ISBN: 978-981-18-3959-7
2022 the 12th International Workshop on Computer Science and Engineering (WCSE 2022)
doi: 10.18178/wcse.2022.06.012

Construction and Analysis of Air Traffic Controller Fatigue Voice Database

Xiaojun Adu[1], Shiyu Huang[1], Xiao Huang[1], Jin Yang[1], Yuting Chen[1], Zetian Zhang[1] and Zhiyuan Shen[1+]

[1]College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China

Abstract. With the rapid increase in civil aviation traffic, the workload of air traffic controllers has also increased. In addition, the characteristics of irregular day and night duty of controllers make controllers more prone to fatigue in the process of duty, thus causing aviation safety hazards and reducing air operation efficiency. Emotional voice recognition is considered to be a non-intrusive fatigue detection method that is more suitable for controllers in daily command than video and eye tracker, etc. In this paper, the first fatigue voice database is constructed based on air-ground communication of controllers. Firstly, the voice data of air-ground communication of controllers are cut and preprocessed, and the traditional voice features are extracted. Then, according to the five specific characteristics of "age group", "gender", "control years", "license level" and "control position", the 700,000 controlled voices in East China are classified. The fatigued speech database for different controllers is constructed by sample library marking by experts. The large database of controller fatigue established in this paper can provide effective data support for experts and scholars to detect fatigue voice and verify corresponding algorithms. By detecting and identifying the fatigue state of controllers, the scheduling plan and training program of controllers can be further optimized, the errors of controllers can be reduced, and the occurrence of air traffic control unsafe incidents can be prevented.

Keywords: air traffic controllers, fatigue, speech, database

1. Introduction

Nowadays, China's economy maintains rapid and stable development, which promotes the vigorous development of China's civil aviation industry. With the further opening of the airspace, the civil aviation industry will be developed further in the future, and the transformation from a civil aviation country to a civil aviation power is just around the corner. As of December 2020, China had about 600 airports, including 76 international airports, and 150 new airports will be built in the next 15 years. Although the COVID-19 pandemic has had a significant impact on the civil aviation industry since 2020, air cargo volumes are still growing.

Air traffic controller plays a vital role in the whole transportation link, to ensure the safety of transportation. The job of an air traffic controller is as follows:

(1) The takeoff, landing, taxiing and airspace flight of the aircraft are under the command of the ATC, who must maintain communication with the pilot as all times.

(2) The ATC is responsible for monitoring aircraft to ensure the certain safety interval must be maintained to prevent collisions between aircraft and aircraft or between aircraft and obstacles.

(3) The ATC needs to iinform the pilot about the weather condition of the forward route.

(4) Assist the pilot to ensure the safe, fast and efficient flight of the aircraft, etc.

Each controller is responsible for certain control area, and all the aircraft in the area are under control . That means the traffic controllers need to control many aircrafts at the same time, which have a high request for their situational awareness, and this will increase the workload of controller to cause them fatigue. Therefore, it is necessary to collect and classify the risk factors of controller fatigue. ATC fatigue databases is

[+] Corresponding author. Tel.: + (13951916587); fax: +(025-84893461).
E-mail address: shenzy@nuaa.edu.cn.

an important method to statistics and analysisATC workload , and is used to judge the fatigue degree and load range of the controller .On this basis, we are able to provide real-time effective controllers fatigue state monitoring and alarm, improve the quality of ATC training, and provide reasonable arrangement of work for ATC.So that we can alleviate the pressure of the controller's work effectively, and keep the controller's workload within the normal pressure range to ensure aviation safety [1].

The ATC fatigue database focuses on voice, and the method to build the fatigue database is to classify and integrate the characteristics of "age group", "gender", "control years", "license level", "control position" and other distinguishing points. Compared with the existing methods for detecting and analyzing controller fatigue, it is more convenient to collect voice samples, with more intuitive data features and richer resources. Besides, different controllers have different voice features, which is conducive to the diversification of fatigue database and more intuitive reflection and detection of controller's state from various aspects. It is convenient for front-line controllers to command and monitor aircraft in the air and on the ground scientifically and effectively, and improve the safety of civil aviation further.

This paper will discuss four aspects: the current situation of emotional speech database, fatigue database data preprocessing, controller data marking, and the role of large database.

2. The State-of-the-art Emotional Speech Database

Fatigue is a comprehensive physiological or psychological state of the human body. Fatigue can be caused by many reasons, such as being overworked, stressed, or exercising vigorously. Fatigue can lead to longer reaction time, poor concentration, reduced work ability and efficiency, and reduced thinking and judgment. In addition, in the fatigue state, the speech rate, intonation, short-term energy, short-term average zero-crossing rate, fundamental frequency, formant, Mel Frequency Cepstrum Coeefficient (MFCC), MFCC logarithmic power ,the spectrum will change [2], in this way the fatigue state can be detected based on these features. Furthermore, language is an important way of expressing emotions, and it is also the most natural, effective and fastest ways for human beings to transmit information, while speech is the sound expression of language. In the state of fatigue, the emotional expression of speech will also be affected. Therefore, speech emotion recognition is of great significance for fatigue detection.

At present, in the world, speech emotion research is still in the stage of exploratory development, and the emotional speech database is the data foundation of speech emotion recognition research. It is particularly important to establish a high-quality emotional speech database. In recent years, scholars in this field at home and abroad have established many emotional speech databases under this background.

The main research abroad is the construction of language emotion speech database, such as the construction of Persian spoken language database., such as the construction of the Persian spoken language database. The database is collected from 1126 native speakers of Persian, which are divided into 6 types of anger, disgust, fear, happiness, sadness and neutrality emotion, and is composed of 34 native Persian speakers to evaluate the effectiveness of the database.

Similar in China, the main research at present is the construction of Chinese emotional speech database.For example, the speech database established by Tsinghua University, which collects 25 boys and 25 girls, uses discrete emotion category labels to describe emotions, which are divided into five emotions: happy, angry, fearful, sad, and neutral. Each emotion has 200 sentences, is the existing large speech database in China. And the database also considers the dual needs of speech emotion recognition and text content recognition, so it is suitable for both emotion recognition and speech recognition research [4]. In addition, a speech database is constructed based on the acoustic characteristics of speech. The database is collected from 3 males and 2 females. Each person speaks 50 sentences, and each sentence is assigned happiness, sadness, anger, fear, surprise and neutral six emotions, a total of 1500 valid recordings were obtained, and 30 acoustic and emotional features were extracted from each recording to build a database [5].

But in general, these existing emotional speech databases at home and abroad are oriented to a wide range of research, or only for the daily use of a certain language. In contrast, the corpus dedicated to fatigue speech is relatively scarce, and there is a lack of large database construction for special groups such as ATC, which has caused great difficulties for scholars engaged in related research.

3. Fatigue Voice Data Processing

3.1. Voice Cutting

Controller works mainly by voice command to establish contact with the pilot, and the voice data during the control process will be retained for a period of time according to the regulations. The voice data contains many physiology characteristic of the human body and information.Therefore we can detect controllers fatigue by extracting speech characteristics and analysis it.t. Compared with traditional controller fatigue detection methods, such as analysis of controller blood, urine, facial recognition and other methods, the detection of controller voice call data is more convenient and effective.

During the control process, voice data is retained for a specified period of time. Since the controller is constantly communicating with the pilot during the work process, we need to separate the controller's voice data from the pilot, and only obtain the controller's voice dataso that the subsequent extraction of voice fatigue features can be carried out. Cut the raw data with GoldWave software, as shown in Figure 1.

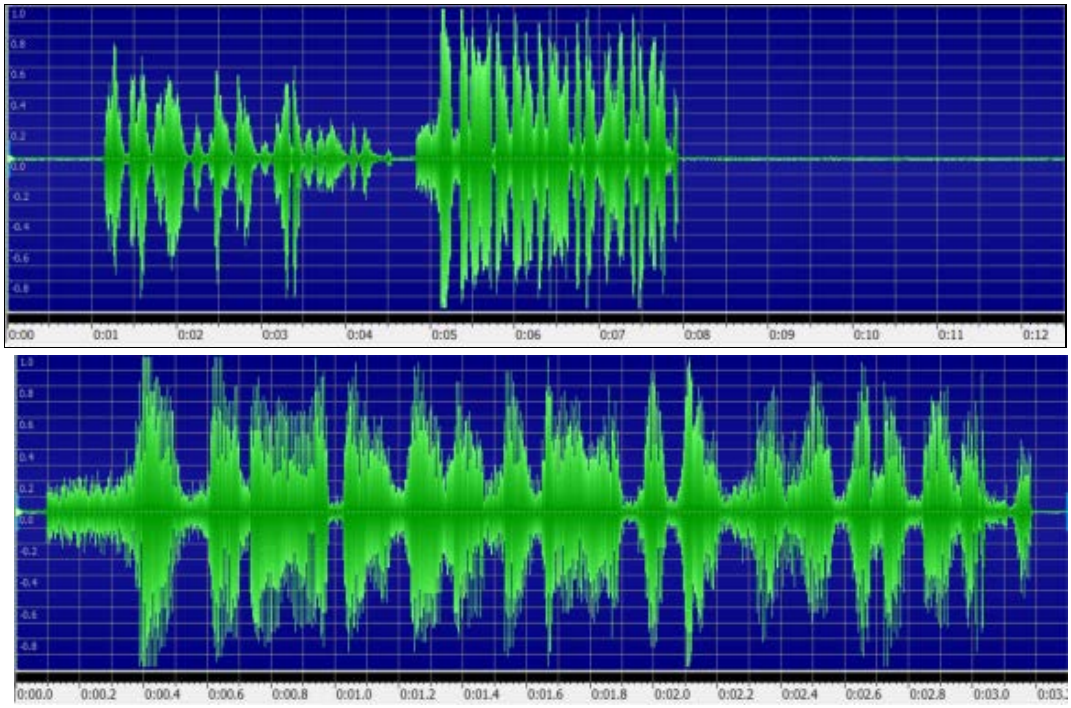

Fig 1: Spectrum diagram comparison of speech data before and after cutting

The left half of Figure 1 is the voice spectrum of the pilot, the right half is the voice spectrum of the controller, which is cut by GoldWave software to obtain the spectrum graph which only contain the voice spectrum of the controller.

3.2. Voice Processing

The audio signal of speech has the characteristics of non-stationary time-varying and short-term stationary [6]. The speech data is segmented into frames, and then the fatigue characteristic information is extracted from each frame. The statement is preprocessed first. Pre-processing mainly includes endpoint detection, pre-weighting and windowing frame processing [7]. Endpoint detection adopts a method based on zero crossing rate and energy double threshold [7].

3.2.1. Framing

High pass and low pass filters were used to remove some low and high frequency noise from the controller's audio signal. Nowadays the control equipment is more and more advanced, the microphone of the controller in the developed area generally has the function of automatic filtering low frequency noise.If the microphone quality is not good ,it will cause a lot of noise interference, and we need to use high and low pass filter.After that, the controller's audio signal is divided into frames. The purpose of frame splitting is to subdivide the audio signal so as to obtain finer voice features, because the audio signal of voice has

long-term instability and short-term stability.For example, if a speech audio signal is 64kHz and the frame length is 25ms, there are 0.025×64000=1600 data points in each frame.

3.2.2. Feature Extraction

After the completion of frame segmentation, feature extraction can be carried out for each frame dataThe commonly used feature extraction methods are as follows: Linear predictive analysis (LPC), which simulates human phonation principle, is obtained by analyzing the model of short tube cascade of vocal tract. Perceptual linear predictive coefficient (PLP), a characteristic parameter based on auditory model. The Tandem feature is the posterior probability vector dimensionality reduction of the corresponding category of nodes in the output layer of the neural network and the splicing of the features such as MFCC or PLP.Bottleneck features are extracted using a special neural network.Linear predictive cepstrum coefficient (LPCC), an important characteristic parameter based on the channel model.Mayer frequency cepstrum coefficient (MFCC) : Based on the auditory characteristics of human ear, mayer frequency cepstrum frequency band division is isometric division on the Mel scale. The logarithmic distribution relationship between the scale value of frequency and the actual frequency is more consistent with the auditory characteristics of human ear, so it can make the speech signal have a better representation.

The transformation method between Mel frequency and linear frequency is as follows:

$$Mel(f) = 2595lg(1+ f / 700) \tag{1}$$

Where f is the frequency, in Hz.

At present, there are several popular methods for the extraction and detection of feature parameters, such as the speech feature extraction method of Yao Hui [7], which extracts five prosodic features and MFCC features: speed of speech, zero crossing rate, energy, fundamental frequency and formant [7]. The speed of speech is determined by the ratio of the number of syllables you pronounce to their duration.

Speed:

$$V = \frac{S}{T} \tag{2}$$

S is the number of syllables and T is the duration.

The frame of energy:

$$E = \log(\sum_{i=1}^{l} x(i)^2) \tag{3}$$

The data of each frame is x_i and the frame length is l.

The zero crossing rate is the number of times that each frame of data passes through zero, the fundamental frequency is the frequency band with the characteristics of the lowest frequency and the largest amplitude in the spectrum, and the resonance peak is some regions in the spectrum where the energy is relatively concentrated.

Li Xiang, Li Guozheng et al. [8] proposed a driving fatigue detection method based on speech feature transfer learning. The proposed method has a significant driving fatigue detection effect, with the highest accuracy rate reaching 86.7% [8].Li Xiang, Li Guozheng et al. [9] proposed a human fatigue detection method based on speech spectrum image features [18], which follows the physical meaning, acoustic characteristics and fatigue characterization mechanism of spectrogram in feature extraction, thus achieving better fatigue detection results than existing methods.The proposed method achieves the highest accuracy of 89.3%, 91.4% and 89.7% in human fatigue detection for all experimental speech samples [9], reaching a relatively high level.In this method, image texture features based on gray co-occurrence matrix (GLCM) are used to describe the fatigue information contained in spectrogram.In the transverse time domain, voiceactive Detection (VAD) is used to segment speech.The Mel scale is evenly divided into 24 overlapping critical frequency subbands using the critical band segmentation idea of triangular filter banks.According to formula (1), the upper and lower bound frequencies of each critical rate subband can be obtained.A total of 15 GLCM feature parameters, such as Angle second moment, contrast and anisotropy [9], are selected for speech fatigue classification.

3.2.3. Training and Testing of Voice Data

After the completion of frame segmentation and speech feature extraction, the BP neural network model of Yang Changqi et al. [10] is applied to train and test it based on the existing database. The number of training sets accounts for 80% and the number of test sets is 20%.When the machine model active learning has the fatigue ability to recognize speech data, it can take the initiative to make more marks in the database.

3.3. Fatigue Voice Label by Expert Assessment Method

To prevent software models from making big mistakes, expert evaluation methods are used. In this way, raw voice data is used, which is the raw data of the communication between the controller and the pilot before the cutting. According to the tone, intonation, speed, voice size, reaction time, and the incidence of control command errors in the cut voice data of the controller, experts judge that the controller is not fatigued, slightly fatigued or seriously fatigued, as shown in Table 1.

Table 1: Subjective fatigue degree determination table

Duty period	The phonetic characteristics	The reaction time	The degree of fatigue
The voice data of duty periods prone to fatigue should be selected to determine the fatigue degree, such as late night, lunch, shift time and peak flight time	Speaking speed coherent smooth, accurate and clear pronunciation	Respond within 2 seconds	No fatigue
	Speaking slightly slower or higher, with slightly unclear pronunciation	Respond within 2 seconds	Slight fatigue
	The speed of speech is obviously slowed down or promoted, ambiguous pronunciation or command error, repetition, interruption, pause and so on increase, the words are not clear, the voice becomes low and turbidness, increased noise	No response within 2 seconds	Severe fatigue

4. Database Construction

4.1. Controllers Information Setting

From the aspects of gender, controllers is a job with higher working load and pressure, and young men controllers is the main force of our country control line, this is because the men's own constitution and shift system adaptability are superior to women.At the same time, studies also show that male cognitive ability and the alert level slightly higher than the female. In terms of posts, as the complexity of regional control sector is greater than that of terminal control area, with the increase of air flow, the workload of regional control is greater than that of approach control, and the workload of approach control is greater than that of tower control. From the perspective of working years and control levels, the occurrence of stress is related to the failure of coping ability to meet the task requirements, that is, controllers with lower working years and ranks may be more prone to stress due to lack of experience, resulting in fatigue [11].

Information for ATC tag in the first place, will all the voice and data preprocessing based on different control categories into the area control, approach control and tower control, in each kind of control under the category of the voice data is divided into different periods, according to the rota will work the same time of the controller's voice data into a class, Finally, the voice data of each time period under each control category are divided into male and female according to gender. To sum up, each piece of voice data obtained contains three characteristic markers of control category, duty period and gender, such as: Zhang SAN/male/approach / 14:00 -- 16:00 /.

As the sound characteristics of controllers are directly related to their performance level, and the latter is affected by the fatigue degree of controllers, the fatigue state of controllers can be detected to a certain extent by marking the sound characteristics [12].So we are with the individual controllers as the research object, to ATC voice response performance characteristics of "quality" as the research key, select a controller of a historical period of fatigue state voice response data (data) as the overall sample space, statistics, summarized the distribution, and the distribution is defined as the atc voice "normal" distribution of the reaction time. The performance of the voice response data in this period was compared with that of the

"normal" distribution data, so as to evaluate the "good or bad" performance of the controller in this period, so as to judge the fatigue level of the controller. Sound features can be divided into features related to prosody (pitch, intensity, rhythm, pause pattern, speed of speech), articulation (ambiguities, reduction, and ellipsis), and speech quality (breathing, tension, sharpness, hoarseness, modal speech) according to the aural-perceptual concept. Based on the existing data, the following physiological changes related to fatigue and sleepiness can affect sound characteristics: reduced cognitive processing speed (impaired speech planning, increased fluency rate) and gentle and slow breathing (hypoglottic depression), corresponding acoustic effects are articulation intensity, articulation accuracy and articulation speed [13,14].In the fatigue state, the voice response speed of the controller slows down significantly, and the stability of the response speed decreases. Meanwhile, in the fatigue state, the voice response ability of the controller will become lower as time goes on, and the reaction ability will decline, that is, the response becomes slow and the stability decreases [15].

4.2. Fatigue Voice Labeling

Secondly, preprocessed speech data were labeled on the basis of the above relevant characteristics, and all data were divided into three categories: non-fatigue, slight fatigue and severe fatigue. After data marking processing, each data is marked in the form of "name/sex/control type/control level/age/control years/fatigue level". Table 2 is obtained after sorting out the data.

Table 2: Statistical table of fatigue speech data marking results

Number	Gender	Operating post	Control level	age	Work year	Fatigue data information		
						non-fatigue	slight fatigue	severe fatigue
①	man	area control	Level 5	24	2	1187	479	99
②	man	area control	Level 4	26	5	1300	360	100
③	man	approach control	Level 3	30	8	1210	456	120
④	man	Approach control	Level 2	40	10	1298	368	223
⑤	woman	tower control	Level 4	27	5	1167	499	50
⑥	woman	tower control	Level 3	31	8	1314	352	67
⑦	woman	area control	Level 2	38	12	1036	630	167
⑧	woman	area control	Level 4	28	6	1200	500	88

5. The Role Analysis of ATC Voice Fatigue Database

5.1. Optimizing Controller Scheduling

With the rapid development of air traffic, the pressure of controllers increases sharply and shifts are fatigued. The traditional manual scheduling mode is time-consuming and labor-intensive [17]. The fatigue analysis data based on the controller's fatigue voice database can be used as an objective constraint condition for controller scheduling to optimize the controller's scheduling. It is used to solve the shortcoming that the head director only relies on personal experience when scheduling shifts and reduces control errors caused by fatigue from shift work.

5.2. Provide Data Support for Fatigue Monitoring

A complete large database plays an important role in training an efficient and accurate controller fatigue state monitoring system [18]. The controller fatigue voice database can be used to form the basic data structure of the controller fatigue monitoring system, which is convenient for further research and establishment of a large fatigue database for each controller in the future, which is helpful for the research of domestic and foreign scholars. The voice data in the controller fatigue voice database almost covers the Chinese and English commands of controllers of different ages and various positions in different scenarios. It provides

comprehensive data support for the follow-up monitoring and identification of controller voice fatigue and the development of controller fatigue management research.

5.3. Analysis of Air Traffic Control Safety Red Line

The voice characteristics of Radiotelephony Communication basically reflect the fatigue level of controllers [19], and the data in the controller fatigue voice database has been distinguished according to different fatigue levels. The change process of the controller's fatigue state in the control work and put forward the control safety red line.

The fatigue voice database is aimed at first-line controllers and is established on the basis of the fatigue level of the controllers in their daily control and command, so the database can be applied to the training of air traffic controllers firstly. During the training process, the control instructor uses the voice data of severe fatigue and mild fatigue in the fatigue voice database to identify the fatigue level of the controllers and test the training quality. At the same time, the voice data of Radiotelephony Communication during the training continue to enrich the database, forming a closed loop. , so that the control level of the controller has been continuously improved. Secondly, the data in the fatigue voice database can be integrated into the air traffic control automation system. According to the method proposed in <A fatigue driving warning method and system> [20], the voice of Radiotelephony Communication is collected and recognized from the controller microphone. Match with the data in the controller's fatigue voice database, and judge whether it is in a fatigue control state according to the matching degree, so as to prevent unsafe air traffic control incidents.

6. Conclusion

This paper presents a new public data set, the ATC Fatigue Large Database.The data set is classified and integrated by marking the controller's "working age", "gender", "post control years", "license level", "control post" and other characteristics as the distinguishing points, and the common controller fatigue characteristics are extracted.The characteristic quantity and characteristic index for each individual or group of controllers and for each control situation were used to identify whether the controller was tired. Finally, the fatigue detection method is given, and the first fatigue speech database for individual controllers is established in China, which provides data support for the further study of fatigue state of controllers by domestic and foreign scholars.

The diversity and comprehensiveness of the database can provide more comprehensive data support for the subsequent monitoring and recognition of controllers' speech fatigue and the study of controllers' fatigue management. The complete large database plays an important role in training the efficient and accurate controller fatigue monitoring system and improving the automation level of air traffic control.

7. Acknowledgements

The authors thank the national Natural Science Foundation of China (NO.71874081) and the Innovation Training Project of Nanjing University of Aeronautics and Astronautics (NO.202110287138Y).

8. Reference

[1] Zhang Ming, Han Songchen, PEI Chenggong.A review of air traffic controller workload research [J]. Human ergonomics,2008,14(4):45-46.

[2] zhang xishi.Speech fatigue detection based on deep learning [D]. Beijing Jiaotong University,2020.

[3] Keshtiari Niloofar et al. Recognizing emotional speech in Persian: a validated database of Persian emotional speech (Persian ESD).[J]. Behavior research methods, 2015, 47(1) : 275-94.

[4] PAN Yuchun, XU Mangling, LIU Linquan, et al. Emotion-detecting based model selection for emotional speech recognition[C]//Proc lMACS Multiconference on Computational Engineering in Systems Applications (CESA) .Beijing,China, 2006:2169-2172.

[5] Zeng G J. The establishment of emotional acoustic feature database of Mandarin speech [D]. Sichuan: University of Electronic Science and Technology of China,2012. DOI:10.7666/ D.d764120.

[6] Sun Ying, Song Chunxiao.Journal of xidian university,2017,44(06):162-168.

[7] Yao Hui, Sun Ying, Zhang Xueying.Journal of xidian university,2016,43(05):167-172.

[8] Li Xiang, Li Guozheng, Peng Liqun, Yan Lixin, Zhang Chi.Journal of the China railway society,2020,42(04):74-81.

[9] Li Xiang, Li Guozheng, Deng Mingjun, Wan Wan, Yan Lixin.Chinese Journal of Scientific Instrument, 201,42(02):123-132.

[10] Yang Changqi, TAN Juan, QIU Zhengping.Research on controller fatigue state recognition based on BP neural network [J]. Aeronautical computing technology,2018,48(01):17-20+25.

[11] Yuan Leping, Sun Ruishan, Liu Lu.Research on controller workload measurement method based on DORATASK [J]. Journal of safety and environment,2014,14(3):76-79.

[12] Yang Changqi, TAN Juan, QIU Zhengping.Research on controller fatigue state recognition based on BP neural network [J]. Aeronautical computing technology,2018,48(01):17-20+25.

[13] Liu Zhen. Quantitative evaluation model of fatigue risk based on controller speech response [J].Science and Technology and Innovation,2018(8):37-39.]

[14] ZHOU G., HANSEN J.H.L., KAISER J.F.. Nonlinear feature based classification of speech under stress[J]. IEEE Transactions on Speech and Audio Proceeding: A Publication of the IEEE Signal Processing Society,2001,9(3):201-216.

[15] CHEN, MEI-LIEN, LU, SHIH-YI, MAO, I-FANG. Subjective symptoms and physiological measures of fatigue in air traffic controllers[J]. International Journal of Industrial Ergonomics,2019,701-8.

[16] Zhang Xingjian, YUAN Leping, Zhao Yi Fei.China safety science journal,2018,28(6):37-42. (in Chinese)

[17] Yan Danqing, Zhang Yong, Jin Shazhou, Research on Air Traffic Controller Scheduling, China Science and Technology Information, 20119.22 :43-44.

[18] Li Haiyan, Meng Yu, Chen Xi.Study on fatigue and workload of air traffic controllers [J]. Youth,2020,71(27):103-104

[19] Sun He, Jia Qi, Liu Chang, Study on speech characteristics of Air-ground Communication under Fatigue Condition, China Safety Science Journal.2020.30(02)

[20] Guangdong Opo Mobile Communication Co., LTD.Fatigue Driving Warning Method and System: CN201410493358.2[P].2015-10-1