International Journal for Housing Science and Its Applications
Publish August 4, 2025. Volume 46, Issue 3 Pages 1466-1476

https://doi.org/10.70517/ijhsa463113

Research on the Application of Gradient Boosting Tree Algorithm in Immersive Theater Experience Design in the Era of Virtual and Reality Convergence

Jun Su[1,*]
[1] College of Humanities and Arts, Xi'an International University, Xi'an, Shaanxi, 710000, China
Corresponding authors: (e-mail: 18220512802@163.com).

Abstract This paper creates an immersive theatrical scene under virtual reality through multimedia technology, and in order to optimize the effect of theatrical experience, the gradient boosting tree algorithm is introduced to recognize the content of theatrical characters' behavior and speech, which lays the foundation for the interactivity of immersive theatrical experience. Starting from speech recognition, the gradient boosting tree algorithm of this paper is compared with other recognition methods to verify the effectiveness of this paper's method. The satisfaction model of theater experience is constructed, and the audience's satisfaction with the immersive theater experience is visualized through the IPA index. The accuracy of this paper's gradient boosting tree algorithm for speech recognition is 93.38%, 92.62%, 99.88%, which is obviously better than other recognition methods and has higher accuracy. The immersive theater based on gradient boosting tree brings a more satisfactory experience to the audience. In immersive theater, the audience is most satisfied with the social experience, and the satisfaction level (IPA index) reaches 0.80. And the stage setting of immersive theater still has room for improvement.

Index Terms virtual reality, gradient boosting tree algorithm, content recognition, theater experience, IPA index

I. Introduction

Under the background of the era of great cultural integration, theater art is undergoing different degrees of transformation in many aspects such as presentation, scene construction and content structure. Immersive theater is a kind of environment-based performance form, which is the correctness and innovation of traditional theater [1]. It not only breaks through the limitations of traditional theatre in the process of evolution and development, optimizes the theatre scene as well as the content structure, but also greatly enhances the relevance between the performers and the audience [2], [3].

In traditional framed theater, the audience only acts as a passive recipient of information in the presence, observing the performance through vision and hearing [4]. For immersive theater performances, however, the audience not only plays the role of a single appreciator, but gradually becomes the key core of the performance and the subject of interaction and experience [5]. As immersive theater transforms "watching" into "experiencing", it expands the participants' way of perception, so that the audience actively participates and understands through multiple ways, generating in vivo cognition, and the experience is more active, deeper, and sustained compared with watching [6], [7]. The experience of active participation leads the audience to change from passive acceptance of information to active exploration, which is exactly where the important difference lies in its value expression and meaning output compared with traditional theater [8], [9]. Based on this, in the process of immersive theater performance, we need to pay attention to the audience's participation, attract the audience to be in the performance scene, and promote the audience to get a good interactive experience in the interaction [10]-[12]. Through the use of artistic techniques to deal with the virtual scene, and the use of a combination of virtual and real scenes can further enhance the audience's sense of immersion experience [13].

In order to enhance the audience experience of virtual reality-based immersive theater, this paper embeds the gradient boosting tree algorithm in the design of the theater to recognize the contents including the actions of the theater actors and speech, so as to improve the audience's immersive theater experience in the virtual reality space. Taking speech recognition as the test object, the gradient boosting tree-based content recognition model in this paper is juxtaposed with other recognition models, and the accuracy and loss function performance of each model on speech recognition are compared, so as to detect the recognition performance of the gradient boosting tree algorithm. Finally, the theater experience satisfaction evaluation model is constructed, and IPA analysis method is introduced to analyze the audience satisfaction in immersive theater experience in detail.

II. Immersive virtual theater scenario creation in virtual reality

II. A. Design aesthetics of spatial mood creation

II. A. 1) Scene: Situational integration, virtual reality

The theater stage space scene is the core of the whole visual effect. In the design process of immersive theatrical stage, the purpose of designing the scene is to create a moody atmosphere with strong emotional color and imagination.

In the spatial design of stage art, different spatial atmospheres of scenes are created by integrating virtual images with actual objects, creating a realm that makes the audience feel as if they are experiencing it personally, thus enhancing the audience's participation and emotional communion. Virtual reality technology can create highly realistic scene effects by combining virtual elements with the real stage environment.

II. A. 2) Structure: building a multidimensional stage space

The application of virtual reality technology in the spatial layout of the immersive stage is one of the important factors in generating the aesthetics of mood. It allows the audience to enter into a unique stage atmosphere by minimizing the distance between the stage performance and the audience, creating a feeling as if they were experiencing it themselves in order to enhance the audience's immersion and participation. Through the multi-dimensional spatial layout of the stage, it creates a virtual environment that is different from reality, and creates a realm under the guidance of contextual aesthetics to build a space that makes the audience's sensory experience more real, profound and delicate, and then promotes and strengthens the audience to enter into a state of immersion.

II. A. 3) Consciousness: multisensory emotional interaction

Virtual reality in immersive stage design incorporates the artist's thoughts and ideas into the stage design, thus achieving a perfect fusion of art and technology. The artist needs to integrate the main idea and feelings of the work, using different image elements, sound effects and colors and other collocations to convey the story emotion. Like light and shadow, all mediums communicate the content of a production directly or indirectly to the audience through the act of perception by one or more of the human senses.

Immersive stage design realizes the organic combination of sound, light, shadow and multi-sensory experience in ideology through virtual reality technology in space, integrating traditional culture, modern technology, dance art and other fields to create a highly immersive aesthetic mood, and the audience is attracted by the visual impact of stage effects.

II. B. Application of multimedia technology

II. B. 1) Innovations in Dynamic Capture and Holographic Projection

Motion capture is a modern digital technology that captures an actor's movements and transforms them into a computer-generated three-dimensional model. Motion capture is a technique for recording dynamic information about an object for analysis, playback and transmission. This technology is increasingly used in the design of theater stages.

Holography technology relies on interference and diffraction to record changes in the phase and amplitude of light waves, ultimately presenting the 3Dized digital image as a virtual reality. It can convert dynamically captured human movements into holograms in real time and project them onto the stage, where the audience can feel and interact with the human movements on stage by immersing themselves in the holograms. Motion capture technology captures the movements of the actors and transforms them into digital numbers. These digital signals can interact with the choreographic elements, such as controlling the projection, lighting and other elements on the stage, to promote the creation of the stage mood in the audience's immersive state.

II. B. 2) Integration of “VR, AR” technology

The application of AR (Augmented Reality) and VR (Virtual Reality) technology can bring a unique immersive experience for the multimedia immersive stage choreography, and improve the audience's participation and emotional resonance.The use of VR and AR technology can bring a richer experience and emotional resonance for the multimedia immersive stage choreography, so that the audience will be more engaged and mesmerized in the stage. At the same time, with the continuous development of technology, the application of AR and VR technology in multimedia immersive stage has great potential and space.

II. B. 3) Innovative interaction of multimedia devices

Immersive installation art in the 21st century transforms the state between stillness and movement. It makes the dynamics produce interactive flux. Multimedia installations in the immersive stage through unique materials, forms, lighting and other elements, creating a unique visual effect, and sound and light and other multimedia elements are integrated with each other, together to build a rich emotional resonance of the art space, the viewer and then produce a sense of immersion. At the same time, combined with the application of digital technology, multimedia devices can further enrich the stage effect, in order to deepen the audience's immersive experience, so that they can produce some kind of strengthening of the immersion experience, the combination of music and light and shadow makes it easier for the viewer to clarify the mood of the work under the creation of the emotional ideas conveyed.

III. Gradient Boosting Tree Based Drama Content Recognition

In order to enhance the immersion of virtual reality drama experience, this paper utilizes the gradient boosting tree algorithm for the recognition of content including drama actor behavior and speech. By enhancing the recognition effect, it leads to better immersive theater design.

III. A. CSI Data Acquisition

In this paper, a TL-WDR6500 model router is selected as the transmitter (TX) and a desktop computer equipped with an IWL 5300 network card with Ubuntu 12.04 system is used as the receiver (RX), and the distance between the transmitter and receiver is 50cm, where there are two antennas within the transmitter and three antennas within the receiver, as a way to obtain the channel status information of the gesture. The current WiFi technology based on 802.11n protocol uses MIMO-OFDM system, when the transmission bandwidth is 20MHz, a channel is modulated into 64 subcarriers, IEEE 802.11n uses 56 of these subcarriers, and then using the CSI tool one can obtain detailed information about 30 of these subcarriers. Thus, a matrix of n × m × 30 can be obtained, where n is the number of transmitting antennas, m is the number of receiving antennas, and 30 is the number of subcarrier information.

The wireless signal propagation model is redefined in vector mode:

$$\vec{Y} = H\vec{X} + \vec{N} \tag{1}$$

$$\vec{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \vec{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}, \vec{N} = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_m \end{bmatrix}, H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nm} \end{bmatrix} \tag{2}$$

where $\vec{X}$ and $\vec{Y}$ are vector representations of signal transmission and reception, respectively. H denotes the complex matrix of CSI and $\vec{N}$ denotes the Gaussian white noise vector.

It can be shown that obtaining channel state information is equivalent to obtaining the channel matrix H. The multipath propagation of the signal is manifested as delay extension in the time domain and selective fading of the signal in the frequency domain, therefore, the multipath propagation of the signal can be described by utilizing the shock response (CIR) of the wireless channel in the time domain with the following expression:

$$h(\tau) = \sum_{i=1}^{L} a_i e^{-j\theta_i} \delta(\tau - \tau_i) \tag{3}$$

where a_i and θ_i denote the magnitude and phase of the $i-th$ multipath component, respectively, τ_i is the time delay, L denotes the total number of multipaths, and $\delta(\tau)$ denotes the Dirac increment function.

In this paper, the research for drama content (behavior, speech) recognition is carried out in an indoor environment, the sender to the receiver is only 50cm apart, the distance is very short, so much so that the packets are sent and received very fast, the signal in the process of sending and receiving the channel basically maintains the same, and due to the human action, speech speed is much slower than the transmission delay, so gesture action for the WiFi signal caused by the phase shift can be largely ignored, then:

$$h(\tau) = \sum_{i=1}^{L} a_i \delta(\tau - \tau_i) \tag{4}$$

In the time domain, the received signal $r(t)$ is the convolution of the transmitted signal $s(t)$ and the CIR, as shown in equation (5):

$$r(t) = s(t) \otimes h(t) \tag{5}$$

Here the CIR is transformed to the frequency domain by Fourier transform, then the CFR is obtained and the frequency response (CFR) of the wireless channel describes the multipath effect of the signal in terms of amplitude-frequency characteristics and phase-frequency characteristics, respectively, as shown in equation (6):

$$H(f) = FT[h(t)] = \sum_{i=1}^{L} \alpha_i e^{-j2\pi f \tau_i} \tag{6}$$

In the frequency domain, the received signal spectrum $R(f)$ is the product of the transmitted signal spectrum $S(f)$ and CFR, as shown in equation (7):

$$R(f) = S(f)H(f) \tag{7}$$

From equation (7), the frequency response of the channel can be solved by calculating the spectrum of the transmitted signal and the spectrum of the received signal. Since the impact response of the channel and the frequency response of the channel are Fourier transform pairs of each other, solving for the frequency response of the channel is also equivalent to solving for the impact response of the channel.

For the phase information, since the distribution of the phase information in the time domain is random, if we want to extract useful information from the phase information, we need to carry out linear transformation and other processing, and the calculation is relatively complicated. Therefore, the system proposed in this paper only extracts the amplitude information as the main processing data.

III. B. Subcarrier Selection

Since different subcarriers have different center frequencies and wavelengths, CSI measurements for the same motion on different subcarriers have different channel responses. Therefore, each subcarrier has a different degree of sensitivity to each behavior and speech.

III. C. Discrete wavelet transform noise reduction

Wavelet transform on the basis of the development of short-time Fourier transform [14], solves the problem that the window size can not be changed under the change of frequency, and can realize the multi-scale analysis and processing of the signal in the time domain and frequency domain. It can amplify the local features of the signal through the transform, get more detailed information of the signal, which aims to decompose the signal into multi-scale signals at different frequency segments in mutually independent frequency bands, providing an effective way for signal filtering and signal-to-noise separation. Because of this, the wavelet transform is used in many fields, and signal noise reduction is one of the more important applications.

For many signals, the low-frequency segments imply the main features of the signal, while the high-frequency segments describe the details and characteristics of the signal. While conventional low-pass filters are effective in filtering out high-frequency noise components, they may also remove useful signal components at the same time. The discrete wavelet transform is a signal analysis method that discretizes wavelet-based scaling and translation [15], which is capable of describing the frequency-domain properties of local time-domain processes as well as the time-domain properties of local frequency-domain processes.

Suppose that a one-dimensional signal with noise $h(t)$ is input:

$$h(t) = \overline{h}(t) + \sigma n(t), t = 1, 2, \cdots, l \tag{8}$$

where $h(t)$ is the original signal, $n(t)$ is the Gaussian noise signal and obeys the N(0,1) distribution, σ is the noise intensity, conventionally σ = 1, and l is the length of the signal sequence. The process of noise reduction is also the process of converging the time $n(t)$ to 0 as much as possible.

III. D. Feature selection

Extracting representative features from CSI data of dramatic content needs to take into account the operational simplicity and computational volume, and this paper decides to use statistical features in the time-frequency domain, which are the mean value, correlation coefficient, four-digit separator, and third-order centroid distance in the time-domain features, as well as the energy and kurtosis in the frequency-domain features, respectively.

(1) The mean value, which reflects the overall numerical situation of the data, is calculated as follows:

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} A(i) \tag{9}$$

(2) Content energy, reflecting signal strength, is calculated as follows:

$$E = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \left|F(\omega)\right|^2 d\omega \tag{10}$$

(3) Correlation coefficient, which indicates the correlation between different antennas, is calculated as follows:

$$\rho_{xy} = \frac{\operatorname{cov}(x,y)}{\sqrt{D(x)}\sqrt{D(y)}} \tag{11}$$

(4) Kurtosis, which reflects the steepness of the data peaks, is calculated as follows:

$$Kurtosis = \frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^4 f_i}{N\delta^4} \tag{12}$$

(5) The four-digit separator, which describes the dispersion of the reflective sample data, is calculated as follows:

$$IQR = \frac{(Q3 - Q1)}{2} \tag{13}$$

(6) The third-order center distance, which responds to the distribution of gesture data, is calculated as follows:

$$E3 = E(X - \overline{x})^3 \tag{14}$$

III. E. Gradient Boosting Decision Tree

Gradient Boosting Decision Tree (GBDT) is an iterative decision tree algorithm [16], [17], which contains multiple decision trees and the conclusions of all the trees are summed up to make the final answer.GBDT has a wide range of applications and achieves good results on a wide variety of datasets.GBDT is mainly consists of the concepts of Regression Decision Tree (DT), Gradient Boosting (GB) and Shrinkage.

III. E. 1) Decision trees

Decision trees are divided into two categories: regression decision trees and classification decision trees. Regression decision trees are used to predict data, while classification decision trees are used to do classification of labeled data. All decision trees in GBDT are regression decision trees rather than classification decision trees.

The workflow of a regression decision tree is generally similar to the C4.5 algorithm, except that each node (middle node or leaf node) in the regression tree represents a predicted value. If a regression tree is used to predict age, then the predicted value of each node in the regression tree represents the average of the ages of everyone in that node. When a node splits, each threshold for each feature data in that node is exhausted and the best split is found.

III. E. 2) Gradient lifting

GB is a machine learning technique for solving regression and classification problems that produces predictive models in the form of an integration of weak predictive models (usually decision trees). It constructs the model in a staged manner like other Boosting algorithms, and it generalizes an arbitrarily optimized differentiable loss function.GB is usually used as the basic learner along with a fixed-size decision tree.

Like other Boosting methods, the gradient boosting algorithm combines multiple weak “learners” in an iterative manner into a single strong learner.The goal of GB is to find the optimal predictive function $F^*(x)$ in sample set $\{x_i, y_i\}_{i=1}^{N}$ that minimizes the expected value of the loss function. That is:

$$F^*(x) = \arg\min E_{x,y}[L(y, F(x))] \tag{15}$$

GB provides the true y and solves for the approximation function $\hat{F}$ in the form of a weighted sum of functions $h_i(x)$ (from the class of Φ known as the base or weak learner), where the predicted value for each tree is:

$$F(x) = \sum_{i=1}^{M} \gamma_i h_i(x) + const \tag{16}$$

According to the principle of empirical risk minimization, an approximate prediction function $\hat{F}(x)$ is found in the training set by finding the minimum average of the loss function. At the beginning of model training, a prediction function $F_0(x)$ is initialized, and then the greedy algorithm is used to find the best prediction function $F_m(x)$:

$$F_0(x) = \arg\min \sum_{i=1}^{N} L(y_i, \gamma) \tag{17}$$

$$F_m(x) = F_{m-1}(x) + \arg\min \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + f(x_i)) \tag{18}$$

where f is is a function of class Φ from the base learning function. However, each step in the selection of the best function f is generally a complex optimization process for any loss function, so we replace the complexity by solving a simpler problem. This approach uses the fastest descent solution to minimize the problem. If we are only concerned with the prediction results at key points in the training set, the function f is unconstrained, and we update the model each time by using the following equation (19), where $L(y, f)$ in the equation is not a function of f, but rather a vector of values $F(x_1), F(x_2), ..., F(x_n)$ functions:

$$F_m(x) = F_{m-1}(x) - \gamma_m \sum_{i=1}^{n} \nabla_{F_{m-1}} L(y_i, F_{m-1}(x_i)) \tag{19}$$

$$\gamma_m = \arg\min \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) - \gamma \frac{\partial L(y_i, F_{m-1}(x_i))}{\partial F_{m-1}(x_i)} \tag{20}$$

However, since f must come from a restricted class of functions, we must choose a function that is closest to the gradient of the function L, and after selecting f, then use a linear search to select the multiplier γ as in equation (20).

III. E. 3) Shrinkage

Typically, fitting the training set too closely results in a loss of model generalization, and regularization can reduce the effects of overfitting by constraining the fitting process.Shrinkage is a commonly used tool for preventing overfitting, and argues that taking a small step can bring a model closer to the true value, whereas taking a large step can bring a model closer to the true value faster, but it is also more likely to result in overfitting. For this reason, Shrinkage proposes an improvement to gradient boosting:

$$F_{m+1}(x) = F_m(x) + v \cdot \gamma_{m+1} h_{m+1}(x), 0 < v \le 1 \tag{21}$$

where the parameter v is the learning rate.

As a rule of thumb, in gradient boosting algorithms, using a small learning rate is significantly more generalizable than not having this learning rate, but the smaller the learning rate the more iterations the algorithm has. Here Shrinkage still uses the residuals as the training target.

IV. Analysis of immersive theatrical experiences

IV. A. Algorithm recognition performance analysis

IV. A. 1) Comparison of accuracy rates

The first key point of this study is the feature extraction and feature fusion of audio, in order to fully validate the performance of the gradient boosting tree based content recognition method on speech recognition, the following will be carried out on three datasets, Urbansound8k, Ali Tianchi Food Speech, and TESS, to validate the effectiveness and feasibility of the scheme.

Table 1: Accuracy of different voice recognition model

Model	Urbansound8k	Ali tianchi food voice	TESS
GMM-HMM	85.46%	76.78%	96.63%
DTW	89.84%	80.87%	95.82%
Wav3Vec	86.06%	87.88%	96.23%
HuBERT	87.02%	84.36%	97.53%
RNN-T	89.28%	85.78%	97.94%
Transformer-Transducer	91.02%	88.25%	96.37%
Whisper	90.34%	89.14%	97.86%
GBDT (Ours)	93.38%	92.62%	99.88%

To ensure the scientificity and reliability of the experiments, the process of experimental validation follows the principle of single variable. The experiments use audio features as independent variables, and feature extraction is performed on each dataset in a way that compares the classification accuracy of the final test set, and the one with the highest classification accuracy is the scheme with the optimal feature processing.

The accuracy data of the test set of different schemes are compared using different speech recognition models, and the comparison results are shown in Table 1. According to Table 1, the recognition accuracy of this paper's gradient boosting tree model on each audio dataset is the highest accuracy (93.38%, 92.62%, 99.88%), which verifies the reliability of the gradient boosting tree-based algorithm on speech recognition.

In this experiment, the extracted features are used as independent variables and fed into the classification network for training, that is, the data is fed into the training network after feature extraction, in which the proportion of training data fed into the recognition model is 75% of the whole dataset, and the remaining 25% is used as a test set to test the results of the training model. The confusion matrix obtained by adopting this paper's Gradient Boosting Tree algorithm for testing in Ali Tianchi food speech dataset is shown in Figure 1. According to the diagonal color depth of this confusion matrix and the corresponding values, it can be seen that using the gradient boosting tree algorithm, the model has a high recognition rate at this time.

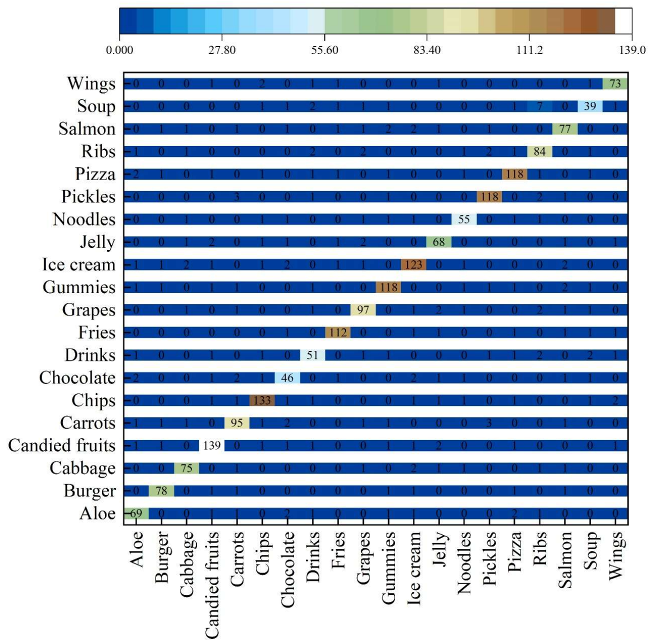
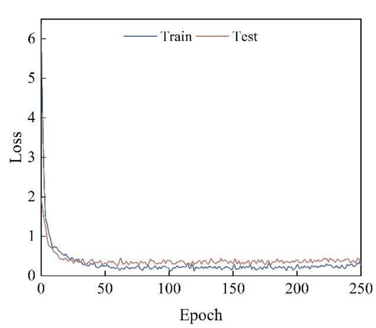


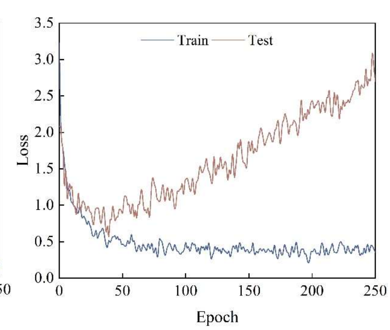
Figure 1: Ali tianchi food voice recognition results

IV. A. 2) Comparison of loss functions

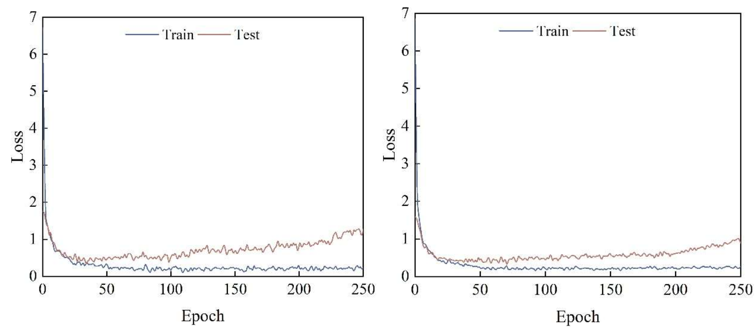
Based on the research objective of this paper is the classification of acoustic scenes, then the classification accuracy is naturally an important comparison index, which has been analyzed in detail in Section 4.1.1. Based on several speech recognition models (GMM-HMM, Wav3Vec, RNN-T, GBDT) in the comparison experiments, the loss functions of the models are investigated, and the results are shown in Fig. 2, and Figs. (a)~(d) are the loss functions of GMM-HMM, Wav3Vec, RNN-T, and the GBDT model in this paper on the UrbanSound8K dataset, respectively. In the experiments using multiple sets of speech recognition models, it can be found that the loss function curves show different decreasing trends in the corresponding loss functions when different speech recognition models are taken.



(a) Loss function of GMM-HMM



(b) Loss function of Wav3Vec



(c) Loss function of RNN-T (d) Loss function of GBDT

Figure 2: Loss function in UrbanSound8K dataset

From the trend of the loss function curve, it is obvious that the recognition models based on the gradient boosting tree all overcome the overfitting defects in the training process, which fully verifies the scientific and reasonable nature of feature fusion.

IV. B. Analysis of Theater Experience Satisfaction

IV. B. 1) Experience Satisfaction Evaluation Model Construction

Based on the evaluation elements and evaluation indexes of immersive theater experience satisfaction, the virtual reality immersive theater experience satisfaction evaluation model is constructed as shown in Table 2.

Table 2: Virtual reality immersion drama experience satisfaction evaluation model

Primary indicator	Secondary indicator	Tertiary indicator
Audience elements (A)	Self-experience (A1)	Comfortable feeling & experience (A11)
	Social experience (A2)	Network interaction (A21)
Space element (B)	Stage setting (B1)	Reasonable setting distance (B11)
		Reasonable setting layout (B12)
		Build a narrative virtual scene (B13)
	Light setting (B2)	Appropriate lighting tone (B21)
		Reasonable use of light (B22)
	Sound setting (B3)	Moderate volume (B31)
Technology element (C)	Digital collection of drama resources (C1)	Visual reconstruction based on data resources (C11)
		Data resource accuracy (C12)
		Collect drama resources (C13)
	Image building (C2)	Build virtual environment (C21)
		Image resolution control (C22)
		Image continuity control (C23)
	Sound design (C3)	Sound and image synchronization (C31)
		Reasonable sound speed setting (C32)
		Sound clarity (C33)
	Color setting (C4)	Color harmony (C41)
		The management and statistics of color data (C42)
Sensory element (D)	Visual experience (D1)	3D virtual space (D11)
		Make dramatic characters (D12)
	Auditory experience (D2)	VR stereo sound (D21)
	Tactile experience (D3)	Virtual tactile interaction (D31)
		Interaction between dramatic character and virtual environment (D32)

IV. B. 2) Satisfaction analysis

The author distributed 100 questionnaires to immersive theater audiences, and 92 valid questionnaires were collected. IPA analysis method is introduced to explore the audience experience in depth.

(1) Paired-sample t-test

The 24 three-level indicators in this questionnaire survey require the audience to score both importance and satisfaction, which are both independent and interconnected. In order to better analyze the differences between the two, a paired-sample t-test is conducted on the mean value of importance and the mean value of satisfaction of the 24 measures of audience satisfaction in the immersive theatre experience to infer whether there is a difference in the means of the two aggregates and to decide whether the data are suitable for IPA analysis. Are suitable for IPA analysis. According to the principle of the paired samples t-test, when the p-value is less than 0.05 at a confidence interval of 95%, the test of significance is passed, and when the p-value is less than 0.01, the difference is highly significant.

The results of the paired samples t-test for importance and satisfaction are shown in Table 3 The two-tailed probability of significance P of the 24 measures is lower than 0.05, and the P value of 20 of them is lower than 0.01, which indicates that the difference in importance and satisfaction of the 24 measures of immersive theater experience passes the test of significance, and is suitable for IPA analysis.

Table 3: The importance and satisfaction matching sample t test

Tertiary indicator	Importance mean	Satisfaction mean	t	p
A11	4.16	4.01	8.137	0.000
A21	4.16	4.02	7.967	0.000
B11	3.86	3.68	8.183	0.019
B12	3.74	3.56	5.639	0.000
B13	3.98	3.81	9.145	0.000
B21	3.91	3.48	2.556	0.034
B22	4.18	4.06	5.796	0.000
B31	4.45	3.57	2.491	0.000
C11	4.15	3.69	7.255	0.000
C12	4.28	3.72	2.352	0.029
C13	3.79	3.71	3.155	0.000
C21	3.96	3.77	8.047	0.000
C22	4.07	3.69	10.22	0.000
C23	4.52	4.17	2.533	0.000
C31	4.25	4.15	6.328	0.000
C32	4.17	4.02	4.458	0.000
C33	4.19	4.14	4.442	0.015
C41	4.27	3.36	5.766	0.000
C42	3.99	3.84	10.081	0.027
D11	3.74	3.42	6.302	0.000
D12	4.06	3.19	7.018	0.000
D21	3.93	3.63	9.571	0.000
D31	4.06	3.12	4.428	0.008
D32	3.78	3.26	5.377	0.000

(2) Overall evaluation of audience satisfaction

Based on the immersive theater experience audience satisfaction measurement system, the primary indicators including audience elements, space elements, technology elements, sensory elements, and the secondary indicators including self-experience, social experience, stage setup, lighting setup, sound setup, etc. were respectively counted and ranked in terms of the mean value of importance and satisfaction, and the difference in the mean value of importance and satisfaction, I-P, was calculated. I-P Mean Difference This indicator mainly reflects the gap between audience expectations and perceived quality, the smaller the gap, the higher the satisfaction of the immersive theater experience, and the larger the gap, the higher the audience satisfaction needs to be improved. Finally, IPA index is calculated, IPA index = (I-P)/I*100. IPA index can also be used to measure the satisfaction of tourists, the larger the number the lower the satisfaction, the smaller the number the higher the satisfaction. The specific statistical results are shown in Table 4.

Audience satisfaction in experiencing immersive theatre ranked in the first tier as color setting, visual experience, self experience, and tactile experience, with a mean satisfaction value of 3.9 or higher. Stage Setting, Lighting Setting, Sound Setting and Social Experience ranked in the second tier, with mean satisfaction values between 3.7

and 3.9. Digital extraction of theatrical resources, sound design, listening experience and image construction ranked in the third tier, with satisfaction mean values below 3.7.

Stage Setting has the largest IPA index of 8.94, indicating that the audience is least satisfied with the stage setting for immersive theater experience. This is followed by sound design with an IPA index of 4.94, which is low in audience satisfaction. The IPA indexes for social experience, lighting setup, image construction, auditory experience, and tactile experience are 0.80, 1.03, 2.72, 2.94, and 1.76 (all less than 3), respectively, indicating high audience satisfaction. Among them, the IPA index of social experience of immersive theater is 0.80, which has the highest level of audience satisfaction.

Table 4: Audience satisfaction evaluation results

Primary indicator	Secondary indicator	Importance		Satisfaction		I-P	IPA index
		Mean	Rank	Mean	Rank		
A	A1	4.17	4	4.03	3	0.14	3.36
	A2	3.77	10	3.74	8	0.03	0.80
B	B1	4.25	3	3.87	5	0.38	8.94
	B2	3.89	7	3.85	6	0.04	1.03
	B3	3.96	6	3.81	7	0.15	3.79
C	C1	3.83	9	3.68	9	0.15	3.92
	C2	3.68	12	3.58	12	0.10	2.72
	C3	3.85	8	3.66	10	0.19	4.94
	C4	4.37	1	4.22	1	0.15	3.43
D	D1	4.29	2	4.13	2	0.16	3.73
	D2	3.74	11	3.63	11	0.11	2.94
	D3	3.98	5	3.91	4	0.07	1.76

V. Conclusion

The author uses the gradient boosting tree algorithm to improve the immersive theater experience under virtual reality, integrating multimedia technology and the gradient boosting tree-based content recognition method in the scene creation of theater performances and other aspects, so as to promote the improvement of the theater experience.

Taking speech recognition as an example, the GBDT model in this paper is validated for its recognition function. The recognition accuracies of this paper's GBDT model on the three datasets are 93.38%, 92.62%, and 99.88%, respectively, all of which are the best recognition performance on the three datasets. The recognition model based on gradient boosting tree overcomes the overfitting defects in the training process, and fully verifies the scientific and rationality of feature fusion. The audience is generally more satisfied with the immersive theater fused with gradient boosting tree. The audience is most satisfied with the social experience of immersive theater (IPA index of 0.80) and least satisfied with the stage setting (IPA index of 8.94). It can be seen that drama content recognition based on the gradient boosting tree algorithm is more helpful in enhancing the social experience of drama.

References

[1] Biggin, R., & Biggin, R. (2017). Immersive theatre, immersive experience. Immersive Theatre and Audience Experience: Space, Game and Story in the Work of Punchdrunk, 1-58.

[2] Turner, R., & Kasperczyk, H. (2022). Space for the unexpected: serendipity in immersive theatre. In The Art of Serendipity (pp. 101-125). Cham: Springer International Publishing.

[3] Yuhui, T., & Zainal, S. (2024). Immersive Theatre: A Comprehensive Review and Future Direction. Pakistan Journal of Life and Social Sciences (PJLSS), 22(2), 3767-3775.

[4] Zhu, M., & Cao, K. (2024, June). Immersion in Theatre: The Emergence of Immersive Theatre Spaces. In International Conference on Human-Computer Interaction (pp. 434-444). Cham: Springer Nature Switzerland.

[5] Punpeng, G., & Yodnane, P. (2023). The route to immersion: a conceptual framework for cross-disciplinary immersive theatre and experiences. Humanities and Social Sciences Communications, 10(1), 1-9.

[6] Hogarth, S., Bramley, E., & Howson-Griffiths, T. (2018). Immersive worlds: an exploration into how performers facilitate the three worlds in immersive performance. Theatre, Dance and Performance Training, 9(2), 189-202.

[7] Belo, M., & Gustavo, N. (2024). Managing motivations and satisfaction in dark immersive theatre experiences. European Journal of Tourism Research, 37, 3709-3709.

[8] Lee, J., Lee, S., Kim, Y., & Noh, J. (2016). Screenx: Public immersive theatres with uniform movie viewing experiences. IEEE transactions on visualization and computer graphics, 23(2), 1124-1138.

[9] Pearlman, E. (2017). Brain opera: exploring surveillance in 360-degree immersive theatre. PAJ: A Journal of Performance and Art, 39(2), 79-85.
[10] Gezgin, Ö., & Imamoğlu, Ç. (2024). Audience behavior in immersive theatre: an environment-behavior studies analysis of Punchdrunk's Sleep No More. Studies in theatre and performance, 44(2), 251-267.
[11] Branch, B., Efstratiou, C., Mirowski, P., Mathewson, K. W., & Allain, P. (2021, May). Tele-immersive improv: Effects of immersive visualisations on rehearsing and performing theatre online. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1-13).
[12] Charitos, D., & Timplalexi, E. (2023). Tracing the impact of the digital virtual ludic on immersive theatre: a case of theatre gamification. Acta Ludologica, 6(1), 4-27.
[13] Melki, H. (2021). Stage-directing the Virtual Reality Experience: Developing a Theatrical Framework for Immersive Literacy. International Journal of Film and Media Arts, 6(2), 129-147.
[14] Bohan Ma,Yushan Xue,Yuan Lu & Jing Chen. (2025). Stockformer: A price–volume factor stock selection model based on wavelet transform and multi-task self-attention networks. Expert Systems With Applications,273,126803-126803.
[15] Yuting Yan,Ruidong Lu,Jian Sun,Jianxin Zhang & Qiang Zhang. (2025). Breast cancer histopathology image classification using transformer with discrete wavelet transform. Medical Engineering and Physics,138,104317-104317.
[16] Yadong Zhang,Shaoping Wang,Enrico Zio,Chao Zhang,Hongyan Dui & Rentong Chen. (2025). Model-guided system operational reliability assessment based on gradient boosting decision trees and dynamic Bayesian networks. Reliability Engineering and System Safety,259,110949-110949.
[17] Qi Zhou,Chenyu Zhang & Ruihuang Liu. (2025). Open-circuit fault diagnosis of AC port of energy router based on wavelet packet decomposition and gradient boosting decision trees. Journal of Physics: Conference Series,2936(1),012026-012026.