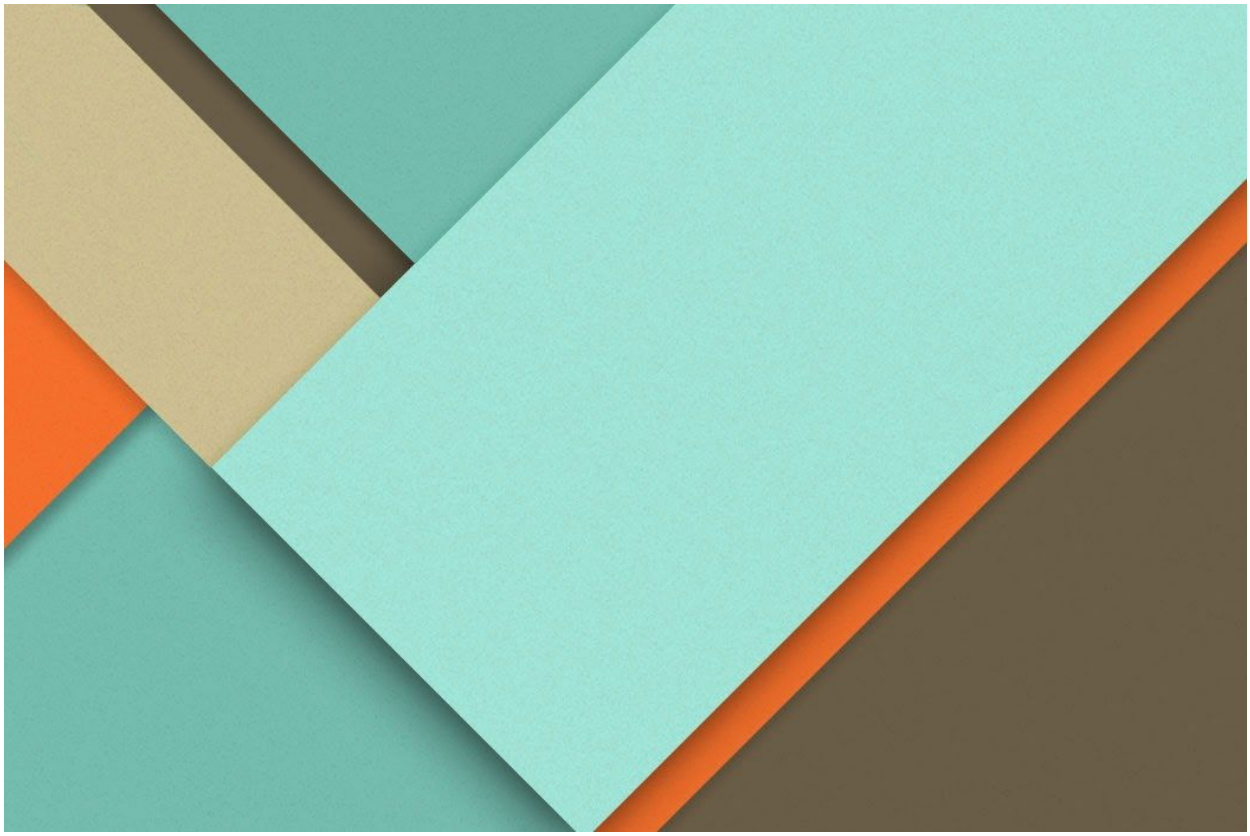


Group Assignment-2

INSTRUCTOR: GOUTAM SAHA

Group Members :-

Harshavardhan Alimi - 18EC10021

Sandeep reddy kasireddy - 18EC10026

Shashank pilla - 18EC10042

Chinni krishna chapala - 18EC10010

Aim:

- To extract the spectrum of audio files and find similarity or dissimilarity amongst them.
- To come with an idea for implementing a lock key system using audio as key.

Description:

- 4 people volunteered to provide their voice samples for analysis.
- Each person recorded 8 samples in 4 days i.e one sample in the morning and one sample after lunch for 4 days.
- We intend to derive a relation among the different audio samples by analysing their spectrum.

Observations & Discussion:

- Top row in each figure indicates morning samples and the bottom row indicates afternoon samples.
- Below four figures represent the frequency response of the samples.

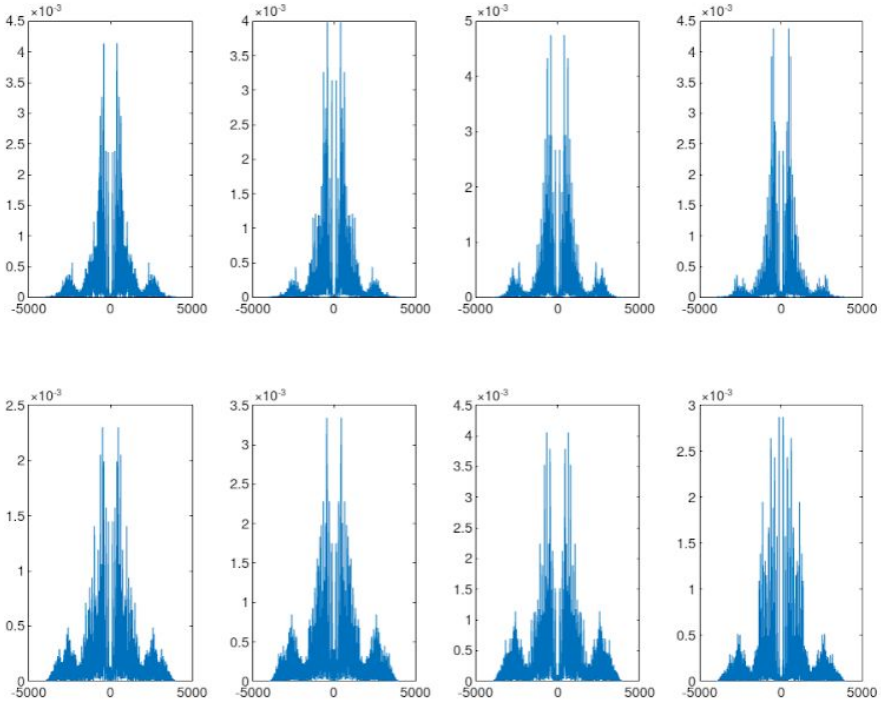


Fig-1: RUDRA AUDIO SPECTRUM

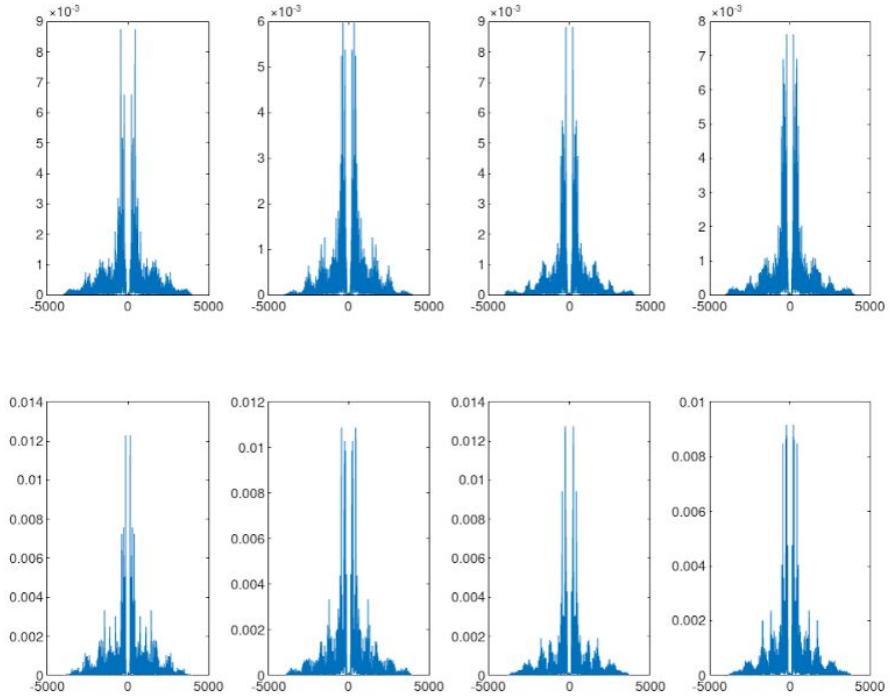


Fig-2: SHOUVIK AUDIO SPECTRUM

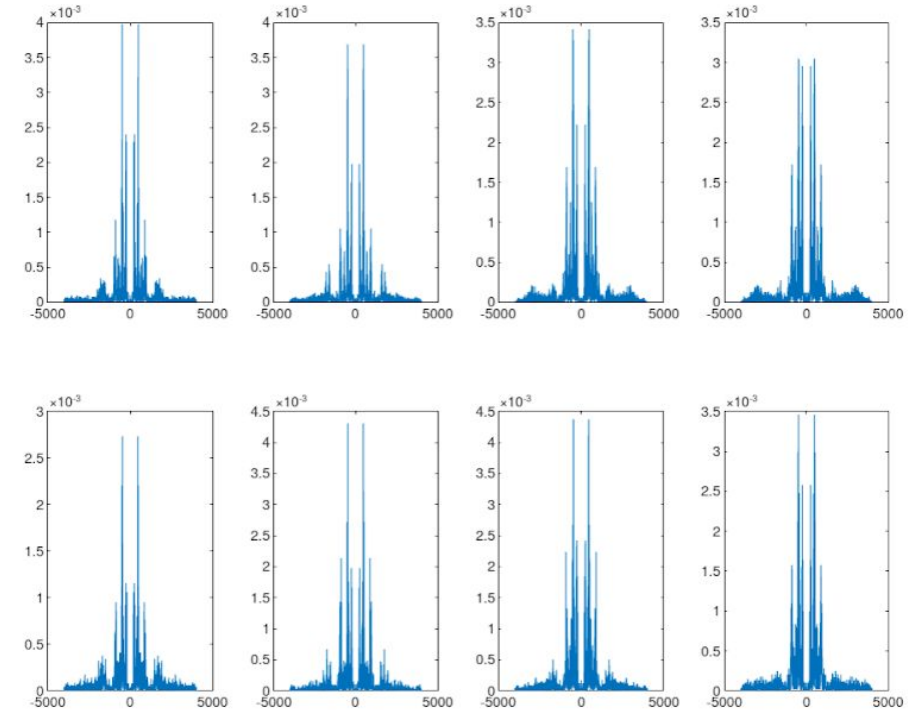


Fig-3: PRERNA AUDIO SPECTRUM

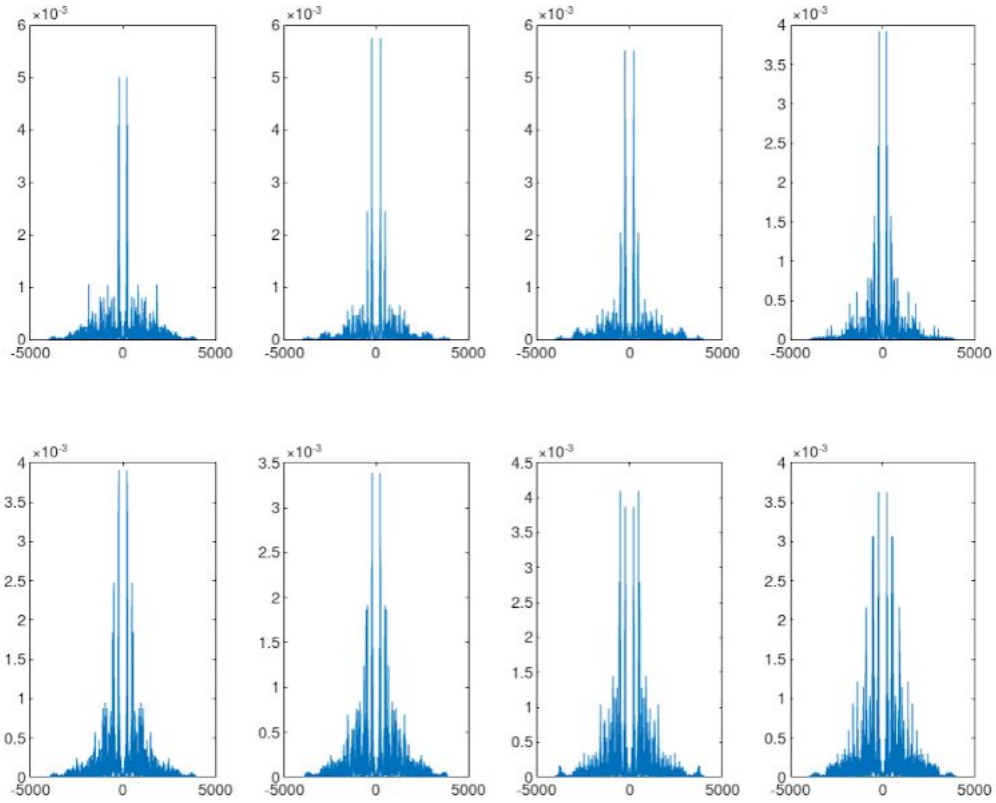


Fig-4: NIVEDITA AUDIO SPECTRUM

- Rudra audio spectrum has the frequency response like falling and increasing and again falling,this property can be used for this particular person to extract an algorithm that allows only a speaker to be able to open the lock by uttering the combination lock number.
- The energy values of the speaker are different in the afternoon when compared to the morning samples.
- The spectrum of Rudra has higher values in the afternoon compared to the morning.
- Shouvik audio spectrum shows that he has higher energy values when compared to others(vary by a great margin).
- The boys(rudra and shouvik) spectrums are similar in terms of spread of energy i.e., here the energy got spread in more no.of frequencies.
- The girls(prerna and nivedita) spectrums are similar in terms of spread of energy i.e., here the energy lies in only a few frequency components.
- The audio signal is passed through equi-spaced filters covering the speech spectrum i.e. if 10 filters are used then 0-400 Hz (LPF), 400-800 Hz(BPF), 800-1200 Hz (BPF) ... etc.
- The below tables show the energy values in each band when 10 filter banks are taken.
- Blackman Window used for this evaluation.

	Energy vector of rudra samples,when banks=10,equally spacing from 0 to 4KHz									
F in Hz	0-400	400-800	800-1200	1200-1600	1600-2000	2000-2400	2400-2800	2800-3200	3200-3600	3600-4000
Day-1,morn	51.0105	167.6455	23.7118	8.4039	0.8207	2.3932	3.574	0.7278	0.0819	0.0092
Day-2,morn	38.4225	108.4323	20.3075	10.1709	0.5244	1.3287	1.4441	0.1818	0.0345	0.0039
Day-3,morn	52.5178	171.4932	40.6155	6.7695	0.4681	3.4822	4.3972	1.4274	0.0506	0.0018
Day-4,morn	33.8749	77.5936	17.7465	2.6619	0.311	0.6297	0.727	0.2149	0.0086	0.0015
Day-1,noon	15.0793	36.3262	13.095	6.9339	1.5282	1.4096	2.6585	0.8599	0.4891	0.0291
Day-2,noon	23.0774	80.3102	32.123	10.369	1.2675	2.6802	6.5392	1.9232	1.0533	0.0649
Day-3,noon	27.2349	143.4005	39.4444	17.6447	2.0707	4.9847	13.7958	5.2075	1.4527	0.095
Day-4,noon	32.4273	92.8004	67.866	29.0438	1.4642	2.9495	8.3879	2.5392	0.7089	0.0978

	Energy vector of shouvik samples, when banks=10,equally spacing from 0 to 4KHz									
F in Hz	0-400	400-800	800-1200	1200-1600	1600-2000	2000-2400	2400-2800	2800-3200	3200-3600	3600-4000
Day-1,morn	360.684	504.9385	17.7179	24.5383	21.5657	6.5128	9.001	1.3566	0.9467	0.4168
Day-2,morn	377.037	347.0862	38.7029	26.7507	22.678	3.4207	8.2394	0.2526	0.395	0.1558
Day-3,morn	394.441	225.7708	22.5095	23.6466	13.2497	1.6392	4.3268	0.208	0.6267	0.6572
Day-4,morn	364.475	241.0218	18.4183	25.464	13.131	1.9745	4.5123	0.4332	0.9984	0.7025
Day-1,noon	762.68	117.3884	81.5178	66.3882	72.789	10.0861	23.6334	2.1233	2.1156	0.1178
Day-2,noon	936.297	529.0998	66.3029	48.0978	37.5907	6.7189	13.1542	2.4866	1.5621	0.4073
Day-3,noon	714.782	460.7944	57.6034	41.8803	48.3043	6.6377	17.1141	4.6396	0.9919	0.0887
Day-4,noon	805.495	359.2747	80.3835	27.5361	51.2023	3.6409	9.4308	2.1791	0.8779	0.1074

	Energy vector of prerna samples,when banks=10,equally spacing from 0 to 4KHz									
F in Hz	0-400	400-800	800-1200	1200-1600	1600-2000	2000-2400	2400-2800	2800-3200	3200-3600	3600-4000
Day-1,morn	29.2684	58.9473	10.1741	1.1664	2.366	0.1539	0.1093	0.1132	0.15	0.107
Day-2,morn	16.716	73.6128	7.8228	0.9513	2.3725	0.4996	0.3774	0.1202	0.0836	0.094
Day-3,morn	24.5934	117.9474	25.5009	0.386	0.8602	0.2971	0.5889	0.9764	0.1474	0.0466
Day-4,morn	42.2961	100.2956	30.9207	0.5844	0.5059	0.4288	0.7465	1.267	0.3769	0.0625
Day-1,noon	10.8921	55.6741	8.1985	1.2636	3.8433	0.6266	0.6918	0.2788	0.1389	0.1495
Day-2,noon	18.3978	109.45	30.985	1.4514	2.9912	0.5013	0.4919	0.1954	0.0936	0.1401
Day-3,noon	29.2049	100.6196	31.2771	0.8526	2.5325	0.7649	0.6409	0.2027	0.1539	0.1125
Day-4,noon	26.5018	117.4521	25.7863	0.4865	1.2852	0.511	0.6331	0.6846	0.3395	0.2102

	Energy vector of nivedita samples,when banks=10,equally spacing from 0 to 4KHz									
F in Hz	0-400	400-800	800-1200	1200-1600	1600-2000	2000-2400	2400-2800	2800-3200	3200-3600	3600-4000
Day-1,morn	98.6887	13.1025	9.8183	6.1196	7.5484	2.1378	0.6751	0.3127	0.0429	0.078
Day-2,morn	79.0565	20.89	3.3188	3.6707	1.2656	0.2107	0.4233	0.218	0.0166	0.0414
Day-3,morn	88.6389	21.1691	5.5134	3.2401	1.8172	0.6691	0.5877	0.5464	0.023	0.0712
Day-4,morn	50.444	12.0155	1.1471	1.443	1.1864	0.1335	0.0663	0.0822	0.0326	0.0109
Day-1,noon	72.6735	45.322	14.5225	4.0748	1.1975	0.5662	0.4627	0.1795	0.0343	0.0353
Day-2,noon	52.5121	50.4365	13.5333	7.6397	1.7237	0.8628	0.8124	0.2377	0.0466	0.0785
Day-3,noon	46.8422	79.7744	19.5697	8.9919	1.5634	0.7721	0.8191	0.1934	0.0635	0.2362
Day-4,noon	54.347	82.9089	31.5127	5.506	3.9333	1.4643	0.9625	0.3969	0.0578	0.0562

- These energy vectors shows that most of the energy of
  - rudra lies in 400-800 Hz band
  - Shouvik lies in 0-800 Hz band
  - Prerna lies in 400-800 Hz band
  - Nivedita lies in 0-800 Hz band
- The similarity and dissimilarity of audio signals can be observed using the euclidean distance,where euclidean distance is the distance between 2 vectors which consists of the Energy value of each band as the vector values.
- As there are 32 samples available so,we got 32C2 of values which should be compared.
- The matlab code consists of the matrix dist which gives euclidean distance between every sample.
- From that matrix we can observe that the girl sample to girl sample have lower euclidean distance(value less than 10) when calculated,which shows they have a nearly similarity in audio speech.
- From the matrix we can observe that when we compare a afternoon sample with a afternoon sample of the same person then we get the euclidean distance very low(value less than 1) i.e., showing the signals are similar,but when an afternoon sample is compared with a morning sample of the same person then we didn't get lower values,we got some moderate values (ranging from 2 to 8) which shows that they are nearly similar but not exactly same.
- When we compare shouvik sample with any other samples then we get larger euclidean distance (value greater than 100) i.e., showing the dissimilarity in the speech.
- The audio signal is passed through equi-spaced filters covering the speech spectrum i.e. if 20 filters are used then 0-200 Hz (LPF), 200-400 Hz(BPF), 400-600 Hz (BPF) ... etc.

- The below tables show the energy values in each band when 20 filter banks are taken,part of that 20 bands is taken in the below tables(which play a crucial role i.e., consists most of the signal).

	Energy vector of rudra samples,when bands=20,equally spacing from 0 to 4KHz									
F in Hz	0-200	200-400	400-600	600-800	800-1000	1000-1200	1200-1400	1400-1600	1600-1800	1800-2000
Day-1,morn	15.1844	35.7929	85.5162	72.7322	11.8554	8.7846	5.2758	2.5934	0.6057	0.1675
Day-2,morn	14.0358	24.3529	58.5825	44.9653	9.9731	8.0985	6.77	2.7175	0.3569	0.1396
Day-3,morn	15.2477	37.2269	63.21	85.8468	27.1246	8.7702	5.4784	1.2052	0.3061	0.1463
Day-4,morn	7.1678	26.6332	44.6636	13.5054	11.8129	5.6321	1.5998	0.8727	0.2434	0.056
Day-1,noon	3.6604	11.3684	22.7236	11.3496	6.8971	3.7988	4.2887	2.4781	1.1658	0.3216
Day-2,noon	6.9959	16.0249	52.7443	24.0598	16.9194	13.6476	6.4704	3.4271	0.6797	0.5267
Day-3,noon	6.6319	20.5226	89.6051	49.6318	19.1186	19.2811	11.2707	5.6166	1.1602	0.8522
Day-4,noon	9.5599	22.8287	27.4066	57.7348	16.8822	42.8739	25.8015	2.8321	0.9254	0.4836

	Energy vector of shouvik samples,when bands=20,equally spacing from 0 to 4KHz									
F in Hz	0-200	200-400	400-600	600-800	800-1000	1000-1200	1200-1400	1400-1600	1600-1800	1800-2000
Day-1,morn	10.6313	291.1282	449.5133	45.1612	9.9601	6.6522	5.9788	17.353	15.6831	4.8178
Day-2,morn	12.3828	305.115	292.8712	44.3163	21.0328	14.1231	6.3566	17.8706	15.0729	5.2094
Day-3,morn	17.4528	290.3781	188.749	29.3222	11.0161	9.3137	5.6596	17.2933	9.546	2.7612
Day-4,morn	14.472	285.139	209.2635	27.0161	10.6642	6.785	9.2662	15.1452	8.0952	4.1941
Day-1,noon	253.624	455.174	53.2881	62.2956	12.8912	64.0265	10.7036	54.754	51.8504	14.1679
Day-2,noon	48.3711	734.8122	444.1591	63.0046	11.862	50.4161	39.7765	5.7217	27.4931	7.3966
Day-3,noon	35.8302	570.113	413.6923	38.7245	14.8049	37.3522	21.1812	20.3879	37.2801	8.001
Day-4,noon	52.7809	591.2661	317.7748	29.0874	20.012	51.9947	13.9049	12.942	40.7442	6.8891

	Energy vector of prerna samples,when bands=20,equally spacing from 0 to 4KHz									
F in Hz	0-200	200-400	400-600	600-800	800-1000	1000-1200	1200-1400	1400-1600	1600-1800	1800-2000
Day-1,morn	0.0975	28.4991	55.8672	2.9747	9.7741	0.1076	0.1208	1.0254	1.5317	0.5705
Day-2,morn	0.0607	16.1137	70.7267	2.8386	7.6555	0.1084	0.0945	0.8136	1.286	0.6338
Day-3,morn	0.0903	23.7134	107.219	10.299	24.5094	0.6701	0.1339	0.2487	0.5651	0.2379
Day-4,morn	0.1515	41.0692	93.0055	7.0824	29.6992	1.1023	0.1341	0.4458	0.2919	0.1422
Day-1,noon	0.1036	9.4282	53.4167	1.6723	8.0442	0.0996	0.1195	1.0976	2.4594	1.1455
Day-2,noon	0.1515	17.1991	106.2442	2.4251	30.0288	0.4071	0.0942	1.3329	2.0012	0.9168
Day-3,noon	0.4578	26.2619	94.2214	3.2198	30.813	0.3535	0.1279	0.706	1.6945	0.7384
Day-4,noon	0.1774	24.8336	110.3889	5.8106	24.6914	0.9425	0.1745	0.2937	0.5919	0.5051

	Energy vector of nivedita samples,when bands=20,equally spacing from 0 to 4KHz									
F in Hz	0-200	200-400	400-600	600-800	800-1000	1000-1200	1200-1400	1400-1600	1600-1800	1800-2000
Day-1,morn	5.1709	56.3447	5.4527	6.6434	3.206	5.8989	2.852	3.0817	3.7333	3.3259
Day-2,morn	1.3093	68.2871	18.5878	2.0764	2.1054	1.0813	1.9358	1.6999	1.1621	0.089
Day-3,morn	2.6175	68.4512	18.9784	1.8397	2.9946	2.3473	1.4744	1.6901	1.5983	0.1971
Day-4,morn	5.9821	19.9364	8.5142	2.6285	0.7321	0.2468	0.3834	0.3582	0.4983	0.2868
Day-1,noon	1.6679	51.4374	39.2428	5.7764	8.3271	5.1282	1.675	2.2972	0.7633	0.3148
Day-2,noon	4.3792	34.1661	44.6958	5.3592	8.0076	4.8717	3.4614	4.0976	1.2321	0.4343
Day-3,noon	0.6528	36.3413	70.983	8.0592	14.5735	3.4038	4.8881	3.7412	0.6501	0.7172
Day-4,noon	0.3088	45.3763	74.9004	7.0689	23.665	7.4644	3.0907	1.56	2.0134	1.8133

- These energy vector shows that most of the energy of
  - rudra lies in 400-600 Hz band
  - Shouvik lies in 200-600 Hz band
  - Prerna lies in 400-600 Hz band
  - Nivedita lies in 200-600 Hz band

- The audio signal is passed through filters covering the speech spectrum equi spaced in 0 to 1200 Hz(BW=150Hz) and equispaced in 1200 to 4000 Hz(BW=700Hz)i.e. if 12 filters are used 0-150 Hz (LPF), 150-300 Hz(BPF), 300-450 Hz (BPF) ... etc.
- The below tables show the energy values in each band when 12 filter banks are taken.

	Energy vector of rudra samples,unequal spacing: 0 to 1200 Hz equal spacing with BW 150Hz & 1200 to 4000 Hz with BW 700 Hz											
F in Hz	0-150	150-300	300-450	450-600	600-750	750-900	900-1050	1050-1200	1200-1900	1900-2600	2600-3300	3300-4000
Day-1,morn	8.565	17.592	52.7851	74.7946	61.1375	13.2417	10.9767	4.2556	9.4932	4.1161	2.6535	0.0342
Day-2,morn	7.486	9.4902	52.3782	41.5243	37.0742	8.9289	7.8994	5.346	10.7881	2.1239	0.8189	0.0165
Day-3,morn	10.07	13.000	62.8307	53.5699	64.8617	24.4871	13.1957	6.4813	7.2057	4.7644	4.6257	0.0278
Day-4,morn	5.893	7.2367	32.2928	38.7204	12.1227	6.692	4.9843	3.795	2.9718	0.8711	0.721	0.0081
Day-1,noon	2.309	6.5167	14.404	16.9196	8.7464	3.9673	6.7292	3.1221	8.7086	2.9845	2.2754	0.3147
Day-2,noon	3.433	6.0989	24.1755	20.9964	16.0635	16.2011	10.5629	7.3805	11.4179	5.7983	5.7065	0.6319
Day-3,noon	1.953	6.0727	24.8648	41.154	37.2101	29.7303	10.9916	9.7882	20.907	12.6402	13.0354	0.776
Day-4,noon	7.131	11.078	25.4089	23.7614	36.9274	13.634	17.3083	39.0823	30.7199	7.1767	7.4807	0.5386

	Energy vector of shouvik samples,,unequal spacing: 0 to 1200 Hz equal spacing with BW 150Hz & 1200 to 4000 Hz with BW 700 Hz											
F in Hz	0-150	150 -300	300-450	450-600	600-750	750-900	900-1050	1050-1200	1200-1900	1900-2600	2600-3300	3300-4000
Day-1,morn	2.583	211.70	241.402	107.503	27.5079	25.7601	5.8617	4.842	46.0342	14.9128	4.9773	1.2414
Day-2,morn	1.607	160.39	279.244	117.0013	27.6193	25.1612	16.3603	10.0397	51.6361	11.6354	2.2393	0.4772
Day-3,morn	1.485	199.29	270.083	80.6494	19.8651	13.3181	9.3081	5.0363	37.9967	5.975	1.3022	1.2156
Day-4,morn	2.041	171.07	302.449	84.5837	21.1168	6.2629	8.2506	3.5599	38.8983	6.7258	1.7674	1.5574
Day-1,noon	68.09	339.00	200.794	24.8865	31.3286	20.6169	13.3197	51.6223	137.5149	29.0438	10.0426	1.4712
Day-2,noon	19.48	806.97	151.080	129.1456	58.4833	7.3387	13.6631	42.0613	84.154	16.4866	9.6504	1.7672
Day-3,noon	15.59	615.27	159.450	127.06	35.8608	8.0572	15.529	29.7885	90.1531	18.5141	12.4796	0.8136
Day-4,noon	17.15	683.63	178.783	95.4413	27.3285	7.5362	22.571	41.3009	79.0883	11.1694	7.0963	0.7059

	Energy vector of prerna samples,,unequal spacing: 0 to 1200 Hz equal spacing with BW 150Hz & 1200 to 4000 Hz with BW 700 Hz											
F in Hz	0-150	150-300	300-450	450-600	600-750	750-900	900-1050	1050-1200	1200-1900	1900-2600	2600-3300	3300-4000
Day-1,morn	0.048	26.216	3.8434	47.2271	2.4436	5.3731	2.2801	0.0629	3.5968	0.4242	0.223	0.2349
Day-2,morn	0.028	14.715	2.8617	54.392	2.6506	1.3579	4.2168	0.0958	4.0654	0.8416	0.29	0.152
Day-3,morn	0.03	20.752	13.1283	65.5746	9.2945	16.1514	3.0829	0.4987	1.1673	0.6987	1.5219	0.1256
Day-4,morn	0.098	35.019	9.3983	66.3376	5.948	16.8162	6.0727	0.656	1.2166	0.886	1.9177	0.2999
Day-1,noon	0.014	8.6983	4.0822	36.901	1.0275	5.9052	1.5953	0.0433	4.5897	1.2774	0.7924	0.2698
Day-2,noon	0.05	15.989	8.5386	83.1195	1.9607	13.2827	13.5245	0.2672	3.9087	1.5553	0.5332	0.1902
Day-3,noon	0.055	23.564	11.2327	61.3516	2.4035	12.9555	8.6458	0.1434	3.1108	1.3318	0.6835	0.2418
Day-4,noon	0.040	22.394	11.4875	93.4885	4.5118	11.983	6.9158	0.7485	1.608	1.0268	1.181	0.5229

	Energy vector of nivedita samples,,unequal spacing: 0 to 1200 Hz equal spacing with BW 150Hz & 1200 to 4000 Hz with BW 700 Hz											
F in Hz	0-150	150-300	300-450	450-600	600-750	750-900	900-1050	1050-1200	1200-1900	1900-2600	2600-3300	3300-4000
Day-1,morn	0.751	96.816	1.061	2.5931	5.0738	3.8342	2.1852	4.5122	13.6531	2.9866	0.5121	0.1157
Day-2,morn	0.106	78.571	0.4189	14.3087	1.8413	1.2098	0.9887	0.7772	4.9982	0.421	0.498	0.0578
Day-3,morn	0.117	88.207	0.577	13.2418	1.7082	1.7672	1.5159	1.8172	5.0358	1.0042	1.0286	0.0918
Day-4,morn	0.146	49.584	2.2943	5.5641	1.8447	2.0873	0.3219	0.2165	2.6532	0.3897	0.1978	0.0352
Day-1,noon	0.135	71.317	1.3178	36.9966	4.5012	3.3144	5.2318	3.3291	5.2721	1.0691	0.3824	0.0659
Day-2,noon	0.067	51.128	1.0782	40.9658	4.9608	3.815	4.4447	2.7457	9.4794	1.8165	0.5533	0.1176
Day-3,noon	0.296	45.486	0.7043	68.9619	5.7551	7.6169	4.0268	2.2732	10.6077	1.808	0.5019	0.3036
Day-4,noon	0.118	52.362	1.3583	71.1006	6.5647	4.941	10.8134	3.0978	11.1271	2.5339	0.8151	0.1146

- These energy vector shows that most of the energy of
  - rudra lies in 450-600 Hz band
  - Shouvik lies in 150-450 Hz band
  - Prerna lies in 150-300 and 450-600 Hz band
  - Nivedita lies in 150-300 450-600 Hz band

Result :-

- We had developed an algorithm using the mean-variance concept which allows only a particular speaker to be able to open the lock by uttering the combination lock number.
- For every speaker we have extracted the mean value,variance and standard deviation of energy of a particular band using the 8 sample signals.
- The concept behind this is we know that 99% of the distribution of any process lies in the range [mean-2.5*standard_deviation,mean+2.5*standard_deviation].
- So for a particular person we will design the lock system such a way:-
  - For example,we have given a random audio signal,then it checks for each band of frequency whether the energy values of that particular band lies in the range [mean-2.5*standard_deviation,mean+2.5*standard_deviation](of a particular band of a particular person),if it satisfies for each band then it will give access.
- This concept depends on the window that we have taken.
- The lowest false acceptance and lowest false rejection is achieved when we take banks=20 and rectangular window.
- When algorithm defined to give access only to rudra,when banks used=20 and rectangular window,Below table shows,to which samples this algorithm will give access(1).

	Day-1,Morn	Day-2,Morn	Day-3,Morn	Day-4,Morn	Day-1,noon	Day-2,noon	Day-3,noon	Day-4,noon
Rudra	1	1	1	1	1	1	1	1
Shouvik	0	0	0	0	0	0	0	0
Prerna	0	0	0	0	0	0	0	0
Nivedita	0	0	0	1	0	0	0	0

False acceptance ratio = 1/32

False rejection ratio = 0

- When algorithm defined to give access only to shouvik,when banks used=20 and rectangular window,Below table shows,to which samples this algorithm will give access(1).

	Day-1,Morn	Day-2,Morn	Day-3,Morn	Day-4,Morn	Day-1,noon	Day-2,noon	Day-3,noon	Day-4,noon
Rudra	0	0	0	0	0	0	0	0
Shouvik	1	1	1	1	1	1	1	1
Prerna	0	0	0	0	0	0	0	0
Nivedita	0	0	0	0	0	0	0	0

False acceptance ratio = 0

False rejection ratio = 0

- When algorithm defined to give access only to prerna,when banks used=20 and rectangular window,Below table shows,to which samples this algorithm will give access(1).

	Day-1,Morn	Day-2,Morn	Day-3,Morn	Day-4,Morn	Day-1,noon	Day-2,noon	Day-3,noon	Day-4,noon
Rudra	0	0	0	0	0	0	0	0
Shouvik	0	0	0	0	0	0	0	0
Prerna	1	1	1	1	1	1	1	1
Nivedita	0	0	0	0	0	0	0	0

False acceptance ratio = 0

False rejection ratio = 0

- When algorithm defined to give access only to nivedita,when banks used=20 and rectangular window,Below table shows,to which samples this algorithm will give access(1).

	Day-1,Morn	Day-2,Morn	Day-3,Morn	Day-4,Morn	Day-1,noon	Day-2,noon	Day-3,noon	Day-4,noon
Rudra	0	0	0	0	0	0	0	0
Shouvik	0	0	0	0	0	0	0	0
Prerna	0	0	0	0	0	0	0	0
Nivedita	1	1	1	1	1	1	1	1

False acceptance ratio = 0

False rejection ratio = 0

- When algorithm defined to give access only to rudra,when banks used=12(unequal spacing as mentioned previously) and Blackman window,Below table shows,to which samples this algorithm will give access(1).

	Day-1,Morn	Day-2,Morn	Day-3,Morn	Day-4,Morn	Day-1,noon	Day-2,noon	Day-3,noon	Day-4,noon
Rudra	1	1	1	1	1	1	1	1
Shouvik	0	0	0	0	0	0	0	0
Prerna	0	1	1	0	1	0	0	0
Nivedita	0	0	0	0	0	0	0	0

False acceptance ratio = 3/32

False rejection ratio = 0

- When algorithm defined to give access only to shouvik,when banks used=12(unequal spacing as mentioned previously) and Blackman window,Below table shows,to which samples this algorithm will give access(1).

	Day-1,Morn	Day-2,Morn	Day-3,Morn	Day-4,Morn	Day-1,noon	Day-2,noon	Day-3,noon	Day-4,noon
Rudra	0	0	0	0	0	0	0	0
Shouvik	1	1	1	1	1	1	1	1
Prerna	0	0	0	0	0	0	0	0
Nivedita	0	0	0	0	0	0	0	0

False acceptance ratio = 0

False rejection ratio = 0

- When algorithm defined to give access only to prerna,when banks used=12(unequal spacing as mentioned previously) and Blackman window,Below table shows,to which samples this algorithm will give access(1).

	Day-1,Morn	Day-2,Morn	Day-3,Morn	Day-4,Morn	Day-1,noon	Day-2,noon	Day-3,noon	Day-4,noon
Rudra	0	0	0	0	0	0	0	0
Shouvik	0	0	0	0	0	0	0	0
Prerna	1	1	1	1	1	1	1	1
Nivedita	0	0	0	0	0	0	0	0

False acceptance ratio = 0

False rejection ratio = 0

- When algorithm defined to give access only to nivedita,when banks used=12(unequal spacing as mentioned previously) and Blackman window,Below table shows,to which samples this algorithm will give access(1).

	Day-1,Morn	Day-2,Morn	Day-3,Morn	Day-4,Morn	Day-1,noon	Day-2,noon	Day-3,noon	Day-4,noon
Rudra	0	0	0	0	0	0	0	0
Shouvik	0	0	0	0	0	0	0	0
Prerna	0	0	0	0	0	0	0	0
Nivedita	1	1	1	1	1	1	1	1

False acceptance ratio = 0

False rejection ratio = 0

PART-B :- pitch

Pitch, in speech, is the relative highness or lowness of a tone as perceived by the ear, which depends on the number of vibrations per second produced by the vocal cords. Pitch is the main acoustic correlate of tone and intonation.

Pitch is sometimes termed as the fundamental frequency of the speech signal .It represents the vibration frequency of the vocal cords during the sound productions.but the above statement is not true, signals having the same frequency of vibration from different sound sources may elicit the perception of different pitches; and on the other hand, signals having different frequencies of vibration from different sound sources may elicit the perception of the same pitch.

Speech is classified into voice, unvoiced and silence. The voice speech is the periodic vibration of vocal folds. Background noise affects the speech signals. In many speech applications calculation of pitch plays a major role. The paper proposes a pitch detection algorithm based on the short-time average magnitude difference function (AMDF) and the short-term autocorrelation function (ACF). Detecting the Pitch within the speech signal is important in most of all the speech related applications. Detection of Pitch is useful in identification of speakers. One solution to get detected with the pitch is by using the time domain algorithms.

Human's primary mode of communication is speech. Through the vocal cords the air is passed which also passes through the vocal tract and the production of speech takes place. Pitch period (T0) is referred to as the time duration of one glottal cycle and the reciprocal is referred to as the fundamental frequency (F0). It has wide applications areas like digital cellular communication, conferencing of video and coding of speech signals etc. Estimation of speech is also important in many applications like synthesis of speech signals, speech recognition, identification of speakers, coding of speech and its verification. In many speech processing algorithms detection of pitch is very important. Recognition of the speech system of the tongue uses pitch chases for tone recognition. Pitch is additionally crucial for speech variation in language systems and text-to-speech systems. The main cue of pitch is its fundamental frequency (F0). However it is tough to make a mathematical model which helps to calculate the pitch of a signal. Therefore it is much more important to design a pitch detection algorithm (PDA) in many of the speech processing applications systems.

The ear of a human responds to a frequency of 20 Hz-20 KHz. A pulse is a summation of a sinusoidal wave of fundamental frequency (F0) and its harmonics. Pitch is important in identification of the characteristics of voice within the human being, the information of the speaker, like the classification of gender that is whether male, female and age whether adults or children .

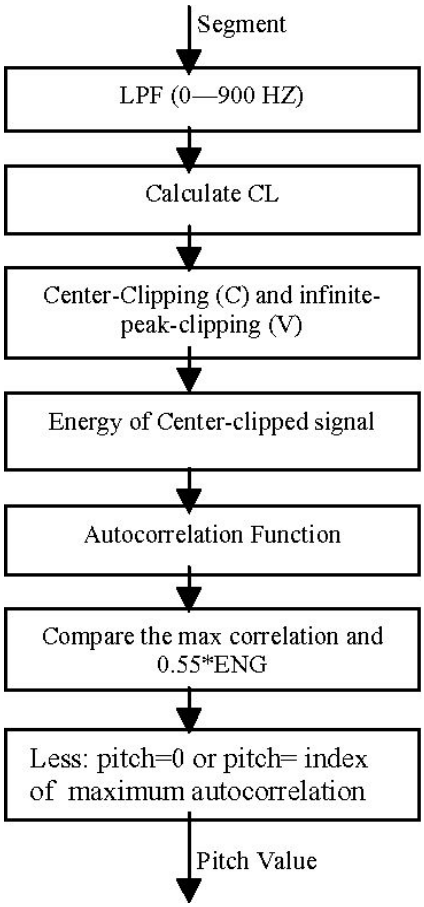
Pitch extraction algorithm:

Speech signal always varies with time that is it always varies with respect to time. To get the exact feature of speech we need to identify the smaller portion of speech signal. The signal is recorded and converted to a wav file so as to process the signal in the Matlab. Time domain pitch analysis includes the Autocorrelation method and AMDF techniques. There are many different ways for detection of pitch signals. Out of this the Autocorrelation is the most reliable pitch detector technique. The requirement for higher processing is needed for this approach. When a signal is transmitted over a telephone line the autocorrelation is a very good method for getting the expected results. It does displays prominent peaks.

Autocorrelation Method:

The Cross-correlation of signal with itself is defined as autocorrelation. The repeated patterns are fined using the autocorrelation function. The pitch periods are estimated for a given signal. The short term analysis technique is used for pitch detection.

fig - Flowchart of Pitch Detection Algorithm using Autocorrelation Method



The detail of the algorithm is depicted in the above flow chart And a summary of the algorithm is given below.

1.The signal is considered and it is filtered at a 900Hz and applied with a sampling frequency of 10 KHz.

2. 10ms intervals are selected from a length of segments of 30msec. Thus the segments overlap by 20msec.

3.The Autocorrelation, Equation (shown below), samples using a hamming window. Overlapping is allowed within the frames. The peak signal level for 30msec determines background noise. If the signal goes above the threshold value then it is termed as a voice otherwise it is termed as silence and no further action is taken.

4. Fixed % is determined by the clipping level.

5. Using clipping level, processing of speech signals takes place.

6.The largest value from the signal is computed and compared to a fixed threshold. If the peak value is below threshold it is termed as an unvoiced signal and when it is above the threshold then it is defined as the largest peak location.

For the appropriate normalization process the autocorrelation function at zero delay is computed. It always searches for maximum peaks. The values are always computed and compared with the threshold. Within the pitch fundamental frequency is computed by,

$$Rx(m) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x(n)x(n+m)$$

Where m is the sample points time lag. Autocorrelation algorithm is relatively impervious to noise.

The time domain method is used for calculation of pitch in which the values are compared to the standard format which is used to classify the gender of a person.

Yes we can use the above mentioned algorithm to extract the pitch of the signal.

MATLAB codes and speech files attested in zip file