	22	rock and roll	23	Diana Ross	3
		night	14	blues	31	Joni Mitchell	4			vintage	21	soul	23	Doors	3
6	104	soft	71	rock	28	Marvin Gaye	4	6	77	rock pop	31	rock	28	Marvin Gaye	4
		romantic	53	rhythm and blues	22	B.B. Kind	3			catchy	30	rhythm and blues	22	B.B. Kind	3
		melancholic	50	soul	18	The Band	3			travel	17	blues rock	18	The Band	3
		slow	33	pop	18	Foreigner	3			vintage	17	pop	18	Foreigner	3
		love	29	blues	17	Aerosmith	2			carefree	16	soul	17	Aerosmith	2
7	86	catchy	29	rhythm and blues	39	Creedence Clearwater Revival	4	7	85	beat	57	hip hop	42	Missy Elliott	8
		movement	28	rock	32	Police	4			repetitive	57	rhythm and blues	40	50 Cent	7
		rock pop	26	soul	22	Gloria Gaynor	3			rap	48	soul	26	Justin Timberlake	6
		vintage	25	pop	20	Jackie Wilson	3			hip-hop	40	pop	20	Beyonce	5
		energy	22	rock and roll	17	Madonna	3			night	32	funk	14	Outkast	5
8	86	repetitive	58	hip hop	42	Missy Elliot	7	8	74	soft	49	rhythm and blues	34	Aaron Neville	4
		beat	53	rhythm and blues	28	Public Enemy	7			romantic	39	pop	33	Al Green	4
		rap	47	soul	25	50 Cent	6			melancholic	29	soul	31	Jackson Browne	4
		hip-hop	34	funk	23	Chic	5			love	27	rock	26	Joni Mitchell	4
		night	33	pop	23	Justin Timberlake	5			slow	25	blues	23	Solomon Burke	4
9	61	catchy	20	rhythm and blues	32	John Lee Hooker	3	9	102	hard rock	51	punk rock	28	Sex Pistols	8
		soft	20	soul	26	Otis Redding	3			chaotic	42	alternative rock	26	Ramones	7
		rock pop	19	pop	16	Al Green	2			fast	41	rock	25	Oasis	6
		rhythm and blues	18	blues	14	Aretha Franklin	2			agitated	37	heavy metal	16	Guns N Roses	5
		vintage	16	rock	14	B.B. King	2			aggressive	36	hard rock	15	Nirvana	5
10	75	repetitive	36	hip hop	25	Africka Bambaataa	5	10	61	rock pop	26	rhythm and blues	20	Amy Winehouse	4
		fast	26	rhythm and blues	25	Beastie Boys	5			melancholic	21	rock	20	Bonnie Raitt	3
		beat	25	soul	23	Eminem	4			soft	21	alternative rock	18	Coldplay	3
		movement	22	disco	17	Gnarls Barkley	3			romantic	18	pop	18	Elvis Costello	3
		rap	19	rock	17	Jackson Five	3			love	16	gospel	15	Radiohead	3
11	84	rock pop	36	pop	31	Righteous Brothers	4	11	50	atmosphere	14	rock and roll	14	Chubby Checker	8
		soft	32	rhythm and blues	27	Bobby Darin	3			acoustic	12	country	10	Hollywood Studio	5
		romantic	31	soul	23	Diana Ross	3			rock & roll	12	rhythm and blues	10	Jerry Lee Lewis	3
		catchy	20	rock	22	The Band	2			melancholic	11	blues	9	Tina Turner	3
		love	26	rock and roll	19	The Beatles	2			soft	11	film score	9	Kanye West	2
12	98	melancholic	41	rock	38	George Harrison	5	12	88	fast	38	rock	35	New York Dolls	6
		rock pop	39	pop	26	Jefferson Airplane	4			hard rock	33	hard rock	24	Jam	4
		soft	35	rhythm and blues	18	U2	4			rock pop	32	pop	22	Alice Cooper	3
		melodic	31	folk rock	16	Del Shannon	3			energy	26	blues rock	17	ABBA	2
		romantic	31	alternative rock	15	Dolly Parton	3			powerful	23	rhythm and blues	16	AC/DC	2
13	61	soft	36	rhythm and blues	27	Tina Turner	4	13	122	soft	78	rhythm and blues	47	Platters	6
		romantic	27	soul	27	Aaron Neville	3			romantic	67	soul	41	Ray Charles	6
		melancholic	25	pop	25	Al Green	3			melancholic	62	pop	35	Johnny Cash	4
		love	19	country	19	Chubby Checker	3			slow	51	rock	35	Randy Newman	4
		slow	18	gospel	19	Isley Brothers	3			love	44	gospel	29	Aretha Franklin	3
14	39	atmosphere	14	rock and roll	13	Ronettes	5	14	81	rock pop	30	rock	32	Allman Brothers	6
		rock & roll	13	pop	10	Hollywood Studio Symphony	4			energy	25	blues rock	18	Del Shannon	4
		movement	9	rhythm and blues	10	Pink Floyd	3			travel	21	alternative rock	16	Foreigner	3
		anxiety	8	blues rock	7	Velvet Underground	3			hard rock	20	hard rock	16	Lynard Skynard	3
		catchy	8	film score	7	Everly Brothers	2			melancholic	19	country rock	11	Tom Petty	3
15	128	rock pop	54	rock	52	Cream	6	15	61	rock & roll	33	rock and roll	32	Carl Perkins	7
		energy	47	hard rock	36	Lynard Skynard	5			catchy	27	rhythm and blues	29	Gene Vincent	5
		hard rock	39	blues rock	33	Verve	5			movement	25	rockabilly	22	The Animals	3
		fast	38	pop	28	Bruce Springsteen	4			fast	21	country	16	Chuck Berry	3
		catchy	33	alternative rock	27	The Byrds	4			ballroom	19	rock	13	Jerry Lee Lewis	3
16	90	soft	63	pop	26	Carpenters	4	16	73	repetitive	29	hip hop	24	Chic	5
		romantic	58	rock	24	Roberta Flack	4			beat	25	soul	23	Gnarls Barkley	5
		melancholic	52	rhythm and blues	23	Bee Gees	3			fast	22	funk	19	Jackson Five	5
		slow	45	folk rock	16	Everly Brothers	3			night	21	disco	16	Eminem	4
		love	43	soft rock	16	Jeff Buckley	3			rap	20	rhythm and blues	15	Jay Z	4

9 References

[1] Schedl, M., Liem, C.C.S., Peeters, G., and Orio, N. A Professionally Annotated and Enriched Multimodal Data Set on Popular Music. Proceedings of the 4th ACM Multimedia Systems Conference (MMSys 2013), Oslo, Norway.

[2] K. Seyerlehner, G. Widmer, and T. Pohle. Fusing Block-Level Features for Music Similarity Estimation. In Proceedings of the 13th International Conference on Digital Audio Effects (DAFx), Graz, Austria, September 2010.

[3] T. Pohle, D. Schnitzer, M. Schedl, P. Knees, and G. Widmer. On Rhythm and General Music Similarity. In Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR), Kobe, Japan, October 2009.