Hindawi
Advances in Multimedia
Volume 2022, Article ID 2196154, 10 pages
https://doi.org/10.1155/2022/2196154

Research Article

Application of Spectrum Analysis Technology in Music Audio Analysis

Yi Li

Jiangxi Technical College of Manufacturing, Nanchang 330095, China

Correspondence should be addressed to Yi Li; 161849021@masu.edu.cn

Received 2 March 2022; Revised 22 March 2022; Accepted 6 April 2022; Published 9 May 2022

Academic Editor: Qiangyi Li

In order to improve the music analysis technology, this paper studies the music analysis technology combined with the spectrum analysis technology and builds an intelligent audio analysis model. In this paper, the nonlinear theoretical method is adopted, and the motion equation of the audio frequency is obtained through variable processing, so as to obtain the mean square fluctuation of the two sound wave models and then obtain the entanglement and compression. Moreover, this paper introduces the combined mode method to describe the interaction of the two laser fields and atomic matter and verifies that both the differential modes are decoupled from the interaction and only the sum mode participates in the interaction. The experiment verifies that the music audio analysis system based on spectrum analysis technology proposed in this paper can play an important role in music analysis.

1. Introduction

The emergence and application of digital technology have brought far-reaching influence on human beings. Moreover, people now live in a digital world, and among many digital technologies, digital audio technology is one of the most widely used digital technologies [1]. At present, all countries in the world are striving to realize the popularization of high-definition digital TV and digital broadcasting, and China is also one of the most active countries. However, the bandwidth provided by terrestrial or satellite-based digital television systems and digital broadcasting systems is very limited. Therefore, the coding technology that can maintain high-quality sound quality under the premise of realizing a large amount of audio signal compression will be widely used [2]. At the same time, the reduction of the bandwidth occupancy rate of the audio signal can also promote the improvement of the video quality, which is very beneficial to the promotion of the audio and video quality of digital TV [3].

In the Internet age, people have more and more needs for multimedia systems on mobile terminals such as mobile phones and tablets, and the demand for multimedia systems is inseparable from multimedia files and multimedia players. Multimedia files are not only large in number but also large in various formats. Therefore, the support for more mainstream multimedia format files and the pursuit of the higher performance of multimedia players have always been the hot spots sought after by major mobile terminal manufacturers. To extend the parser in the media framework, there are generally two overall ideas: self-developed or borrowed from the parser scheme in the existing mature media framework. At present, there are many mature and stable media frameworks, many of which are open-source, such as FFmpeg, GStreamer, and OSMF (Open Source Media Framework). Compared with independent research and development, using existing mature parsers to expand the parsers in the media framework not only can shorten the development cycle and reduce project costs but also can ensure product reliability and stability.

In order to improve the music analysis technology, this paper combines the spectrum analysis technology to study the music analysis technology, constructs an intelligent audio analysis model, and improves the music audio analysis effect.

2. Related Work

In recent years, communication tools have developed rapidly, and people can conduct long-distance and seamless

information exchange. There are many ways to exchange information, such as voice communication, video communication, e-mail, and telephone communication, among which voice communication. Communication is one of the most basic and important communication methods. At present, the requirements for the transmission quality of the voice signal are constantly improving, so the voice compression coding technology has become more and more important [4]. In the past 10 years, people have obtained a lot of related research results, and some speech coding standards have also been formulated by relevant standardization organizations [5] and applied to speech communication. Speech coding and decoding technology mainly focus on maximizing speech quality under low bit rate conditions, and a lot of research has been done on this. Generally speaking, speech coding techniques can be divided into four categories [6], which are divided by signal bandwidth, coding method, processing domain, and coding rate. Currently, the most widely used coding types are time- and frequency-domain coding. In the current field of speech communication, these coding types receive more and more attention and research. The current speech coding technology is mainly used in mobile communication and Internet communication [7]. For the mobile communication industry, the third-generation mobile communication technology has been widely used in many countries and regions, and the fourth-generation mobile communication is also constantly developing; for the Internet communication industry, due to the instability of the network, the voice coding techniques tend to be flexible and automatically adapt to the needs of the network's real-time environment. The current popular speech codec technologies mainly include SpeeX and SILK [8]. The SpeeX [9] codec is a voice codec technology for Voice over Internet Protocol (VoIP) applications and is mainly used to transmit high voice quality at low bit rates. In order to achieve this goal, SpeeX adopts code-excited linear prediction (CELP) algorithm; meanwhile, SpeeX also has the function of sound preprocessing and acoustic echo cancellation [10]. SpeeX performs well between low bitrates (4.8 kbps) and high bitrates (16 kbps), making it suitable for applications such as blogging, VoIP, and more. The SILK [6] codec is a speech codec technology developed by the Skype team, which is mainly based on linear predictive coding [11]. Compared with the traditional linear prediction coding technology, SILK combines the advantages of iLBC and SpeeX and can perform coding and decoding work adaptively in the network environment at low bit rates. SILK supports 8, 12, 16, or 24 kHz [12] sampling rate and 6–40 kbps bit rate, and the algorithm delay can be as low as 25 ms, which can maintain stable call quality even in a network environment with a packet loss rate of 10%. At present, SILK has been widely used in various mobile devices and personal computer applications [13].

ggVorbis is a free and open-source project led by Xiph.org, which is mainly used to develop a lossy audio compression standard with flexible coding and free open source. Although compared with popular audio formats such as AAC and MP3, the coding complexity is slightly higher, but the decoding complexity is lower, and the decoded sound quality is very good, so it is very suitable for mobile devices such as audio players [14]. As a perfect free audio standard, OggVorbis will have very broad application prospects in the future. CELT [15] is a next-generation audio coding standard that inherits Vorbis. CELT, as a universal and low-latency coding standard, not only inherits the excellent performance of Vorbis but also has very low latency and low CPU and memory requirements. CELT supports stereo; its overall algorithmic delay can be as low as 5 ms and also has a lower bit rate than Vorbis, and CELT also provides superior audio fidelity to many audio input signals that are not the most pristine. In stereo audio at 24 kbps to 64 kbps and 48 kHz sample rate, CELT is comparable to HE-AACv1, and compared to AAC-LD, CELT can provide considerably higher audio fidelity at a lower latency degree [16].

The decoding and playback of network audio data need to design a corresponding decoder according to the different data formats received, and then play it. The data of general network music stations are in MP3 format, while most of the radio stations are ASF streaming media data. The research on MP3 decoding by domestic and foreign scholars has been quite mature [17]. In the literature [18], the decoding of MP3 and its application on special platforms are studied in depth. WMA is a new audio format with the same name as MP3 format, launched by Microsoft. At a low bit rate, the sound quality and compression ratio of WMA are better than MP3, and it is far better than RA (RealAudio) audio compression method. It can also produce good sound quality when the sampling frequency is low. It is small in size and very suitable for application in web streaming and mobile devices [19]. The optimization research of WMA decoding is mainly to achieve the effect of smooth playback of WMA files under special platforms, such as floating-point to fixed-point conversion, the use of look-up tables, and the optimization of MDCT (discrete cosine transform) algorithms commonly used in the transplant process [20].

3. Audio Waveform Spectral Analysis

To analyze the entanglement of two audio modes a_1 and a_2 in Gaussian state, we adopt the sufficient and necessary criterion for the entanglement of two-body continuous variables.

$$M \equiv \langle (\delta u)^2 \rangle + \langle (\delta v)^2 \rangle < 2, \tag{1}$$

where

$$\begin{aligned} u &= X_1 + X_2, \\ v &= P_1 - P_2. \end{aligned} \tag{2}$$

According to this criterion, if the mean square fluctuations of the two EPR-like operators u and v of the two audio modes satisfy the above equation, then the two audio modes are entangled. Here, we introduce two orthogonal operators modulo a_1 and a_2.

$$X_j = \frac{1}{\sqrt{2}}\left(a_j e^{-i\phi_j} + a_j^{\dagger} e^{i\phi_j}\right),$$
$$P_j = \frac{-i}{\sqrt{2}}\left(a_j e^{-i\phi_j} - a_j^{\dagger} e^{i\phi_j}\right), \tag{3}$$

where $\phi_j\,(j = 1, 2)$ is the phase of the a_j-mode, and we choose the phase ϕ_j, and it is such that X_j and P_j are the amplitude and phase operators of the field a_j when it varies around its steady-state value, respectively.

When describing quantum beat audio or correlated emission audio, it is convenient to use the combined mode method. In general, for two audio modes, the combined mode is defined as follows:

$$A = \frac{1}{\sqrt{2}}\left(a_1 e^{-i\theta_1} + a_2 e^{-i\theta_2}\right),$$
$$B = \frac{1}{\sqrt{2}}\left(-a_1 e^{-i\theta_1} + a_2 e^{-i\theta_2}\right), \tag{4}$$

where $\theta_{1,2}$ are two constants determined by the external field. The Hermitian operators corresponding to the amplitude and phase of the sum modulo are proportional to the real and imaginary parts of the operator A, respectively. Therefore, we generally refer to the combined mode A as the sum mode. Similarly, there are Hermitian operators corresponding to the amplitude and relative phase of the differential mode, which are determined by the real and imaginary parts of operator B, respectively. We generally refer to the combined mode B as the differential mode. The commutation relations between these two combined modes are $[A, A^{\dagger}] = [B, B^{\dagger}] = 1$ and $[A, B] = [A, B^{\dagger}] = 0$.

Taking the correlated emission tone as an example, one of the biggest advantages of using combined modes is that the differential modes are generally decoupled from the interaction. Therefore, differential mode B only experiences absorption caused by the vacuum reservoir. However, the sum modulo A is scaled up and works above the threshold, as we will see later. Under this condition, the steady-state value of the amplitude of the differential mode is zero, that is, $\langle B\rangle = 0$. From this, $(a_1)e^{-i\theta_1} = \langle a_2\rangle e^{-i\theta_2}$ can be obtained. This formula shows that the phase difference of the two audio modes A and S is locked to a fixed value, that is, $\phi_1 - \phi_2 = \theta_1 - \theta_2$. Among them, the phase ϕ_1 or ϕ_2 of the two fields can be selected by injecting a weak probe field that does not change the steady-state value of the field amplitude. However, it can cause the phases of the two fields to each lock to a fixed value.

The operators u and v are expressed in combination modulo A and B as follows:

$$u = Ae^{-i(\phi_1 - \theta_1)} + A^{\dagger} e^{i(\phi_1 - \theta_1)},$$
$$v = i\left(Be^{-i(\phi_1 - \theta_1)} - B^{\dagger} e^{i(\phi_1 - \theta_1)}\right). \tag{5}$$

They correspond to the amplitude operator for the sum mode and the phase operator for the differential mode, respectively. The differential mode operator B is decoupled from the audio medium, indicating that the differential mode is in its vacuum state at steady state. From this, it can be obtained that the mean square fluctuation of operator v is at the level of vacuum quantum noise, that is, the mean square fluctuation of operator v is $\langle(\delta v)^2\rangle = 1$. When the audio operates well above the threshold, its amplitude fluctuations are almost negligible relative to the amplitude itself. In this case, the fluctuation of the operator u can be related to the Mandel factor Q, that is, $\langle(\delta u)^2\rangle = 1 + Q$, where $Q = \langle: (\delta u)^2: \rangle / \langle I\rangle$. Among them, the intensity I is defined as $I = A^{\dagger}{}_{A^{\circ}}$, and the constant 1 represents the vacuum quantum noise level. It is already known that $Q = 0$ for fields where the coherent states obey the sonic statistics of the outlier states. When $Q < 0$, the sonic statistics representing the compression of the field in the cavity obey the sub-Poisson sonic statistics. Using the Q factor, we can get the output fluctuation spectrum of the normal order as follows:

$$S(\omega) = 2\int_0^{\infty} d\tau\, \cos(\omega\tau)\frac{\langle: i(t+\tau), i(t): \rangle}{\langle i(t)\rangle}$$
$$= 1 + 2Q\left(\frac{2\kappa}{\lambda}\right)\frac{\lambda^2}{\lambda^2 + \omega^2}, \tag{6}$$

where $i(t) = kA^{\dagger}(t)A(t)$ is the output sonic flow operator, and the reciprocal of the intensity correlation time λ is proportional to the differential gain. $S(\omega) = 0$ corresponds to shot noise, while $0 > S(\omega) \geq -1$ corresponds to sub-Poisson sonic statistics.

Using the above relation, we can get the output spectrum as follows:

$$M(\omega) = 2 + S(\omega). \tag{7}$$

It can be seen from this expression that the two audio modes are entangled if the sum modes are sub-Poisson sonic statistics. It should be emphasized that the condition for this relationship to be established is that the associated emission audio frequency works and the system works above the min value.

Using the expressions of the variances of the sum and differential modes, we can obtain the Q-factor $Q_{1,2} = Q/2$ of each mode and the strength of each mode $I_l = a_l^{\dagger} a_l$. This shows that the sub-Poisson sonic statistics of each mode and the sub-Poisson statistics of the sum mode exist at the same time, and the noise suppression degree of each mode is half that of the sum mode. Correspondingly, the output fluctuation spectrum of the two fields in the output field is also half of the output sum norm, that is, $S_l(\omega) = S(\omega)/2$. Therefore, when the sum mode is a sub-Poisson sonic statistic $S(\omega) < 0$, the two audio modes $a_{1,2}$ are also compressed.

It is easy to see from the above analysis that the entanglement $M(\omega) < 2$ between the two audio modes and the compression $Q_{1,2} = Q/2 < 0$ of each mode are based on the following two reasons. One is the associated emission audio effect, that is, $\langle(\delta v)^2\rangle = 1$; the other is the noise suppression of the sum mode, that is, $S(\omega) < 0$. In the following content, we will give two specific models [154, 155] and use nonlinear theory to analyze their strength fluctuations to obtain the

compression and entanglement properties of the two audio modes.

Consider an ensemble consisting of N four-level tones placed in a dual-mode cavity, the energy levels and transitions of each tone are shown in Figure 1(a). For simplicity, we assume that the decay rates of spontaneous radiation from audio level $|3\rangle$ to levels $|1\rangle$ and $|2\rangle$ are both γ_1, and the decay rates of spontaneous radiation from levels $|1\rangle$ and $|2\rangle$ to audio level $|0\rangle$ are both γ_2. Moreover, the incoherent pumping rate from level $|0\rangle$ to level $|3\rangle$ is Λ_0, and the microwave field with frequency ν_μ is resonantly coupled at the Rabi frequency $\Omega_\mu e^{-i\theta_\mu}$, that is, $|1\rangle \leftrightarrow |2\rangle$, Two audio modes $\omega_{1,2}$ at frequency $a_{1,2}$ couple the transitions separately with the same coupling constant g, that is, $|1,2\rangle \leftrightarrow |3\rangle$. The resonant frequencies of these two transitions are $\omega_{3l}\,(l = 1, 2)$, respectively, and for simplicity, we assume that the decay rates of both cavity fields are κ.

Under the dipole approximation and the rotating wave approximation, the Hamiltonian of the system in the interaction picture is

$$H = H_0 + V, \tag{8}$$

where

$$\begin{aligned} H_0 &= \frac{\hbar}{2}\sum_{j=1}^{N}\Omega_\mu\left(e^{-i\phi_\mu}\sigma_{12}^{j} + e^{i\phi_\mu}\sigma_{21}^{j}\right), \\ V &= i\hbar\sum_{j=1}^{N}\left(g_1 a_1^{\dagger}\sigma_{13}^{j}e^{-i\Delta_1 t} + g_2 a_2^{\dagger}\sigma_{23}^{j}e^{-i\Delta_2 t}\right) + H.c., \end{aligned} \tag{9}$$

where $H.c.$ denotes Hermitian conjugation. The audio operator $\sigma_{kl}^{j} = (|k\rangle\langle l|)^j\,(k \neq l; k, l = 0-3)$ is the dipole transition operator of the j-th audio frequency, and the audio field detuning is defined as $\Delta_j = \omega_{3j} - \omega_j\,(j = 1, 2)$.

We assume that the coupling of the microwave field is strong enough, that is, $\Omega_\mu \gg (g_j\langle a_j\rangle, \Lambda, \gamma_j)\,(j = 1, 2)$, and we turn to the ornamental representation of audio. First, we diagonalize the part of the Hamiltonian related to the microwave field to obtain the ornamental state of the audio, as follows:

$$\begin{aligned} |+\rangle &= \frac{1}{\sqrt{2}}\left(e^{-i(\theta\mu/2)}|1\rangle + e^{i(\theta\mu/2)}|2\rangle\right), \\ |-\rangle &= \frac{1}{\sqrt{2}}\left(-e^{-i(\theta\mu/2)}|1\rangle + e^{i(\theta\mu/2)}|2\rangle\right). \end{aligned} \tag{10}$$

The eigenvalues corresponding to these two eigenstates $|\pm\rangle$ are $\lambda_{\pm} = \pm 1/2\hbar\Omega_\mu$, respectively. The frequency of the cavity field is selected to be $\omega_l = \omega_{3l} - \Omega_\mu/2$ in the decorative state representation of the audio frequency, and the fast oscillation term oscillating with $e^{\pm i\omega t}$ is ignored, and the sum-modular operator A and the differential-modular operator B corresponding to $\theta_1 = -\theta_2 = \theta_\mu/2$ are introduced. We can get the following effective Hamiltonian:

$$\widetilde{V} = i\hbar g\sum_{j=1}^{N}\left(A^{\dagger}\sigma_{+3}^{j} - A\sigma_{3+}^{j}\right). \tag{11}$$

From this expression, it can be seen that the differential mode B is decoupled from the interaction, and only the mode A acts on the transition $|+\rangle \leftrightarrow |3\rangle$, as shown in Figure 1(b). From this effective Hamiltonian, we can find the audio frequency and the equation satisfied by the density operator modulo A. In order to obtain the required noise properties of the sum modulo A, we adopt the c-number Langevin method, adiabatically eliminate the audio variable, and obtain the Langevin equation of the c-number amplitude corresponding to the sum modulo A. Thus, the noise properties of the mode we need are obtained, and then the expression for the amount of judging the compression and entanglement properties of the two audio modes is obtained.

Our numerical results are given below. All the parameters in our numerical calculation include the incoherent pumping rate Λ, the atomic spontaneous emission decay rate γ_1, and the cavity loss rate κ, and these parameters are in the unit of the spontaneous emission decay rate of the tone γ_2. The definition of the set parameter C is $C = 2g^2N/\gamma_2\kappa_0$. In Figure 2, the parameter corresponding to each curve in the figure is the cooperation parameter $C = 200$: the solid line corresponds to $\gamma_1 = 0.02$; the yellow dotted line corresponds to $\gamma_1 = 0.03$; and the green dotted line corresponds to $\gamma_1 = 0.05$. The graph in the figure is the variation curve of the intensity $\langle I_1\rangle = \langle I_2\rangle$ of each mode with the incoherent hundred-pull rate in units of γ_2^2/g^2, and the value of the parameter is the same as that of the zero-frequency output spectrum. It can be seen from the figure that the output spectrum first decreases and then increases gradually with the increase of the incoherent pump rate. At the same time, the intensities of both modes increase with the incoherent pumping rate.

Another point is that when the spontaneous emission decay rate γ_1 of the audio frequency decreases, the zero-frequency output spectrum decreases, and the intensities of the two modes increase. When in the limit case $\gamma_1 \longrightarrow 0$, that is, in the rate matching condition, the zero-frequency output spectrum M(0) has a minimum value 3/2, which indicates that the dimension correction condition can still be satisfied when the quantum beat audio works far from the threshold. At the same time, the output spectrum of each mode also has a minimum value of $S_l(0) = -1/4$, which indicates that both modes have 25% noise rejection. There are two mechanisms that contribute to entanglement and shrinkage in this process. One is correlated emission, only the sum mode A interacts with the audio, and the differential mode B is decoupled from the interaction with the audio. Another is the dynamic noise suppression mechanism. The four-level system in Figure 1(b) can be reduced to a three-level system consisting of energy levels $|0\rangle, |+\rangle$ and $|3\rangle$. In this reduced three-level system, the audio transition $|+\rangle \leftrightarrow |3\rangle$ and the two-step incoherent transition process $|+\rangle \longrightarrow |0\rangle \longrightarrow |3\rangle$ constitute a channel that circulates the electrons in the audio, resulting in dynamic noise suppression. It is the combination of these two mechanisms that leads to the compression of the entanglement of the two audio modes.

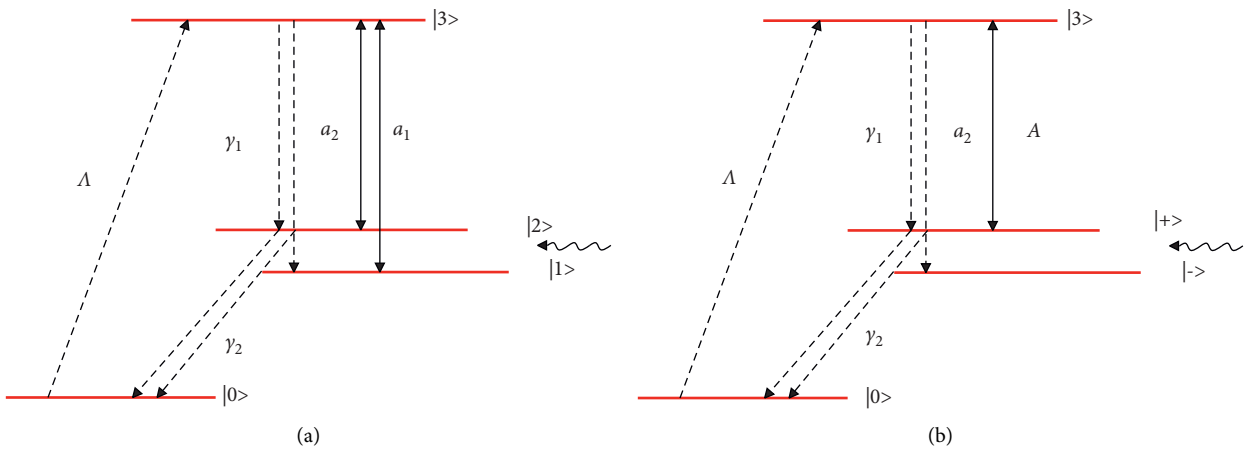

FIGURE 1: (a) Schematic diagram of the audio and field effects of incoherently pumped quantum beat audio and (b) equivalent audio field interaction diagram in the decorated state of the microwave field.

Next, we present the model and equations for coherently driven quantum beat audio. N four-level audio frequencies are captured in a dual-mode cavity, and the energy levels of the audio frequencies are shown in Figure 3(a). The decay rates from energy levels $|1\rangle$ and $|2\rangle$ to $|3\rangle$ ($|0\rangle$) are $\gamma_1(\gamma_2)$, respectively. The incoherent pumping rates from level $|0\rangle$ to levels $|1\rangle$ and $|2\rangle$ are both Λ. The two coherent fields with frequency $\nu_{1,2}$ are coupled to transition at the Rabi frequency $\Omega_{1,2}$, that is, $|1,2\rangle \leftrightarrow |3\rangle$, and the microwave field with frequency ν_μ is resonantly coupled with the Rabi frequency $\Omega_\mu e^{-i\theta_\mu}$, that is, $|1\rangle \leftrightarrow |2\rangle_0$, and the audio modes $a_{1,2}$ at frequency $\omega_{1,2}$ are generated from transitions, that is, $|1,2\rangle \leftrightarrow |3\rangle$. For simplicity, we also assume that the coupling constants of the two cavity modes are equal, and both are g.

Under the dipole approximation and the rotating wave approximation, the Hamiltonian of the system in the interaction picture is

$$H' = H_0' + V', \tag{12}$$

where

$$H_0' = \frac{\hbar}{2}\sum_{j=1}^{N}\Omega_\mu\left(e^{-i\phi_\mu\sigma_{12}^j + e^{i\phi_\mu}\sigma_{21}^j}\right),$$

$$V' = \frac{\hbar}{2}\sum_{j=1}^{N}\left(\Omega_1 e^{-i\delta_1 t}\sigma_{31}^j + \Omega_2 e^{-i\delta_2 t}\sigma_{32}^j\right) \tag{13}$$

$$+ i\hbar\sum_{j=1}^{N}\left(g_1 a_1^\dagger e^{-i\Delta_1 t}\sigma_{01}^j + g_2 a_2^\dagger e^{-i\Delta_2 t}\sigma_{02}^j\right) + \text{H.c.}$$

Audio-field detuning here is defined as $\Delta_l = \omega_{l0} - \omega_l (l = 1,2)$ and $\delta_l = \omega_{l3} - \nu_l (l = 1,2)$. In this paper, we adopt a method similar to that used in the quantum beat audio of incoherent pumping and transfer it to the representation of the rotation of the embellished state with the interaction Hamiltonian of the microwave field; choosing the frequency of the cavity field as $\omega_l = \omega_{3l} + \Omega_\mu/2$ and neglecting the term

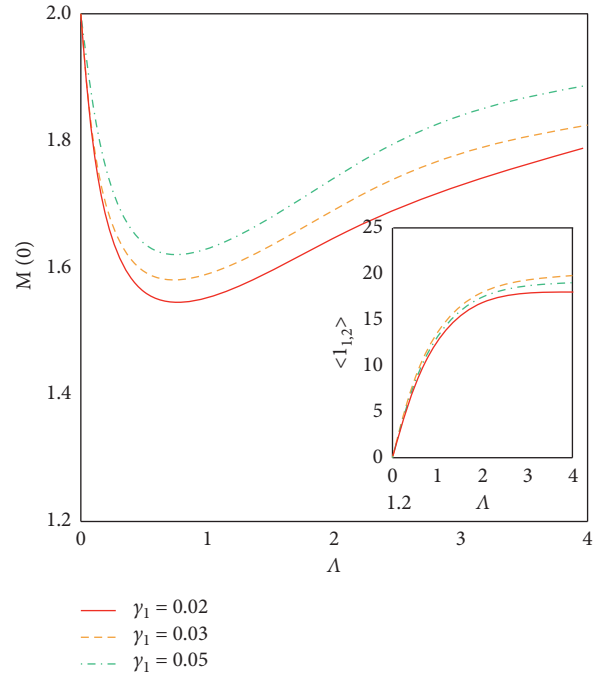

FIGURE 2: Variation curve of zero-frequency output spectrum $M(0)$ with incoherent pumping rate Λ.

of fast oscillation with $e^{\pm i\omega t}$, we can obtain the following effective Hamiltonian:

$$\tilde{V}' = i\hbar g\sum_{j=1}^{N}\left(A^\dagger\sigma_{0+}^j - A\sigma_{+0}^j\right) + \frac{\hbar}{2}\sum_{j=1}^{N}\left(\Omega\sigma_{3+}^j + \Omega^*\sigma_{+3}^j\right), \tag{14}$$

where the effective coupling constant is defined as $\Omega = 1/\sqrt{2}(\Omega_1 e^{-i(\phi\mu/2)} + \Omega_2 e^{i(\phi\mu/2)})$, and here, we introduce the sum-modular operators A and B corresponding to $\theta_1 = -\theta_2 = -1/2\theta_\mu$. From the effective Hamiltonian, we can see that the differential mode B is decoupled from the interaction, and both mode A and the effective coupling field Ω interact with the audio resonance, as shown in Figure 3(b).

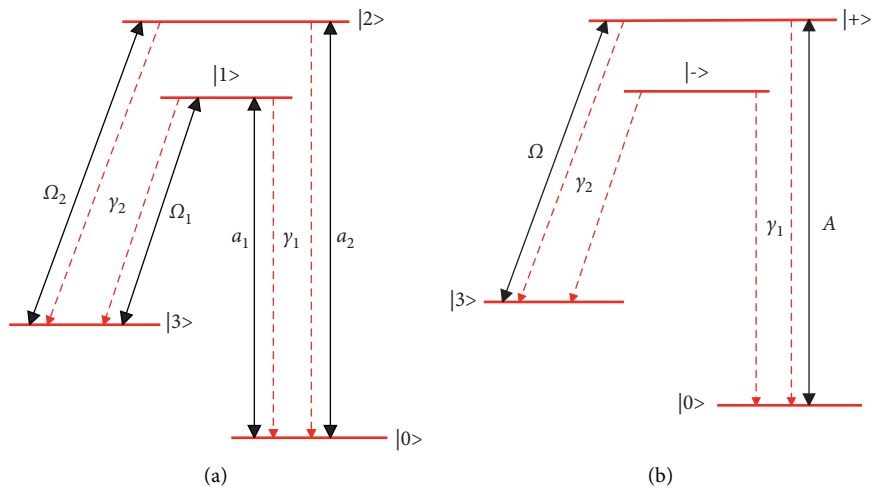

FIGURE 3: (a) Schematic diagram of the interaction between the audio and field of coherently driven quantum beat audio and (b) the equivalent audio and field interaction diagram in the decorated state of the microwave field.

The processing method here is similar to the method in the previous section, using the c-number Langevin method to obtain the equations corresponding to the modulo and audio operators. Furthermore, the atomic variables are eliminated adiabatically to obtain the Langevin equation for the sum mode, which in turn obtains the expression we need to determine the amount of compression and entanglement of two audio modes.

Our numerical results are presented in Figure 4. All the parameters in our numerical calculations include the incoherent pump rate Λ, the audio frequency spontaneous emission decay rate γ_1, and the audio frequency decay rate κ, and these are all in units of the audio's spontaneous emission decay rate γ_2. The cooperation parameter C is defined as $C = 2g^2N/\gamma_2\kappa$. In Figure 4, we present the corresponding zero-frequency output fluctuation spectrum $M(0)$ as a function of the cooperative parameter C when different coherent fields are coupled. The parameters corresponding to each curve in the figure are: the solid line corresponds to $|\Omega| = 3.0$; the green dotted line corresponds to $|\Omega| = 2.5$; and the red dotted line corresponds to $|\Omega| = 2.0$, and the remaining parameters take the values of $\gamma_1 = 4.0$ and $\Lambda = 0.5$. The inset in the figure is the variation curve of the strength $\langle I_1 \rangle = \langle I_2 \rangle$ of each mode with γ_2^2/g^2 as the unit with the cooperation parameter under the same parameter condition. The $\Lambda < \gamma_1$ condition indicates that the system works in the case of no inversion of the population.

It can be seen from the figure that the output spectrum M(0) decreases with the increase of the cooperation parameter C, and it also decreases with the increase of the $|\Omega|$ of the absolute value of the effective Rabi frequency of the coherent field. In the limiting case of $\gamma_1 \ll \gamma_2$, the audio operates well above the threshold value, and the output spectrum is close to 3/2. The fluctuations in the intensity of both modes of the output are 25% noise rejection. The combination of the incoherent population transfer channel $|0\rangle \longrightarrow |+\rangle \longrightarrow |3\rangle$ and the coherent population transfer channel $|3\rangle \longrightarrow |+\rangle \longrightarrow |0\rangle$ makes the electrons in the

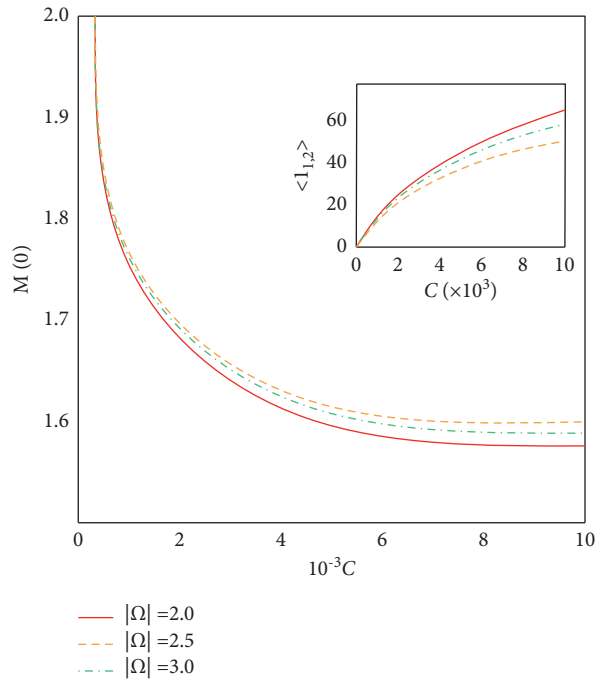

FIGURE 4: Variation curve of zero-frequency output spectrum M(0) with the change of cooperation parameter C.

audio material circulate, resulting in the suppression of the intensity fluctuations of the sum mode. Therefore, it is also the combination of correlated radiation and dynamic noise suppression mechanisms that lead to the entanglement and compression of the bimodal audio field in coherently driven quantum beat audio.

4. Application of Spectrum Analysis Technology in Music Audio Analysis

In order to achieve the goal of evaluating the sound quality of electronic music, this paper combines the characteristics of

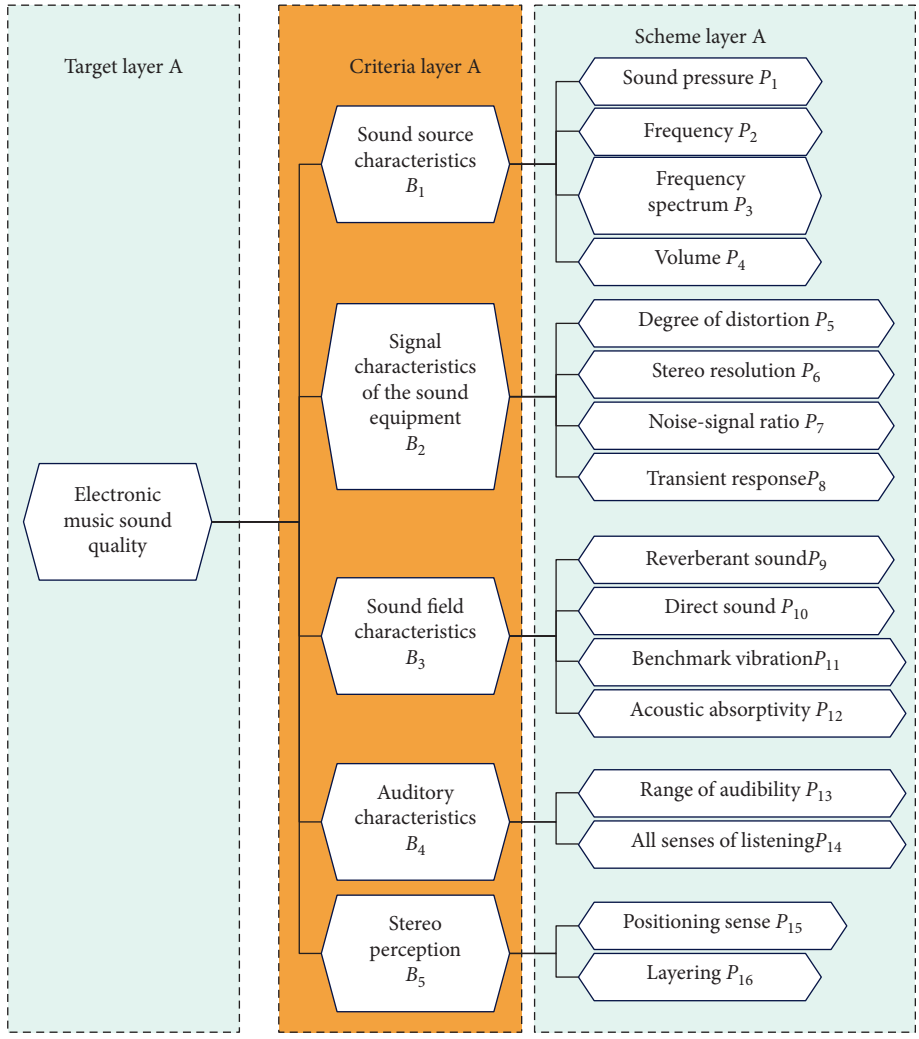

Figure 5: Music audio analysis model.

the sound quality of electronic music and divides the sound quality evaluation of electronic music into three levels: the target layer, the criterion layer, and the scheme layer, and the upper-level elements dominate the lower-level elements. The target layer is electronic music sound quality. The criterion layer includes electronic music sound quality sound source characteristics, signal characteristics of audio equipment, sound field characteristics, auditory characteristics, and stereoscopic evaluation indicators. Each index in the criterion layer is subdivided into multiple sound quality influencing elements, that is, the scheme layer is formed. The structure of the electronic music sound quality evaluation index system based on AHP is shown in Figure 5.

As shown in Figure 6, a music sentiment model is established, which divides music into four categories from a sentimental perspective. Adjacent categories have a certain sentimental similarity, and sentiments can be converted to adjacent categories. The two opposite sentimental categories express opposite sentimental connotations, which is also an important research content in music audio analysis.

Sentiment analysis of melody or lyrics alone is inaccurate, mainly for two points. (1) Lyrics lacking melody cannot determine what the composer wants to highlight, so the analysis of sentiment is inaccurate. (2) Lyrics without a melody background tend to ignore the sense of rhythm in the lyrics, and the sense of rhythm in music is very important for sentimental expression. As this paper argues, sentiment is still largely expressed through musical rhythm. Therefore, the algorithm designed in this section takes both sentiment and rhythm into consideration and designs a word-song matching algorithm model as shown in Figure 7.

On the basis of the above research, the effect of the music audio analysis model based on spectrum analysis technology

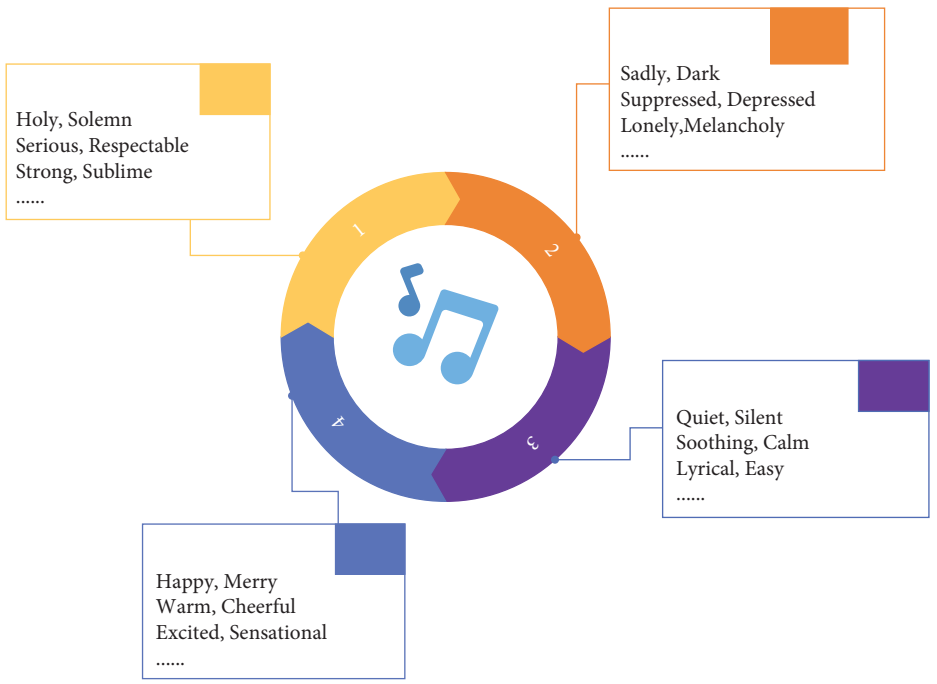

Figure 6: Music sentiment analysis model.

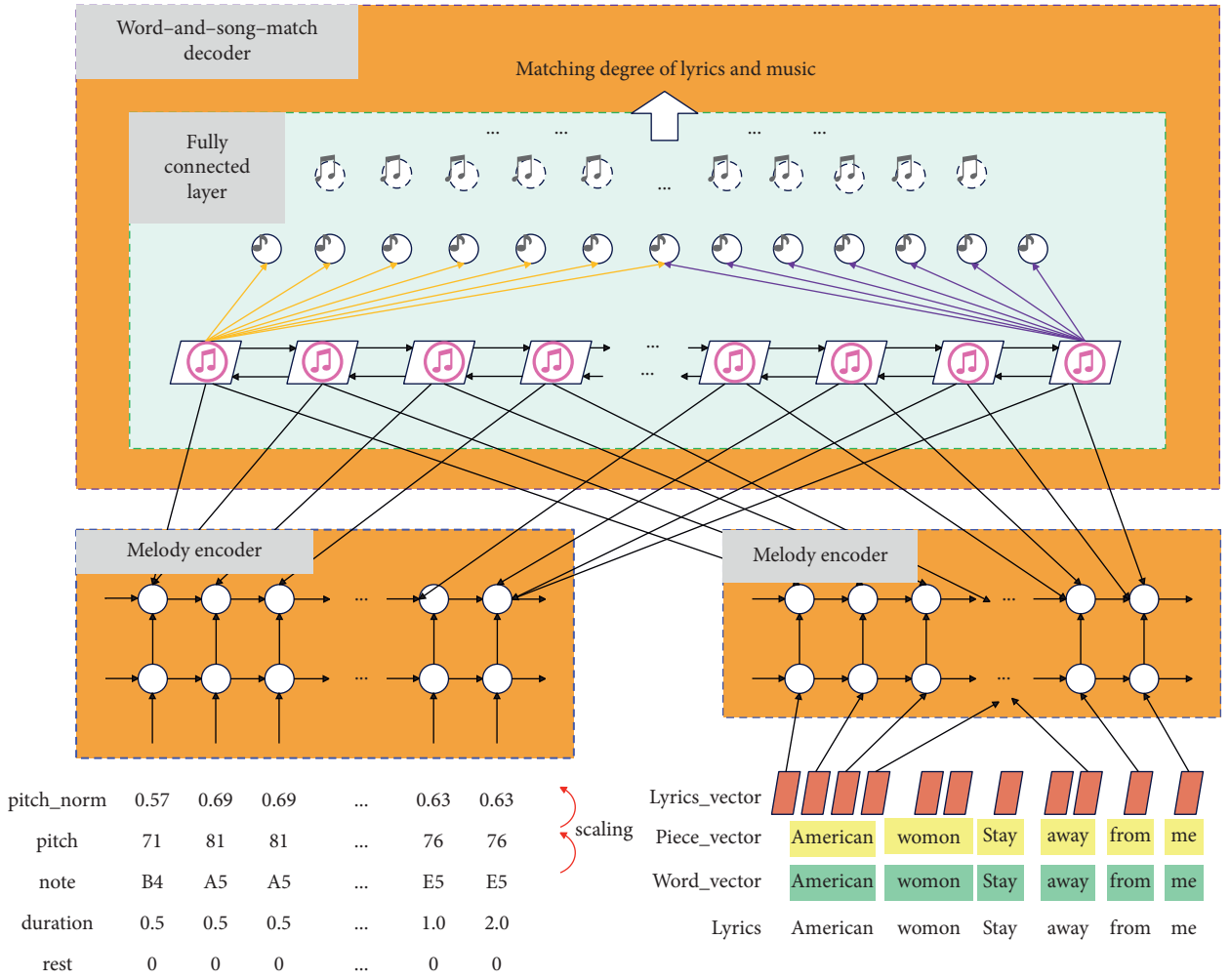

Figure 7: The framework of the word-song matching algorithm.

Table 1: Statistical table of music audio analysis effect of the music audio analysis model based on spectrum analysis technology.

Number	Audio analysis	Number	Audio analysis	Number	Audio analysis
1	74.59	18	82.23	35	67.21
2	69.90	19	74.66	36	71.34
3	78.73	20	66.35	37	71.02
4	80.09	21	76.85	38	71.29
5	86.39	22	78.23	39	80.54
6	69.50	23	73.22	40	78.51
7	85.83	24	71.81	41	79.57
8	78.44	25	70.40	42	83.23
9	67.49	26	71.99	43	86.98
10	76.90	27	77.06	44	82.74
11	66.99	28	72.43	45	66.04
12	85.86	29	84.93	46	85.45
13	69.82	30	81.36	47	78.19
14	68.56	31	85.59	48	67.63
15	70.14	32	70.73	49	77.54
16	70.57	33	66.60	50	72.42
17	76.70	34	69.60	51	81.02

Table 2: The music sentiment analysis effect of the music audio analysis model based on spectrum analysis technology.

Number	Emotion analysis	Number	Emotion analysis	Number	Emotion analysis
1	71.73	18	67.70	35	62.37
2	66.81	19	62.05	36	62.59
3	66.27	20	62.52	37	61.27
4	61.99	21	63.21	38	67.33
5	73.31	22	70.58	39	69.59
6	63.21	23	71.36	40	71.12
7	61.46	24	63.42	41	63.27
8	73.07	25	65.89	42	71.84
9	63.31	26	69.35	43	69.22
10	72.42	27	61.19	44	67.40
11	67.91	28	62.73	45	66.48
12	61.77	29	68.43	46	63.32
13	69.70	30	72.50	47	71.16
14	69.00	31	72.90	48	61.29
15	69.81	32	67.65	49	71.20
16	68.44	33	61.01	50	73.99
17	66.50	34	64.77	51	68.22

proposed in this paper is evaluated, and the music analysis effect and sentiment analysis are verified, and the results shown in Tables 1 and 2 are finally obtained.

From the above research, we can see that the music audio analysis system based on spectrum analysis technology proposed in this paper can play an important role in music analysis.

5. Conclusion

Audio coding is digitized by means of MUSICAM coding, which converts traditional analog signals into digital 0/1 data information. It uses the sound masking effect to quantize and encode sound data according to the characteristics of inear hearing, so as to reduce the resource space occupied by useless information. In the new generation of audio compression technology HE-AAC V2, it is more efficient in encoding, and at the same time, it can also obtain good sound quality when reconstructing audio signals. The DAB system can transmit multiple programs at one time. The service data needs to conform to the service data interface protocol when the channel coding data is input, and the channel coding divides, codes, and integrates the transmitted data. The receiver is assisted in decoding by carrying service information data, identification data, and so on in the Fast Information Channel (FIC). In order to improve the music analysis technology, this paper studies the music analysis technology combined with the spectrum analysis technology. The experimental results verify that the music audio analysis system based on spectrum analysis technology proposed in this paper can play an important role in music analysis.

Data Availability

The labeled data sets used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was sponsored by Jiangxi Technical College of Manufacturing.

References

[1] D. Tomašević, S. Wells, I. Y. Ren, A. Volk, and M. Pesek, "Exploring annotations for musical pattern discovery gathered with digital annotation tools," *Journal of Mathematics and Music*, vol. 15, no. 2, pp. 194–207, 2021.

[2] X. Serra, "The computational study of a musical culture through its digital traces," *Acta Musicologica*, vol. 89, no. 1, pp. 24–44, 2017.

[3] I. B. Gorbunova and N. N. Petrova, "Digital sets of instruments in the system of contemporary artistic education in music: socio-cultural aspect," *Journal of Critical Reviews*, vol. 7, no. 19, pp. 982–989, 2020.

[4] E. Partesotti, A. Peñalba, and J. Manzolli, "Digital instruments and their uses in music therapy," *Nordic Journal of Music Therapy*, vol. 27, no. 5, pp. 399–418, 2018.

[5] B. Babich, "Musical "covers" and the culture industry," *Research in Phenomenology*, vol. 48, no. 3, pp. 385–407, 2018.

[6] L. L. Gonçalves and F. L. Schiavoni, "Creating digital musical instruments with libmosaic-sound and mosaicode," *Revista de Informática Teórica e Aplicada*, vol. 27, no. 4, pp. 95–107, 2020.

[7] I. B. Gorbunova, "Music computer technologies in the perspective of digital humanities, arts, and researches," *Opción*, vol. 35, no. 24, pp. 360–375, 2019.

[8] A. Dickens, C. Greenhalgh, and B. Koleva, "Facilitating accessibility in performance: participatory design for digital musical instruments," *Journal of the Audio Engineering Society*, vol. 66, no. 4, pp. 211–219, 2018.

[9] O. Y. Vereshchahina-Biliavska, O. V. Cherkashyna, Y. O. Moskvichova, O. M. Yakymchuk, and O. V. Lys, "Anthropological view on the history of musical art," *Linguistics and Culture Review*, vol. 5, no. S2, pp. 108–120, 2021.

[10] A. C. Tabuena, "Chord-interval, direct-familiarization, musical instrument digital interface, circle of fifths, and functions as basic piano accompaniment transposition techniques," *International Journal of Research Publications*, vol. 66, no. 1, pp. 1–11, 2020.

[11] L. Turchet and M. Barthet, "An ubiquitous smart guitar system for collaborative musical practice," *Journal of New Music Research*, vol. 48, no. 4, pp. 352–365, 2019.

[12] R. Khulusi, J. Kusnick, C. Meinecke, C. Gillmann, J. Focht, and S. Jänicke, "A survey on visualizations for musical data," *Computer Graphics Forum*, vol. 39, no. 6, pp. 82–110, 2020.

[13] E. Cano, D. FitzGerald, A. Liutkus, M. D. Plumbley, and F. R. Stöter, "Musical source separation: an introduction," *IEEE Signal Processing Magazine*, vol. 36, no. 1, pp. 31–40, 2018.

[14] T. Magnusson, "The migration of musical instruments: on the socio-technological conditions of musical evolution," *Journal of New Music Research*, vol. 50, no. 2, pp. 175–183, 2021.

[15] I. B. Gorbunova and N. N. Petrova, "Music computer technologies, supply chain strategy and transformation processes in socio-cultural paradigm of performing art: using digital button accordion," *International Journal of Supply Chain Management*, vol. 8, no. 6, pp. 436–445, 2019.

[16] J. A. Anaya Amarillas, "Marketing musical: música, industria y promoción en la era digital," *INTERdisciplina*, vol. 9, no. 25, pp. 333–335.

[17] G. Scavone and J. O. Smith, "A landmark article on nonlinear time-domain modeling in musical acoustics," *Journal of the Acoustical Society of America*, vol. 150, no. 2, pp. R3–R4, 2021.

[18] L. Turchet, T. West, and M. M. Wanderley, "Touching the audience: musical haptic wearables for augmented and participatory live music performances," *Personal and Ubiquitous Computing*, vol. 25, no. 4, pp. 749–769, 2021.

[19] L. C. S. Way, "Populism in musical mash ups: recontextualising Brexit," *Social Semiotics*, vol. 31, no. 3, pp. 489–506, 2021.

[20] E. Costa-Giomi and L. Benetti, "Through a baby's ears: musical interactions in a family community," *International Journal of Community Music*, vol. 10, no. 3, pp. 289–303, 2017.