Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 4572904, 11 pages
https://doi.org/10.1155/2022/4572904

Research Article

Vocal Music Teaching Pharyngeal Training Method Based on Audio Extraction by Big Data Analysis

Chunyuan Huang

Music College of Zhaoqing University, Zhaoqing, 526061 Guangdong, China

Correspondence should be addressed to Chunyuan Huang; 2007020006@zqu.edu.cn

Received 24 March 2022; Revised 14 April 2022; Accepted 21 April 2022; Published 6 May 2022

Academic Editor: Chia-Huei Wu

In the process of vocal music learning, incorrect vocalization methods and excessive use of voice have brought many problems to the voice and accumulated a lot of inflammation, so that the level of vocal music learning stagnated or even declined. How to find a way to improve yourself without damaging your voice has become a problem that we have been pursuing. Therefore, it is of great practical significance for vocal music teaching in normal universities to conduct in-depth research and discussion on "pharyngeal singing." Based on audio extraction, this paper studies the vocal music teaching pharyngeal training method. Different methods of vocal music teaching pharyngeal training have different times. When the recognition amount is 3, the average recognition time of vocal music teaching pharyngeal training based on data mining is 0.010 seconds, the average recognition time of vocal music teaching pharyngeal training based on Internet of Things is 0.011 seconds, and the average recognition time of vocal music teaching pharyngeal training based on audio extraction is 0.006 seconds. The recognition time of the audio extraction method is much shorter than that of the other two traditional methods, because the audio extraction method can perform segmented training according to the changing trend of physical characteristics of notes, effectively extract the characteristics of vocal music teaching pharyngeal training, and shorten the recognition time. The learning of "pharyngeal singing" in vocal music teaching based on audio extraction is different from general vocal music training. It has its unique theory, concept, law, and sound image. In order to "liberate your voice," it adopts large-capacity and large-scale training methods.

1. Introduction

In the current vocal music teaching in normal universities, pharyngeal singing, as a unique teaching theory, is gradually accepted by some teachers and students. In the process of vocal music learning, incorrect vocal methods and excessive use of voice have brought a lot of problems to the voice and accumulated a lot of inflammation, so that the level of vocal music learning stagnates or even retreats [1]. How to find a way to improve ourselves without damaging our voice has become a problem we have been looking for. In recent years, a technique "pharyngeal voice," which has been recognized by more and more singers and vocal music teachers, is a golden key to solve these problems. Therefore, in-depth research and discussion on "pharyngeal voice singing" is of great practical significance for vocal music teaching in normal universities [2, 3]. "Pharyngeal sound" practice has two functions: one is to treat voice diseases. Actors, teachers, commentators, salespeople, and other people who use more voice suffer from vocal nodules, congestion, edema, vocal cord paralysis, incomplete closure, and hoarseness. Through "pharyngeal sound" practice, they have good curative effects [4]. As we all know, vocal music teaching can effectively improve students' vocal music quality. Compared with ordinary voice training methods, the biggest feature of "pharyngeal sound" is that it pays great attention to the overall scientific adjustment and training of human vocal organs, which can enable students to grasp the coordination and unity of their own vocal organs faster. Moreover, pharyngeal sound training is applied to various vocal methods, which also plays a good role in the health care and treatment of voice [5, 6].

The essence of event detection is audio classification. Since we want to separate these audio, the most natural idea is whether we can simply extract one or several features and then use these features to separate all audio at once. Each

category has a relatively high recognition rate [7, 8]. Audio extraction is a nonstationary random process, whose characteristics change with time, but this change is very slow. Based on this characteristic, the audio signal can be divided into some continuous short segments to ensure that the audio characteristics remain relatively stable in short time intervals, and then, the audio unit is further processed, which is the short-time processing technology. The above-mentioned audio units are called audio frames, and the length of the frames is generally between one, and there may be partial overlap between adjacent frames, which is conducive to the smooth transition of audio characteristics between frames. For audio extraction, it is the feature of good distinguishability and good stability for different kinds of voice teaching and pharyngeal training extraction [9, 10]. Compared with other audio features, the characteristics of the target event must have its particularity. Therefore, for the target event, we should try our best to select newer and more effective features, that is, features that can distinguish the event from other categories to the maximum extent. At the same time, we can also put forward new universal features that are effective in classifying the target audio events by observing the waveforms in time domain, frequency domain, and cepstrum domain [11]. This method extracts the same feature set for all audio, which is equivalent to universal feature extraction. If you want to classify all audio once, you can get a high recognition rate. These features need to be able to reflect the essential differences between these classes, and the differences within each class should be as small as possible [12].

Vocal music teaching is a main course of music major in normal universities. Different from music theory course, it has strong performance and imitation. The traditional vocal music teaching mode focuses on oral heart transmission. In fact, this teaching method focusing on sensibility is more abstract, which makes many students confused and at a loss in the process of vocal music learning [13, 14]. The study of "pharyngeal singing method" in vocal music teaching based on audio extraction is different from general vocal music training. It has its unique theory, concept, law, and sound image. In order to "liberate your voice," it adopts large-scale training methods. The complexity of features means that it is difficult to determine the target audio category to be classified. Fixed audio features can distinguish a variety of audio categories. Audio is different from vocal music teaching pharyngeal training recognition. Compared with audio event detection, vocal music teaching pharyngeal training recognition can be regarded as a distinction within a single event set, because many years of research can well describe the generation process of vocal music teaching pharyngeal training through vocal mechanism, so the proposed features have good stability [15]. In order to make "pharyngeal singing" produce better effects on learners, it should be based on two abilities: one is more skilled abdominal breathing with large functions, and the other is that it can eliminate all kinds of adverse factors in sound operation and always maintain a relatively stable open throat state. If these two abilities are not good, then "pharyngeal singing" is not easy to master.

This paper studies and innovates the above problems from the following aspects:

(1) A classifier model of vocal music teaching pharyngeal training method based on audio extraction is proposed. The design of classification model is different from the previous speech recognition. In speech recognition, because the same feature set is used in feature selection, the design of classification model is relatively simple. For different types of audio event detection, because of the different feature sets of different categories, the design of classifier should be reasonably constructed according to this point

(2) The vocal music teaching pharyngeal training system based on audio extraction is constructed. The teacher's model singing and explanation, as well as the better part of the students' classroom exercises, are collected by audio, and then, the students can refer to the exercises after class, which can deepen the students' understanding and consolidation of knowledge points and help to form the interaction between teachers and students. This kind of method has been used in some students of undergraduate musicology class and achieved remarkable results. Students can have a better grasp of the content of the class, and self-practice after class can be carried out well, and the progress of learning vocal music has been greatly improved

The paper is divided into five parts, and the organizational structure is as follows:

The first chapter introduces the research background and present situation of vocal music teaching pharyngeal training and puts forward and summarizes the main tasks of this paper. The second chapter introduces the related work of pharyngeal training in vocal music teaching. The third chapter introduces the algorithm and model of audio extraction. The fourth chapter introduces the implementation of vocal music teaching pharyngeal training method and compares the performance of the system through experiments. The fifth chapter is the full text summary.

2. Related Work

2.1. Research Status at Home and Abroad. Lin et al. proposed that in the article "the application and effect of pharyngeal sound in vocal music teaching" combined with vocal music teaching practice, they introduced their own experience in teaching pharyngeal sound training skills in vocal music teaching and proposed to carry out targeted pharyngeal sound technique teaching based on Bel Canto and combined with the specific situation of students [16]. Craciun et al. proposed "pharyngeal sound" as the core singing method of traditional Italian bel canto, and its vocal mechanism is to complete the vocal function through the cooperation of multiple organs in the pharynx. The training of pharyngeal voice can not only quickly improve the function of human voice but also play a role in health care and treatment of voice [17]. Lianxiu puts forward the specific methods of pharyngeal training in vocal music teaching; in "using" the

pharynx "to solve students' difficulties in national vocal music teaching," it focuses on how to use pharynx training to solve students' problems such as treble and timbre [18]. Sun puts forward the functional state of pharyngeal sound production and then expounds in detail and systematically the eight key steps of how to effectively carry out pharyngeal sound practice in vocal music teaching [19]. Shan proposed that "pharyngeal sound" is not a throat sound, not a false sound, not a nasal sound. It is a crisp, bright, solid, and metal sound with high pronunciation position, penetration and durability, beautiful and pleasant timbre, both rigid and soft, flexible and flexible, and deep artistic expression [20]. Jin proposed that the "pharyngeal sound" has a crisp, bright, and metallic tone and strong burst force, penetration, and persistence. If you integrate the "pharyngeal sound" into the sound, it will make your song shine [21]. Lang pointed out that the vocal music circles have different views on "pharyngeal sound", so not many people use "pharyngeal sound" practice method in vocal music teaching in art colleges, which may be related to some comrades' lack of familiarity with "pharyngeal sound" [22]. Yu puts forward that the premise of practicing pharyngeal sounds is to have a good breath and a stable throat, because the throat is the key part of singing pronunciation organs, throat placement is the basis of pronunciation, and breath is the driving force of sound. In singing practice and teaching, the throat muscles are tense due to the throat lifting or pressing, which leads to the phenomenon that the voice is tight, stiff, or the voice is empty and cannot meet the standard of artistic voice, all because the correct throat placement has not been solved [23]. Yu Xiang proposed that in the Teaching Discussion on the Application of Pharyngeal Techniques in Singing Art, the source of pharyngeal sounds was introduced, and then, the pronunciation principle and characteristics of pharyngeal sounds were analyzed. Finally, the basic technical methods and detailed teaching steps of pharyngeal sounds practice in vocal music teaching were analyzed and discussed pertinently, and the results were more practical [24]. Nick et al. put forward that there are a certain number of researches on pharyngeal training at present. Different musicians, combined with their own vocal music teaching practice, study pharyngeal training, and more and more related workers realize the importance of pharyngeal training in vocal music teaching, which enriches the research literature on pharyngeal training in China. For example, in "Talking about Pharyngeal Voice and Singing Art", the basic concepts of "pharyngeal voice" are defined in detail, and the scientificity of pharyngeal voice is proved from the perspective of physics, further pointing out that pharyngeal voice training is the basis of vocal music training [25].

2.2. Research Status of Throat Training in Vocal Music Teaching Based on Audio Extraction. The above related research has important reference and practical guiding significance for the teaching and application of pharyngeal techniques in Chinese vocal music teaching and also has important theoretical and practical significance for the development of Chinese vocal music art. In view of the above content, this paper proposes to study the pharyngeal training in vocal music teaching based on audio extraction. Pharyngeal training method is a fast and effective method to improve students' sound quality. In China, pharyngeal training method is generally only used by professional actors engaged in singing, and few people apply it to vocal music teaching in normal universities. In the practice of audio extraction, we should have the determination of dropping water through the stone. Pharyngeal training can effectively enhance the comprehensive vocal ability of the human body, cultivate the special advantages of singers, correct the wrong vocal methods, and treat various voice diseases. At present, the focus of research and practice in the field of vocal music theory is how to better practice and use the pharyngeal singing method. When learning the pharyngeal sound in vocal music teaching based on audio extraction, we should not only learn it objectively, scientifically, and rationally according to the requirements and methods of practice but also fully realize that vocal music is an overall singing art; because any subject has its significance and inevitability, so we can study and learn from the good things of pharyngeal sound at the same time of vocal music teaching for our own use. We should eliminate distractions, overcome difficulties, and arrange time scientifically, so as to effectively master the essentials and methods of practice. In pharyngeal practice, the practice of various functional actions of audio extraction is monotonous and boring, but singers must not ignore it. If the above exercises can be mastered accurately, they will play an unimaginable role and effect in your artistic singing. In the long development process of "pharyngeal sound system" audio extraction, there have also been controversial periods. However, with the development of science and technology and the progress of society, people gradually realize the scientificity of pharyngeal sound practice. Finally, it has been widely recognized in the vocal music industry and promoted the development of vocal music art.

3. Algorithm and Model of Audio Extraction

Universal feature extraction is also the most difficult and meaningful work focus in feature extraction. The definition of new features can also be considered from three aspects of feature classification: physical features, structural features, and mathematical features. People usually use physical features and structural features to identify objects, while computer identification of objects depends on mathematical features. For the sampled and quantized audio signal, people can observe its envelope in the two-dimensional plane and reflect the change of the signal with time, but it is impossible to classify the audio signal by simple observation in the time domain, so it is necessary to extract the features of the audio signal and complete the classification of the audio according to the features extracted from different signals. At first, the audio features proposed by people may be considered based on physical meaning and occurrence model, but in the end, they are all transformed into mathematical features to enable computers to recognize and calculate. When designing new features, starting from the analysis of the particularity of events, the optimal distinguishing features for specific events are proposed through the time-frequency domain

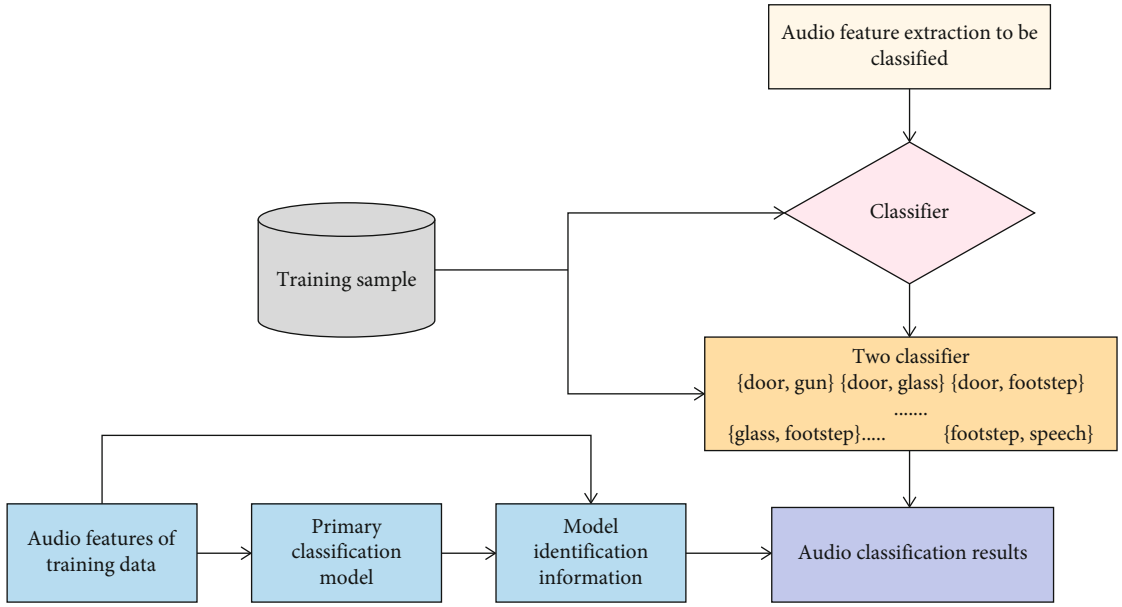

Figure 1: Structure diagram of classifier model.

characteristics of signals, and the existing features can also be considered to be improved and optimized. According to different methods of parameter extraction, audio signal analysis can generally be divided into time domain analysis, frequency domain analysis, cepstrum domain analysis, and other domain analysis methods. According to different analysis methods, audio signal analysis can be divided into model analysis method and nonmodel analysis method. The feature extraction stage is the processing of the divided audio frames. Because the obtained frame lengths are short-term, the extracted frame features are short-term information of the reflected target audio. In the feature extraction stage, the method adopted is to classify and extract the target audio events. After obtaining the audio features of audio events through the above process, the next step is to train the model.

In the final analysis, abnormal audio is still a classification problem. Compared with continuous speech recognition, the detection of audio events is intuitively easier to understand. However, compared with vocal recognition, abnormal audio events also have their own complexity. Many years of research results have developed a quite mature vocal sound model for the vocal sound mechanism, which can also explain why some features can bring good recognition results to speech recognition. The design idea adopted in this paper is the design of secondary classification model. The reason why the two-level classifier in this paper can solve the above two deficiencies is that the design of the two-level classifier is not too large from the perspective of error transmission, and the complexity of the model has been controlled to a certain extent. The second is that the design of the second level feature-dependent classifier well combines the different characteristics of event categories, so that the design of the second classifier depends on the selection of event features. The structure of classifier model is shown in Figure 1.

As can be seen from the above figure, the multiclassifier is the first-class classifier of the model. After the primary rough classification results are obtained by this classifier, they are sent to the second-class classifier for the second-class classification test, and finally, the final audio classification results are output. In this model design, two cepstrum features can better classify the target category preliminarily in rough classification. In addition, besides the short-time frame features, the feature extraction also considers the extraction of long-time audio features. The trial features are obtained after a certain calculation of short-time feature values, and the general calculation method is to obtain the long-time features of audio events by statistical average calculation. Basically, the classification of classifiers is to find the differences between feature data sets. Whether using distance measurement or similarity calculation, in the final analysis, the category labels with large similarity are used as output categories. However, a lot of previous research work focused on model optimization or feature dimension reduction. It is undeniable that both of them can bring optimization to recognition results, but there is often limited room for improvement. In addition, many features are put forward for speaker recognition. Because of the pertinence of the features, the applicability will be greatly reduced if we switch to the direction of audio event detection. The second point is error transmission in model design. Many recognition methods use multilevel classifiers for recognition, but the multilevel classifiers have their own defects. The above feature extraction and transformation are different for the feature sets selected by each category, so this puts forward new requirements for the construction of the classifier.

The design of classification model is different from the previous speech recognition. In speech recognition, because the same feature set is selected in the selection of features, the design of classification model is relatively simple. For different types of audio event detection, due to the different

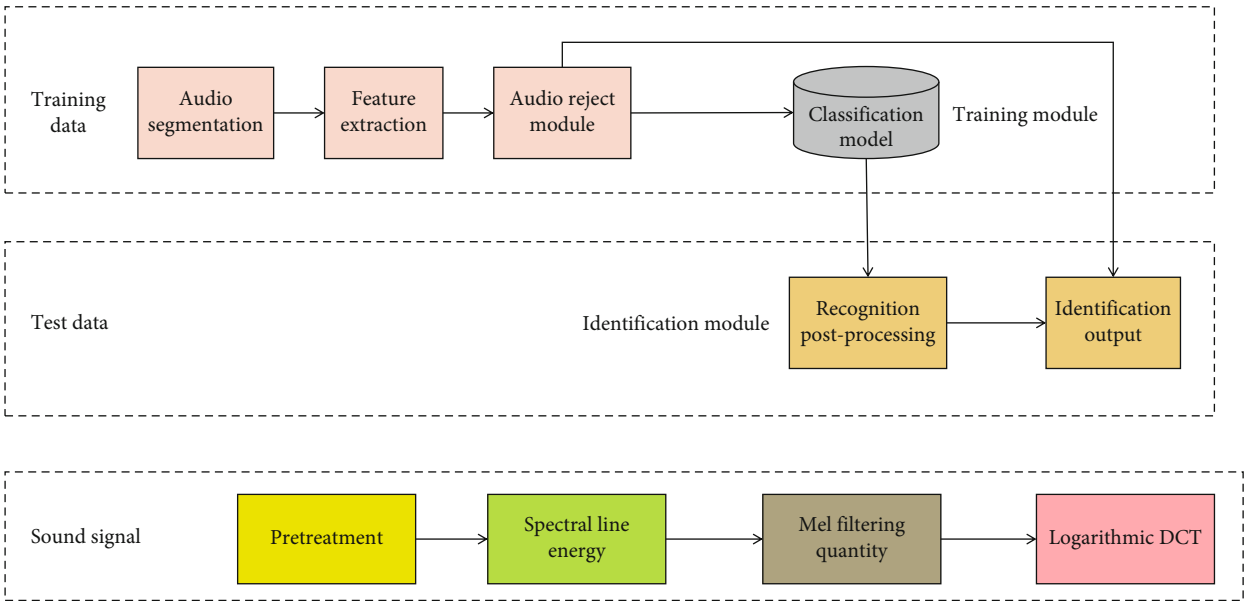

Figure 2: Flow chart of audio extraction system.

feature sets of different categories, the design of classifier should be reasonably constructed according to this point. Feature extraction: after windowing and framing the data, extract the frame-based audio features, and adjust the window length or carry out multiframe statistical calculation to obtain the long-term features of audio. In system optimization and postprocessing, smooth the corresponding recognition results and detect the reliability, so as to improve the recognition accuracy. In addition, the system further adds dimension reduction processing to reduce the required feature dimension and improve the recognition efficiency and accuracy of the system. The flowchart of audio extraction system is shown in Figure 2.

The linear predictive cepstrum coefficient is generally used to represent the spectral envelope information of the signal. The calculation process of audio cepstrum is roughly as follows: firstly, Fourier transform is performed on the input signal; then, the obtained signal spectrum is moduled; then, logarithm is taken, and finally, inverse Fourier transform is performed. According to the solution of linear prediction coefficient, because the frequency response $H(e^{j\omega})$ of the system reflects the frequency response of the vocal tract and the spectral envelope of the analyzed target signal, the linear prediction cepstrum coefficient of $\log|H(e^{j\omega})|$ can be calculated by inverse Fourier transform.

Assuming that $H(z) = 1/(1 - \sum^{p} a_i z^{-1})$ is the system function of the filter obtained through linear prediction analysis and its impulse response is $h(n)$, the method of calculating the cepstrum $h^{\wedge}(n)$ of $h(n)$ is as follows.

According to homomorphism, $H^{\wedge}(z) = \log H(z)$. Since $H(z)$ is analytical in the unit circle, it can be expanded into a series form as shown in the formula

$$\hat{H}(z) = \sum_{n=1}^{+\infty} \hat{h}(n) z^{-n}. \tag{1}$$

Let $h^{\wedge}(0) = 0$, take the derivative of z^{-1} on both sides of the formula at the same time, and get

$$\frac{\partial}{\partial z^{-1}} \lg \frac{1}{1 - \sum_{i=1}^{p} a_i z^{-1}} = \frac{\partial}{\partial z^{-1}} \sum_{n=1}^{+\infty} \hat{h}(n) z^{-n}. \tag{2}$$

Obtained after simplification

$$\left(1 - \sum_{i=1}^{p} a_i z^{-i}\right) \sum_{n=1}^{+\infty} n \hat{h}(n) z^{-n+1} = \sum_{i=1}^{+\infty} i a_i z^{-i+1}. \tag{3}$$

The power coefficients of z on both sides of the equal sign are equal, respectively, that is, the formula for calculating cepstrum $h^{\wedge}(n)$ from the prediction coefficient a_i is obtained.

$$\begin{cases} \hat{h}(1) = a_1, \\ \hat{h}(n) = a_n + \sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) a_i \hat{h}(n-i), 1 < n \le p, \\ \hat{h}(n) = \sum_{i=1}^{p} \left(1 - \frac{i}{n}\right) a_i \hat{h}(n-i), p < n. \end{cases} \tag{4}$$

Loudness, that is, the loudness of sound, reflects the energy of audio signal. The change of sound intensity and pitch will affect loudness. It can be calculated by the amplitude of the signal, and the calculation formula is

$$XD(g) = \sqrt{\frac{\sum_{n'=1}^{N} \left(x_g\left(n'\right)\right)^2}{N}}. \tag{5}$$

Expressed as the inner product of a square integrable function $\varphi(t)$ and a wavelet function $\psi(t)$, as shown in the formula.

$$\langle \varphi, \psi_{am,bm} \rangle = \frac{1}{\sqrt{am}} \int_{-\infty}^{\infty} \varphi(t) \psi^* \left(\frac{t - bm}{am} \right) dt, \tag{6}$$

where $\langle *, * \rangle$ represents the inner product, am is the scale factor, bm is the displacement factor, ψ^* represents the conjugate negative number of ψ. $\psi_{am,bm}(t)$ is the basis function, and its calculation formula is shown in formula.

$$\psi_{am,bm}(t) = \frac{1}{\sqrt{am}} \psi \left(\frac{t - bm}{am} \right). \tag{7}$$

It can be found that function $\psi_{am,bm}(t)$ is a cluster of functions generated by displacement and expansion of function $\psi(t)$.

If the linear prediction coefficient is known as a_i, the reflection coefficient KM_i can be obtained by the following recursive relation.

$$\begin{cases} a_j^{(p)} = a_j, j = 1, 2, \cdots, p, \\ KM_i = a_i^{(i)}, i = p, p-1, \ldots 1, \\ a_j^{(i-1)} = \dfrac{\left(a_j^{(i)} + a_j^{(i)} a_{i-j}^{(i)} \right)}{\left(1 - KM_i^2 \right)}, j = 0, 2, \cdots, i-1. \end{cases} \tag{8}$$

From the point of view of acoustic theory, the reflection coefficient KM_i can reflect the reflection of sound waves at the boundary of each acoustic tube segment to a certain extent, as shown in formula, where sa_i represents the area function of the acoustic tube in the i segment.

$$KM_i = \frac{sa_{i+1} - sa_i}{sa_{i+1} + sa_i}. \tag{9}$$

After further transformation, it can be concluded that

$$\frac{sa_i}{sa_{i+1}} = \frac{1 - KM_i}{1 + KM_i}. \tag{10}$$

It can be seen that the acoustic tube area of each segment can be obtained by knowing the reflection coefficient and the boundary acoustic tube area. According to different methods of parameter extraction, audio signal analysis can generally be divided into time domain analysis, frequency domain analysis, cepstrum domain analysis, and other domain analysis methods. According to different analysis methods, audio signal analysis can be divided into model analysis method and nonmodel analysis method.

4. Realization of Pharyngeal Training Method in Vocal Music Teaching

4.1. Vocal Music Teaching Pharyngeal Sound Training System Based on Audio Extraction. In the process of vocal music teaching, teachers should first have solid theoretical knowledge as guidance. At the same time, their own auditory system is also good, and they should have high musical talent. When teaching students vocal music, we can clearly observe the students' sense of music and their respective abilities. Vocal music teachers only have strong judgment and sharp understanding ability, can excavate students' strengths in time, and can use clear and clear language when expressing their views to students, so as to help students understand better. Based on the audio extraction, a vocal music teaching pharyngeal training system is constructed. After the audio collection of the teacher's singing and explanation and the better part completed in the students' classroom practice, the students can refer to the practice after class, which can deepen the students' understanding and consolidation of knowledge points and is conducive to the formation of the interaction between teachers and students. This kind of method is adopted by some students in the undergraduate class of musicology and has achieved obvious results. Students can better master the content of the class and carry out self-practice after class. The progress of learning vocal music has been greatly improved. In vocal music teaching, open the mouth and throat to practice, so as to solve the problems of weak throat function and weak chest resonance. Because the vocal cord is long in the throat chamber, when performing vocal practice in the throat training system, it is necessary to form the muscles of the throat, open the singing state of the throat and connect it with the chest, and make a sound with certain laryngopharyngeal resonance and chest resonance. This method can solve the problems of students' voice disconnection, make the upper and lower singing channels unobstructed, and achieve the stability and unity of singing state. Stage students' singing in the pharyngeal sound training system for audio acquisition and extraction, so that students can listen to their singing effect under the teacher's explanation. In this way, compared with the teaching means of only listening to the teacher's explanation and singing, students are more likely to find their own problems and deficiencies. More importantly, audio files can be circulated among students or uploaded to the network for communication, which can greatly improve students' enthusiasm in learning vocal music.

Table 1: Audio database.

Audio category	Total time/min	Number of audio tracks/audio segments
Pure vocal song	57	15
Electronic music	155	45
Ballad	174	40
Pop music	142	33
Light music	56	17
Rap	77	23
Beijing opera	92	24

Under the pharyngeal training system in vocal music teaching, each learner's range and sound quality and color are different, and some people think that this range and

Table 2: Statistics of original feature set.

Feature type	Characteristic term	Dimension	Sequence number of item in feature set
Time domain feature	Short-term energy	1	2
	Short-time zero crossing rate	1	1
	Short-term average amplitude difference	1	2
Frequency domain feature	Frequency domain energy	1	5
	Subband energy ratio	4	7-9
	Frame-based spectral centroid	1	11
Cepstrum domain feature	Mel cepstrum coefficient	12	28-40
	Linear prediction cepstrum coefficient	12	15-25
Other features	Loudness range	1	53
	Reflectance	1	58-70

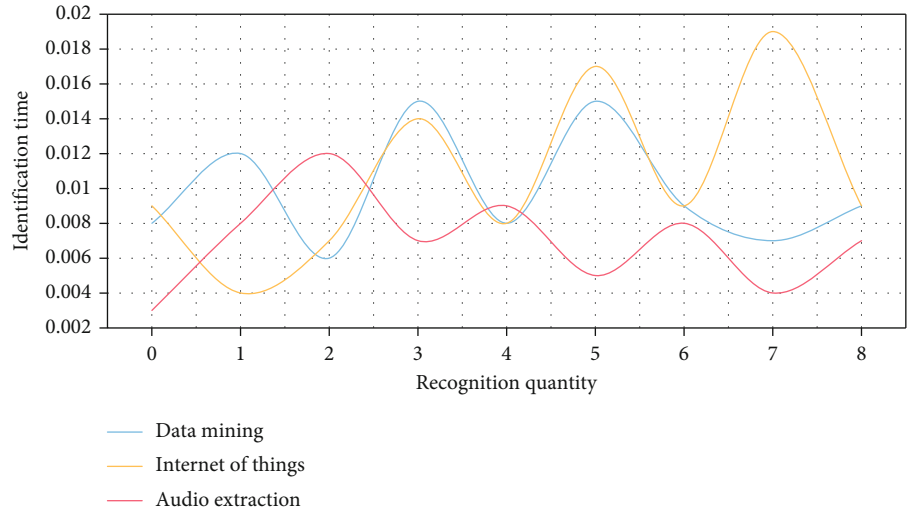

Figure 3: Experimental results of the first identification time.

sound quality and color are unique to everyone and cannot be changed by external interference factors. In fact, the result is just the opposite. If you master and apply pharyngeal vocalization skills through professional training and training, the result can be imagined that everyone's range and sound quality and color can be changed. If the vocal music teaching is reasonably applied with the pharyngeal training system, the audio is extracted, the vocal music teaching is visualized, the vocal music teaching is expanded, and a vocal music teaching evaluation model with music professional characteristics is formed, the above problems can be well alleviated. Open your mouth, you can practice the strength of the hypopharynx cavity and chest cavity resonance. Requirements: During practice, keep the chin and hyoid bone relaxed all the time. When the mouth is opened up, it must be triangular, and the corners of the mouth must be slightly upturned, just like the feeling that colleagues will smile when yawning. When the mouth is opened up, the hyoid bone must move up and back, and stay as still as possible during vocalization. The vocal music teaching should grasp the teaching principle of "teaching students in accordance with their aptitude and adapting to their needs" and find a simple and convenient way to guide students, so as to give students a good teaching impression. It is not necessary for vocal music teachers to be very successful and famous singers, but it is necessary to ensure the most correct singers, so as to use the correct methods to guide students in vocal music training.

4.2. Experimental Results and Analysis. In this paper, the audio classification experiment mainly includes three aspects: the classification of speech and music, the classification of music genres, and the classification of musical instruments. In the classification experiment of speech and music, the speech signal is pure vocal audio, and the music signal is the audio used in the classification experiment of music genres; the music genre classification experiment includes music audio of seven genres: pure vocal songs, electronic music, ballads, pop music, light music, rap, and Beijing opera. The audio database is shown in Table 1.

In the audio classification experiment, the audio segment features are generally divided into two types: one is the audio segment features based on audio frame, that is, the audio segment with a certain length is processed by frame, and the average value of the features of different frames is taken as the feature value of this segment of audio signal; the other

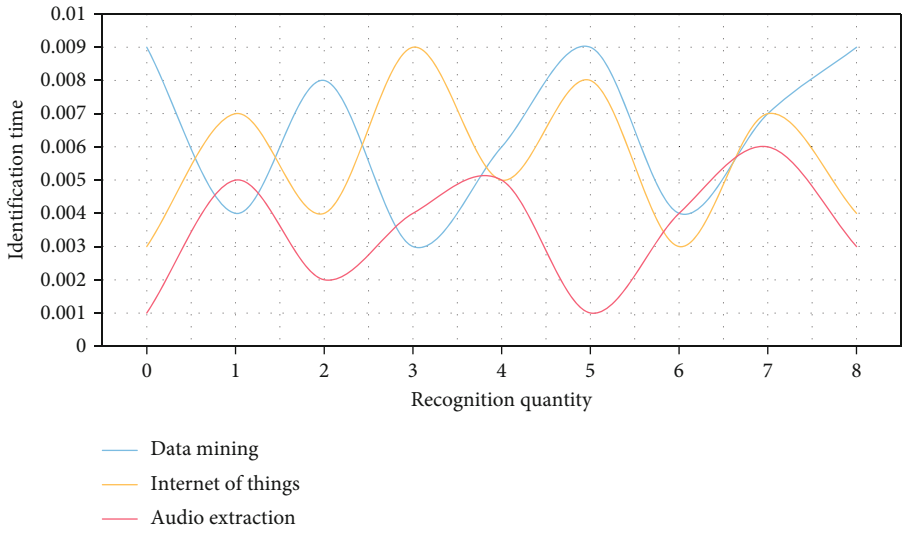

Figure 4: Experimental results of the second identification time.

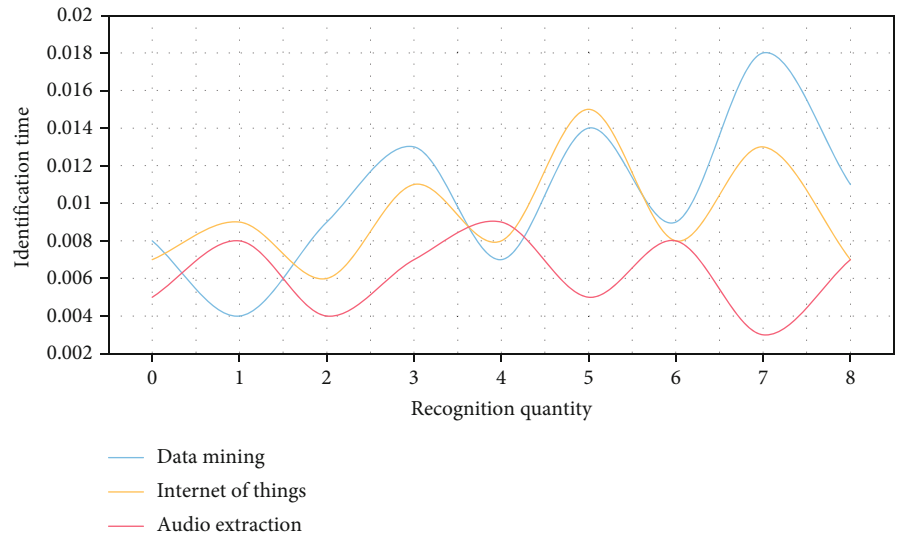

Figure 5: Experimental results of the third recognition time.

is the audio segment feature based on audio envelope, that is, directly extract the feature of the segmented audio segment, and take the feature mean of each segment of audio as the feature value of this segment of audio signal. The feature types include time domain features, frequency domain features, cepstrum domain features, and other features. The corresponding feature categories are 5, 8, 3, and 14, respectively. The total number of feature categories is 30 and the number of feature dimensions is 88. The original feature set of audio signal is shown in Table 2.

In order to ensure the rationality and scientificity of this experiment, this paper designs vocal music teaching pharyngeal training method based on data mining, vocal music teaching pharyngeal training method based on Internet of Things, and vocal music teaching pharyngeal training method based on audio extraction as the traditional comparison methods of the experiment and identifies the same music melody according to three methods. In the process of each method, the background computer will record the notes, and the staff will record the time spent by the three methods, respectively, and finally balance the processing time of the three methods and the effect of vocal music teaching pharyngeal training as the final experimental result. This experiment was compared for three times, and the experimental results are shown in Figure 3, Figure 4, and Figure 5.

Analysis of Figures 3–5 shows that different methods of vocal music teaching have different times for pharyngeal training. When the recognition amount is 3, the average recognition time of vocal music teaching pharyngeal training based on data mining is 0.010 seconds, the average recognition time of vocal music teaching pharyngeal training based on Internet of Things is 0.011 seconds, and the average recognition time of vocal music teaching pharyngeal training

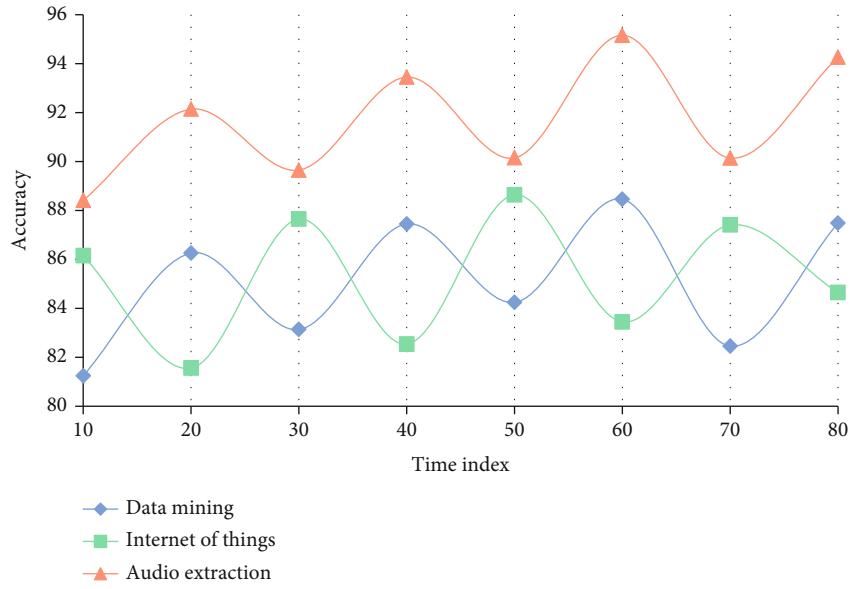

Figure 6: Experimental results of recognition accuracy.

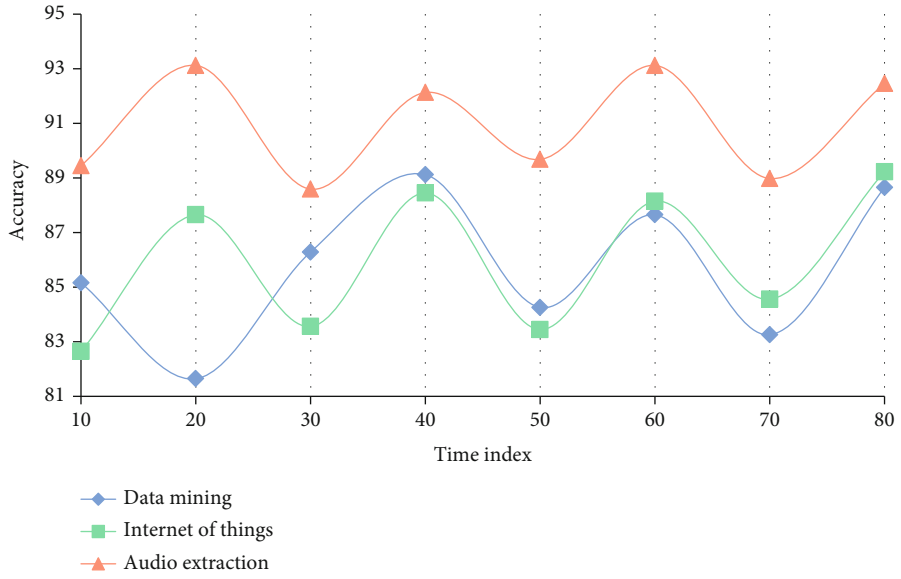

Figure 7: Experimental results of recognition accuracy.

based on audio extraction is 0.006 seconds. The recognition time of the audio extraction method is much shorter than that of the other two traditional methods, because the audio extraction method can perform segmented training according to the changing trend of physical characteristics of notes, effectively extract the characteristics of vocal music teaching pharyngeal training, and shorten the recognition time.

Based on the above contents, this paper has carried out experiments on the recognition accuracy, respectively. Similarly, the vocal music teaching pharyngeal training method based on data mining, the vocal music teaching pharyngeal training method based on the Internet of things, and the vocal music teaching pharyngeal training method based on audio extraction are used as the traditional comparison methods of the experiment. Two experiments are carried out to compare the recognition accuracy, respectively. The experimental results are shown in Figures 6 and 7.

According to the analysis of Figures 6 and 7, the final result of data summary is that the throat training method of vocal music teaching based on audio extraction studied in this paper not only takes less time than the traditional recognition method but also has better processing effect. Using pharyngeal method to control breath is a positive migration for breath. Compared with natural pronunciation methods, the exhaust volume of pharyngeal training is small, and the waste of breath will be appropriately reduced, so as to better control the singer's breath. If pharyngeal training is used in vocal music education for a long time, the singer's

abdominal muscles will have strong elasticity, and it will be handy in controlling its breath.

5. Conclusions

From the above, it can be seen that the "pharyngeal voice" practice method is actually a "labor-saving and healthy" practice method, which is simple to operate and has direct and obvious effect from the perspective of human physiological operation. Firstly, this paper analyzes the characteristics of vocal music teaching pharyngeal training and then further extracts and analyzes the characteristics of audio features. Finally, based on the identification technology of vocal music teaching pharyngeal training system and audio feature technology, it summarizes the identification workflow of vocal music teaching pharyngeal training based on audio feature technology. The application of "pharyngeal training method" in vocal music teaching in normal universities is beneficial to carry forward national culture, develop national vocal music, train students to establish unique sound concepts, and adapt to the diversified development of today's society. Different methods of vocal music teaching and pharyngeal training have different times. When the recognition amount is 3, the average recognition time of vocal music teaching pharyngeal training based on data mining is 0.010 seconds, the average recognition time of vocal music teaching pharyngeal training based on Internet of Things is 0.011 seconds, and the average recognition time of vocal music teaching pharyngeal training based on audio extraction is 0.006 seconds. Based on audio extraction technology, pharyngeal training is introduced into vocal music teaching in China, which complements traditional vocal music teaching methods. It not only enhances the effect of vocal music teaching in China but also promotes the research of "pharyngeal" teaching theory, and at the same time lays a solid foundation for improving Chinese vocal music teaching methods and research on teaching theory. It is hoped that these practice methods can arouse the deep research and attention of singing artists and theorists on "pharyngeal singing" in order to better apply it to vocal music learning and practice.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

No competing interests exist concerning this study.

Acknowledgments

This work was supported by Project Name: Special Research Project of Philosophy and Social Sciences in Colleges and Universities in the 13th Five-Year Plan of Guangdong Province. The function construction and Practice Research of "pharyngeal technique" in singing training system. Approval No.: 2019GXJk179.

References

[1] Y. Zhang, "Enlightenment on vocal music classroom teaching from the perspective of neuroscience," *NeuroQuantology*, vol. 16, no. 6, pp. 132–137, 2018.

[2] X. Zhao, "Application of situational cognition theory in teaching of vocal music performance," *NeuroQuantology*, vol. 16, no. 6, pp. 308–313, 2018.

[3] M. L. Doherty and M. van Mersbergen, "Personal and professional characteristics of music educators: one size does not fit all," *Journal of Voice*, vol. 31, no. 1, pp. 128.e7–128.e13, 2017.

[4] L. Havrilova and N. Voronova, "Online training course for teaching master-degree music students," *Science & Education*, vol. 26, no. 6, pp. 100–105, 2017.

[5] E. U. Grillo, "A nonrandomized trial for student teachers of an in-person and Telepractice global voice prevention and therapy model with Estill Voice Training assessed by the VoiceEvalU8 app," *American Journal of Speech-Language Pathology*, vol. 30, no. 2, pp. 566–583, 2021.

[6] C. Long and S. Wang, "Music classroom assistant teaching system based on intelligent speech recognition," *Journal of Intelligent and Fuzzy Systems*, vol. 14, pp. 1–10, 2021.

[7] A. Remacle, S. Bouchard, A. M. Etienne, M. C. Rivard, and D. Morsomme, "A virtual classroom can elicit teachers' speech characteristics: evidence from acoustic measurements during in vivo and in virtuo lessons, compared to a free speech control situation," *Virtual Reality*, vol. 25, no. 4, pp. 935–944, 2021.

[8] V. I. Jónsdóttir, "Cordless amplifying system in classrooms. A descriptive study of teachers' and students' opinions," *Logopedics, phoniatrics, vocology*, vol. 27, no. 1, pp. 29–36, 2002.

[9] D. Wang, "Discussion on the construction of multi-style singing teaching mode of vocal music major in colleges and universities, pp. 1083–1087, 2019.

[10] X. Ling, "Training on musical literacy of students in vocal teaching in universities, pp. 2373–2377, 2019.

[11] C. Jia, "Cultivation and application of singing thinking in vocal music teaching in normal colleges, pp. 1127–1131, 2019.

[12] R. V. Ventura, R. Z. Lopes, L. T. Andrietta et al., "402 audio information retrieval for describing gait patterns in Brazilian horses," *Journal of animal science*, vol. 98, Supplement_4, p. 27, 2020.

[13] K. Arumugam, "Feature selection based on Mbfoa for audio signal classification under consideration of Gaussian white noise," *IET Signal Processing*, vol. 12, no. 6, pp. 777–785, 2018.

[14] Q. Zhiguo, H. Huangxing, and L. Tao, "Novel quantum watermarking algorithm based on improved least significant qubit modification for quantum audio," *Chinese Physics B*, vol. 27, no. 1, p. 010306, 2018.

[15] Y. Ma, Y. Hao, M. Chen, J. Chen, P. Lu, and A. Košir, "Audio visual emotional fusion: a deep and effective weighting method," *Information fusion*, vol. 46, pp. 184–192, 2019.

[16] L. Wu, L. Hao, Y. Bo et al., "Adaptive localization method of audio binaural speech extraction source based on multi-scale extended convolution network," *Complexity*no. 3, 7 pages.

[17] S. Craciun, Kirchgessner, A. D. George et al., "A real-time and energy-saving vocal music teaching system," *Real time vocal music teaching*, vol. 14, no. 2, pp. 379–394, 2018.

[18] D. Lianxiu, "Research on the innovation of vocal teaching course based on multicultural perspective, pp. 563–567, 2019.

[19] J. Sun, "Research on model innovations of vocal teaching under the background of Internet plus, pp. 389–393, 2018.

[20] L. Shan, "The status quo and development ways of vocal music teaching in colleges, pp. 102–106, 2018.

[21] B. Jin, "Study on the application of the idea "teaching students in accordance of their aptitude" to vocal teaching in colleges and universities, pp. 603–608, 2016.

[22] Q. Lang, "Application of psychological regulation ability in vocal teaching, pp. 120–124, 2016.

[23] X. Yu, "Application of performance assessment in the vocal teaching of normal universities, pp. 1326–1329, 2016.

[24] X. Yu, "A study on current status of vocal teaching by view of the vocal teaching papers summary from year 2000 to year 2014 a case study of the core music journals, pp. 1330–1333, 2016.

[25] B. Nick, E. Julie, and S. Gary, *Making Music in the Primary School:Whole Class Instrumental and Vocal Teaching*, Taylor and NNN, 2020.