International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Age Detection of Singer using KNN Algorithm

Sumeet S. Andhalkar

S.P. Pune University, G.H.R.C.E.M, Pune, Maharashtra, India

Abstract: ***In this paper we are going to study implementation age detection of singer through KNN algorithm. We can study that, it is possible to implement this age detection application through KNN algorithm. KNN is a type of instance-based learning or lazy learning, where the role is only approximated locally and all computation is overdue until classification. The KNN algorithm is one of the simplest of all machine learning algorithms. We can use timbre values for training and testing and can use KNN algorithm. The system can be trained so it will give correct output based on training parameters and testing parameters. The training testing methods we can use to assess the age. The algorithm KNN and other timbre values and voice features are also important key factors.***

Keywords: Timbre, Pitch, Vocals, Sinusoidal, Vibratos, Tremolos.

1. Introduction

The paper elaborates the age detection method that may be one of the possible methods in voice processing. Pitch, Timbre, Frequency, RMI energy are the basic features of voice. These features have some different values for samples of different age group people. By analyzing these different group people samples training of samples is done. On the basis of training samples and applied algorithm we can assess the age group of the person. Now Timbre is very important characteristic of voice and by detection it we can develop other special research mechanism in voice processing field. We also study the advantages of Timbre and its limitations. We can also study the future application that we can develop by using Timbre. Timbre detection also play important role in Musical information retrieval, Identification of the singer and in lyric recognition. Timbre can identify gender of singing voice. It means it can identify the gender of the singer. Timbre also can examine or find the time of voice recording or the time slot of voice recording. So it is very much necessary to study Timbre and its future work. To assess age we require some values of features for this purpose MFCC (Mel Frequency Cepstrum Coefficient) algorithm is used. The Mel frequency Cepstrum Coefficient (MFCC) feature has been used for designing a text dependent speaker identification system. The extracted speech features (MFCC's) of a speaker are quantized to a number of centroids using vector quantization algorithm. These centroids form the information table of that speaker. MFCC's are calculated in training phase and again in testing phase. Then KNN algorithm assess the age in groups.

2. Need of the Work

As we know the era is information technology era. The information can be processed to get different outputs, we can process the input samples to the system to assess age or different other application development.

The research in this area is required to develop many applications related to the work. Many applications can be created for the enrichment of mankind. The applications like decease identification gender identification or for some security purposes we can use these applications.

3. Concepts

3.1 Pitch

The Voice Processing is the study of voice signals and the processing approaches of these signals. The signals are usually treated in a digital depiction, so voice processing can be viewed as a extraordinary case of digital signal processing.

Pitch can be defined as the extent to which sound is high or low. It is the level of sound which may be high or low. It is also known as particular level of intensity of sound. Pitch detection is known as determining the level of intensity of voice. Pitch detection is determining the extent of voice to which it is high or low. Pitch detection is very important in some related tasks of voice processing. Pitch detection is crucial task in singing voice separation also. Pitch detection also play important role in Musical information retrieval, Identification of the singer and in lyric recognition. Pitch can identify gender of singing voice. It means it can identify the gender of the singer. Pitch also can examine or find the time of voice recording or the time slot of voice recording.

3.2 Timbre

The Timbre can be a set of subjective opinions of individuals Toward a sound that is independent from the frequency or the amplitude. Timbre is also known as color or tone of sounds. In music, timbre also known as tone color or tone quality from psychoacoustics, is the worth of a musical note, sound, or tone that distinguishes dissimilar types of sound production, such as voices and musical instruments, string instruments, wind instruments, and percussion instruments. The substantial characteristics of sound that find out the insight of timbre include spectrum and envelope. Timbre is very special character of voice. It identifies an object by its sound or voice. Timbre is special thing that separated or distinguish two or many sounds from each other. This timbre can be computed by using pitch factor.

4. MFCC

Many algorithms are developed by the researchers for feature extraction. In this work, the Mel Frequency Cepstrum Coefficient (MFCC) feature has been used for designing a text dependent speaker identification system. Some modifications to the existing technique of MFCC for feature extraction are also suggested to improve the speaker recognition efficiency. Mel-frequency cepstral coefficients (MFCCs) are coefficients that jointly make up an MFC. They are derived from a type of cepstral depiction of the audio clip (a nonlinear "spectrum-of-a-spectrum"). The difference between the cepstrum and the mel-frequency cepstrum is that in the MFC, the frequency bands are similarly spaced on the mel scale, which approximates the human auditory system's reply more closely than the linearly-spaced frequency bands used in the normal cepstrum. This frequency warping be able to allow for better illustration of sound.
MFCCs are commonly derived as follows

1) Take the Fourier transform of a signal.
2) Map the powers of the spectrum obtained above onto the mel scale, using triangular overlapping windows.
3) catch the logs of the powers at each of the mel frequencies.
4) Take the discrete cosine transform of the catalog of mel log powers.
5) The MFCCs are the amplitudes of the resulting spectrum.

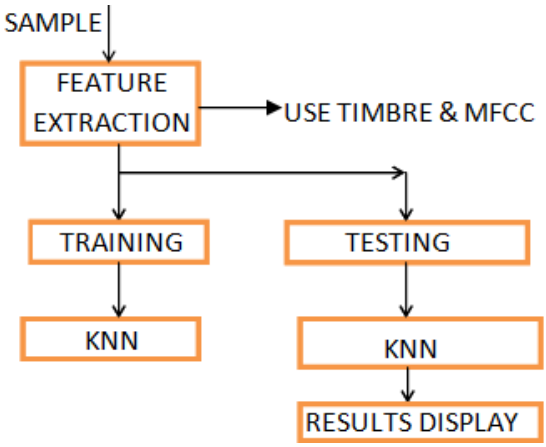
5. KNN Algorithm

k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. The *k*-NN algorithm is among the simplest of all machine learning algorithms. Both for classification and regression, it can be useful to assign weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones.

The neighbors are taken from a set of objects for which the class (for *k*-NN classification) or the object property value (for *k*-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. A shortcoming of the *k*-NN algorithm is that it is sensitive to the local structure of the data. The algorithm has nothing to do with and is not to be confused with *k*-means, another popular machine learning technique. The best choice of *k* depends upon the data; generally, larger values of *k* reduce the outcome of noise on the classification, but make boundaries between classes less distinct. A good *k* can be selected by various heuristic techniques (see hyper parameter optimization). The extraordinary case where the class is predicted to be the class of the nearby training sample (i.e. when $k = 1$) is called the nearest neighbor algorithm.
The accuracy of the *k*-NN algorithm can be degraded by the presence of noisy or irrelevant features, or if the feature scales are not consistent with their significance. Much research attempt has been put into selecting or scaling features to get better classification. A particularly popular approach is the use of evolutionary algorithms to optimize feature scaling. Another popular approach is to scale features by the mutual information of the training data with the training classes.

In binary (two class) classification problems, it is helpful to choose *k* to be an odd number as this avoids tied votes. One popular way of choosing the empirically optimal *k* in this setting is via bootstrap method.

6. Model



In this model we can see, the samples are given input to the system. These samples are then processed in training and testing by MFCC classifier and timbre is the basic feature of classification. Then as the training results are then stored and testing results are then made by comparison. We first do the training at the beginning for some samples. Then future results are based on those samples. After testing phase we can get all the results accurately.

7. Merits of the System

- **Separation of voice and music**- If we have the Timbre then we can separate the human voice and musical parts. We have to decide Timbre range for each alphabet or for each word. Then by analyzing various Timbre ranges from the musical record we can decide the words and separate them out to the paper.
- **Gender Identification** - As we have the Timbre, we can easily detect the gender of the singer as male and female voices have different pitch range. As male and female Timbre/ Pitch ranges or levels are different as differentiable, then we can definitely identify the gender of Singer.
- **Lyric Separation/Extraction**- We can separate out the words of the song sung by the singer. Pitch/Timbre is an vital feature here. Lyric is the words spoken by singer or words of the song. These words have different pitch/ Timbre ranges. So by detecting pitch/ Timbre, we can easily spot the of the song. So we con separate out words or lyric from the record.
- **Information Retrieval**- We can retrieve large amount of information from these musical records, if we have the Pitch/Timbre. Various instruments have various Pitch/Timbre ranges. So they can give more relative musical information. Large amount of characteristics can be retrieved from music or song. Pitch/Timbre and other

characters of voice can be retrieved from the voice easily by using Pitch/Timbre.

- **Instrument Detection:** Various instruments have various Pitch/Timbre ranges. Vocal partials are known as vibratos whereas Musical partials are known as Tremolos. There are two types of musical instruments, first one is harmonic instrument and second one is percussive. Harmonic instruments are mostly key pressing or tense wired instruments those includes flute, guitar. Percussive instruments includes tabla, drum etc. So, these instruments have different Pitch/Timbre range, by estimating Pitch/Timbre we can identify the instrument easily.
- **Identification of deceases-** We can identify deceases related to the vocal system like Lungs, throat, other organs included in production of voice. We also can determine the problems related breathing organs and some vocal organs. This research will give vital benefit to medical science research and will work for benefit of mankind.

8. Future Applications Based on Model

- **Confirmation of voice:** Many times controversies related to recorded voice, happening in real time world. These controversies can be resolved by using voice identification system. This system will compare person's voice with recorded voice and check for similarities.
- **Lyric Identification System:** This system identifies words from the song or speech or voice. The Lyric are separated from the record by using pitch/Timbre range. As each word has its own range.
- **Gender Identification System:** By using pitch/Timbre, we can identify the gender of singer or person. Male and female voices have different pitch/Timbre ranges. So by comparing these ranges we can determine the gender from voice of the person.
- **Decease detection system**: Various deceases related to the lung and breathing system can be identified by using voice processing techniques. As lungs and entire breathing system works to produce voice, by detecting and analyzing various aspects of voice we can detect deceases related to breathing system of human body.
- **Security Applications-** We can create various devices for authentication of person, similar to biometric devices. We can identify suspected enemy by his voice. We can use these records as evidence in judicial matters. These are definitely the future scopes that can benefit mankind.

9. Conclusion

This paper studies, how anybody can assess the age of singer through voice processing method. We can get age group values from the model. This paper studies use of KNN and MFCC and timbre combination for assessing the age of singer. This paper also studies advantages, applications and future scope of the model. We mainly studied that there is a method that can detect age of singer from singing voice and we can obtain its range by the respective method. The paper also studies KNN, MFCC and some basic concepts of voice processing.

10. Acknowledgement

The Singing voice pitch detection and voice separation and supporting documents are the work of many people, including Sang Hyun Park, Chao Ling Hsu, DeLiang Wang, Jyh-Shig Roger Jang. I also thank teachers and authors of different papers which are referred for study. I also thank my project guide and my teachers who always guided me. I also thank my classmates and friends who helped me in my study. Last but very important I thank my parents who always supported me in my good and bad times.

References

[1] Musical Instrument Extraction through Timbre Classification, Sang Hyun Park, NVIDIA Corporation, Santa Clara, Email: andyp@nvidia.com
[2] MFCC and its applications in speaker recognition Vibha Tiwari, Deptt. of Electronics Engg., Gyan Ganga Institute of Technology and Management, Bhopal, (MP) INDIA.
[3] Chao Ling Hsu, Deling Wang,Jhy Shing"A Trend Estimation Algorithm For Singing Pitch Detection In Musical Recordings".
[4] Chao Ling Hsu, Deling Wang"A Tandem Algorithm For Singing Pitch Extraction and Voice Separation From Musical Accompaniment".
[5] L. Regnier, G. Peeters "singing voice detection in music tracks using direct voice vibrato detection".
[6] Yipeng Li and DeLiang Wang "detecting pitch of singing voice in polyphonic audio".
[7] L. Regnier, G. Peeters, "singing voice detection in music tracks using direct voice vibrato detection".

Author Profile

Andhalkar Sumeet Sudhakar, studying Masters of Engineering (Computer Engineering) in G.H.R.C.E.M., Wagholi, Pune. He has completed Bachelor of Engineering in Information Technology from B.V.C.O.E. Kolhapur.