**INTRODUCTION:**

Sound is the vibration that travels through air or any medium and these vibrations are audible when they reach an individual’s ear and sound is formed by the unbroken and consistent vibrations, The first ever sound that was noted was assembled by a gadget called a phonautograph, invented by Édouard-Léon Scott de Martinville in 1857. Phonautograph write out sound waves into a line that is drawn on paper but with these waves there are some features through which sound can be categorized in many classes or categories were extracted, let’s take an example when we hear any kind of sound our brain start processing on it and categorize that sound like we can predict that this is the voice of a female without seeing that female because we know which value range belongs to which category, but the major challenge is to extract the features and their different ways of doing it such as MFCC, RASTA, LPCC, Cepstral Analysis, LPC and many others [1] The majority of these proposed frameworks consolidate two handling stages. The ﬁrst stage studies the received sound wave and extracts parameters (features) from it. The feature extraction process ordinarily includes a vast data reduction. Second stage does classiﬁcation based on the extracted features and both of these stages are defined briefly blow.

[2, 3, 4] Many set of feature extraction are proposed earlier for audio classification. Largest portion has been covered by low-level signal features and then second important feature set consist of Mel-frequency cesptral coefficient (MFCC) [5] and then those remaining features come.

All of the features are used audio classification and are very powerful in classifying the audio class but it gradually decrease when amount of classes increase. So, using which feature set with which amount of classes is an issue which can create further more issues if we select wrong feature set with respect to the problem description, result will help you with comparison done which will guide you when to use which set. [6, 7]

REFRENCES:

[1]<https://www.researchgate.net/publication/250008991_Features_for_Audio_Classification?fbclid=IwAR0jhPNvxN0jAwuewqzGFKgK1HjqqbF3YdMFFgPzLkDz-Mo9Xd8OH05Z19c>

[2] [S. B. Davis and P. Mermelstein. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Transactions on acoustics, speech and signal processing, ASSP-28:357– 366, 1980](S.B.%20Davis%20and%20P.%20Meermelstein).

[3] [E. Wold, T. Blum, D. Keislar, and J. Wheaton. Content-based classification, search, and retrieval of audio. IEEE Multimedia, Fall:27–36, 1996. [3] M. S. Spina and V. W. Zue. Automatic transcription of general audio data: Preliminary analysis. In Proc. 4th Int. Conf. on spoken language processing, Philadelphia, PA, 1997.](E.Wold%20T.Blum%20..)

[4] E. Scheirer and M. Slaney. Construction and evaluation of a robust multifeature speech/music discriminator. In Proc. ICASSP, pages 1331–1334, Munich, Germany, 1997.

[5] H. Hermansky and N. Malayath. Spectral basis functions from discriminant analysis. In International Conference on Spoken Language Processing, 1998

[6] M. Zhang, K. Tan, and M. H. Er. Three-dimensional sound synthesis based on head-related transfer functions. J. Audio. Eng. Soc., 146:836–844, 1998.[7] T. Zhang and C. C. J. Kuo. Audio content analysis for online audiovisual data segmentation and classification.

IEEE Transactions on speech and audio processing, 2001.