Leveraging Intelligent Big Data Technology for Optimizing Vocal Singing Training based on Embedded Systems Framework

Zhigang Ding[1*] and Beibei Zhang[2]

[1]College of Marxism, Anhui Public Security College, Hefei, Anhui, 230000, China
Anhui, 230000, China, Dzg998996@163.com
[2]Department of Science and Technology, Anhui Institute of Public Security Education, Hefei, China, BeibeiZhang@outlook.com

Corresponding author: Zhigang Ding, Dzg998996@163.com

Abstract. In order to improve the effect of vocal music singing training, this paper studies the waveform intelligent analysis of the vocal music singing process, and uses the vector wave research method to carry out the intelligent identification and analysis of the vocal music singing waveform. In order to realize the real-time positioning of vocal input notes in the musical score, this paper adopts the dynamic programming score tracking algorithm based on the extended cave. In addition, this paper constructs a vocal singing training system based on intelligent big data technology. After constructing the intelligent training system, this paper evaluates the effect of vocal singing training method based on intelligent big data technology. It can be seen from the research that the vocal singing training method based on intelligent big data technology proposed in this paper can play a good auxiliary role in modern vocal music training.

Keywords: intelligent big data; vocal singing; training method; optimization: Embedded Systems Framework
DOI: https://doi.org/10.14733/cadaps.2024.S8.196-211

1 INTRODUCTION

If music is likened to a language, then the phrases are like the pieces of poetry that make up the music. Chinese classical poetry genres have their own creative combination forms, such as five-character rhythm poems, where five characters form a sentence, and each sentence has its own meaning. If we interpret the structure of the verse, it will not answer the natural poetry. Moreover, like appreciating a poem, in order to fully understand the author's thoughts, it is necessary to understand the meaning of each verse. The phrases in music are like this, and the phrases are the logical arrangement of the composer's overall emotion. Therefore, in order to sing a work that clearly expresses the author's thoughts, we need to deal with the phrases, understand the breathing, division, and ups and downs of the phrases, and clarify the logical relationship between the phrases. In our commonly used vocal music etude textbooks, most of the works are simple and clear in terms of notation, and are marked with clauses connecting lines, which is convenient for practice. Beginning practitioners can establish basic phrases and phrases concepts through repeated practice under the principle of following the notation requirements, and master the breathing points between phrases. Students with basic knowledge should have a deeper understanding of the phrases, and not stick to small phrases so as to break the musical integrity of the work. At the same time, it is necessary to pay attention to the coherence of the phrases, and to look at the relationship between the phrases from a long-term perspective, which helps to grasp the overall style of the work and make the phrases have direction. For example, as the mood of the music changes, phrases include rising mood, crescendo ascending, low mood or contracting, weakening downside, narrative, eloquent parallel forms, and even emotional rest forms. These phrases appear in vocal etudes, and the practice should follow the notation requirements and sing strictly according to the score to achieve the real purpose of practice.

Vowels, phrases, technical training and accompaniment cooperation awareness are all for the improvement of the final score expression. As a student, it is not easy to express the requirements of the score map in place. Moreover, doing the above points is not easy. It requires the practitioners to practice patiently, as well as long-term learning and practice accumulation, so that the singers can sing with ease. In other words, it requires learners to sing a large number of works, learn from the works, accumulate and become familiar with various musical styles, strengthen the memory of various musical styles, and make them memorize by heart. In singing, when encountering the practiced musical motives or phrases, it is easy to handle the musical emotional expression in place. Moreover, as students, they should start from the study of charts and pay attention to the musical terms and emojis of charts. Our commonly used vocal music etudes are carefully arranged by famous vocal music educators. The works have rich musical styles, which are both interesting and technical in music. Moreover, each etude has a different mode key. As the difficulty increases, the tonality changes several times in an etude, along with complex rhythm patterns, which all require repeated practice to sing well. Composers are marked with musical terms and emoticons in each work, which helps us learn to understand the composer's creative intentions and thoughts, and also requires students to sing according to the requirements of these musical terms and emoticons. When the students learn to sing according to the requirements of the music chart, the students can improve their musical literacy while consolidating their sound foundation, so as to achieve the purpose of making the sound technology practically serve the music, and achieve a perfect artistic effect. During the learning process of vocal etudes, Professor Xianyu Yuege asked the students to translate and read the musical terms on the notation before class, and use a metronome to proofread the tempo of the work. During the study period, it is necessary to check the recitation of musical terms from time to time, and it is required not to ignore every emoticon, and to sing strictly according to the requirements of the music chart. For a long time, it has cultivated students' attention to the expression of music charts, thereby improving the singing level. Furthermore, Embedded Systems can be employed in live performance systems for real-time audio processing, mixing, and synchronization. These systems can enable seamless integration of multiple audio sources, automated sound adjustments, and synchronization with visual elements or other performance aspects.

This paper studies the optimization of vocal music singing training method combined with intelligent big data technology, and improves the vocal music singing effect through the intelligent vocal music singing method.

2 RELATED WORK

Literature [10] gives the concept of vocal audio moments and the construction of the audio matrix (7 moment invariants). The audio matrix is calculated by the structural formula for the target vocal audio. When the vocal audio changes in translation, rotation and scaling, the seven invariants of the audio matrix maintain relatively unchanged characteristics, but the anti-noise ability of the audio matrix is poor. If the vocal audio moment is analysed from a mathematical point of view, then the vocal audio moment can be understood as the projection of the vocal audio function f(x, y) on its basis function system, and the basis function system used by Hu when constructing the audio matrix is a non-orthogonal function It is known from the orthogonal theory that it is difficult to restore the original vocal audio using the vocal audio matrix. The invariants of complex moments and rotation moments are invariant to translation, rotation and scaling, but the kernel functions of complex moments and rotation moments are also not orthogonal, so they are not suitable for reconstructing vocal audio [3]. Literature [16] uses orthogonal polynomials to replace the conventional transformation kernel of geometric moments, firstly proposes the concept of orthogonal vocal audio moments, and gives Legendre and Zernike two orthogonal vocal audio moments, both of which can be used. To reconstruct the original vocal audio, the main difference between the two is that the moment invariant of the Legendre moment has no invariant characteristics to the vocal audio rotation, but has invariant characteristics to the vocal audio translation and scaling, while the moment invariant of the audio matrix has the vocal audio frequency invariant. The characteristics of audio translation, rotation, and scaling are invariant; the literature [14] proposes the complex moment, and derives the complex moment invariant; the literature [9] analyses the audio matrix, complex moment, Rotation moment, audio matrix and Legendre moment are evaluated, and the results show that the audio matrix has the best performance among these moments. Later scholars have carried out a lot of research on the fast calculation of audio matrix, including recursive method [12], coefficient method [2] and so on. Literature [7] conducted an error analysis on the traditional calculation method of Zernike moments, and it was confirmed that there is an integral discretization calculation error of transformation; Literature [8] proposed a polar coordinate-based audio matrix definition and calculation based on the analysis of these two error sources. The simulation results show that the reconstruction effect is better than the calculation method in the Cartesian coordinate system. In addition, many other optimizations and solutions have been proposed for the calculation of the audio matrix [6].

The saliency detection method based on the spatiotemporal domain is to fuse the saliency detection results in the spatial domain with the saliency detection results in the time domain, and finally generate the detection results in the spatiotemporal domain. This class of methods captures information in both the spatial and temporal domains of a task by learning a sequence of objects or behaviors of interest. For example, through neural network, support vector machine [19] and other machine learning methods, the learning of the mapping relationship between vocal audio-visual salient features and artificially marked saliency is completed, thereby generating a saliency detection model. The advantage of this machine learning-based saliency detection method is that it does not require prior assumptions about vocal audio features, but it is computationally less efficient. Literature [15] uses a regression classifier to learn the dependencies between a given scene and auditory attention points; Literature [18] uses two vocal audios marked saliency by the user for training, and generates machine learning-based joint saliency. Detection model. Literature [11] is based on the data of people's auditory area of vocal audio, and uses different levels of vocal audio features to train support vector machines, so as to predict human detection of vocal audio auditory.

The saliency detection method based on the frequency domain is to use vocal audio transformation, transfer the components of a certain color mode of vocal audio from the spatial domain to the frequency domain, and use the vocal audio features in the transformed domain to complete the saliency detection. Literature [4] found that the spectrum of similar areas in the original vocal audio is similar, and the time-space domain vocal audio corresponding to the abnormal part of the spectrum is the significant area of the original vocal audio. Operations such as searching for abnormal spectral components and inverse Fourier transform can obtain the corresponding salient positions in the space-time domain. Literature [1] uses the redundancy and anomalous components of the phase spectrum to find the abrupt part of the vocal audio in the spatiotemporal domain, and achieves saliency detection. Literature [13] selects the sign of the impulse cosine transform coefficient as the visual feature for the saliency calculation of the target vocal audio; Literature [17] uses the absolute value of the discrete cosine transform coefficient to calculate the saliency of the target vocal audio.

According to the way the computer processes the machine auditory information, the modelling of saliency detection can be divided into bottom-up and top-down [20]. The bottom-up detection method is mainly driven by target data, while the top-down detection method is mainly driven by the purpose and task of detection [5]. In general, when the application scenario of the visual saliency model does not have prior knowledge, the bottom-up visual saliency model is usually used in combination with the basic characteristics of the target vocal audio; on the contrary, in the application scenario with prior knowledge, it is often used. A top-down visual saliency model.

3 VOCAL AUDIO WAVE DETECTION ALGORITHM

The wave equation for vocal audio waves is derived from Maxwell's equations. First, the following classical form of Maxwell's equations is observed:

$$\nabla \times E = -\mu \frac{\partial H}{\partial t} \tag{1}$$

$$\nabla \times H = J + \varepsilon \frac{\partial H}{\partial t} \tag{2}$$

$$\nabla \bullet H = 0 \tag{3}$$

$$\nabla \bullet E = \frac{\rho}{\varepsilon} \tag{4}$$

The algorithm takes the curl on both sides of formula (1), and uses the vector identity $\nabla \times \nabla \times E = \nabla(\nabla \bullet \mathrm{E}) - \nabla^2 E$ to get:

$$\nabla(\nabla \bullet E) - \nabla^2 E = -\mu \frac{\partial}{\partial t}(\nabla \times H) \tag{5}$$

The algorithm then substitutes formulas (2) and (4) into the above formula to get:

$$\nabla^2 E - \mu\varepsilon \frac{\partial^2 E}{\partial t^2} = -\mu \frac{\partial J}{\partial t} + \frac{\nabla \rho}{\varepsilon} \tag{6}$$

Similarly, the algorithm takes the curl on both sides of (2) and makes use of the vector identities. Then, there is:

$$\nabla(\nabla \bullet \mathrm{H}) - \nabla^2 H = \nabla \times J + \varepsilon \frac{\partial}{\partial t}(\nabla \times \mathrm{E}) \tag{7}$$

By substituting formula (1)(3), we get:

$$\nabla^2 H - \mu\varepsilon \frac{\partial^2 H}{\partial t^2} = -\nabla \times J \tag{8}$$

Formulas (6) and (8) are the active vector wave equations satisfied by the instantaneous vector E and the sum, or are called non-secondary vector wave equations. In particular, in the passive region, the following homogeneous vector wave equation can be obtained:

$$\nabla^2 E - \mu\varepsilon \frac{\partial^2 E}{\partial t^2} = 0 \tag{9}$$

$$\nabla^2 H - \mu\varepsilon \frac{\partial^2 H}{\partial t^2} = 0 \tag{10}$$

For time-harmonic vocal audio waves, the above equation satisfies the following homogeneous vector Helmholtz equation:

$$\nabla^2 E + k^2 E = 0 \tag{11}$$

$$\nabla^2 H + k^2 H = 0 \tag{12}$$

Among them, the wavenumber is $k = \omega\sqrt{\mu\varepsilon}$. Vocal audio waves that satisfy the above vector equations can be called "vector waves". Further, if the polarization direction of the vocal audio wave is unified, for example, the polarization direction of the sound wave field is polarized along the x direction of the Cartesian coordinate system, the wave equation of the sound wave field can be simplified to the scalar Helmholtz equation of the following form:

$$\nabla^2 E_x + k^2 E_x = 0 \tag{13}$$

Vocal audio waves that satisfy such an equation can be called "scalar waves". Scalar waves naturally also satisfy the vector wave equation.

The scalar wave is equivalent to moving the vector characteristics out of the equation, so that only the amplitude and phase distribution characteristics are concerned, and a scalar solution that is only related to the amplitude and phase is obtained.

Of course, vocal audio waves are essentially vector, and scalar waves are just a particular solution or component solution. Due to the vector characteristics of vocal audio waves, when solving, it has to go through two "decompositions". One is about the decomposition of the vector, and the other is about the decomposition of the amplitude and phase. The so-called vector decomposition is to project the polarization direction of the vocal audio wave in the reference coordinate system (rectangular coordinate system, cylindrical coordinate, spherical coordinate). For example, it can be decomposed into $(\mathrm{E}_x, \mathrm{E}_y, \mathrm{E}_Z)$ in the Cartesian coordinate system, which turns the vector equation into a scalar equation. Then, when using the separation variable method to solve the scalar equation, it is necessary to decompose the amplitude and phase again. For example, the amplitude and phase components of Ex can be expanded in a cylindrical coordinate system: $E_x(\rho, \varphi, \mathrm{z})$.

It is worth emphasizing that in this paper, when unifying the polarization directions of scalar waves (vector decomposition), the Cartesian coordinate system is used, because the direction of the base coordinate vector of the cylindrical coordinate system or the spherical coordinate system itself changes with space. That is to say, the scalar wave is not only the solution of the H scalar wave equation, but also emphasizes that its sound wave vector decomposition is completed in the Cartesian coordinate system. Therefore, the component solution in the rectangular coordinate system is also called the scalar modulus, and the component solution in the cylindrical or spherical coordinate system is called the vector modulus (which is also the solution of the scalar equation).

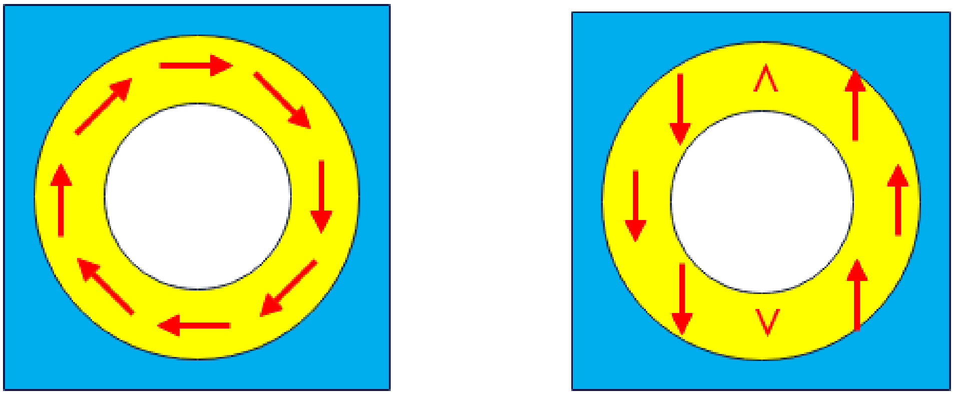
Since the polarization mainly studied in this paper is circular polarization, for the convenience of expression, the Jones vector is added here. The Jones vector is a vector (column vector) that can be used to unify circularly polarized waves. For a uniform circularly polarized planar vocal audio wave propagating in the +z direction, the Jones vector can be expressed as:

$$E_{LHCP}=\begin{bmatrix}E_x\\E_y\end{bmatrix}=E_0\begin{bmatrix}1\\j\end{bmatrix}(\text{Left-handed}) \tag{14}$$

$$E_{RHCP}=\begin{bmatrix}E_x\\E_y\end{bmatrix}=E_0\begin{bmatrix}1\\-j\end{bmatrix}(\text{Right-handed}) \tag{15}$$

Whether it is a vector vortex wave or a scalar vortex wave, there is a common feature of the wave pattern that there is an energy singularity at the axis of the propagation direction. The energy singularity of the vector vortex wave is formed due to the superposition of its sound wave vector at the axis center to zero. The characteristics of such a beam can be characterized by a polarization factor: $\alpha(\varphi)=\mathrm{m}\varphi+\alpha_0$, where m is the polarization angle, φ is the azimuth angle, and is the polarization angle at $\varphi=0$.

These sound wave vectors are of equal amplitude on each concentric circle, but the polarization directions are different. The amplitude distribution of the sound wave vectors superimposed and canceled to form a hollow doughnut shape is shown in Figure 1(a):



(a) (b)

Figure 1: Schematic Diagram of the Amplitude Distribution of the Sound Wave Field: (a) Schematic Diagram of the Sound Vector and Sound Field Amplitude Distribution of the Vector Vortex Wave, (b)

Schematic Diagram of the Sound Vector and Sound Field Amplitude Distribution of the Scalar Vortex Wave.

The energy singularity at the center of the main axis of the scalar vortex wave is formed by the superposition of the same polarized sound wave vector to zero due to the difference in phase. A typical indicator is that they have an azimuth-dependent spatial phase factor: $e^{jl\varphi}$, where l is an integer. The energy flow of this beam spirals along the central axis and has the characteristics of orbital angular momentum, so it is also called orbital angular momentum beam. The sound wave receiving system finally designed can be called the orbital angular momentum sound wave receiving system. Taking linearly polarized waves as an example, a possible instantaneous amplitude and phase distribution of an orbital angular momentum beam is shown in Figure 1(b):

It can be seen that at the same time due to the different spatial phase factors, the sound wave vector has different directions and sizes, so that the superposition is zero at the center of the main axis.

The scalar OAM excitation model is a model that performs phase control on the basis of the original scalar wave model to generate OAM beams. For example, a uniform plane wave (scalar wave) is passed through (transmitted or reflected) a phase modulation surface to generate a helical phase, as shown in Figure 2.

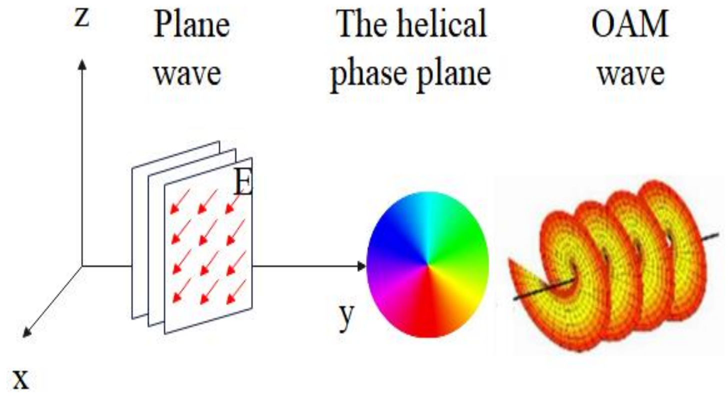


Figure 2: Schematic Diagram of the Scalar Analysis Model.

On the one hand, the helical phase surface can be controlled by a wavefront device, which is called the scalar model 1 here. This model can be considered as a sampling of the phase plane of the helical wave front. Similar to the sampling theorem, the number of OAM modes of the UCA sound wave receiving system has a certain limit relationship with the number of array elements: $|l_{OAM}| < N/2$. As a quasi-scalar system, all units of UCA have the same polarization direction, as shown in Figure 3. The scalar model is the most straightforward, and modern meta-surface technology and array technology also allow people to precisely tune the vocal audio wavefront to generate orbital angular momentum beams of any polarization and any OAM mode. The disadvantage of this model is that the precise control of the wavefront often requires a larger wavefront control surface or a large array. This is complex and time-consuming to design simulations, such as the large number of elements of the meta-surface, and the complex feed network of the UCA.

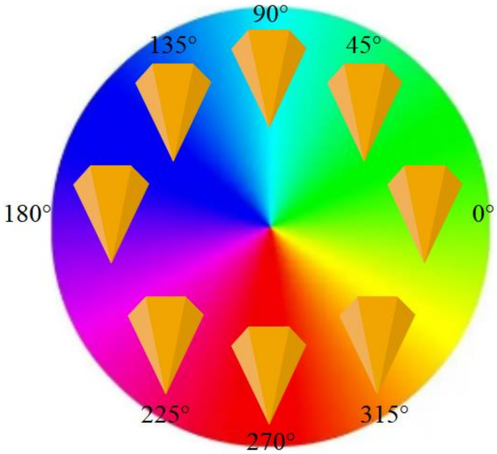


Figure 3: Schematic Diagram of UCA.

The analytical model shown in Figure 4 below is the toroidal phase gradient dipole model:

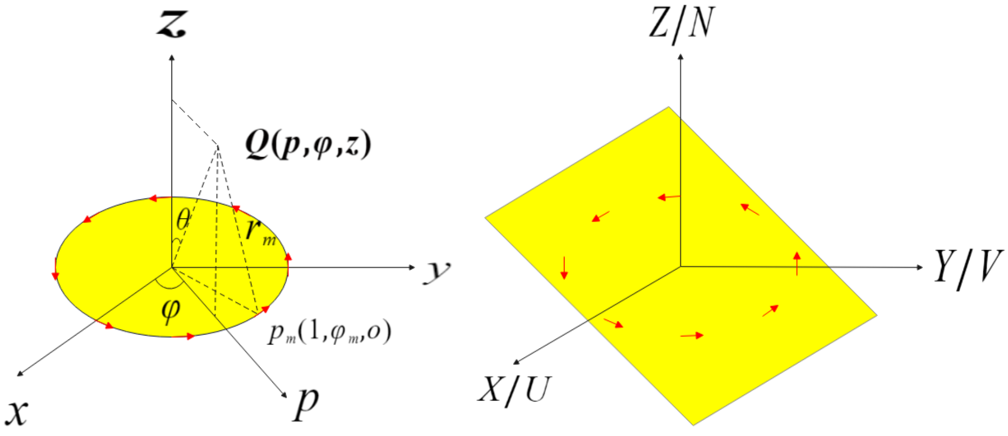


Figure 4: Vector Analysis Model and its FEKO Modeling Diagram.

There is a unit circle with a radius of 1 (R=1) placed on the xoy plane, and electric dipoles with the same modulus value of the polarized electric dipole moment in the φ direction are evenly placed on the unit circle along the circumference. The m-th element has the corresponding initial phase $e^{jl\varphi m}$, the electric dipole moment of each dipole in the cylindrical coordinate system can be expressed as $P_m = \hat{\varphi} p_0 e^{jl\varphi m}$, and P0 is the modulus value of the electric dipole moment. The position of the observation point is Q(ρ,φ,z), and the distance of the m-th dipole element from the observation point is rm. If the number of dipole elements is assumed to be q, the constants v=2πR/λ and A=P0/(4πε0) are introduced, then the total superimposed sound wave field of these electric dipoles at the observation point Q is:

$$E_l(\rho,\varphi,\mathrm{z}) * \left[(\frac{v^2}{r_m} - \mathrm{j}\frac{v}{r_m^2} - \frac{1}{r_m^3})(\hat{\mathrm{r}}_m \times \hat{\varphi}_m) \times \hat{\mathrm{r}}_m - (\mathrm{j}\frac{2v}{r_m^2} + \frac{2}{r_m^3})(\hat{\mathrm{r}}_m \bullet \hat{\varphi}_m) \bullet \hat{r}_m \right] = \frac{A}{R^3} \sum_{m=1}^{q} \exp(\mathrm{jvr}_m)\, \mathrm{e}^{jl\varphi_m} \tag{16}$$

The vector mode components in the cylindrical coordinate system (both polarization components and amplitude and phase components are expanded in cylindrical coordinates) are:

$$E_{\rho,1}(\rho,\varphi,\mathrm{z}) = \mathrm{j}^l \frac{Alq}{R^3} \frac{v}{\rho} \phi(\rho,\mathrm{z})\, \mathrm{e}^{jl\varphi} J_l \tag{17}$$

$$E_{\varphi,l}(\rho,\varphi,\mathrm{z}) = \mathrm{j}^{l-1} \frac{Aq}{R^3} \frac{v^2}{z} \phi(\rho,\mathrm{z})\, \mathrm{e}^{jl\varphi} J_l \tag{18}$$

$$E_{z,l}(\rho,\varphi,\mathrm{z}) = -\mathrm{j}^l \frac{Alq}{R^3} \frac{v}{z} \phi(\rho,\mathrm{z})\, \mathrm{e}^{jl\varphi} J_l \tag{19}$$

Among them, $\phi(\rho,\mathrm{z}) = \exp\left\{ jv\left[\mathrm{z} + (\rho^2+1)/2\mathrm{z} \right] \right\}$ is the propagation phase factor related to ρ, z, which is related to the spherical wavefront. In addition, $\theta = \tan^{-1}(\rho, z), \mathrm{J}_i = J_i(-\mathrm{vtan}\,\theta)$, and J_{l}、 $\mathrm{J}_{\mathrm{l}}^{'}$ are the l-th order Bessel functions of the first kind and their derivatives, respectively.

This model defines that the wave propagates along +z, so the $\mathrm{E_Z}$ component will decrease drastically as the wave propagates (will be much smaller than the lateral components E_ρ and E_φ). Therefore, the focus here is on its lateral component. It can be found that the E_ρ component leads the E_φ component by 90°. When the pitch angle θ is small, the magnitudes of the two moduli are similar, so it can be first determined that this is a left-handed circularly polarized wave. In addition, it can be found that both carry the same spatial phase factor $\mathrm{e}^{jl\varphi}$, but this phase factor follows the column vector component. In order to observe its orbital angular momentum characteristics, the transverse sound wave field is represented by a scalar wave with the Jones vector to observe the situation of the spatial phase factor on the scalar wave component:

$$E_{T,l} = \begin{bmatrix} E_{x,l} \\ E_{y,l} \end{bmatrix} = -j\frac{Aq}{R^3}\frac{v^2}{2z}\phi(\rho,\mathrm{z})\left\{ J_{l-1}\exp\left[j(\mathrm{l}-1)\varphi \right] \begin{bmatrix} 1 \\ j \end{bmatrix} + J_{l+1}\exp\left[j(\mathrm{l}+1)\varphi \right] \begin{bmatrix} 1 \\ -j \end{bmatrix} \right\} \tag{20}$$

When l>0, the value of the l+1-order Bessel function in the above formula is much smaller than the l-1-order Bessel function. Therefore, it can be approximately considered that the system produces a left-handed circularly polarized wave carrying the orbital angular momentum of l-1 topological charge number. When l<0, the opposite is true, that is, a right-handed circularly polarized wave carrying the orbital angular momentum of l+1 topological charge is generated. This model can generate OAM beams of any number of orbital angular momentum. However, there is a limitation that when the OAM topological charge is positive, it can only be left-handed circularly polarized waves, and when the OAM topological charge is negative, it can only be right-handed circularly polarized waves.

At this time, the dialectical relationship between vector vortex and scalar vortex can be properly discussed. This analytical model forms a hollow beam when the OAM topological charge is not equal to zero (l>1 or l<-1). This hollow beam can be regarded as either a vector vortex beam formed by superposition of equal-amplitude vectors to zero, or a scalar vortex beam formed by superposition of circularly polarized waves due to different phases. That is to say, the circularly polarized orbital angular momentum beam has both the characteristics of scalar vortex waves and the characteristics of vector vortex waves, and has rich electromagnetic diversity (EM diversity).

For an ideal circularly polarized plane wave (|Ex|=|Ey|), its polarization decomposition in the rectangular coordinate system and the cylindrical coordinate system has the following relationship:

$$E_x \pm jE_y = (\cos\varphi \mathrm{E}_\rho - \sin\varphi \mathrm{E}_\varphi) \pm \mathrm{j}(\sin\varphi \mathrm{E}_\rho + \cos\varphi E_\varphi) = \mathrm{e}^{\pm j\varphi}(\mathrm{E}_\rho \pm jE_\varphi) \tag{21}$$

This shows that when the circularly polarized wave in the rectangular coordinate system is represented by the cylindrical coordinate system, there will be an additional spatial phase factor $e^{\pm j\varphi}$ related to the azimuth angle. ±1 of the exponential term of this phase factor can be considered as the topological charge of the SAM s: $e^{js\varphi}$ (left-handed s=+1, right-handed s=-1). According to the expression of vector model 1, it can be summarized as follows:

When it is known that a bunch of vocal audio waves contain both SAM and OAM, the l in the exponential term of the phase factor carried by the transverse vector modulo component (such as E_ρ or E_φ) represented by its cylindrical coordinates will represent its total topological charge (total angular momentum) $l = l_{tot} = s + l_{OAM}$. For example, in the vector model 1, when $l = l_{tot} = 1$, due to the left-handed circular polarization $s = 1$, the orbital angular momentum number is 0($l_{OAM} = l_{tot} - s = 0$), that is, the vector model at this time simply produces a left-handed circularly polarized wave.

When the algorithm goes back to formula (21), by multiplying the exponential term to the left, the formula is written as:

$$E_\rho \pm jE_\varphi = e^{\mp j\varphi}(\mathrm{E}_x \pm \mathrm{jE}_y) \tag{22}$$

The right side of the equation represents the circular polarization in Cartesian coordinates, while also carrying the spatial phase factor $e^{\mp j\varphi}$. This shows that ideal circularly polarized waves (propagating in the z direction) in cylindrical coordinates naturally carry orbital angular momentum. The circularly polarized wave generated in the cylindrical coordinate system can use the analytical model shown in Figure 5 below:

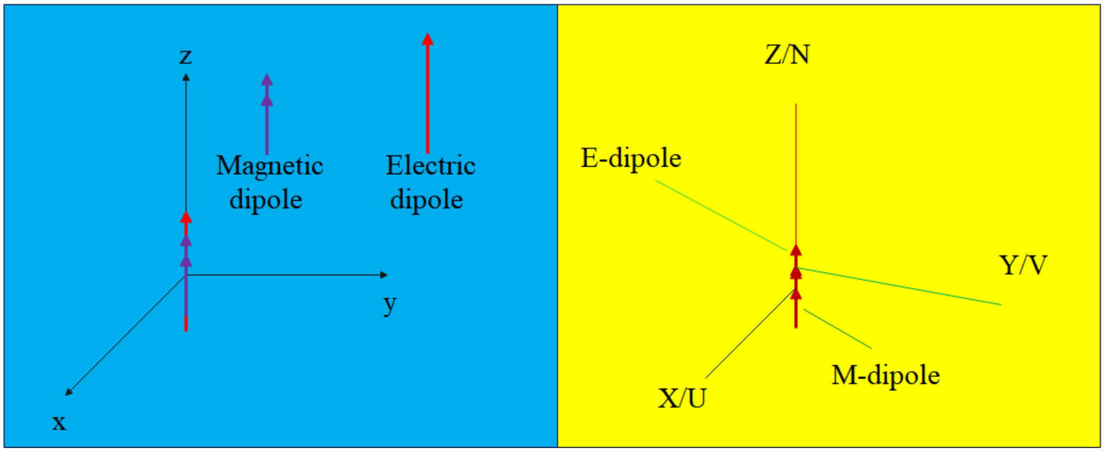


Figure 5: Vector Analysis Model 2 and its FEKO Modeling Diagram.

The algorithm places an electric dipole and a magnetic dipole parallel to the z-axis at the origin of the coordinate system (parallel magnetoelectric dipole model). In spherical coordinates, electric dipoles provide the E_θ -component, and magnetic dipoles provide the E_φ -component:

Electric Dipole: $E_\theta = j\frac{\eta_0 I_e l}{2\lambda R}\sin\theta_e^{-jkR}$

Magnetic Dipole: $E\varphi = -j\frac{I_m l}{2\lambda R}\sin\theta e^{-jkR}$

Among them, I_e and I_m represent the current and magnetic current, respectively, l is the dipole length, R is the distance from the observation point to the origin, λ is the vacuum wavelength, and η_0 is the vacuum wave impedance. If we set $\eta_0 I_e l = I_m l$ and set the observation plane to be perpendicular to the z-axis, when the pitch angle θ is small, the cylindrical coordinate system is observed to have $|E_\varphi| = |E_\theta| \approx |E_\rho|$ (that is, the electric dipole is considered to generate E_ρ) and the two have a 180° phase difference. The circularly polarized wave in the cylindrical coordinate system can be generated only when the electric dipole and the magnetic dipole have a phase difference of 90° when feeding. At the same time, when viewed in Cartesian coordinates, this is a circle carrying orbital angular momentum.

4 OPTIMIZATION OF VOCAL SINGING TRAINING METHOD BASED ON INTELLIGENT BIG DATA TECHNOLOGY

In order to realize the real-time positioning of vocal input notes in the score, a dynamic programming score tracking algorithm based on the extended cave is adopted. The flowchart of the algorithm is shown in Figure 6.

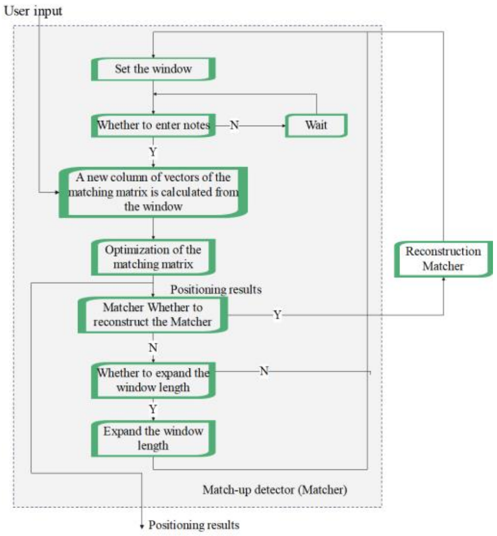


Figure 6: Flowchart of Vocal Score Tracking Algorithm.

The feature extraction of user singing is another difficulty of this system. Only by correctly extracting features useful for evaluation, can a singer's singing level be correctly measured. Therefore, user singing feature extraction and singing evaluation algorithm are the two focuses of this paper. Figure 7 shows the implementation framework of the real-time singing scoring system.

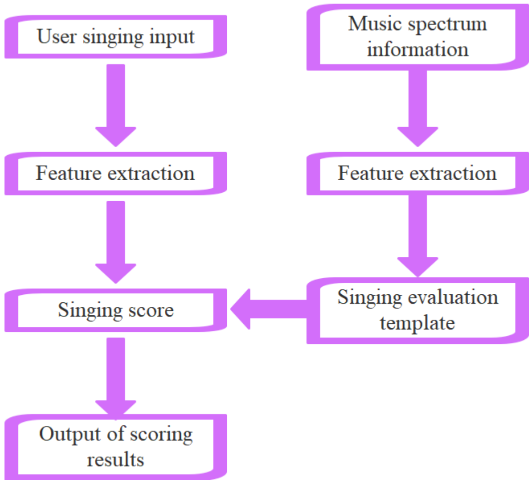


Figure 7: Real-Time Singing Scoring System Implementation Framework.

The block diagram of the parallel processing method is shown in Figure 8. After the speech signal is preprocessed, a series of pulse signals are formed. This series of pulse signals retains the periodic characteristics of the signal, while ignoring the signals irrelevant to the pitch detection, and then the period is estimated by some parallel detectors.

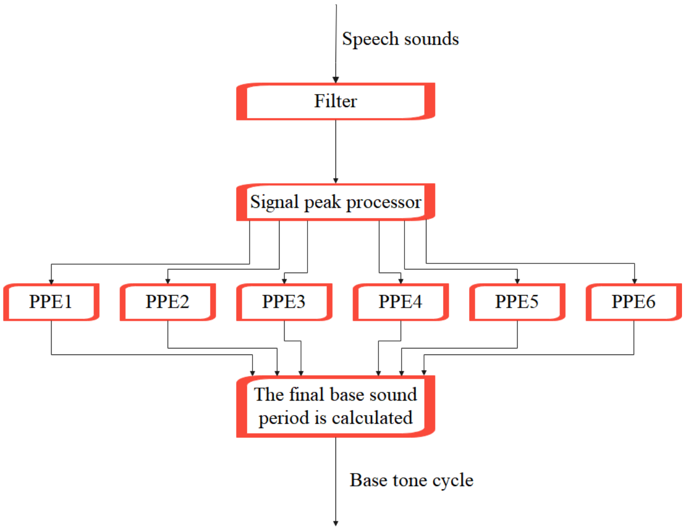


Figure 8: Pitch Detection of Parallel Processing Method.

The vocal music audio detection algorithm in this paper is used to conduct audio detection experiments, and the results shown in Figure 9 are obtained.

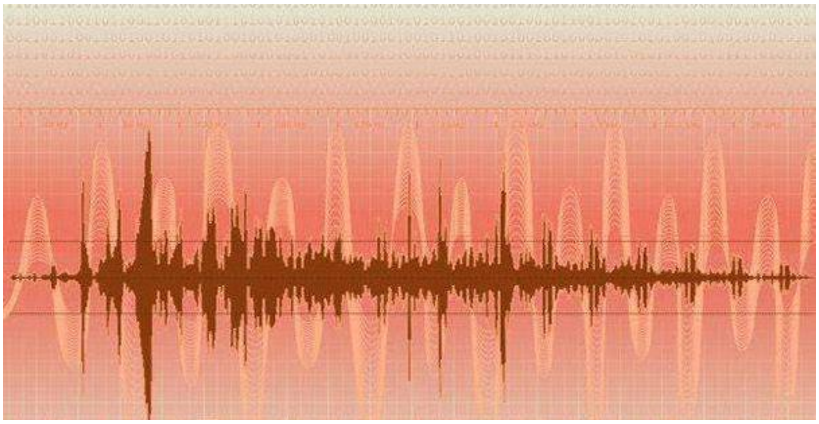


Figure 9: Audio Detection Image of Vocal Music Based on Intelligent Big Data Technology.

On the basis of the above research, the effect of vocal singing training method based on intelligent big data technology is evaluated, and the research results shown in the following table 1 are obtained.

Number	*Singing training*	*Number*	*Singing training*	*Number*	*Singing training*
1	*90.21*	*23*	*89.46*	*45*	*89.42*

2	*92.11*	*24*	*92.70*	*46*	*91.95*
3	*93.25*	*25*	*88.90*	*47*	*89.00*
4	*90.35*	*26*	*88.06*	*48*	*88.13*
5	*93.15*	*27*	*87.11*	*49*	*88.47*
6	*90.52*	*28*	*87.76*	*50*	*92.46*
7	*90.79*	*29*	*93.63*	*51*	*92.37*
8	*91.03*	*30*	*93.40*	*52*	*87.77*
9	*87.96*	*31*	*92.06*	*53*	*92.18*
10	*87.33*	*32*	*92.04*	*54*	*89.79*
11	*92.82*	*33*	*93.73*	*55*	*92.66*
12	*93.43*	*34*	*92.53*	*56*	*88.67*
13	*87.30*	*35*	*90.39*	*57*	*91.53*
14	*93.12*	*36*	*92.10*	*58*	*87.71*
15	*92.92*	*37*	*88.27*	*59*	*92.67*
16	*87.77*	*38*	*91.48*	*60*	*88.44*
17	*90.88*	*39*	*92.88*	*61*	*88.89*
18	*88.59*	*40*	*92.75*	*62*	*91.48*
19	*92.14*	*41*	*93.00*	*63*	*92.74*
20	*87.55*	*42*	*91.98*	*64*	*90.58*
21	*88.21*	*43*	*88.25*	*65*	*90.25*
22	*91.29*	*44*	*87.73*	*66*	*88.84*

Table 1: Effect Evaluation of Vocal Singing Training Method Based on Intelligent Big Data Technology.

It can be seen from the above research that the vocal singing training method based on intelligent big data technology proposed in this paper can play a good auxiliary role in modern vocal music training.

5 CONCLUSION

The main goal of the real-time vocal music singing training system is to automatically, objectively and accurately measure and evaluate the real-time vocal music singing situation. With the accompaniment of song melody, vocal singers are subject to certain constraints when singing specific songs. That is to say, only in the correct time period, the pitch and rhythm of vocal singing that meet the requirements of the song will be correct, which is fundamentally different from the humming retrieval system that has been studied more. Obviously, the humming retrieval system is more arbitrary, and the vocal singer only needs to hum a section of the song, and the requirements for the speed and rhythm accuracy of vocal singing are not strict. This paper studies the optimization of vocal music singing training method combined with intelligent big data technology, and improves the effect of vocal music singing through intelligent vocal music singing method. The research results show that the vocal singing training method based on intelligent big data technology proposed in this paper can play a good auxiliary role in modern vocal music training.

Zhigang Ding, https://orcid.org/0009-0009-8428-4592
Beibei Zhang, https://orcid.org/0009-0005-2301-8160

ACKNOWLEDGEMENT
Anhui Higher Education Quality Project 2022: University Aesthetic Teaching Innovation Team(2022cxtd024).

REFERENCES

[1] Anaya Amarillas, J.A.: Marketing Musical: Música, Industria Y Promoción En La Era Digital, Interdisciplina, 9(25), 333-335.
[2] Babich, B.: Musical "Covers" and the Culture Industry: From Antiquity to the Age of Digital Reproducibility, Research in Phenomenology, 48(3), 2018, 385-407. https://doi.org/10.1163/15691640-12341403
[3] Calegario, F.; Wanderley, M.M.; Huot, S.; Cabral, G.; Ramalho, G.: A Method and Toolkit for Digital Musical Instruments: Generating Ideas and Prototypes, IEEE MultiMedia, 24(1), 2017, 63-71. https://doi.org/10.1109/MMUL.2017.18
[4] Cano, E.; FitzGerald, D.; Liutkus, A.; Plumbley, M. D.; Stöter, F. R.: Musical Source Separation: An Introduction, IEEE Signal Processing Magazine, 36(1), 2018, 31-40. https://doi.org/10.1109/MSP.2018.2874719
[5] Costa-Giomi, E.; Benetti, L.: Through a Baby's Ears: Musical Interactions in a Family Community, International Journal of Community Music, 10(3), 2017, 289-303. https://doi.org/10.1386/ijcm.10.3.289_1
[6] Dickens, A.; Greenhalgh, C.; Koleva, B.: Facilitating Accessibility in Performance: Participatory Design for Digital Musical Instruments, Journal of the Audio Engineering Society, 66(4), 2018, 211-219. https://doi.org/10.17743/jaes.2018.0010
[7] Gonçalves, L.L.; Schiavoni, F.L.: Creating Digital Musical Instruments with Libmosaic-Sound and Mosaicode, Revista de Informática Teórica e Aplicada, 27(4), 2020, 95-107. https://doi.org/10.22456/2175-2745.104342
[8] Gorbunova, I. B.; Petrova, N.N.: Digital Sets of Instruments in the System of Contemporary Artistic Education in Music: Socio-Cultural Aspect, Journal of Critical Reviews, 7(19), 2020, 982-989.
[9] Gorbunova, I.B.: Music computer Technologies in the Perspective of Digital Humanities, Arts, and Researches, Opcion, 35(SpecialEdition24), 2019, 360-375.
[10] Jack, R. H.; Mehrabi, A.; Stockman, T.; McPherson, A.: Action-Sound Latency and the Perceived Quality of Digital Musical Instruments: Comparing Professional Percussionists and

Amateur Musicians, Music Perception: An Interdisciplinary Journal, 36(1), 2018, 109-128. https://doi.org/10.1525/mp.2018.36.1.109
[11] Khulusi, R.; Kusnick, J.; Meinecke, C.; Gillmann, C.; Focht, J.; Jänicke, S.: A Survey on Visualizations for Musical Data, In Computer Graphics Forum, 39(6), 2020, 82-110. https://doi.org/10.1111/cgf.13905
[12] Partesotti, E.; Peñalba, A.; Manzolli, J.: Digital Instruments and Their Uses in Music Therapy, Nordic Journal of Music Therapy, 27(5), 2018, 399-418. https://doi.org/10.1080/08098131.2018.1490919
[13] Scavone, G.; Smith, J. O.: A Landmark Article on Nonlinear Time-Domain Modeling in Musical Acoustics, The Journal of the Acoustical Society of America, 150(2), 2021, R3-R4. https://doi.org/10.1121/10.0005725
[14] Serra, X.: The Computational Study of a Musical Culture Through its Digital Traces, Acta Musicologica, 89(1), 2017, 24-44.
[15] Tabuena, A. C.: Chord-Interval, Direct-Familiarization, Musical Instrument Digital Interface, Circle of Fifths, and Functions as Basic Piano Accompaniment Transposition Techniques, International Journal of Research Publications, 66(1), 2020, 1-11. https://doi.org/10.47119/IJRP1006611220201595
[16] Tomašević, D.; Wells, S.; Ren, I.Y.; Volk, A.;Pesek, M.: Exploring Annotations for Musical Pattern Discovery Gathered With Digital Annotation Tools, Journal of Mathematics and Music, 15(2), 2021, 194-207. https://doi.org/10.1080/17459737.2021.1943026
[17] Turchet, L.; West, T.; Wanderley, M.M.: Touching the Audience: Musical Haptic Wearables for Augmented and Participatory Live Music Performances, Personal and Ubiquitous Computing, 25(4), 2021, 749-769. https://doi.org/10.1007/s00779-020-01395-2
[18] Turchet, L.; Barthet, M.: An Ubiquitous Smart Guitar System for Collaborative Musical Practice, Journal of New Music Research, 48(4), 2019, 352-365. https://doi.org/10.1080/09298215.2019.1637439
[19] Vereshchahina-Biliavska, O. Y.; Cherkashyna, O.V.; Moskvichova, Y. O.; Yakymchuk, O. M.;Lys, O. V.: Anthropological view on the History of Musical Art, Linguistics and Culture Review, 5(S2), 2021, 108-120. https://doi.org/10.21744/lingcure.v5nS2.1334
[20] Way, L. C.: Populism in Musical Mash Ups: Recontextualising Brexit. Social Semiotics, 31(3), 2021, 489-506. https://doi.org/10.1080/10350330.2021.1930857