International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-9S3, July 2019

Analysis of Spectral Features for Speaker Clustering

Badhe Sanjay S., Gulhane S. R., Shirbahadurkar S. D.

Abstract—In this paper Spectral feature like Spectral Roll off, Spectral Centroid, RMS (Root Mean Square) energy, Zero crossing Rate, Spectral irregularity, Brightness, of speech audio signals are extracted and analyzed. From analysis, prominent features are selected. These prominent features are used for speaker identification. For performing feature analysis, database of seven speakers is created. By using features, speakers are divided into two groups or clusters.

Keywords— Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity & Brightness.

I. INTRODUCTION

In clustering set of objects are grouped in such a way that objects in the same group are more similar to each other than to those in other groups (clusters). It is a main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, bioinformatics, data compression, and computer graphics [1][2].

We can create different groups or clusters by using various features of audio signal. The process of obtaining features of audio signal is called feature extraction. Various features are developed for the analysis of audio sound signals in the last few decades. These features are Temporal features (also called as time domain features), Spectral features (also called as Frequency domain features) and Cepstral features.

We can use Spectral features (Timbral features) like Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity, Brightness for clustering of audio signals. Timber is nothing but the quality of sound or tone. Features related to timbre are called as spectral features or Timbral features. Zero Crossing Rate: Noisiness of sound is represented by Zero Crossing Rate. If ZCR is more means more noise.

Revised Manuscript Received on July 25, 2019.

Mr. Badhe Sanjay S. , Research Scholar (DYPIT), DYPCOE (SPPU), Pune,India, sanjubadhe@gmail.com

Mr. Gulhane S. R., Research Scholar (DYPIT) , DYPCOE (SPPU), Pune,India, sushenrgulhane1@rediffmail.com.com

Dr. Shirbahadurkar S. D., Research Guide (DYPIT), Zeel COE (SPPU) , Pune,India, shirsd112@yahoo.in

Centroid: Centroid represented the location of gravity i.e. magnitude spectrum. Rolloff: Roll-off is a measure of way to measure the amount of high frequencies in the sound signal. Brightness is similar to roll-off. The cut-off frequency is fixed first & brightness is calculated by measuring amount of energy above cut-off frequency values between 0 and 1. Irregularity is the measure of degree of variation of the sequential peaks of the spectrum. RMS Energy is the measure of the maximum of absolute amplitude in each frame of a signal.

II. THEORY

Spectral features (Timbral features) like Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity, Brightness of audio signals are extracted. These features are used for clustering.

Audio or speech signals of seven speakers are recorded for database creation. Same speech file is recorded fifteen times per speaker. Each file contains 69 words. Hence 7245 samples analyzed.

A. Block Diagram:

Following figure shows the block diagram of feature extraction.

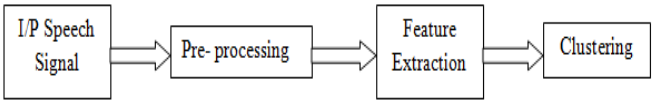


Fig.1 Block diagram of feature extraction

Recorded the audio signal are used as i/p signals. In pre-processing stage, input signals are processed. After preprocessing, various features are extracted. Features like Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity, Brightness are compared for clustering or grouping of speakers.

III. EXPERIMENT RESULTS

For experimentation database is created for seven speakers. Fifteen samples per speaker are recorded. Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity & Brightness of speech audio signals are extracted for each sample. Mean of Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity and Brightness is calculated for each speaker. Mean of Spectral Roll off, Spectral Centroid, RMS

Retrieval Number: I30270789S319/2019©BEIESP
DOI: 10.35940/ijitee.I3027.0789S319

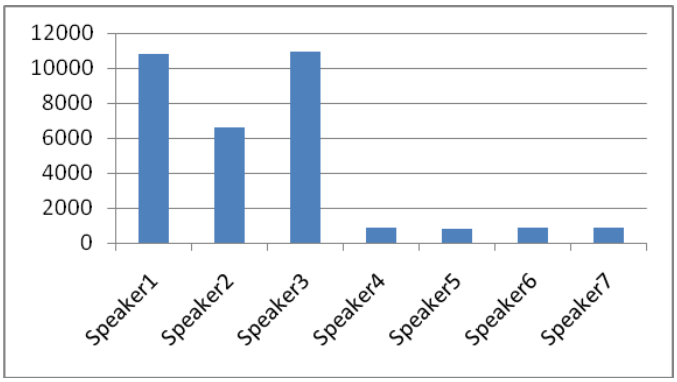
Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

energy, Zero crossing Rate, Spectral irregularity and Brightness can be classified into two groups or clusters. Similarly, mean of each cluster can be calculated and mean of two clusters can be compared. Mean values are given in table number 1.

TABLE I. COMPARISON OF MEAN OF FEATURES

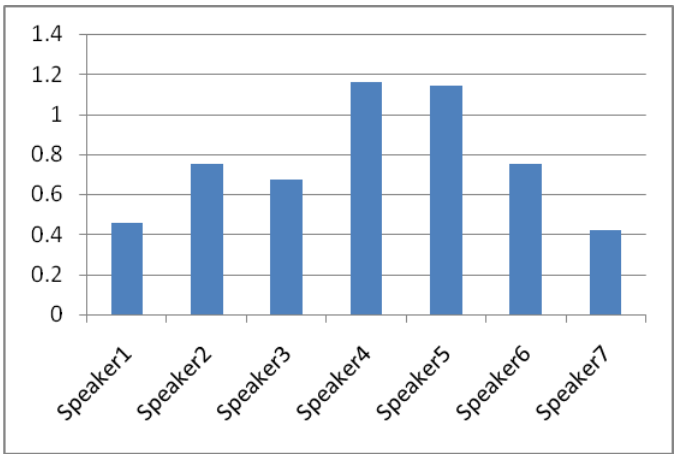
Comparison of Mean of features of All Speakers						
Sr. No.	***Spea-ker***	***Roll off***	***Spectral Irregul-arity***	***Bright-ness***	***RMS energy***	***Spectral Centroid***
1	Sp_1	10797.1	0.45479	0.59039	0.92595	5241
2	Sp_2	6580.70	0.75370	0.48021	0.96764	3911
3	Sp_3	10963.8	0.67326	0.70961	0.61785	6533
4	Sp_4	867.819	1.16041	0.08608	0.03460	886
5	Sp_5	805.561	1.14205	0.08317	0.03916	673
6	Sp_6	878.936	0.75016	0.13473	0.04931	673
7	Sp_7	874.326	0.42069	0.06243	0.03820	633

A. Roll off of speakers is compared from mean values, as follows:



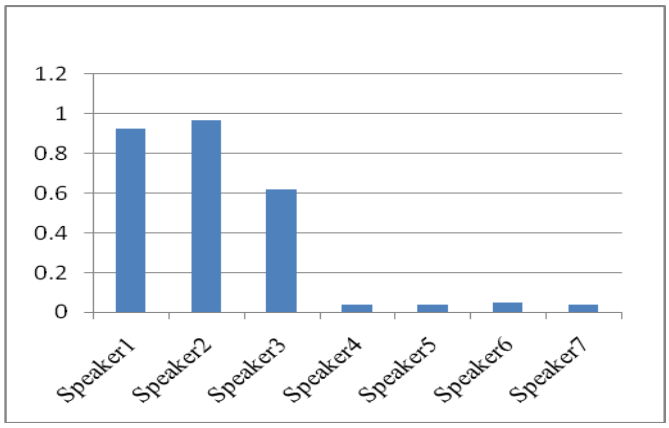
Graph. 1 Rolloff Comparison of All Speakers

B. Spectral Irregularity of speakers is compared from mean values, as follows:



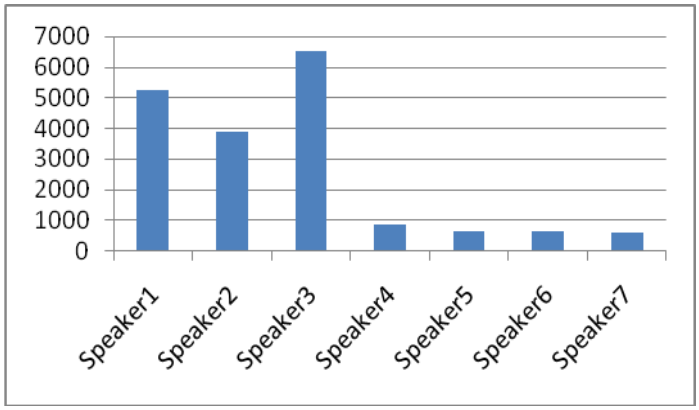
Graph. 2 Spectral irregularity comparison of all speakers

C. Spectral Irregularity of speakers is compared from mean values, as follows:



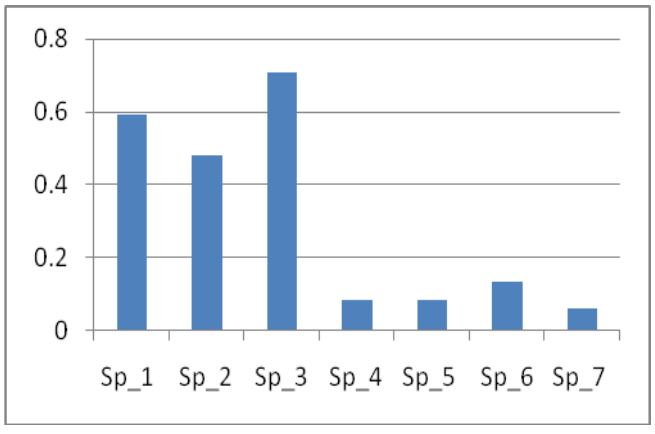
Graph. 3 RMS energy comparison of all speakers

D. Spectral Centroid of speakers is compared from mean values, as follows:



Graph. 4 Spectral Centroid comparison of all speakers

E. Brightness of speakers is compared from mean values, as follows:



Graph. 5 Brightness comparison of all speakers

Retrieval Number: I30270789S319/2019©BEIESP
DOI: 10.35940/ijitee.I3027.0789S319

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

TABLE II. COMPARISON OF MEAN FEATURES OF TWO CLUSTERS

Sr. No.	Cluster	Rolloff	Spectral irregular-rity	Brightn-ess	RMS energy	Spectral Centroid
1	Cl_1	9447.23	0.627248	0.593404	0.83715	5228.23
2	Cl_ 2	856.66	0.868328	0.091603	0.04051	716.281

Graph. 6 Mean of Roll off comparison of two clusters

Graph. 7 Mean of Spectral irregularity comparison of two clusters

Graph. 8 Mean of Brightness comparison of two clusters

Graph. 9 Mean of RMS energy comparison of two clusters

Graph. 10 Mean of Spectral Centroid comparison of two clusters

IV. ANALYSIS:

Mean values of Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity and Brightness are compared in Table number 1 and graphs are plotted. Spectral Roll off mean values of all speakers is compared in graph number 1. Spectral irregularity is compared for mean values of all speakers in graph number 2. RMS energy is compared for mean values of all speakers in graph number 3. Spectral Centroid is compared for mean values of all speakers in graph number 4.

Table 2 gives mean features comparison of two clusters. From table 2, if mean of roll off is calculated for each cluster, we can easily differentiate one cluster from other cluster. Similarly, mean of Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity and Brightness of two clusters can be compared. By using mean of Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity and Brightness, two clusters are differentiated.

V. CONCLUSION

From above analysis it is concluded that Timbral features like Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral

Retrieval Number: I30270789S319/2019©BEIESP
DOI: 10.35940/ijitee.I3027.0789S319
Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

irregularity, Brightness can be used for dividing speakers into two clusters.

VI. FUTURE SCOPE

Timbral features like Spectral Roll off, Spectral Centroid, RMS energy, Zero crossing Rate, Spectral irregularity, Brightness can be used for dividing speakers into many clusters.

VII. ACKNOWLEDGMENT

We are thankful to our guide, who provided expertise that greatly, assisted the research. This research was supported by D. Y. Patil Institute of Technology-Pimpri and D. Y. Patil College of Engineering-Ambi. We are also grateful to Dr. S. N. Mali and Dr. A. V. Kulkarni for guidance.

REFERENCES

[1] Bailey, Ken (1994). "Numerical Taxonomy and Cluster Analysis". Typologies and Taxonomies. p. 34. ISBN 9780803952591.

[2] Tryon, Robert C. (1939). Cluster Analysis: Correlation Profile and Orthometric (factor) Analysis for the Isolation of Unities in Mind and Personality. Edwards Brothers.

[3] A De Cheveigné and H. Kawahara, "Yin, a fundamental frequency estimator for speech and music," J. Acoust. Soc. Amer., vol. 111, p. 1917, 2002.

[4] B. Gold and L. R. Rabiner, "Parallel processing techniques for estimating pitch periods of speech in the time domain," J. Acoust. SOC. Amer., vol. 46, pp. 442-448, Aug. 1969.

[5] F. Itakura, "Minimum prediction residual principle applied to speech recognition," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-23, pp. 67-72, Peb. 1975

[6] B. S. Atal and L. R. Rabiner, "A pattern recognition approach to voiced-unvoiced-silence classification with applications to speech recognition," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-24, pp. 201-212, June 1976

[7] N. C. Bui, J. J. Monbaron, and J. G. Michel, "An integrated voice recognition system," IEEE Trans., Acoust., Speech, Signal Processing, vol. ASSP-31, pp. 323-328, Feb. 1983.

[8] H. L. Kwok, T. C. Tai, and Y. M. Fung, "Machine recognition of the Cantonese digits using bandpass fiters,"ZEEE Trans, Acoust., Speech, Signal Processing, vol. ASSP-31, pp. 220-222, Feb. 1983.

[9] YIU-KEI LAU AND CHOK-KI CHAN," Speech Recognition Based on Zero Crossing Rate and Energy", IEEE Transactions On Acoustics, Speech, And Signal Processing, Vol. Assp-33, No. 1, February 1985

[10] C. J. Heinselman and J. F. Vickrey," On the spectral analysis and interpretation of incoherent scatter plasma line echoe", Radio Science, Volume 27, Number 2, Pages 221-230, March-April 1992.

[11] S. Demircan and H. Kahramanlı, "Feature Extraction from speech data for emotion recognition," Journal of Advances in Computer Networks, vol. 2, No. 1, pp. 28–30, Mar. 2014

[12] B.H.Juanag and T. h. Chen,"The past, present and future of speech processing", IEEE signal processing magazine, vol.15,no.3,pp. 24-48, May-1998.

[13] Alexandr Arkhipov, Glazunov Nikolaj and Khyzhnyak Anna, "Remote sensing, spectral brightness and heuristic criterion for class recognition", 2017 IEEE Microwaves, Radar and Remote Sensing Symposium (MRRS).

[14] Dalibor Mitrovi´C, Matthias Zeppelzauer, and Christian Breiteneder, "Features For Content-Based Audio Retrieval", Advances in Computers Vol. 78, pp. 71-150, 2010.

[15] Guodong Guo and Stan Z. Li, "Content-Based Audio Classification and Retrieval by Support Vector Machines", IEEE Transactions on Neural Networks, Vol. 14, No. 1, January 2003. B c c dc

[16] Dongge Li, Ishwar K.Sethi, "Classification of General Audio Data for Content Based Retrieval", Pattern Recognition Letter, Elsevier 2001.

[17] Hyoung-Gook Kim, Nicolas Moreau, and Thomas Sikora, "Audio Classification Based on MPEG-7 Spectral Basis Representations", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 5, May 2004.

[18] R.Christopher Praveen Kumar, D.Abraham Chandy," Audio Retrieval using Timbral Feature", 2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN 2013)

[19] R. O. Duda, P. E. Hart, and David G. Stork, Pattern Classification (2nd Edition), 2nd ed. 2000.

[20] L. R. Rabiner and R. W. Schafer, "Introduction to Digital Speech Processing," Foundations and Trends in Signal Processing, vol. 1, no. 1–2, pp. 1–194, 2007.

Retrieval Number: I30270789S319/2019©BEIESP
DOI: 10.35940/ijitee.I3027.0789S319

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication