

# SUPERVISED LEARNING FOR GENRE CLASSIFICATION OF AUDIO TRACKS

Author: Angel Bergantiños Yeste

Director: Dr. Sergio Escalera Guerrero

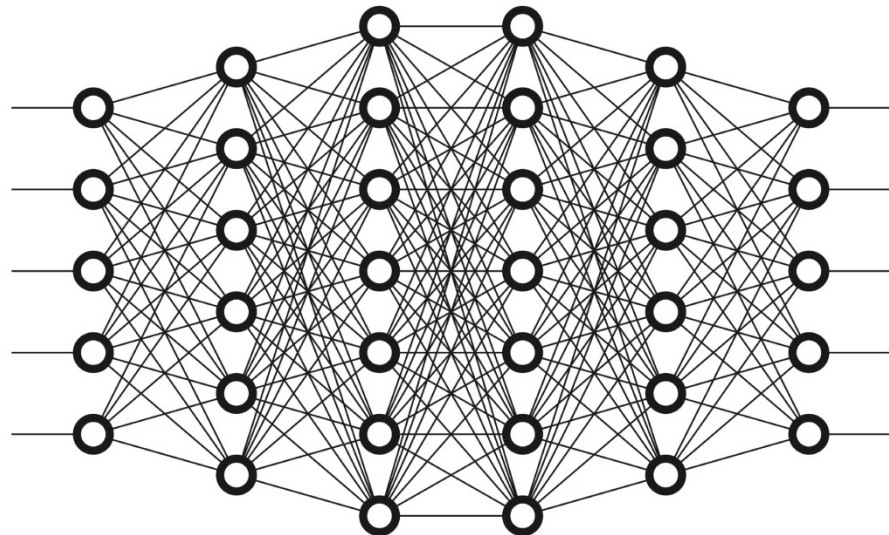
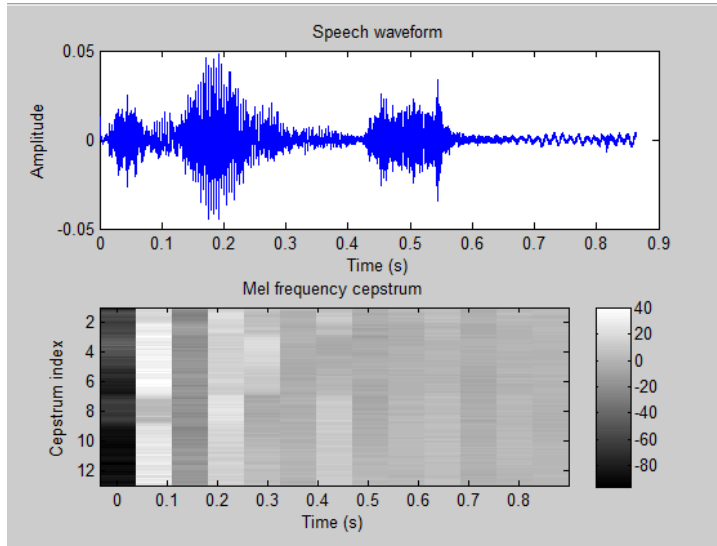
# Context and motivation



# History



# State of the art

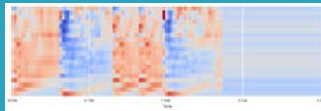


# Method whole overview

RAW SONG



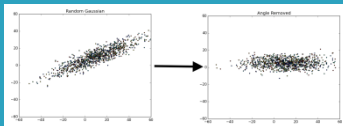
MFCC



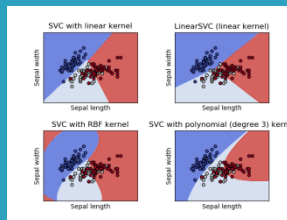
HISTOGRAMS



PCA



LEARNING



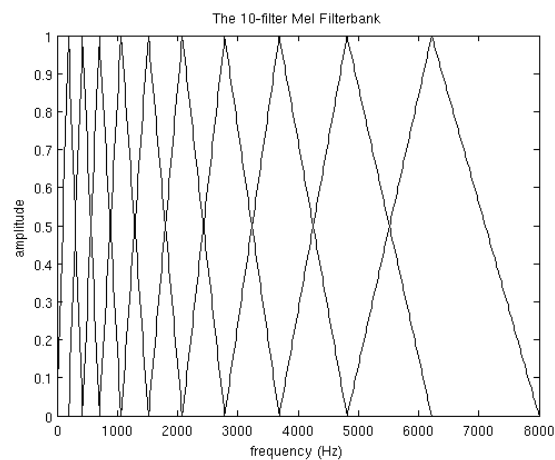
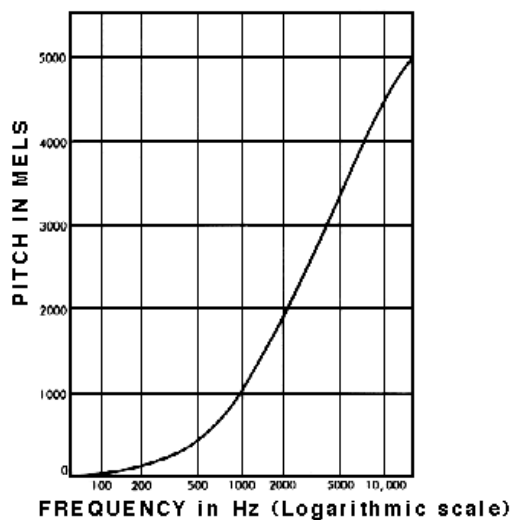
GENRE



# Method: MFCC

## Mel Scale

$$m = 2595 \cdot \log_{10} \left( 1 + \frac{F}{7000} \right)$$



# Method: MFCC

$$C_n = \sum_{k=1}^k (\log D_k) \cos \left[ m(k - \frac{1}{2}) \frac{\pi}{k} \right]$$

-0.44902	-0.22222	0.07720	1.11410	1.06772	1.01322	0.94002	-0.91341	-0.92722	0.92255	-0.10950	-0.02599
-5.06668	-4.64326	4.06634	-2.24359	1.40211	0.59463	-1.47944	1.59902	-0.21553	0.60075	-0.15000	-0.42364
-1.00555	-0.92077	3.72055	-0.97051	1.02377	-1.23002	-1.44634	0.95555	-1.25980	0.29004	-0.67255	0.16455
-0.21097	-0.70117	3.14213	0.69700	1.77082	-0.52446	-1.03357	3.44669	-0.47513	0.39944	-0.12312	0.71217
-1.30931	-0.26693	2.01033	-2.34411	-0.36008	0.99333	-0.54446	0.99771	0.14564	0.60997	-0.15044	-0.14097
-1.72596	-6.45350	4.16380	-2.09122	-0.52133	0.98557	-0.74300	0.58644	0.23550	1.19280	0.14677	-0.04559
-1.06460	-3.05455	3.55466	-0.96316	0.94316	0.32003	-1.14448	1.03118	-0.40677	0.11223	0.12637	-0.42339
-0.40112	-2.38324	1.30892	-0.49051	-0.21711	0.11642	-0.30155	1.27855	0.06400	0.69255	0.30316	-0.14455
-0.07713	-0.08699	1.14411	-0.14211	0.00588	0.14440	-0.24446	1.14441	0.25244	0.00076	-0.44316	-0.44316
0.49467	-4.42002	0.92755	-2.31688	-1.17777	1.11550	0.44566	2.34468	0.06312	-0.00079	0.63280	-0.55211
0.12077	-3.37991	0.77377	-0.55946	-0.30177	0.99226	0.24668	1.00993	0.28448	0.44119	0.50000	-0.29216
-0.07210	-3.27543	0.26699	-0.04659	0.22349	0.47516	0.49992	1.40312	0.14552	0.01122	-0.23011	-0.03312
-0.33392	-0.46377	0.76744	0.06656	-0.41044	0.90464	0.24077	0.00000	0.72111	0.00007	0.19316	-0.03700
-0.00211	-2.43116	0.65544	0.21228	0.22855	1.71551	0.00026	0.90311	0.31377	0.64477	-0.20318	-0.77781
1.20808	-4.74609	1.47711	-1.90316	0.18806	0.79117	-1.90556	0.00885	-0.01044	-0.65300	-0.55596	-0.70888
4.34084	-7.08500	2.20555	-3.03011	0.40777	0.30009	-2.41066	1.05220	-0.41955	-0.35566	-1.47660	-1.02800
5.11011	-0.26319	1.41111	-2.05519	-0.07477	-0.97555	-2.30026	0.65277	-1.00880	-0.02400	-1.42310	-1.00990
5.34070	-6.30008	1.72493	-3.16411	-0.46697	-0.94444	-2.40077	0.40039	-1.31750	-0.07213	-1.57914	-1.20514
4.42099	-2.27955	2.08916	-1.79447	0.12999	-0.67778	-2.09999	0.74555	-0.52077	0.15553	-0.28449	-0.56699
0.30947	-2.22947	1.53119	-2.03883	1.22813	0.77549	-1.25312	0.77727	-0.70755	0.71440	0.22446	-0.45444
0.75311	1.11000	0.67794	-0.23002	1.00890	0.43003	-0.77911	0.50868	-0.07442	-0.11000	0.33099	-0.67712
0.99889	-0.04440	3.24700	-0.44448	0.41999	-0.86778	-3.07448	-0.24000	-1.07442	-0.00011	0.07000	-1.33393
7.40310	-1.14000	3.22446	-0.04418	0.25312	-1.00338	-3.12315	-0.32227	-2.44447	-0.55005	0.47044	-1.70112
7.59310	-0.03154	2.70655	-1.10442	0.45119	-1.02700	-2.93999	-0.42999	-2.20244	-0.43300	-2.52444	-1.70990
7.49050	-0.09319	3.13211	-0.38005	0.72555	-0.77617	-2.77811	-0.34211	-2.22310	-0.44442	0.45002	-1.33880
4.42446	1.10916	1.71112	0.34688	0.17911	-0.24000	-1.70443	-1.14517	-1.40006	-0.70995	0.10116	-1.10914
6.10544	-3.09916	5.00314	-0.16110	-0.05008	0.23311	-0.00215	1.11855	-0.66606	0.04780	0.35311	-0.05002
7.13577	-1.45330	4.97773	-0.21223	0.38814	-0.41999	0.25900	0.03210	-0.40991	0.30400	-0.21312	0.24003
0.20067	-1.14740	4.12077	-0.55311	-0.97029	0.33617	-0.11199	0.53400	-0.47611	0.02446	-0.07100	-0.59211

Each row represents the Coefficients of one frame

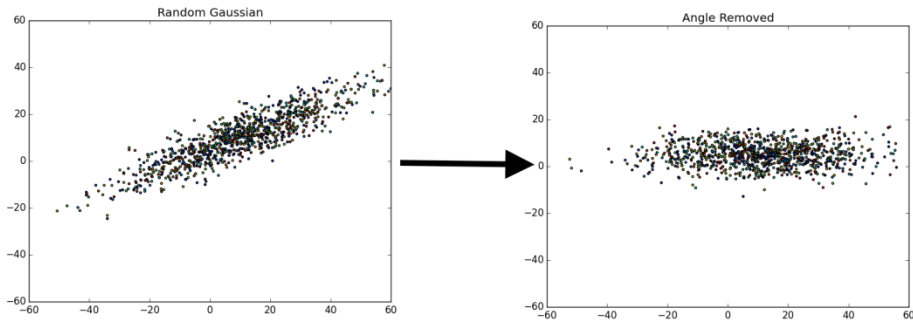
1

Each column represents one Extracted Coefficient

12

7

# Method: PCA and Feature representation





# Method: PCA and Feature representation

# Method: learning schemes

# Results: Data

## Genres

Blues	Jazz
Classical	Metal
Country	Pop
Disco	Reggae
Hip hop	Rock

## Properties

- **Sample rate:** 22050Hz
- **Channels:** 1 (Mono)
- **Frame rate:** 22050 fps

Results: Evaluation protocol and method parameters.

# Results: Discussion

# Results: Discussion

# Results: Discussion

# Conclussion and future work





**Thanks for your attention!**