

Multimodal Interactive Intelligent Voice Interaction System based on Artificial Intelligence Technology

Ke Chen^a

School of Telecommunication, Tianjin Railway Technical and Vocational College, Tianjin
300240, China

^a670611395@qq.com

Abstract

In addressing the accessibility issues faced by the elderly and individuals with disabilities in the utilization of metropolitan railway ticketing systems, a multimodal interactive intelligent voice ticketing system has been conceptualized. This system is an all-encompassing self-service apparatus that integrates intelligent ticket procurement, vocal inquiry, and facial recognition-based account replenishment functionalities. The seamless amalgamation of touch screen technology and voice command operations facilitates a user-friendly experience for passengers, enabling them to effortlessly execute ticket transactions, route inquiries, and benefit from a design that prioritizes human-centric principles. This innovative system leverages cutting-edge multimodal interaction techniques, allowing for effective communication through a variety of sensory modalities, thereby enhancing accessibility and inclusivity. The design philosophy underscores the importance of accommodating the diverse needs and preferences of its user base, ensuring that the system is not only technologically advanced but also universally accessible. The integration of these features represents a significant stride towards creating an inclusive public transportation ecosystem that transcends the barriers of age and physical ability.

Keywords

Ticketing System; Multimodal; Intelligent Voice Interaction System; Artificial Intelligence.

1. Introduction

With the continuous development of science and technology and the acceleration of urbanization, artificial intelligence technology has gradually penetrated into all aspects of our lives [1]. In order to implement the "State Council on the issuance of the "14th Five-Year Plan" for the protection and development of people with disabilities notice", to consolidate the results of poverty alleviation of people with disabilities with the effective linkage of the revitalization of the countryside.

Improve the social security system and caring service system for people with disabilities, and promote the comprehensive development of the cause of people with disabilities, the emergence of multimodal interactive intelligent voice ticketing system is of great significance [2]. Disabled people are a special group, each with a different type of disability, different degree of impairment, and different difficulties that need to be solved urgently. Some of them are physically disabled, mobility; some are hearing impaired, unable to communicate normally. Data from the China Disabled Persons' Federation show that the total number of people with various types of disabilities in China has now reached 85 million [3]. For this part of the group, take the bus, subway and other public transportation travel smooth or not directly related to their happiness index.

At this time, the generation of voice interaction [4] can quickly and efficiently solve the dilemma faced by special populations in the purchase of tickets, voice interaction in addition to the advantages of almost no learning costs, especially for the elderly group, freeing the human hands as input, can be synchronized to carry out other tasks, and the mobilization of the human auditory sensory experience without being completely confined to the visual [5]. For this reason, we have developed a multimodal interactive intelligent voice ticketing system, so that they will not be inconvenient to use the ticketing machine, as well as the location of unfamiliar and illiterate to make a solution for the special populations to provide a simpler way to purchase tickets, so that it is convenient for everyone.

2. System Analysis

2.1 User Needs

Times are progressing, technology is innovating, the industry is developing, the new trend requires the railroad units in the service of special travelers, and gradually realize the transformation from meeting the basic needs to provide accurate services, different groups of special travelers have different needs. According to the survey, most of the reluctance to buy tickets in the machine for the following reasons: crowded, inconvenient, will not be able to use, etc., as well as special populations, such as deaf and mute will not be able to buy tickets, people with disabilities can not be used normally to buy tickets, and so on. For this reason, we not only need to ensure that the recognition accuracy reaches a high level on the premise of improving the response speed, to give the user a better interactive experience. Accurately recognize the user's intention to buy tickets, and give the user accurate information such as: origin and destination, time, vehicle time information, and the nearest route. Provide accurate responses to various situations that may arise during the ticketing process.

Times are progressing, technology is innovating, the industry is developing, the new trend requires the railroad units in the service of special travelers, and gradually realize the transformation from meeting the basic needs to provide accurate services, different groups of special travelers have different needs. According to the survey, most of the reluctance to buy tickets in the machine for the following reasons: crowded, inconvenient, will not be able to use, etc., as well as special populations, such as deaf and mute will not be able to buy tickets, people with disabilities can not be used normally to buy tickets, and so on. For this reason, we not only need to ensure that the recognition accuracy reaches a high level on the premise of improving the response speed, to give the user a better interactive experience. Accurately recognize the user's intention to buy tickets, and give the user accurate information such as: origin and destination, time, vehicle time information, and the nearest route. Provide accurate responses to various situations that may arise during the ticketing process.

2.2 Comparison with Existing Ticketing Methods

(1) Traditional manual ticketing

During peak periods, manual ticketing may have long waiting lines, user emotional dissatisfaction or other behaviors that lead to human factors influence on the conductor.

(2) Self-service ticket machines

Passengers need to operate manually, which may have certain learning cost and consumption time cost for passengers who are unfamiliar with the operation process or not subject to cultural level.

(3) Network Payment Platform

In the subway ticketing scenario, passengers are required to check the station, route, train number, fare and other information before making payment, which is relatively cumbersome.

(4) Other intelligent voice technologies

It is necessary to have a high voice recognition accuracy, and in the subway, which is full of people and noisy environment, it may be interfered, and the voice recognition can not be accurate to the specific command.

In summary, the multimodal interactive voice ticket machine has a greater competitive advantage in the subway voice ticketing scenario, but it also needs to continuously optimize and improve its technology and services to adapt to the changing market demand. We will continue to invest in research and development, and continue to introduce new products and services in order to satisfy our customers' needs and maintain our market leading position. Through continuous innovation, we can maintain our competitive advantage and continue to expand our market.

3. Multimodal Interactive Intelligent Voice Interaction System

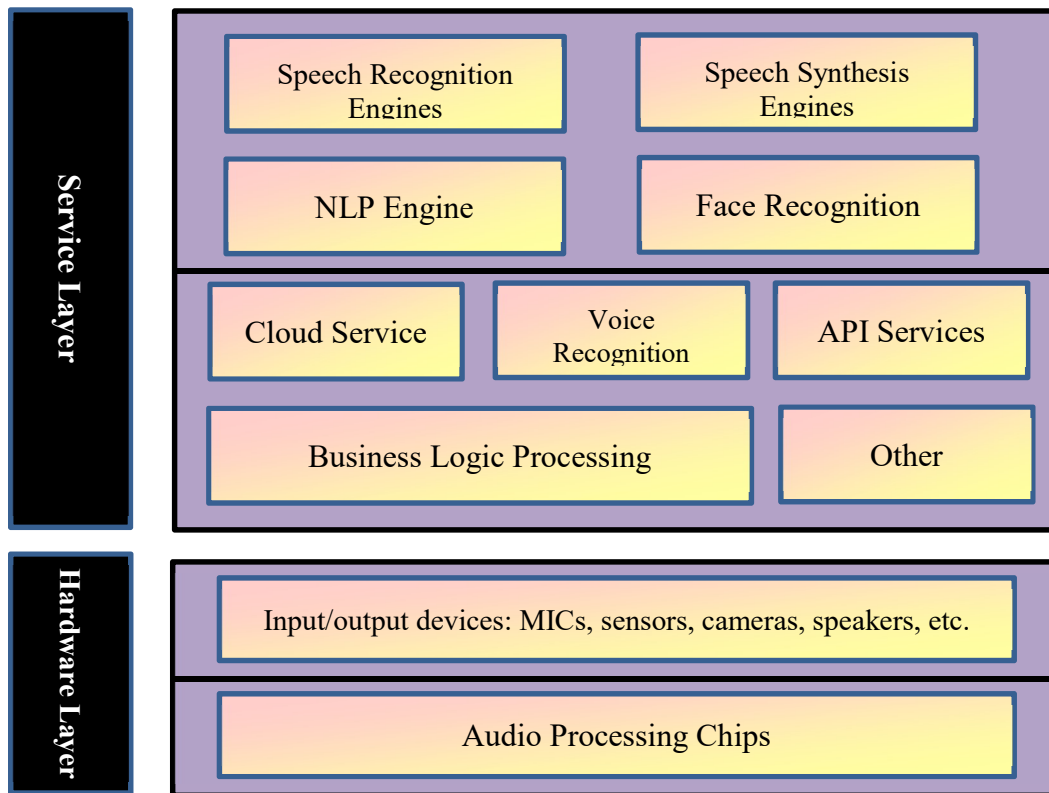


Fig. 1 A basic framework for voice interaction systems

Combining voice signal preprocessing and feature extraction technology, adding manual input destinations, and retaining some of the original machine's functions to ensure that all users can get started faster and more conveniently. The whole system can be divided into four components: hardware layer, service layer, application layer and operation and management platform.

(1) Hardware layer in the voice interaction system, the key hardware devices include input/output devices and audio processing chips, of which the input/output devices are responsible for collecting the user's voice commands and feedback information, mainly including: microphone arrays, speakers, cameras, sensors, lights, etc., the chip part of the main responsibility for the processing and analysis of the audio signals, which mainly includes the digital signal processor (DSP), audio decoder, audio amplifier, etc. Intelligent voice ticketing system interaction system is a highly complex integrated system, which not only requires sophisticated hardware and software to work together, but also requires professional operation and management to ensure its reliability and stability.

(2) The service layer is the core of the voice interaction system, which integrates the processing and parsing of multimodal data such as voice, image and video, and provides the important task of necessary feedback and response. It mainly includes modules such as natural language processing (NLP) engine, speech recognition engine, speech synthesis engine, voiceprint recognition, cloud service, API service, and business logic processing service.

(3) Application layer refers to the specific application programs provided to users based on the combination of the core capabilities provided by the service layer and the actual needs of the users, in order to help the users realize specific function control through voice interaction.

(4) Operation and management platform, users in the process of using the voice interaction system, will produce a large number of user behavioral data and audio, text, image data, operation and management platform through the statistical analysis of these data, to provide data support for the continuous optimization of the voice interaction system and various AI models. The basic framework of the voice interaction system is shown in Fig. 1.

4. Analysis of Experimental Arguments

For the elderly, deaf and mute people, foreigners and other difficulties in purchasing tickets, we have done two tests, the test results are shown in Table 1, Table 2.

(1) Tests affected by different scene factors at the time of speech recognition.

It performs well in the recognition of proprietary words such as station name, nearby location, and path query recommendation, which can be applied to subway voice ticketing scenarios. Although there is a small part of the words recognition error, but does not affect the user to buy tickets operation, in addition to the model itself can not completely accurate recognition of all the words of the intrinsic factors, but also by the speaker's accent, speaking environment and other external factors.

Table 1. Test 1

Serial No.	language input	Actual input	Identification
One	Where do I need to transfer to Marina International Airport	Where do I need to transfer money to Marina International Airport	"Station" is identified as "account"
Two	Here's a ticket to the cancer hospital	Here's a ticket to Oncology One	"Medical" is recognized as "one."
Three	How to get to Tianjin West	How to get to Tianjin West	error-free
Four	Where is Tianjin East Railway Station	Where is Tianjin East Railway Station	error-free
Five	Can my kid get a free ticket for 125	Can my kid get a free ticket for 125	error-free
Six	Where do military personnel purchase tickets	Where do military personnel purchase tickets	error-free

(2) Entry of different language types

Various speech recognition algorithms can solve most of the foreigner's in entering the subway because of a variety of language problems caused by a variety of uncharacteristic problems, all the words of the intrinsic factors in addition to the situation is also subject to the speaker's accent, the speaking environment and other extrinsic factors.

Table 2. Test 2

Serial No.	Language Input	Actual Input	Identification
One	Give me a ticket to the airport	Give me a take to the airport	"ticket" is recognized as "take"
Two	West station	West station	error-free
Three	How to transfer station	How transfer station	"to" is not recognized

The customers of the multimodal interactive voice ticketing system are mainly subway passengers. It provides intelligent and efficient ticketing services to the public by realizing operations such as passengers purchasing tickets at designated stations and purchasing tickets by fuzzy place name search through complete voice interaction. It also helps the elderly and visually impaired, as well as foreigners who have just arrived in the country, to purchase tickets more conveniently.

(3) Ticket Purchase Process

The customers of the multimodal interactive voice ticketing system are mainly subway passengers. The purpose is to provide intelligent and efficient ticketing services to the public by realizing operations such as ticketing at designated stations and ticketing by fuzzy place name search through completely natural voice interaction. In addition, it mainly helps the elderly and the visually impaired as well as foreigners who have just arrived in our country to purchase tickets more conveniently. The ticketing process is shown in Fig. 2.

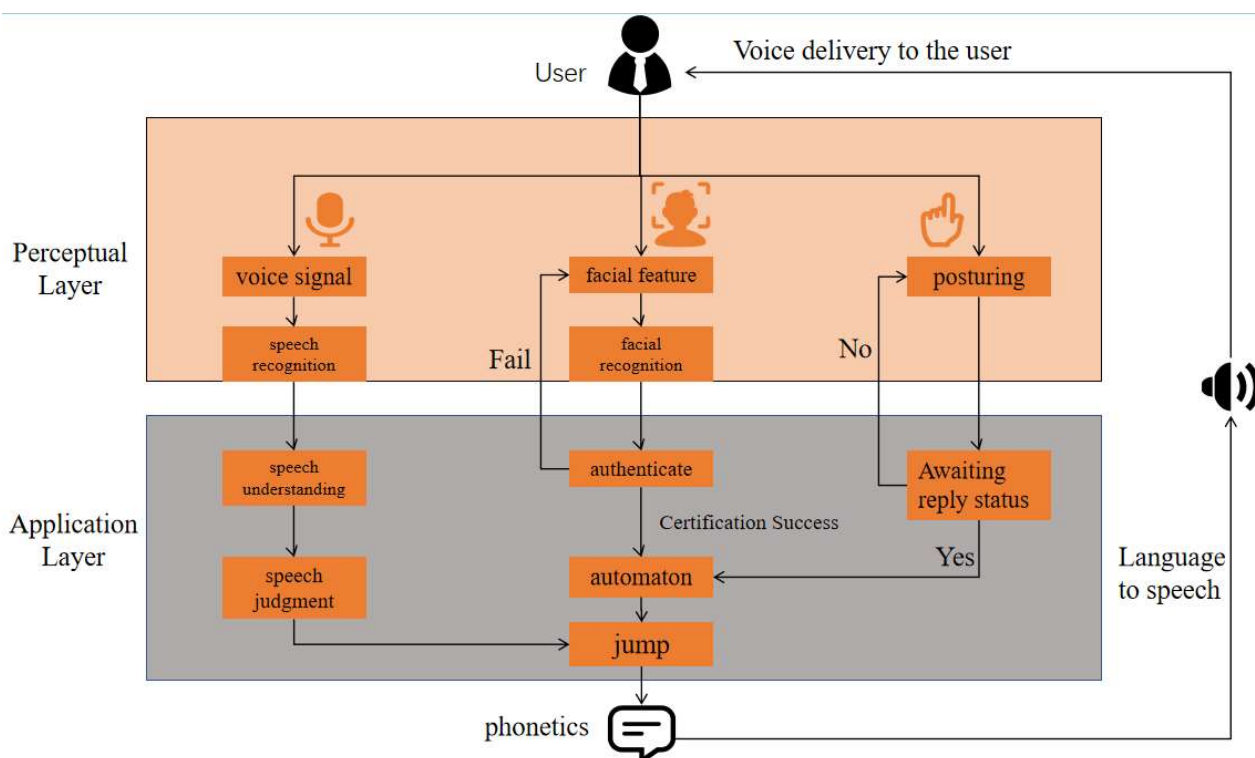


Fig. 2 Ticketing Process

The Multimodal Interactive Voice Ticketing System utilizes network technology and advanced automation technology to provide ticketing services to passengers through on-screen operation, which works as follows.

a) Structural composition

Multi-modal interactive voice ticketing system generally consists of host computer, interactive interface, payment module, printing module, including: autonomous ticketing, voice ticketing.

b) Software system

The automatic ticketing system realizes the communication with the backstage data system by installing the automatic ticketing software running on the host computer, so that it can complete the operations of ticketing, querying, refunding and exchanging tickets.

c) Payment Module

The automatic ticketing system supports a variety of payment methods, such as cash, bank card payment, scanning QR code and so on.

d) Printing Module

The automatic ticketing system will automatically generate tickets and print them through the printing module, and finally the tickets. The automatic ticketing system usually supports multiple languages, and passengers can choose the language that suits them according to their needs.

e) Select Ticket Type

Passengers can choose different types of tickets such as one-way ticket, round-trip ticket, student ticket, etc. as needed.

f) Input Ticket Information

Passengers need to enter the ticket information, including the date of travel, station, and number of trains.

g) Select payment method

The automated ticketing system will provide a variety of payment methods for passengers to choose from, including cash payment and bank card payment.

h) Waiting for ticket printing

After the operation is completed, the automatic ticketing system will automatically generate the ticket and print it through the printing module, and the passenger can pick up the ticket within a certain period of time and submit it to the passenger.

5. Conclusion

Compared with most of the existing manufacturers of the same type of products in the market, the multimodal interactive intelligent voice ticketing system can realize the passenger's designated station ticketing, fuzzy place name query ticketing and other operations through completely natural voice interaction, providing the public with intelligent and efficient ticketing services. Voice ticketing machine can realize accurate long-distance voice interaction in the public environment, and quickly give the recommended station through AI and other technologies, allowing passengers to select the station in less than one second. The time for passengers to select a station is shortened from more than ten seconds to one second. For visually impaired, hearing impaired and other special groups, the intelligent voice ticketing system should provide appropriate barrier-free services. Intelligent voice ticketing system should be equipped with intelligent customer service functions, able to answer the problems encountered by passengers in the process of purchasing tickets. Under the premise of protecting passengers' privacy, the intelligent voice ticketing system should have a certain degree of security protection, and can also be connected to the cloud server to achieve rapid software updates and continuous optimization of the function, which has a certain application value.

References

- [1] T.H. Zhang. A review of research progress in artificial intelligence technology[J]. *Electronic Components and Information Technology*, 2023, 7(08):69-72+76.
- [2] Y.Q. Chen, X.Y. Qian. Research on key technology and application of intelligent voice ticket machine for urban rail transit[J]. *Urban Rail Transportation Research*, 2022,25(10):149-153.
- [3] Annual Overview of Development and Social Progress of Persons with Disabilities in China (2023)[J]. *Disability Studies*, 2023(04):97.
- [4] L.X. Cai, J. Zhang. Research on the application of voice interaction technology in smart grid[J]. *Electroacoustic Technology*, 2023, 47(12):40-42+66.
- [5] J.K. ZHAO, Z.Y. KANG. Design scheme of intelligent voice-activated walking stick for the blind[J]. *Internet of Things Technology*, 2024, 14(02):149-151.