

PITCH DETECTION FROM SINGING VOICE, ADVANTAGES, LIMITATIONS AND APPLICATIONS OF PITCH IN VOICE PROCESSING

Sumeet S. Andhalkar¹

¹Student, Department of Computer Engineering, G.H.R.C.E.M., Wagholi, Pune, Maharashtra, India

Abstract

Detecting the Pitch of singing voice in the presence of music is a challenging work to do. In this paper we study the problem of finding pitch range or value by analyzing the voice. We also review the vibratos and tremolos used to detect the singing voice. We just take review of detecting pitch from musical record and merits and demerits of pitch. Pitch is very important characteristic of voice and by detection it we can develop other different research works in voice processing field. We also study the advantages of pitch and its limitations. We can also study the future application that we can develop by using pitch. 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. By using pitch we can detect various other characteristics of voice. We can develop many applications those are very useful for the benefit of human being or mankind. 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. So it is very much necessary to study pitch and its future work.

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

1. INTRODUCTION

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.

One method to find pitch is extract melody from the music record. Then find the pitch of that particular singing voice. Alternative way suggests separation of voice and music and then only find sinusoidal wave of singing voice. Then extract necessary data from wave and then we can estimate the pitch.

The pitch ranges are usually large to cover most of the possible pitches of singing voice such as from 80 Hz to 800

Hz. However, it is unlikely that pitch changes in such a wide range in a short period of time. Furthermore, the upper pitch boundary of singing can be as high as 1400 Hz for sharp singers.

The detection of the pitch of a harmonic sound in the presence of other sounds has been mainly studied in the context of speech segregation, speech enhancement and other related fields. As per the study, before applying methods developed for pitch detection for speech, it is helpful to make a comparison between singing and speech. We can develop large number of applications in computer technology by using this pitch range. These applications may be useful in technology, medical science, security applications etc. which are enlisted in this paper. So pitch is a crucial factor in voice processing field. So in later part of the paper we will study the method to estimate the pitch and advantages and disadvantages and applications and future works of pitch and pitch detection.

1.1 Need of the Work

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 disease identification gender identification or for some security purposes we can use these applications.

Timbre – 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.

2. PITCH DETECTION

We can detect trend of the pitch by Trend estimation algorithm. There are basically four steps.

- 1] Vocal Improvement.
- 2] Sinusoidal Mining.
- 3] Instrumental Pruning.
- 4] Pitch estimation.

2.1 Vocal Improvement

Here, The Vocal improvement methods are used to improve the voice from musical record. It also increases the quality of the musical record. They use Harmonics and percussive sound separation. Harmonic instruments are like flute, Guitar. Percussive instruments are like Tabla, Drum. These methods reduce the energy of harmonic instrument and reduce the energy of percussive instruments. This means that the signals of these harmonic and percussive devices are made weaker. So that we can easily catch voice signal more rich and loud and clean.

Here, we use first step only. The signals of harmonic instruments are reduced only. The Percussive instruments do not make any difficulty to sense the voice and its pitch. So, this is the first step in Pitch Detection.

2.2 Sinusoidal Mining

The sinusoidal partials can be mined by using this step. The Waves of the musical record and the vocal and musical parts can be found in this step. Some defective parts or peaks from these waves can be detached. These peaks do not belong to the musical file. Then, regular peaks can be used to form Sinusoidal partials. After peak selection one grouping algorithm is used to reduce the gaps.

It consists of three steps: early grouping, re-grouping and purifying.

- i) Early grouping: It starts by selecting any ungrouped peak as the first peak in the group and recursively groups other peaks neighboring to the group until each peak fits to a group. The closest peak is selected if the difference between it and the previous peak is not larger than a semitone. If no peak exists, the group is divided into two.
- ii) Re-grouping: The objective of this step is to unite the partials that are possible to be originated from the same.
- iii) Purifying: Here study says that we require filling the gap in between the partials. This gap filling is the basic concept of purifying. Because the re-grouping step introduces gaps in a group, this step chooses one peak from overlapping peaks and fills the gaps for each partial. Thus we can purify the signal. For filling the gaps, values of frequency and magnitude are interpolated according to the adjacent peaks. As we fill all the gaps, then we can say we have purified or enhanced these partials effectively.

2.3 Instrumental Pruning

First we should know the actual meanings of some technical words and then we can start up with pruning.

Vibrato-Vibratos are the vocal partials.

Tremolos-Tremolos are the instrumental partials.

This Pruning step categorizes vibratos and tremolos. This means it simply classifies the vocal partials and instrumental partials. As vibratos are vocal partials, Vibratos refer to the variation of pitch. As Tremolos are instrumental partials, Tremolo refers to the periodic vibration of intensity. As per the study and survey, this whole step used to detect singing voice. Two qualities are computed to describe vibrato and tremolo the rate and the extent of vibrato or tremolo. According to study and literature, for human singing voice, the average rate is around 6Hz for both vibrato and tremolo. Hence we determine the relative extent around 6Hz by using the Fourier transform for both vibrato and tremolo.

2.4 Pitch Estimation

According to literature study, some graphs are obtained by the above step. By analyzing those graphs the series or variety of the singing voice can be estimated. This is an important step for actual picking of range of the pitch variation. Pitch varies as singer proceeds with the song. This extent to which the voice is high or low as singer sings a song is pitch variety or range. This is the basic thing to study. If we can get the pitch range then we can develop variety of applications based on voice processing. By analyzing those graphs and some figures we can easily estimate the range of the pitch.

The pitch obtained in this process/method is not a value, but it a range or variety of values that can be used for further processes. The following/ below model can be used to extract these pitch ranges from vocal partials.

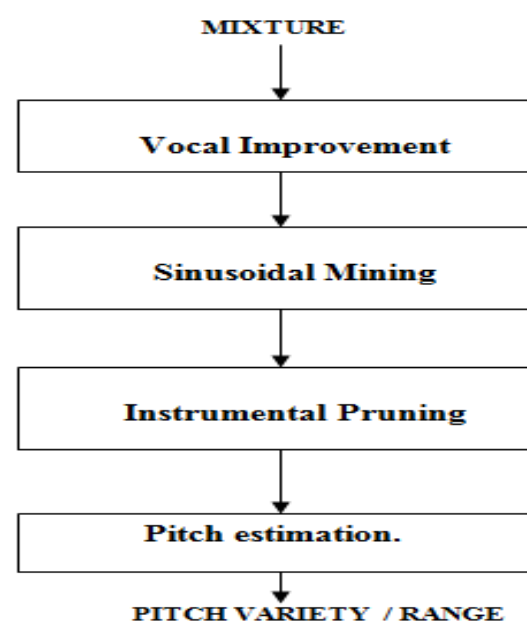


Fig 1 Steps in Pitch Detection

3. MERITS OF PITCH DETECTION.

There are many advantages of pitch. By detection of pitch we can enhance the research area and enrich research in voice processing. The advantages of the pitch are elaborated below.

i) Pitch is the most basic characteristic of voice. So by detecting the pitch we can do many other things like intensity and timbre detection.

ii) Separation of voice and music - If we have the pitch then we can separate the human voice and musical parts. We have to decide pitch range for each alphabet or for each word. Then by analyzing various pitch ranges from the musical record we can decide the words and separate them out to the paper.

iii) Gender Identification - As we have the pitch, we can easily detect the gender of the singer as male and female voices have different pitch range. As male and female pitch ranges or levels are different as differentiable, then we can definitely identify the gender of

iv) Lyric Separation/Extraction- We can separate out the words of the song sung by the singer. Pitch is an important aspect here. Lyric is the words spoken by singer or words of the song. These words have different pitch ranges. So by detecting pitch, we can easily detect the of the song. So we can separate out words or lyric from the record.

v) Information Retrieval- We can retrieve large amount of information from these musical records, if we have the pitch. Various instruments have various pitch 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.

vi) Instrument Detection: Various instruments have various pitch 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 range, by estimating pitch we can identify the instrument easily.

vii) Raag Identification- Identifying Raag is definitely beneficial for new classical singers or classical music students. As Raag is digitized here, they can be stored and easily studied by the students.

viii) Identification of diseases- We can identify diseases 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.

4. LIMITATIONS OF PITCH DETECTION

There are some limitations also which are discussed below.

4.1 Complexity

The entire process of detection of pitch is very complex and difficult to implement. The method requires some electronic devices and accurate software. It is very difficult to develop software which is very complex.

4.2 Accuracy

Any kind of mistake or wrong conversions leads to an error. These are not allowed. Any wrong conversion or input may change entire output and system may work wrongly. Accuracy level must be very high. Any mistake or bug or error can lead to failure of application or will give wrong output.

4.4 Separation and comparison:

The separation of voice, noise, music, words, vocal parts are difficult to implement. Here, high level of accuracy required. Then to compare that data and to produce result more complex methods are required. So, this comparison must not be error prone.

5. APPLICATIONS OF PITCH DETECTION

By extracting pitch range or trend or variety we can assess many things. By using or analyzing those things we can develop many soft wares and applications those may add value to the society.

i) Singer Identification System: We can identify singer by estimating pitch range of his voice. This application can be used for various purposes related to music. We can identify the singer by using pitch variety in his singing. The pitch range can identify the singer who is performing. This can be done at real time also if system have the historic samples of singer's voice.

ii) 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.

iii) Voice Separation System: Separation of voice from musical record is required in special cases. To make only music file, the vocal parts are removed. Karawaks are made by using this technique.

iv) Lyric Identification System: This system identifies words from the song or speech or voice. The Lyric are separated from the record by using pitch range. As each word has its own pitch range.

v) Gender Identification System: By using pitch we can identify the gender of singer or person. Male and female voices have different pitch ranges. So by comparing these ranges we can determine the gender from voice of the person.

vi) Raag Identification System: As pitch is basic characteristic of voice, by using pitch we can estimate raag from the classical music or classical song. Raag' detection is also needful when application deals with classical music. Raag is central structure in Indian classical music. Raag is also known as unique set of complex melodic gestures. In one song similar bits and trend is repeated constantly and continuously. That repeated pattern is known as 'Raag'.

Same pattern of wave forms are repeated that can be collected and stored to identify the raag. If we can separate that repeated pattern then we can easily separate, identify and store that pattern as raag.

vii) 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.

vii) Composing Song: By analyzing and understanding digital raag, music composer can compose many songs by using same raag. This is more important advantage.

vii) Identification of Time of Recording of Voice: We can identify the time slot in which the voice is recorded. The pitch and timbre combination can lead to identification of recording time. The system can identify time slot of recording like morning, evening or afternoon recording.

viii) Age Identification System- Age of the singer can be assessed by using this pitch and timbre. Age of singer or age group can be identified by the software using pitch and timbre values.

ix) 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.

6. CONCLUSION

This paper studies method for singing pitch detection in musical recordings. This paper studies method of detecting pitch range from voice. The studied method says that by following some steps it can extract pitch from voice. We studied that 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. This paper also studies advantages and disadvantages of the pitch. This paper also studies applications and future scope of the pitch i.e. uses of pitch. We mainly studied that there is a method that can detect pitch from singing voice and we can obtain its range by the respective method.

ACKNOWLEDGEMENTS

The Singing voice pitch detection and voice separation and supporting documents are the work of many people, including 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.

REFERENCES

- [1]. Chao Ling Hsu, Deling Wang, Jhy Shing "A Trend Estimation Algorithm For Singing Pitch Detection In Musical Recordings".
- [2]. Chao Ling Hsu, Deling Wang "A Tandem Algorithm For Singing Pitch Extraction and Voice Separation From Musical Accompaniment".
- [3]. L. Regnier, G. Peeters "singing voice detection in music tracks using direct voice vibrato detection".
- [4]. Yipeng Li and DeLiang Wang "detecting pitch of singing voice in polyphonic audio".
- [5]. L. Regnier, G. Peeters, "singing voice detection in music tracks using direct voice vibrato detection".

BIOGRAPHIES



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