Characterization of Pitch Intonation of Beijing Opera

Kainan Chen

MASTER THESIS UPF / 2013

Master in Sound and Music Computing

Master thesis supervisor:

Xavier Serra

Department of Information and Communication Technologies

Universitat Pompeu Fabra, Barcelona

UNIVERSITAT POMPEU FABRA

Abstract

Beijing Opera is a Chinese traditional art that has been developing for several hundred years. The music of the Beijing Opera is quite distinct, without much influence from the west. One of its distinct characteristics is the pitch intonation that the singers do and in this thesis we have developed an analysis methodology for studying it. From a collection of audio recordings of sung arias we extracted the vocal pitch contours of the automatically detected vocal segments of the recordings and computed a pitch histogram for each aria. From these pitch histograms we studied the intonation by analyzing the peak distributions. Since all the melodies are organized in reference to a tonic, or keynote, we also had to identify the tonic in each aria and normalize the histograms to it in order to compare all the histograms. We were able to study the distinct intonation of the notes fa and si, comparing it with the western equal tempered intonations. We also studied how the pitch histograms can be used to characterize the different role types of the Beijing Opera.

Acknowledgements

I would like to thank all the people who helped me. The first should be my supervisor, Prof. Xavier Serra. I experienced a great academic improvement working with him in this project.

Thanks to Compmusic project and all the people working there. Gopala Koduri, who did a related research on Carnatic music. During this time, he has given me a lot of useful advise, and even spent an important amount of time on my questions, saving me plenty of time. Rafael Caro is a Spanish researcher on Chinese traditional music, and has helped me with theoretical issues. I am also grateful to Post-doc researcher Dr. Mohamed Sordo, PhD candidates, Ajay Srinivasamurthy, Sankalp Gulati, Sertan Senturk, researcher Alastair Porter, and collaborator Shuo Zhang.

Also Prof. Emilia Gómez, hinted me when I was working on the keynote estimation. When I had questions in the algorithm Melodia, her PhD candidate Justin Salamon, helped me analyze the problem.

Every time I faced a problem with Essentia, a powerful Audio DSP tool, Dmitry Bogdanov, one of its developers, was always willing to help me.

Nadine Kroher, my classmate, who will be the new PhD student in MTG, and did the peer review of the state of the art, shared a lot of helpful ideas with me.

There are so many people who helped me during this year; for those I forgot to mention here, please accept my apologies.

Contents

List of Figures

Chapter 1

Introduction

Beijing Opera is a form of traditional Chinese theatre that combines Chinese traditional Poetry, singing, instrumental music, dance, acrobatics.... It arose in the late 18th century in the city Beijing and developed to be very popular in the middle of 19th century and spread to the whole country, receiving influences from each of the different local traditions until being celebrated nowadays as China national opera [1].

1.1 Musical factors in Beijing Opera

We introduce several important concepts from the traditional theory of Beijing Opera that are relevant for our research, namely, scale, shengqiang and role type.

1.1.1 Scale

Beijing Opera uses a anhemitonic pentatonic scale as the main melody structure, and a hexachord scale as the secondary. The intervals are similar to the Western scale. For example, the distance between the 1st and 2nd degree is the same to note C and D. The notation we use is called “numbered music notation” (The more detail is in the Appendix). The keynote is not fix in the score, and also singers change it to the most fitting in performances or recordings.

	Male Style	Female Style
Male Character	Lao Sheng	Xiao Sheng
Female Character	Lao Dan	Zheng Dan

TABLE 1.1: Role types we concern

1.1.2 Shengqiang

Shengqiang is like the concept "mode", the structure of the pitch material, but it is bigger than that. Besides of the pitch, it also defines the type of singing, main accompanying instrument and associated emotional content. There are many shengqiang in Chinese traditional music, and in Beijing Opera, most of arias are in two shengqiang, Xipi and Erhuang, so in our research and experiments, we mainly focus on these two. And we are also interested in the associated emotional content of them, because Xipi represents happy and angry, Erhuang shows peaceful and sad. And in the result section, we talked about shengqiang identification from pitch histogram.

1.1.3 Role Type

All the characters in Beijing opera scripts are classified into four different categories. These categories imply specific gender and age, psychological pattern, singing style and scenic behaviour. Role type can be categorized male and female characters, but the singing styles of male and female in Beijing Opera are not always the same to the gender of role characters, like it is shown in the table 1.1.

1.2 Problem description

We define intonation as the pitches used by a performer in a given musical piece. From this definition our approach will consider a performance of a piece as our intonation unit, as the same as the definition in Carnatic pitch intonation research [3]. From the scale, we can see that when 1st degree is fixed onto a Western note, all the other notes on the pentatonic scale can be mapped to the notes on Western scale — the notes interval is similar to the tonality music [1]. But to characterize the additional notes, the 4th and 7th degree of the Chinese hexachord, the intervals are not semitones like B and D in Western scale. Like it is written, "The Beijing Opera fa is between a sharp and a natural Western fa, and the ti is slightly lower than a Western ti." [1]. It is very interesting, but we couldn't find any evidence or proof about it. So we take this question to be our first task.

The study of intonation differs from that of tuning in its fundamental emphasis. As the role type and shengqiang are so important in Beijing Opera, we also take their identification from pitch histogram into account.

To characterize this, we should not only detect a single note, or notes in some tracks, but to evaluate it through a dataset includes most of the popular works by famous artists as our final proposal in this intonation feature. We did these steps below:

- Extract voice melody
- Specify the intonation feature from the melody
- Implement it through the dataset
- Compare result by Role type...

1.3 Related work

As my supervisor, Xavier Serra, is leading the project Compmusic, which explores the world culture music, including Chinese music, in a view of computing, like audio digital signal processing, machine learning... Since it has been started years, there are several topics have been finished and my thesis referred a lot from them.

1.3.1 Intonation research in Carnatic music

This research is done by G. Koduri et al, with the full name of Characterization of Intonation in Carnatic Music by Parameterizing Pitch Histograms [3]. Our methodology is the most related to their work, mainly on the segment classification and pitch analysis. Even though they mainly focus on violin, mine is on voice, the processes were succeed in the tests and results.

[1] Elizabeth Wichmann, *Listening to Theatre: The Aural Dimension of Beijing Opera*, University of Hawaii Press, 1991, 82

1.3.2 Tonic identification in Indian music

For some Chinese music's feature, the keynote identification becomes to necessary. For the tonic or keynote identifications in western music are very well developed, but not so useful in other traditional music, like Chinese and Indian. Our keynote identification process is like the way S. Gulati et al, discover it in Indian music [4] in their research A Tonic Identification Approach for Indian Art Music.

1.3.3 Pitch analysis in Maqam music

Both G. Koduri et al's and mine pitch analysis are related to B. Bozkurt's research Pitch Histogram based analysis of Makam Music in Turkey [5][6], especially the pitch histogram part.

Chapter 2

General concepts

The main knowledge of our intonation research can be categorized to audio signal processing and machine learning. In the following sections, some concepts related to our work are described.

2.1 Audio signal processing

Audio signal processing we used is about Music Information Retrieval (MIR), read music information like tempo, pitch... from audio signals. From digital signal processing point of view, to get higher level information, like intonation, we have to first extract low-level features [7][8]. These features can be classified to temporal, spectral, perceptual and cepstral features.

2.1.1 Temporal features

Temporal features are amplitude features in time domain. Here are four features I used, and how they work in my project are shown in the list below:

- Log-attack time: Time length of the transient (attack).
- Temporal centroid: Sustaining analysis.
- Zero-crossing rate: Correlated to Spectral centroid, but fast computing.
- Energy: How hard it is played, filter out silence in classification

2.1.2 Spectral features

Spectral features are the analysis in frequency domain. This is a very important part in our intonation research. As we know, our voice is monophonic sound, with fundamental frequency and harmonics, and our interest, the pitch, is just the fundamental frequency of the voice, which is spectral feature. in the section of voice segment classification, I used several spectral features to help, like Spectral spread, Spectral centroid... In the list below, it shows the features worked in our case and some explanations.

- Spectral centroid: Very related to the brightness. Both higher pitch and bright timbre (trumpet) will return a higher value. (In the spectrum of trumpet, the harmonics are with more energy, even the first, with higher energy than f0, so it is very bright.)
- Spectral spread: Because it minus the spectral centroid, the pitch change doesn't return a very different value. It more concerns the timbre feature.
- Spectral skewness: From the equation, comparing to the spectral spread, the weight of the difference to the spectral centroid is higher.
- Spectral kurtosis: From the equation, comparing to the spectral spread, the weight of the difference to the spectral centroid is higher, higher than skewness.
- Spectral slope: Amount of decreasing. (How fast the value of harmonics down to about 0)
- Spectral roll-off: f / 95% of the signal energy is below this frequency (related to the slope).
- Spectral flatness: Noisiness, the higher noise, the higher value return. [0, 1]
- Spectral flux(variation): Both the timbre feature and the pitch change will return a high value. For example, noise is very unstable, so the flux is changing fast and with a very high average value.

In the pitch identification, I used Melodia [9] to extract predominant pitch in tracks, polyphonic audio. It can find the predominant fundamental frequency from all kinds of instruments. Also this algorithm considers the smoothy of the frequency between current and previous frames, so that when sometimes the pitch from the predominant melody isn't predominant in a frame, this algorithm will still return its pitch.

2.1.3 Cepstral features

Cepstral feature is in "quefrency" domain, which we use Mel-frequency cepstrum coefficient(MFCC) to extract. It use Discrete Fourier Transform (DFT) to extract spectral information and in MFC, the frequency bands are equally spaced on Mel scale, which is a non-linear scale based on human's perception. So MFCC is a timbre descriptor, which can be very useful on classification of timbre, which can be identified by human. It talks these four steps:

- Take the DFT of (a windowed excerpt of) a sampled digital audio signal.
- Map the powers of the spectrum obtained onto the Mel scale.
- Take the logs of the powers at each of the Mel frequencies.
- Take the discrete cosine transform of the list of Mel log powers, as if it were a signal.

2.1.4 Perceptual features

This kind of features can be felt by human directly, like loudness. For example, if we record someone playing piano loudly, and play the recording with lower volume, the audience will still feel it is loud. So the loudness here is a kind of timbre feature, not simply volume. The list below shows the two features I used:

- Loudness: Use the algorithm of Bark bands.
- Sharpness: Spectral centroid using bark-band loudness. (all the brightness timbre, high pitch and the loudness together)

2.2 Machine learning

Machine learning is quite hot topic today as a very important part of artificial intelligence and it is also very useful in MIR world like genre classification. Like Tzanetakis famous research [10], using machine learning to classify 10 genres.

2.2.1 Why we need machine learning?

Since audio signal computing has been developed long, many musical features can be described by computer into data, like we use more than 300 features to describe a second of music in our research. With this much of data, we need some help to analysis and implement the result to process, and the most popular tool is machine learning. There are many algorithms to filter instances, attributes and classification, the we use the tool WEKA [11], implemented these algorithms in our voice segment classification.

2.2.2 Attributes and instances filtering

Some of the attributes and instances should be filtered, like over qualified, if we only have a small dataset N to classify M clusters, the attribute number A should be limited, shown in function below:

$$A \leqslant \frac{N}{10} * M$$

So we always need to filter the attributes, and the basic strategy is to keep the most classification contributed ones. And also compare with other aspects, like when our training dataset is too limited, theoretical working attributes should be kept.

Not all instances should be kept. When the training data is large, there could be wrong in the "truth" and we can pick them up by their values in the contributes, like outlier, extreme values... Sometimes, the instance is the outlier in one attribute but, normal in others, and if we filter them all, there would waste a lot of instances. So with high correction required machine learning tasks, we have to analysis attributes and instances together to decide which one to be filtered, like in our case, and I will explain it in the voice segment classification section.

Chapter 3

Experiments

As our purpose is to analyze the intonation, which is a high-level feature, and we started from the audio digital signal, this chapter is about how we defined the methodology and the implementation.

3.1 Related experiments

From the most related work, Intonation research on Carnatic music [3], pitch histogram is a very good statistic result demonstrating the pitch intonation feature, shown in figure 3.1. Peak positions denote notes by cents. From this histogram, we can find a lot of features, like the kurtosis of the peak from hundreds cents. For example, 700 is G in the figure (normalized), from how much the 4th peak offset describes how much this note is played offset to western music in average.

We can base on this methodology to research the intonation of Beijing Opera, if it works well in the experiments to find the pitch feature of fa and si. And it was successful in the result. The following section is about how it was.

3.2 Methodology test

The purpose of this implementation is to find whether the voice can be demonstrated well by pitch histogram, if it is successful, we will research based on this methodology, with much specified for Beijing Opera.

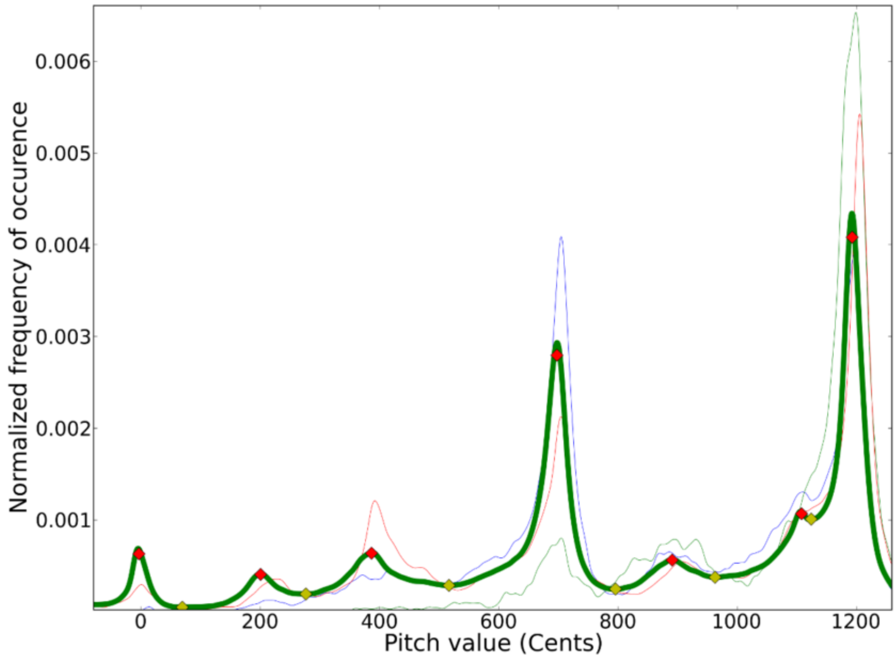

FIGURE 3.1: Pitch histogram sample for Carnatic music [3]

In the implementation, we just simply processed function in the list below without changing many parameters.

- Choose a track, *White Gate Tower - Song Xiaochuan*
- Select voice part (without instrument-only part)
- Extract predominant melody
- Convert pitch frequency to cents (equivalent scale)
- Draw pitch histogram
- Move keynote to 0

Figure 3.2 shows the pitches for a section of extracted melody. Some parts are missing because the predominant melody is not clear and because I cut off instrument-only part, the melody is not so continuously any more, not all the pitches are extracted. But our first purpose is to prove if it exists, we only need to keep what we found is true, not necessary to collect all the truth data. And after comparing by listening, the demonstrated curve is almost correct, so we confirm this step and keep on.

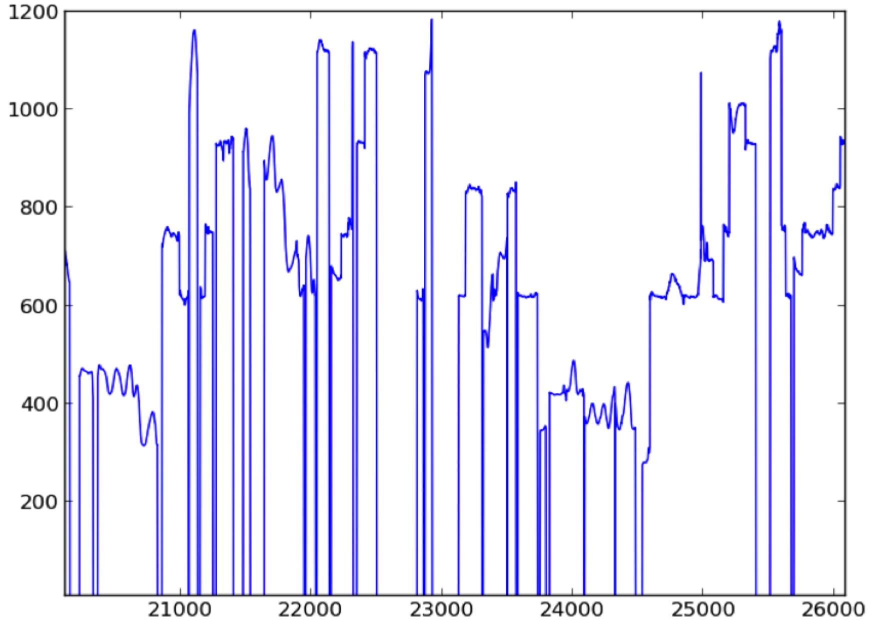

FIGURE 3.2: Example identified voice pitch

Figure 3.3 is the final pitch histogram we got in this experiment and notes correlated peaks are figured out. After zoom in figure 3.4, we can find this is clearly showing our proposal exists in this case - fa is higher than it supposed to be in normal. Hence based on this method, we expended it to our large dataset to find an overall result.

3.3 Audio collection

As Beijing Opera was popular about 100 years ago, at the time the recording equipment was not well developed, and not so popular nowadays, there are not as many as tracks as western music. Furthermore, there are not many new opera since the end of last century, the amount of studio recorded tracks are very limited. So we don't need to buy so much to cover most kinds of Beijing Opera.

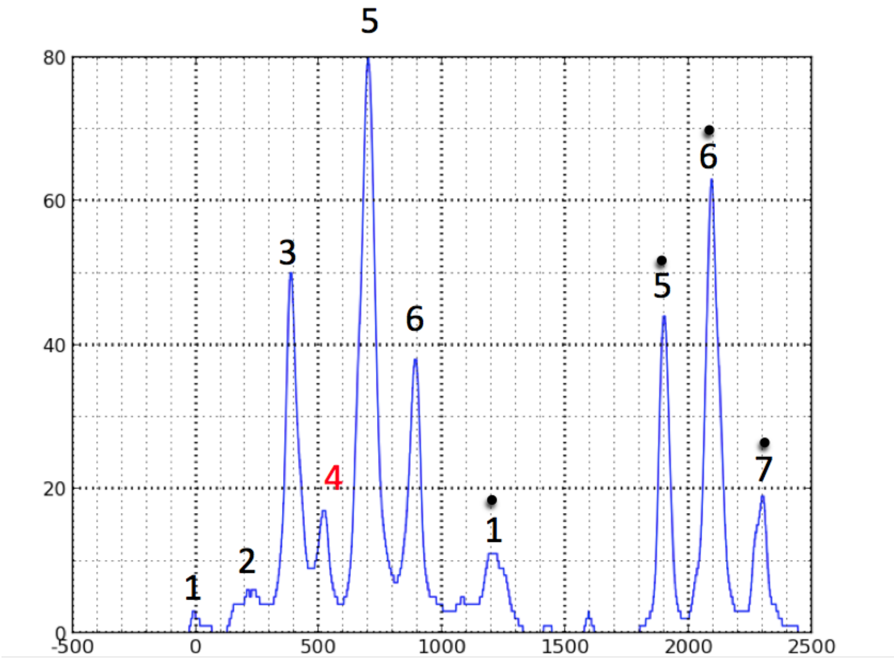

FIGURE 3.3: Pitch histogram with notes figured out

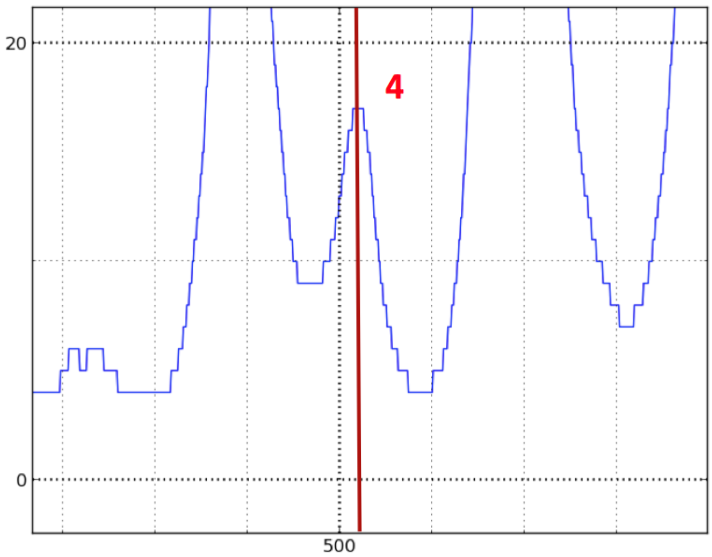

FIGURE 3.4: Pitch histogram with fa zoomed in

3.3.1 Track selection

For choosing the tracks, because we research the Beijing Opera by a computing view, the quality is considered very important. When we were choosing the track, we looked more into the label and year, new recorded from professional company. And finally, we found two good companies, which are Chinese state-owned business.

Based on these requirements, we would like them to cover most popular artists, schools, role types and shengqiang. And we found there is already a big collection of them, called Beijing Opera Star. This collection includes 30 famous Beijing Opera artists nowadays and recorded with modern studio environment.

Besides of these professional artists, there are several good artists very good but not famous enough to be collected to Beijing Opera Star, we were glad to include them as well. Whole opera is a very interesting point to research, but unfortunately, singers today only play small tracks, like a scene, we can only find whole opera recorded more than 20 years ago or live version. So we just chose three for further research.

Summing them up, there are 56 CD, include 558 tracks, 51artists, 7 role types, 18 schools. The details are attached in Appendix. The best way to access the Beijing Opera still popular recordings, is to buy CD/album directly from China. It is hard to buy them in Amazon or some other e-commercial companies, which have contract with both China and Spain customs, and the only way we found is to buy them from the audio shop, who had no idea about how to export CD out of China.

We purchased the CD from a shop, which the owner knew a lot and gave me much information and recommendations. After we got them, we posted them but blocked at the Chinese custom for two months, and finally we got them two more months later, taken by my friends and my parents directly.

3.3.2 Metadata maintenance

Because of our computer point of view, we can hardly do anything if the metadata is printed on the boxes of CD, except we copy them to computer. To store them in a better structure and related way, we chose to upload them to MusicBrainz [1].

Besides of the metadata we found on the CD, we also uploaded some more important information that were tested to correct, or referred to somewhere official. Like translations, alias... The translation as a prominent metadata, we used Pinyin, with the method [12], when we could not find one official translation. (This part of work is done by me and Rafael Caro together.)

As a problem, there is much metadata, but still missed some important ones for our research, like Role type, shengqiang, school... But they are obvious in the wearing of the artists and the melody in the song. So we added them ourselves when they were missing and compiled them into a excel file, so that easy to read or look up. For the experiment point of view, it was better to store the translation as pinyin, no matter if there was an official English translation.

I finished most of the experiments in python, a programming language and in this way, excel file was not so readable, and the better way was to restore them into a YAML file. How easy it is for computer to read was the most important factor this time.

In Beijing Opera, it is very common that many artists play the same scene (track with the same name), so there would be many ambiguousness in the title of tracks and they were all in Chinese characters, which made me not convenience when doing experiments. So we convert the title into MusicBrainz ID, using MusicBrainz Picard. an MBID is a 36 character Universally Unique Identifier that is permanently assigned to each entity in the database, i.e. artists, release groups, releases, recordings, works and labels. For example, the artist Queen has an artist MBID of 0383dadf-2a4e-4d10-a46a-e9e041da8eb3, and their song Bohemian Rhapsody has a recording MBID of ebf79ba5-085e-48d2-9eb8-2d992fbf0f6d. So now, the MBID is our new title, and the original Chinese title is now in the metadata.

I didn't use the pinyin stored in the metadata excel, because it was better if all the letters are without accents, and all converted by the same algorithm. To translate them into new standard, I used the algorithm, based on a python lib [2], added punctuations conversion.

The YAML file contained:

[1]MusicBrainz is an open music encyclopedia that collects music metadata and makes it available to the public. URL: www.musicbrainz.org

[2]Pinyin Conversion Python Lib, Auther: Lx Yu, URL: http://lxyu.github.io/pinyin/

- MBID
- Title (original + pinyin)
- Artist (original + pinyin)
- Role type (pinyin)
- Shengqiang (pinyin)
- School (pinyin)

3.4 Vocal, non-vocal segmentation

As our research interest is mainly focused on the singing part in Beijing Opera. we have to extract the singing melody from the tracks. Currently, separating the voice from source generally is difficult, but as we only need the melody (pitches), it is much more convenient to extract the predominant melody of the singing part, which is almost the voice melody. But for sure, when there is not a voice, only instruments, there will be a kind of instrument plays as the predominant melody that will be extracted out, which we don't want to. So before extracting melody, we have to segment the section with predominant voice.[3]

Features extraction, train model building and classification are talked in this section. The tools I used are Weka[11], a Data Mining Software in Java, developed by the University of Waikato and also use the audio DSP library Essentia to extract low-level features, which is developed by here in Music Technology Group in Universitat Pompeu Fabra.

3.4.1 Segmentation method

Transparently, we can easily segment a track by listening and cut-off, and the other way is to do it by machine, which is much more complex. The question of which method to choose is mostly depends on the volume of our dataset and the quality of machine segmentation. As we have 558 tracks which is too many to do it all by listening. With the good result of violin segmentation which was

[3]Code is available at `https://github.com/MTG/beijing-opera-intonation/tree/master/segmentation`

finished in Carnatic music research [3], done by Gopala Koduri, we decide to first equally segment it to small pieces and classify them to vocal or instrumental using machine learning with the algorithm implemented by Weka [11].

3.4.2 Segment length specification

As we determined to first equally segment audios, the length should be specified properly. Transparently, the smaller, the less errors, but the time taken and data will be huge if we set it too small (because we have to extract features for all the segment). And if it is set to too big, like 5 seconds, we will probably get a segment with from 0 to 2 seconds instrumental part combined, which doesn't make sense to do further classification any more.

Theoretically, no matter how small we set the length, there should be segments with partial instrumental-only, so the task is to find the longest length can be tolerated. With our analysis, the tolerance mainly depends on the music feature and melody extraction algorithm [9].

- Music feature: Beijing Opera is not a kind of choppy music, and very continuously. And it is not easy to find a single section shorter than a second.
- Algorithm tolerance: The algorithm I used is published by J. Salamon in 2012, which can tolerant some missed part with the segment. How much it can be tolerated is specified by the parameter *Tolerance*. The instrumental length is better within a second to guarantee the quality.

We assumed the classified segment with at least 75% vocal section, so the maximum is 5 seconds. But considering in Beijing Opera, it is very common for a singing section with 2 or 3 seconds and we don't want to waste them all, so we set the length to just one second, with the parameter to 1.2 (maximum 1.4). And at last, this decision was tested to be fine, about 3 days extracting segments low-level features and the data was about 17GB.

3.4.3 Create training data

To set the size of training data, it depends on two factors: the size of the whole database and machine learning overfitting.

- As our database includes 58 CD, and there are 558 tracks, the whole length is longer than 24 hours. So we should create a comparable training set. We assume the average length is 5 minutes, hence the whole length is 2790 minutes, about 2 days long. The training data is better larger than 1% , which is half an hour.
- We have two clusters C to classify, and 50 attributes A maximum to train, so as $I = A * C * 10$ we need 1000 instances each. For one instance a second, the amount is about twenty minutes each.

Sum up two factors and with the theory that the larger training set the better, we decided to create the training set 30 minutes each cluster.

The basic strategy of selecting is to collect vocal segment to a single wave file and move instrumental, silence... everything we don't want to another one using software *Audacity*. Because we cannot cover 558 tracks in this one hour dataset. When choosing the tracks, as the voice in Beijing Opera is complex with male, female and male sings the female character and the opposite in different Shengqiang (mode), We used 20 tracks in a large range, covering most of popular Shengqiang, role type, school and popular artists. The list of this twenty tracks are added in the appendix.

When picking up the voice, there was sometimes having ambiguous myself. What we decided is only choose the ones we are confident with and also when the voice is not predominant, like fading out, we cut them as well, to keep our picked up sections are correct.

3.4.4 Audio feature extraction

Before extract features, we should determine the length to extract. It is meaningless to extract features from the whole song, and also not necessary to do that by each frame. Considering the predominant melody extraction function can detect some unvoiced, the unvoiced length within 1 second is acceptable, we finally decide to segment it by each second.

To extract features, we basically followed Gopala Koduri's violin segmentation in Carnatic intonation research and also a instrument classification research [13], extracted about 300 more features - overlap most of low-level descriptors at first using Essentia, and then filter them by attribute, which will be explained in the following section.

To do the machine learning, we used the tool called Weka, and we have to make our data into the *.arff format, which can be read by the tool. And to make more sense between the training data and the target database, we had to normalize it. So we cluster the data from the same track into the same arff file, and normalize them within the song.

3.4.5 Attributes and instances filtering

This is a very important part as the preparation of the classification. Take out outlier instances and not helpful attributes. As we have large dataset and large enough useful features, the good factor is, the machine learning result will be closed to the folder test result but the difficult part would be balancing filtering between instances and attributes.

For example, if we first take out outlier and extreme instances and then filter the attributes, we would waste a lot useful instances. Like in the testing, 1548 of 3096 was taken out in instrumental section. The reason is, the instance may be only the outlier for the useless attributes, and they are not supposed to be taken out. On the other hand, if we filter attributes first, many valuable attributes were taken out because of the outliers.

For the work would not take too much time, I filtered attributes one by one as the first step. It was to filter outliers for one attribute, and to see if this attributes was useful, if not, took instances back and remove the attribute.

As the result, only 35 features chosen from 338, and only 435 of 6192 instances removed. The attributes supporting a lot are MFCC, which describe much in timbre, pitch confidence, a parameter of Melodia, describe the ratio of predominant, and also some other spectral features helped a lot. This attributes include covariance, skewness... to the same feature, to combine them, there are only 11 features left listed below:

- Inharmonicity: How much it is not so harmonic.

- Pitch confidence: By product of Melodia, to see whether the predominant melody is clear and continuous.
- Spectral flatness: spectral shape descriptor
- Spectral flux: spectral shape descriptor
- Spectral rms: spectral shape descriptor
- Spectral rolloff: spectral shape descriptor
- Spectral strong peak: spectral shape descriptor
- Spectral contrast: spectral shape descriptor
- Spectral valley: spectral shape descriptor
- Zero crossing rate: related to spectral centroid
- MFCC: Timber descriptor

3.4.6 Segment classification

We first trained the model, using 10-fold cross validation to choose the algorithm. As the result of validation, both algorithm k Nearest Neighbor and Support Vector Machine returned very high correction, about 95% , very little influenced when I changed the parameters, like k, and the kernel of SVM. With this result, we analyzed that it is because the classification is very clear by these attributes. All the tested results are shown in the talbe 3.1 below, and finally we chose kNN and k equals to 2, trained the model, the 10-cross fold validation result was 96.5%, we can see that, the results with advanced classifiers got almost the same high value, so the voice features are clear extracted out.

Classifier	Parameter(value)	Result(%)
ZeroR		51.1094
NaiveBayes		88.8457
Trees.J48	Confidence(0.25)	90.0407
SVM	RBFKernel(G0.13)	95.2713
K-NN	K(2)	96.4587

TABLE 3.1: Machine learn classification results

With the trained model, we classified segments by their features, and stored them into an YAML file. After we got this file, we considered whether we have to cut and connect the audio file or not. As mentioned above, the melody extraction algorithm. Melodia, would be more stable if the melody is continuous, we decided not to cut it.

3.4.7 Segment classification evaluation

Though we got 96.5% correction in the 10 cross fold validation, it doesn't mean the segmentation result will be as high as this correction. The method we evaluated was to select 20 tracks in very wide role types, shengqiang (mode), artists as well, but different from the 20 training tracks, and listen whether the seconds selected have predominant singing voice.

We were worrying, maybe the women voice is easy to be confused by Jinghu, both with similar high pitch range and difficult sometimes to distinguish by listening myself. Fortunately, tested from these 20 samples, 100% selected vocal segments are with significant voice and in the instrumental segments, some of them with not clear voice. It proves our proposal were correct, the features make sense and the classification is clear. It was successful enough to have this result to do the following section, faster and stabler than I listen to it myself.

3.5 Pitch histogram

As it is talked, pitch histogram is a way to analysis pitch intonation, in this section, we used the result we had already got from the segment classification to extract and modify the voice pitch data to histogram. [4]

3.5.1 Predominant melody extraction

We just extracted the predominant melody so that we could get the melody of voice. I used tool Melodia (Salamon & Gomez, 2012) [9][14], implemented in the

[4] The code is available at `https://github.com/MTG/beijing-opera-intonation/tree/master/intonation`

library Essentia , and then select the pitches by the segment classification result. And here is a sample of the result is shown below 3.5.

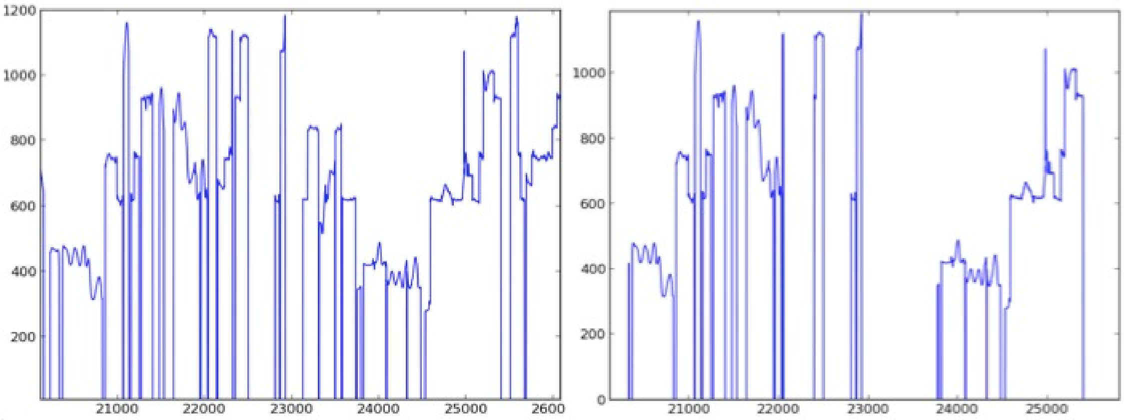

FIGURE 3.5: pitch contour before (left) and after vocal segmentation (right).

But this is not perfect correct, sometimes with an octave error. This problem is currently still unsolved, and with the topic of my thesis, I am not improving this algorithm. As known, there is another algorithm, YIN [15], which have a very high precision, but it is obvious that doesn't work well on polyphonic music. I can only try to set parameters fitting the best and evaluate how much this problem will influence my result. In the figure 3.6, the original wave is shown on the top, which is continuously singing wave from the track "Diao Jin Gui", by the artist Guo Yuejin. This work is Erhuang and with the role type of Sheng. From the figure in the middle, the colorful line shows the melody pitches by Melodia, without changing parameters. From the truth (compared by listening), the first higher part are with an octave higher error, which takes about 60% of this piece, unable to be accepted. The figure at the bottom, using Beijing Opera specified parameters, the correction is much higher, only a very little piece wrong.

With tested by several pieces, the octave error is between 1% to 10%, and all of the errors are with one octave higher. So from this result, at least the notes in the first low octave, are always right, guaranteed to analysis. The parameters are shown below:

- Bin resolution = 1
- Guess unvoiced = False
- Min frequency = 120

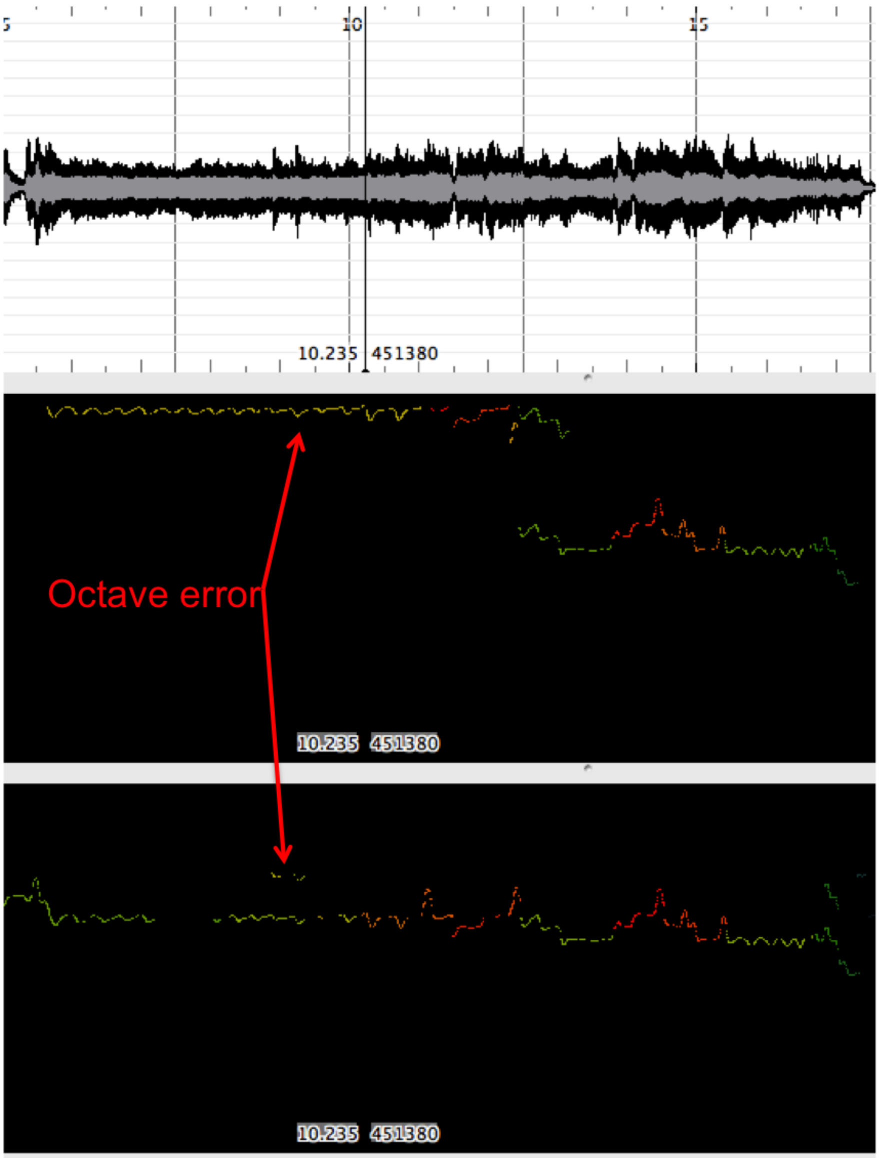

FIGURE 3.6: Original wave (top), extracted pitch by default setting(mid) and pitch by using specified parameters(bottom).

- Max frequency = 900
- Voice tolerance = 1.2

3.5.2 Histogram from pitch

To have the good pitch histogram to analysis, it is not enough just convert the pitch contour matrix to a histogram, but also doing some modify on it. As the default, the bin is at the left edge of the bar, but as the basic visual intuition, we moved it to the center. We can see that from the picture below on the left, the peaks are not clear. To better analyze the peak features, we smoothed it using Gaussian filter, and with the resolution parameter of 7.

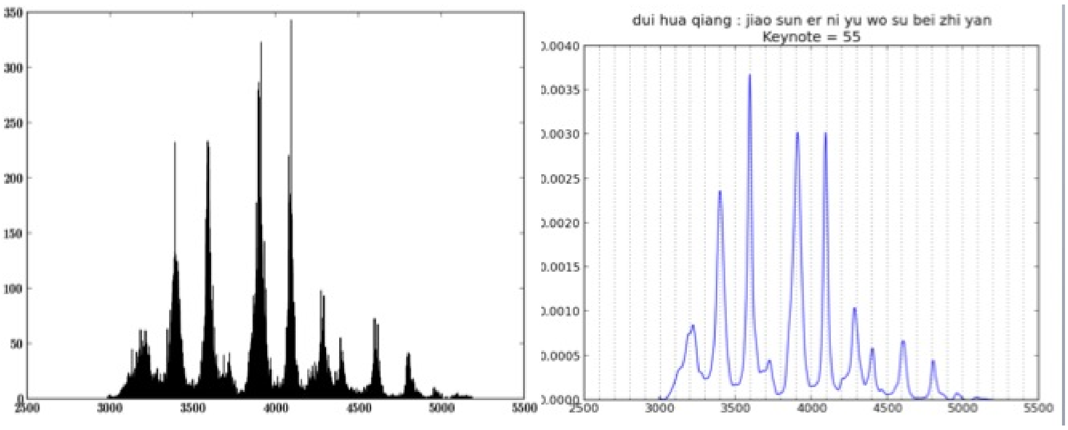

FIGURE 3.7: Pitch histogram before (left) and after modified (right).

As an interesting result, the highest peak is changed after smoothing, but it is not an error. The frequent of the sung note is not only shown by the highest bar, but also the pitches in a small range. The figure on the right, The peaks are integrated by all the bars around the highest peak, and that is the one we need to show how frequently the notes are sung.

The algorithm to evaluate and detect peak is written by G. Koduri, it not only concerns the points higher than their neighbors, but also filters the valleys. These parameters are not simply determined, but dynamic in the keynote identification section, which will be explained below.

3.6 Keynote identification

Traditional Chinese music uses the anhemitonic pentatonic scale, and in the score, the key is not clearly defined, which means, it is a modal music. in modal music, only the relevant scale is important, and how high the singers sing depends on their own characterization. So though we found the pitch, we cannot simply find the note from a single frequency. To find the note, we have to know what the reference frequency, the keynote is. The Chinese notation using the numbered music notation, which means, when the keynote is C, the 1,2,3... 7, map to C, D, E... B, (explained in the figure below). For example, if we want to find how much offset the 4 is, we need to know what the frequency of 1 is. So before we characterize the pitch intonation, we need to find the keynote frequency first.

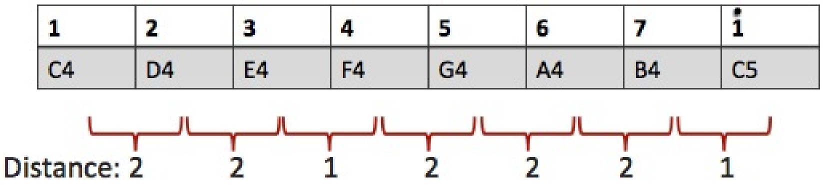

FIGURE 3.8: the notes distribution when 1=C.

3.6.1 Filter peaks by peak-valley distance

There are several peaks in the pitch histogram, but not all of them are what we wanted. So we setup a parameter peak-valley distance to filter them, which describes the peak and its closest valley's vertical distance (figure 3.9). Generally, the longer distance, the more probability the peak is we want. But it doesn't mean the peak with low distance cannot represent a note. So we set this filtering as a factor.

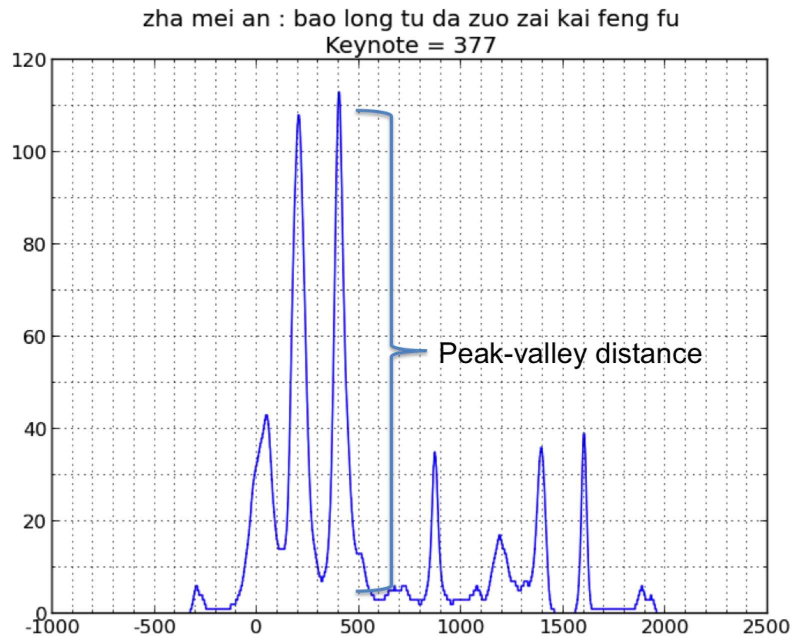

FIGURE 3.9: The longest distance for the peak to its closest valley

This way is better than only calculate the amplitude, and how it works is shown in figure 3.10. With two different threshold, the more peaks are selected when the threshold is set to lower. Obviously, the higher distance, the more significantly, it is the note sung, but we cannot say the low distance peak, like the first black star one, is not a note, and actually, it is. The further filtering will be talked in the following section.

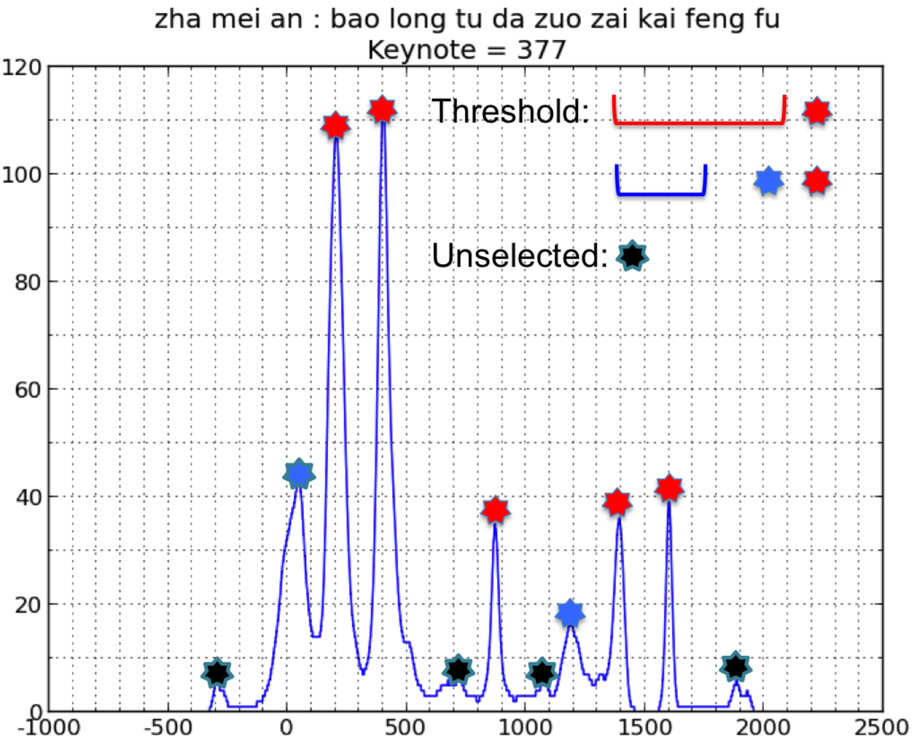

FIGURE 3.10: Select peaks by different peak-valley distance threshold

3.6.2 Filter peaks by horizontal position

How to find out peaks that with low peak-valley distances, but represent notes. As all the tracks are professional recorded from famous artists, we suppose they sing generally on the pentatonic scale. If there are some cases like two semitone closed to each other, one of the peak should not be the structural note. So we also use the peak position information as a factor to filter the peaks.

Still use the example in figure 3.10, why this first black star should be in account. In figure 3.11, the horizontal distances between confident peaks are notified, and we can find the structure can be guessed out. In this horizontal distances, the reasonable note guess are notified with 1, 2, 3... And from the distance, we find that the black star peak, is reasonable to be a 6 in lower octave, so it can be selected.

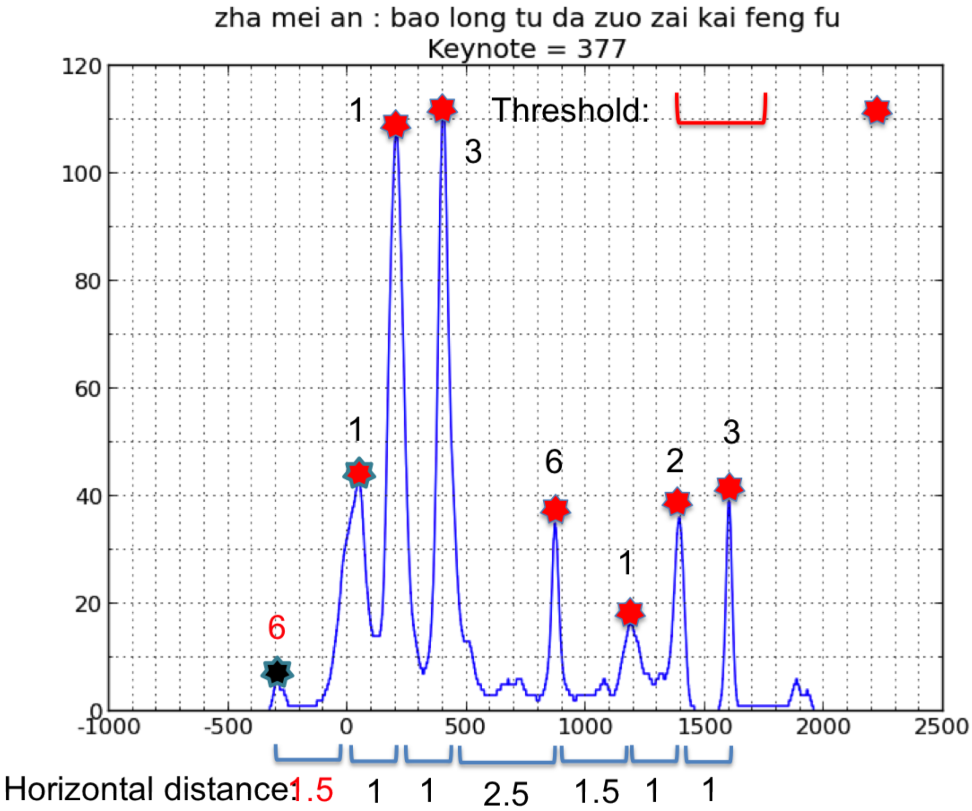

FIGURE 3.11: Select reasonable peaks by horizontal distances

3.6.3 Find keynote candidates

This part is inspired by the paper [16] and [17], a Boundary Search Algorithm (BSA) for determining points of modulation in a piece of music using a geometric model for tonality called the Spiral Array.For a given number of key changes, the computational complexity of the algorithm is polynomial in the number of pitch events. But here in our case, we don't use western scale, but the Chinese pentatonic scale.

We should first define the interval of peaks. As the absolute note pitch is nothing important in Beijing Opera, we cannot map them to the western scale. We choose the most frequent sung note as the reference, and find the distances to it as the relative distances. This is described in the figure 3.12.

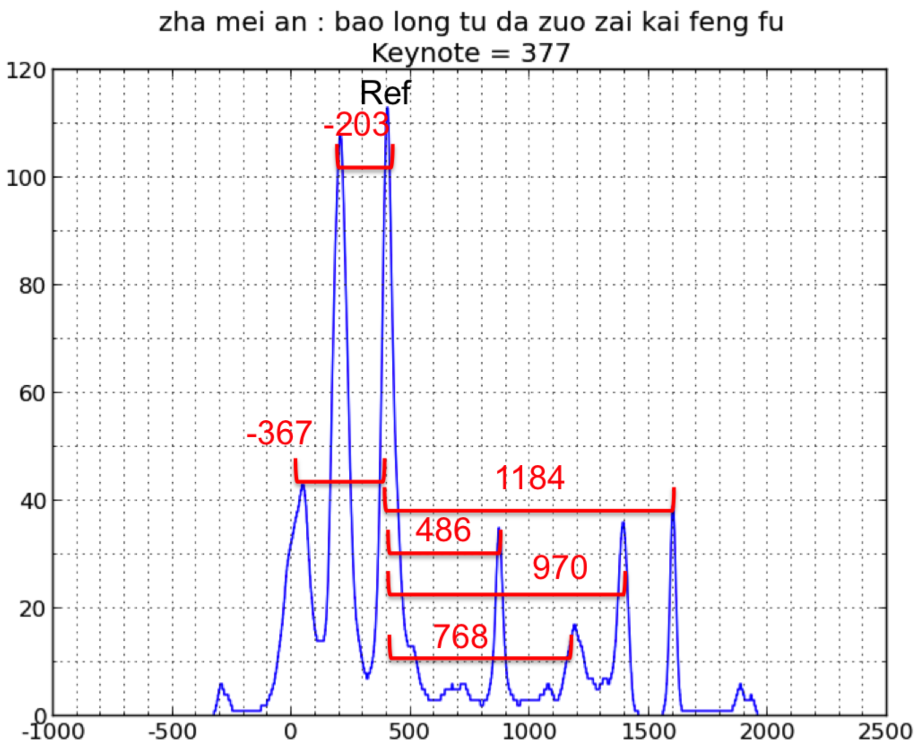

FIGURE 3.12: How we define intervals

WIth the relative distances, we integrate it to integer by $int(d/1200)$. So the mission left is to find the match with the scale. It is shown in figure 3.13, the notes and their intervals are spiraling together to find match with peak intervals.

Once the peak positions match the pentatonic scale, at least 3 notes in an octave, we can find the 1 as a keynote candidate. But this method can often return several candidates. The idea to get more information from the matching, is to suppose the candidate as the keynote, and extend the scale to more octaves to find how many peaks match the new scale, and sum up all the matched peak valleys.

3.6.4 Keynote identification work flow

The main part of the algorithm is shown in the figure below. It starts from the left top corner. The valley threshold is k, which is 0.5 as default to find peaks. The found peaks have two parameters, position and valley.

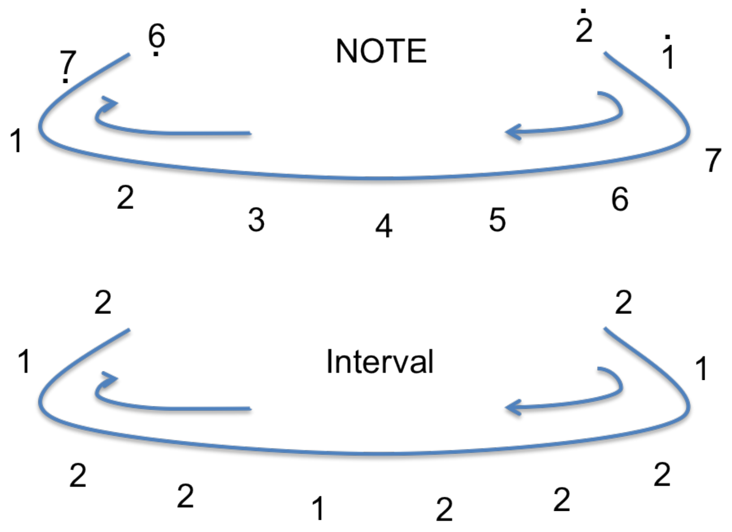

FIGURE 3.13: Spiral Chinese scales and intervals array

Calculate distances between peak positions, and check if they are theoretically correct. The position error means, as the artists are professional, the significant peaks should be on pentatonic scale, if not, there could be an error of detection or the peak is with a very low valley and just a grace note, not representing structural note. For example, if we find there are three semitone locate closing to, the system will report error.

When there isn't any error, the system will continue to guess keynote. The method is using pentatonic scale to match peaks to find which peaks have possibilities to be keynotes.

We suppose the keynote is correct and expand the theoretical note to higher and lower octaves to create a matrix. If the proposal is true, most of the peaks, especially the high ones can match to the matrix. So in this step, we sum up the matched peaks' valley together as a confidence factor, and the next step, find the highest as the keynote.

This process sometimes cannot find keynote, even keynote candidates, and most of times were because the valley threshold initially was too high, and only 3 or even 1 peak was found. So if this happen, the system will reduce the valley threshold to find more peaks, as a loop, again and again until, there is an position error or the keynote is found.

If the keynote is not found but there happens an error, the system will force to reduce valley only once to find the keynote, but will report a list store this kind of instances. The keynotes of Beijing Opera are mostly in a range, if the keynote is not in that normal range, the system will report it into the list as well. Our final strategy is to check them manually by listening or look up from the score. The result of the twenty tests are shown in the appendix.

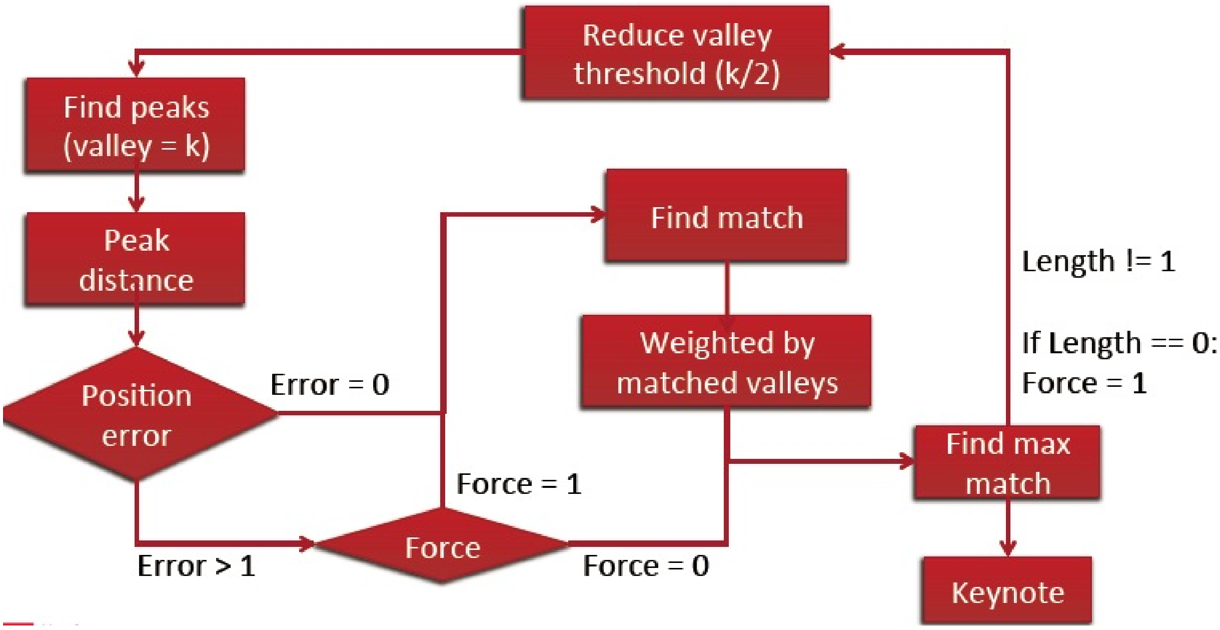

Figure 3.14: Flow chart.

Chapter 4

Results

The result is the intonation analysis of what we found from the pitch histogram. It includes how the intonation performs comparing to the scale is, the 4th and 7th degree of Beijing Opera scale characterization, which is the tuning research, and some factor identification, like role type, from pitch histograms.

4.1 Notes and scales identification

This is a fundamental section of the pitch histogram analysis of Beijing Opera. From the previous chapter, the Keynote identification, we got the keynotes of the arias. Refer to the keynote, we look into, how are the notes, and scales.

4.1.1 Pentatonic scale structure identification

Beijing Opera is mainly on the structure using pentatonic scale. It is reasonable that if all the steps we did are right, the notes on the scale can be more significant in pitch histogram. From the figure below 4.1, the notes are pointed out, by the distances refer to the keynote. By analyzing the distances, we can find this is exactly the pentatonic scale and our progress is proved to be correct.

Also we were curious about what the peaks are after the first 5 peaks. From their positions, it is shown that they are high pitch notes. We first compare whether it matches to the notes on a higher octave, and the answer is, most of they are

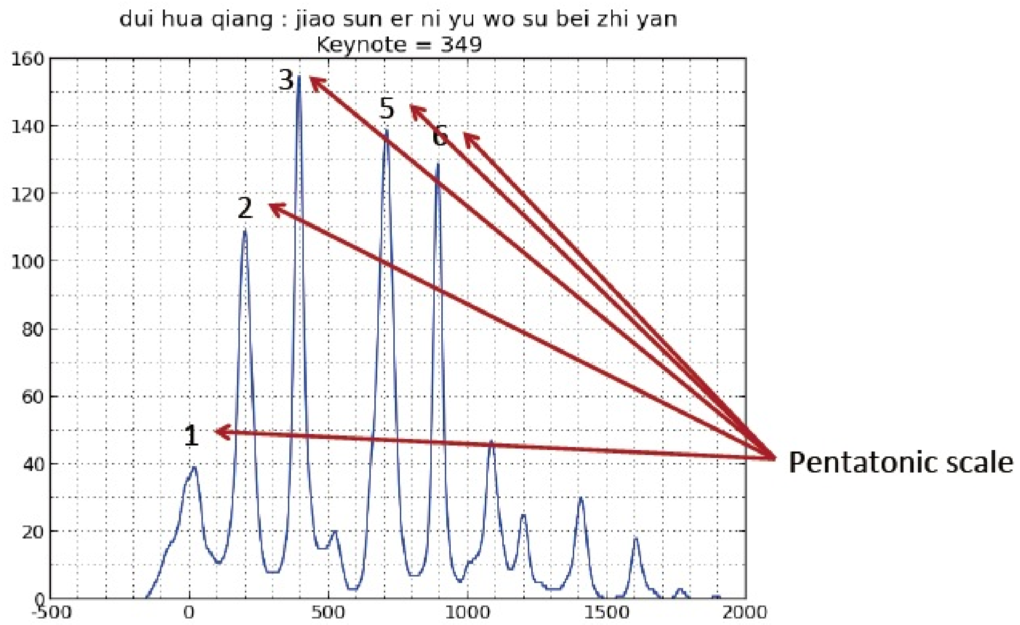

FIGURE 4.1: Main scale structure.

matched, some are on the additional notes, as we mentioned, the 4th and 7th degree.

In Beijing Opera, the 4th and 7th degree notes are not in the main scale, so the amplitudes usually are small. In the second octave, the amplitudes distributions are strange. With this wondering, we listened to this aria, and we found that it is true, the notes are sung like that distribution, but because of a section of modulation, changed the key.

The modulation is very common in Beijing Opera, more frequently than western music. A very popular modulation, from male singing style to female, is shift the key 4 degrees higher, for example, from C shift to G, in Beijing Opera, from Gong to Zhi, like this case. The peaks from the 6th, are the consequence of the combination of modulation and a higher octave of original scale. The modulated scaled are notified in the figure 4.2.

4.1.2 4th and 7th degrees characterization

4 and 7 are not main structural notes, mostly as grace notes. As one of our interest of researching Beijing Opera, the 4 and 7 in some records written that they are not on the equal temperament scale, 4 is a little sharp, and 7 is a little flat.

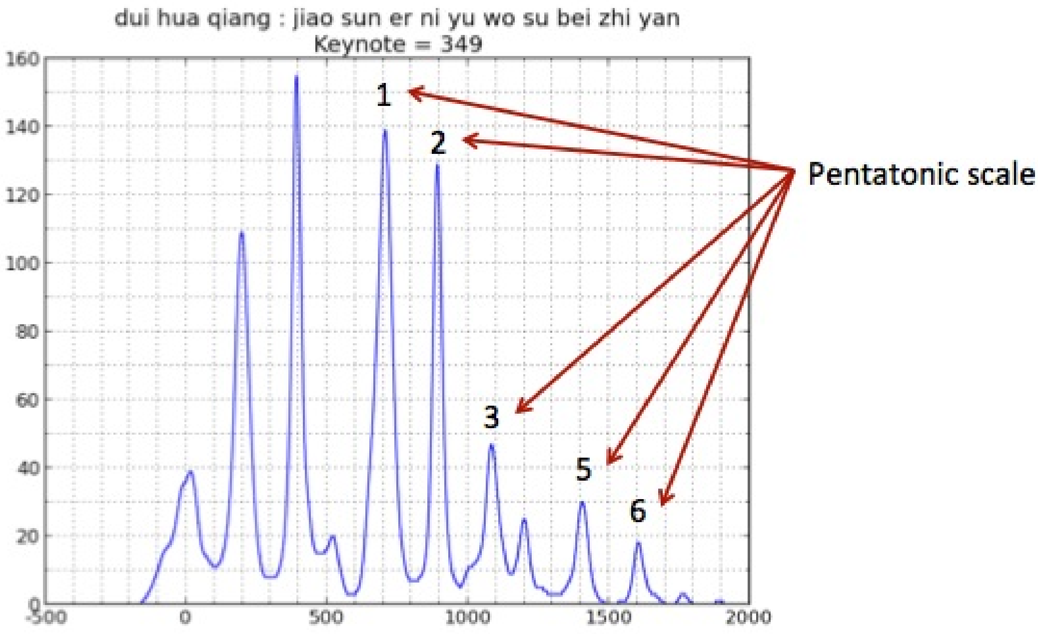

FIGURE 4.2: Scale structure after key transfer.

From many samples, we did find this feature, and like an example shown in figure 4.3. The pitch histogram has already been normalized to the keynote. After we zoom in to see 4 and 7 clearer. In figure 4.4, from the picture on the left, 3 and 5 are almost on the equal tempered scale, but 4 is not, between normal and sharp. And from the picture on the right, 6 and 1 are on the equal tempered scale, but 7 is between flat and normal 7. From the samples we analyzed, we proved that this phenomenon happens, but could not prove it is common everywhere in Beijing Opera.

4.2 Shengqiang and role type identification

With not much experience of listening to Beijing Opera, people can easily distinguish it from western music, with its own feature. But within this kind of music, different role type and shengqiang (mode) have their own clear characterization and important for our research to help understand Beijing Opera.

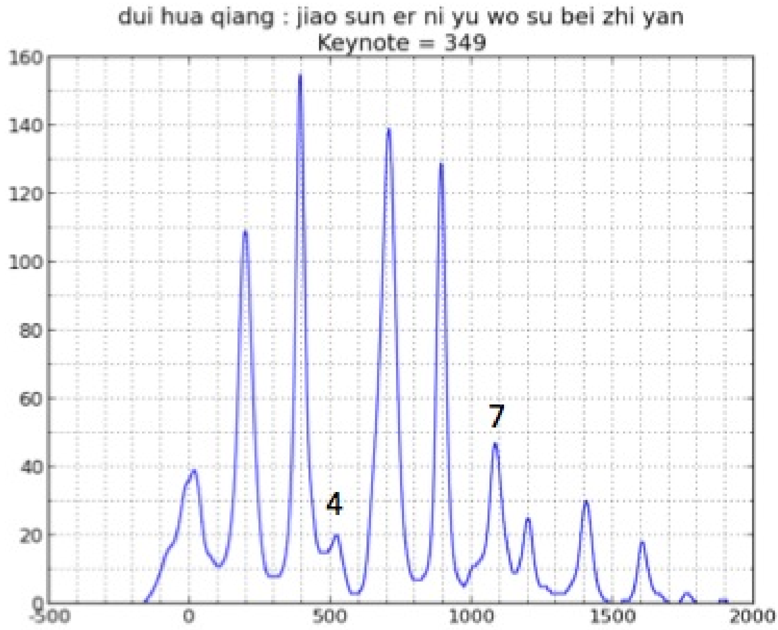

FIGURE 4.3: with 4 and 7 figured out.

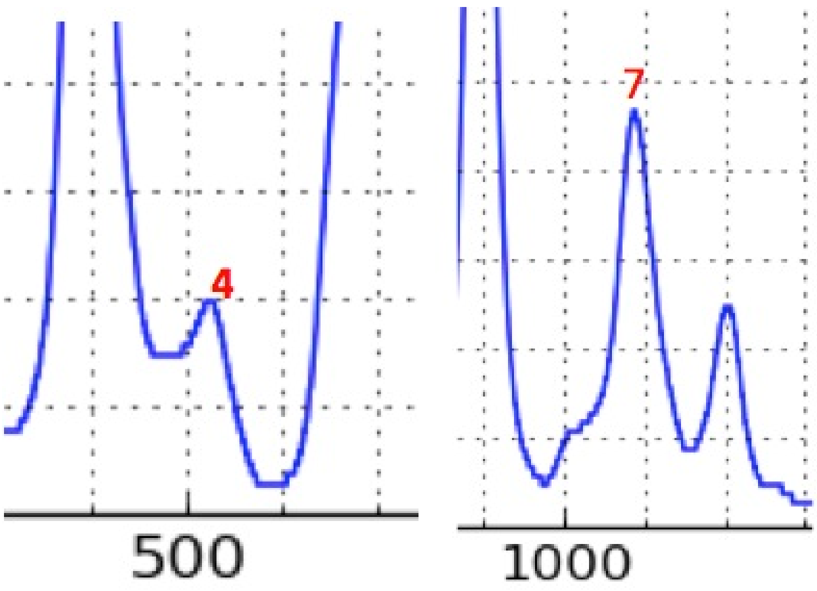

FIGURE 4.4: zoomed to notes 4 and 7.

4.2.1 Shengqiang and role types' selection

There are many shengqiang and 4 role types in Beijing Opera. In the first experiments, we chose two of the most important shengqiang, Xipi and Erhuang, and the role types of Sheng and Dan, which are categorized by male and female singing style. The experimental arias are listed in the appendix, 20 arias selected to do qualitative analysis.

4.2.2 Regular method for identifications

In the lyrics, the sentences can be called opening line and closing line, and combine them together to build a whole sentence. For musicians to learn which role type or shengqiang the aria is, they often take care of the final notes of both opening line and closing line by listening. The theoretical differences are shown in the table 4.5 below.

Shengqiang	Role type	Final note	
		Opening Line	Closing Line
Xipi	Sheng (male)	2	1
	Dan (female)	6	5
Erhuang	Sheng (male)	1	2
	Dan (female)	$\dot{1}$	5

FIGURE 4.5: final notes for different shengqiang and role types.

The differences are well defined, but our methodology — analysis by pitch histogram, doesn't include the opening and closing lines information, that means, we cannot simulate this process perfectly by computer.

4.2.3 Problems for regular method

Our basic idea was to implement this method into algorithm to find shengqiang and role type automatically, as the final notes are supposed to be sung frequently. But this is failed when we looked into the pitch histogram. From the figure 4.6, the final note is 1, but it is not high at all. After we listen to the recording and look into the score. We got the conclusion to explain:

- Final note is important in structure
- Note length can be short, little present in pitch histogram
- Difficult to identify mode/role from pitch histogram only by the amplitude of final note.

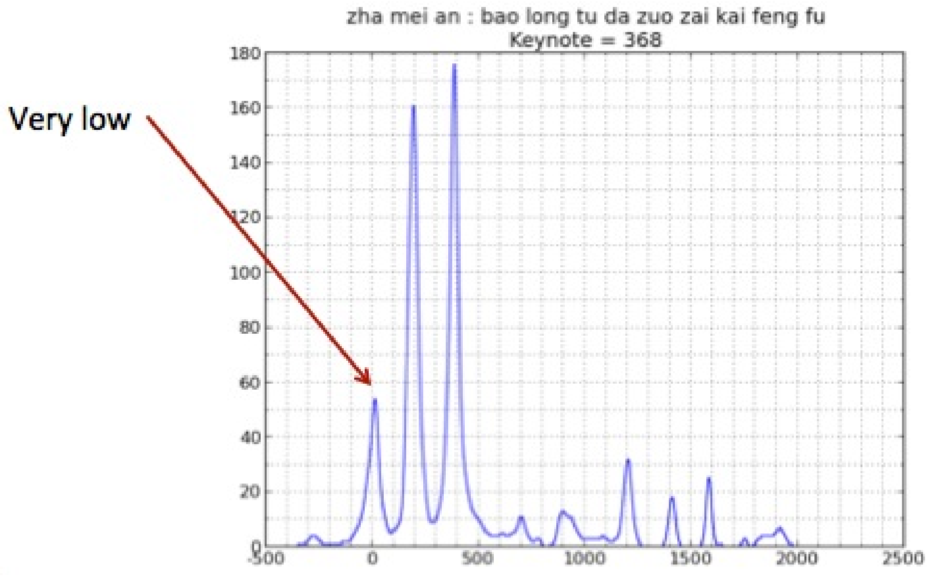

FIGURE 4.6: final note in pitch histogram.

4.2.4 New possible identification method

Even though the method of implementing what often the musician do to identify the shengqiang and role type, we found some other characterization can help. From the figure 4.7, we find the most frequent note and second most one are different in different role type. In sheng, the highest is 3 and 2nd one is 2, but in dan, the highest is 5, and 2nd is 6.

After found this interesting phenomenon, we summarized these two peaks in all this 20 samples, into the table below 4.8. First 3 are very clear in identification, but the last one is not. By analyzing this, we guess this maybe because the male sings dan role type can be in a very wide range of pitch and they transfer keys a lot in this shengqiang, so the most frequent and second frequent notes are not quite clear.

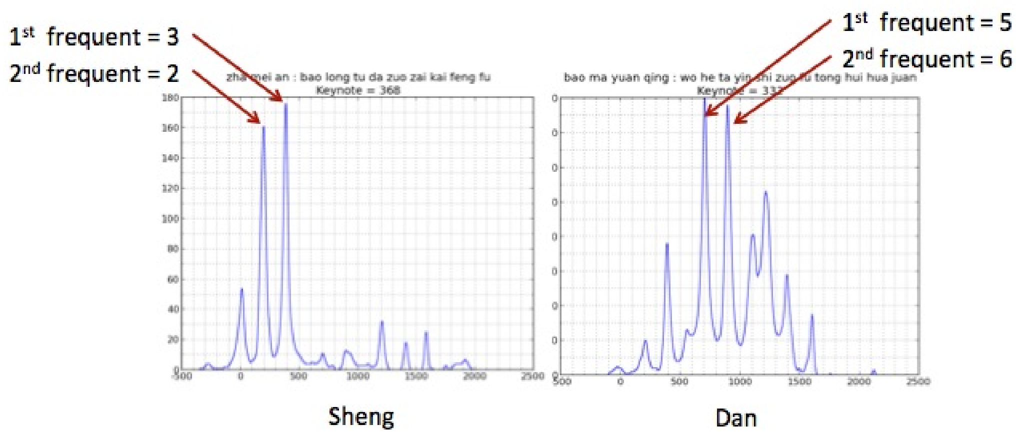

FIGURE 4.7: sheng on left, dan on right.

Shengqiang(mode)	Role type	Most frequent	2nd frequent
Xipi	Sheng (male)	3	2
	Dan (female)	5	6
Erhuang	Sheng (male)	2	3
	Dan (female)	?	?

FIGURE 4.8: summarization of most and second frequent note in different shengqiang and role types.

Chapter 5

Conclusion

With our research on pitch intonation in Beijing Opera from recordings of aria performances through an audio digital signal computing method, in the first place, we validated the assertion found in literature that the 4th degree usually is higher than its equivalent in the equal-tempered scale and 7th degree is lower. Secondly, we carried out a quantitative analysis of pitch distribution properties for role type and shengqiang, developing a way to identify them from the pitch histogram.

In our computing process, we first implemented a voice/non-voice segmentation by machine learning. After that we extracted the predominant melody, i.e. the pitches from the voice segment, and we computed histograms from them. Keynote is a very important factor in Beijing Opera, and we developed an algorithm to find it from the pitch histogram and normalized the histogram to it.

The precision of our pitch result is also a limitation of the future analysis, and in each of our processes, many details are remained to be improved well.

Regarding our audio collection, we have 558 arias, that cover the analyzed properties comprehensively. But in order to have a larger collection to obtain the data, we can enlarge the volume by taking account of the less popular but well recorded arias, and further more, include old recordings covering early stages of the tradition.

It has been proved by ourselves, that the result of voice/non-voice segmentation is useful for now, the time resolution can be improved higher, better with 200ms, to reduce the unvoiced tolerant parameter in the melody extraction part, but in this case, we need a server with a larger volume.

By using the predominant melody extraction algorithm, we still have octave error in the melody extraction part and the error ratio is not small. The possible way to enhance it is to add more correction algorithms based on Beijing Opera's features, including both voice and instruments.

The keynote identification is a very important mid-level feature analysis. Although we have already got an algorithm with a very high precision, we have to face that the error tolerance is very low, the better constrain the error rate into 1% . A possible solution is to have an additional parameter describes the "confidence", and re-check the low ones manually. For a further difficult idea is to compute it time variantly.

Being one of the first computational approaches to the analysis of pitch intonation in Beijing Opera, there is much future work to be done. Based on the pitch result we obtained so far, we can design algorithms to compare to find more potential features, like the differences between role types. The pattern in Beijing Opera is very significant, and it is meaningful to add time variance analysis to recognize it.

During our research we have experienced that the tools developed for the analysis of Western music are not directly applicable to Chinese traditional music, since it is based on a different traditional musical culture. Consequently, these tools need to be adapted to such features in order to obtain results that are relevant to this tradition. Our research has achieved some of these possible adaptation, some new trial and possible direction of future researches on it.

Bibliography

[1] Xu Chengbei. *Peking Opera (Cultural China Series)*. China Intercontinental Press, 2003.

[2] Elizabeth Wichmann. *Listening to Theatre: The Aural Dimension of Beijing Opera*. University of Hawaii Press, 1991.

[3] Koduri Krishna Gopala, Joan Serrà, and Xavier Serra. Characterization of intonation in carnatic music by parametrizing pitch histograms. *ISMIR*, pages 199–204, 2012.

[4] Sankalp Gulati. A tonic identification approach for indian art music. *Masters' Thesis at UPF*, 2012.

[5] Barıs Bozkurt. An automatic pitch analysis method for turkish maqam music. *Journal of New Music Research*, 37(1):1–13, August 2008.

[6] Gedik Ali and Bozkurt Baris. Pitch-frequency histogram-based music information retrieval for turkish music. *Signal Processing*, 90(4):1049–1063, 2010.

[7] Fabien Gouyon, Perfecto Herrera, Emilia Gómez, Pedro Cano, Jordi Bonada, Àlex Loscos, Xavier Amatriain, and Xavier Serra. *Content Processing of Music Audio Signals*, chapter Sound to Sense, Sense to Sound: A State of the Art in Sound and Music Computing, pages 83–160. Logos Verlag Berlin GmbH, Berlin, 2008. ISBN 9783832516000.

[8] Geoffroy Peeters. A large set of audio features for sound description (similarity and classification) in the cuidado project. *CUIDADO I.S.T. Project Report*, 2004.

[9] Justin Salamon and Emilia Gómez. Melody extraction from polyphonic music signals using pitch contour characteristics. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(6):1759–1770, 2012.

[10] George Tzanetakis and Perry Cook. Musical genre classification of audio signals. *IEEE Transactions on Audio, Speech, and Language Processing*, 10 (5):293–302, July 2002.

[11] Remco R. Bouckaert, Eibe Frank, Mark A. Hall, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. Weka-experiences with a java open-source project. *Journal of Machine Learning Research*, pages 2533–2541, 2010.

[12] Yin Binyong. *Chinese Romanization: Pronunciation and Orthography*. Sinolingua, January 1991.

[13] Da Deng. Melody transcription from music audio: Approaches and evaluation. *IEEE Transactions on Audio, Speech, and Language Processing*, 38(2):429–438, April 2008.

[14] Graham E. Poliner, Dan P. W. Ellis, Andreas F. Ehmann, Emilia Gómez, Streich Sebastian, and Beesuan Ong. Melody transcription from music audio: Approaches and evaluation. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(4):1247–1256, 2007.

[15] Alain de Cheveigne and Hideki Kawahara. Yin, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917–1930, 2002.

[16] Elaine Chew. The spiral array: An algorithm for determining key boundaries. In *Proceedings of the Second International Conference, ICMAI 2002*, pages 18–31. Springer, 2002.

[17] Ching-Hua Chuan. Polyphonic audio key finding using the spiral array ceg algorithm. *IEEE Transactions on Audio, Speech, and Language Processing*, pages 21–24, July 2005.

Chapter 6

Appendix

6.1 Numbered music notation

Here are only five notes in an octave used as the main structure, so it is pentatonic music. there is no semi-tone between notes, so Beijing Opera is anhemitonic pentatonic [2]. Like other traditional Chinese music and from 19th century, started to use numbered notation as the standard score. As an example of numbered notation, when $1 = C$, 1, 2, 3, 4, 5, 6 and 7, are equal to C, D, E, F, G, A and B. As an anhemitonic pentatonic, Beijing Opera mainly use 1, 2, 3, 5 and 6 as the normal note structure, and use 4 and 7 as secondary notes. The absolute pitch is nothing important, but the relative, so it is also called modal music. Here is an example shown in figure 6.1. the 1, which is the first degree in Beijing Opera, is the keynote, called Gong note. In the notation, we often specify the keynote to a frequency if we want to fix the absolute pitch.

6.2 Track list

Releases	Disc no.	Role Type	School
Beijing Opera Star–刁	1	Dan 旦角	Mei 梅芳梅派
Beijing Opera Star–董	1	Dan 旦角	Mei 梅芳梅派
Beijing Opera Star–史依弘	1	Dan 旦角	Mei 梅芳梅派
Beijing Opera Star–王蓉蓉	1	Dan 旦角	Zhang 君秋派
Beijing Opera Star–秀君	1	Dan 旦角	Zhang 君秋派

Beijing Opera Star–李佩	1	Dan 旦角	Cheng 程秋程派
Beijing Opera Star–小秋	1	Dan 旦角	Cheng 程秋程派
Beijing Opera Star–桂娟	1	Dan 旦角	Cheng 程秋程派
Beijing Opera Star–李海燕	1	Dan 旦角	Cheng 程秋程派
Beijing Opera Star–管波	1	Huadan 花旦	Gou 荀慧生荀派
Beijing Opera Star–耿巧云	1	Huadan 花旦	Gou 荀慧生荀派
Beijing Opera Star–炳	1	Laosheng 老生	Yu 余叔岩余派
Beijing Opera Star–杜	1	Laosheng 老生	Yu 余叔岩余派
Beijing Opera Star–王佩瑜	1	Laosheng 老生	Yu 余叔岩余派
Beijing Opera Star–克	1	Laosheng 老生	Yang 森派
Beijing Opera Star–李	1	Laosheng 老生	Yang 森派
Beijing Opera Star–杜杰	1	Laosheng 老生	Yang 森派
Beijing Opera Star–建	1	Laosheng 老生	Xi 奚伯奚派
Beijing Opera Star–朱	1	Laosheng 老生	Ma 良派
Beijing Opera Star–子茹	1	Laodan 老旦	Li 李多奎李派
Beijing Opera Star–郭	1	Laodan 老旦	Li 李多奎李派
Beijing Opera Star–袁慧琴	1	Laodan 老旦	Li 李多奎李派
Beijing Opera Star–赤	1	Hualian 花	Yuan 袁世海袁派
Beijing Opera Star–俊杰	1	Hualian 花	Qiu 裘盛戎裘派
Beijing Opera Star–安平	1	Hualian 花	Qiu 裘盛戎裘派
Beijing Opera Star–包	1	Xiaosheng 小生	Jiang 姜妙香姜派
Beijing Opera Star–江其虎	1	Xiaosheng 小生	Jiang 姜妙香姜派
Beijing Opera Star–靳斌	1	Xiaosheng 小生	Ye 盛派
Beijing Opera Star–李宏	1	Xiaosheng 小生	Ye 盛派
Beijing Opera Star–宋小川	1	Xiaosheng 小生	Ye 盛派
耿其昌唱段	1	Laosheng 老生	Yu 余叔岩余派
于魁智唱腔	1	Laosheng 老生	Yang 森派
李唱腔	2	Laosheng 老生	Yang 森派
言朋唱腔	1	Laosheng 老生	Yan 言菊朋言派
少云唱腔	1	Laosheng 老生	Lin 周信芳麒派
梁一唱腔	1	Laosheng 老生	Gao 高奎高派
李康唱腔	1	Dan 旦角	Mei & Zhang梅派.派
火丁唱腔	1	Dan 旦角	Cheng 程秋程派
方翔唱腔	2	Hualian 花	Qiu 裘盛戎裘派
惠唱腔	2	Dan 旦角	Mei 梅芳梅派
唱腔	2	Dan 旦角	Shang 尚小云尚派

三起解全（君秋）Whole	1	Dan 旦角	Zhang 君秋派
盆全（天）Whole	1	Laosheng 老生	Yu 余叔岩余派
李逵探母（尚.王云）Whole	1	Comp	Qiu 裘派李派合作
白蛇全唱腔（李炳淑）Whole	1	Dan 旦角	Mei 梅芳梅派
二（李炳淑.沐.乃）	1	Comp	Mei&Qiu&Yang 梅派.裘派.派
玉堂春-三堂（李世）	1	Dan 旦角	Cheng 程秋程派
建峰演唱	1	Laosheng 老生	Xi 奚伯奚派
童芷苓唱腔精粹	2	Dan 旦角	Gou 荀慧生荀派

6.3 Experiment track and keynote list

Name	Artist	Shengqiang	Role type	Keynote (frequency)
02 元媒：自那日与六郎前相	秀君	Erhuang	Dan	D(291.8)
06 成叫：勒停蹄站城道	靳斌	Erhuang	Dan	G(389.4)
07 成托兆：天下荒荒真主呀	江其虎	Erhuang	Dan	G(388.2)
16 妻嫁妹：素秋女只哭得如雨降	管波	Erhuang	Dan	A(441.8)
05 金：叫我的儿啊	郭	Erhuang	Sheng	G#(410)
05 行路子：自昏只走得金叫	子茹	Erhuang	Sheng	G(392.9)
06 相和：在金殿定官是非辨	安平	Erhuang	Sheng	F#(381.1)
08 探皇陵：听打了初更候	俊杰	Erhuang	Sheng	F(358.1)
11 文昭：心中有事合眼	炳	Erhuang	Sheng	F(349)
14 捉放曹：一明月照窗下	李	Erhuang	Sheng	F(349.5)
02 柳：山翠如含笑	包	Xipi	Dan	E(338.7)
02 麟囊：日里好光忽	桂娟	Xipi	Dan	E(329)
04 情：我和他吟作同卷	李海燕	Xipi	Dan	E(332.3)
09 小宴：那一日在虎牢大	宋小川	Xipi	Dan	F(348.3)
01 珠帘寨：太保令把收	李	Xipi	Sheng	E(332)
02 花：叫儿你与我速	子茹	Xipi	Sheng	F(349)
02 槊：自起兵把	赤	Xipi	Sheng	F(350.9)
03 女：一句得我火燃鬃	袁慧琴	Xipi	Sheng	F(352.1)
04 美案：包打坐在封府	俊杰	Xipi	Sheng	F(354)
16 伐：昔年沙	王佩瑜	Xipi	Sheng	F# (368.9)

.

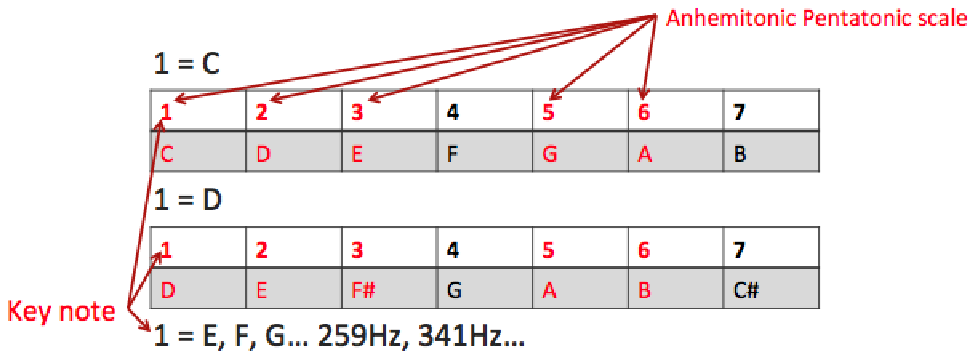

Figure 6.1: Example of modal anhemitonic pentatonic scale