Research Article

A High-Precision Fatigue Detecting Method for Air Traffic Controllers Based on Revised Fractal Dimension Feature

Zhiyuan Shen ¹, Guozhuang Pan ¹ and Yonggang Yan^{1,2}

¹College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China

²Air Traffic Management Bureau of Civil Aviation Administration of China, Beijing 100022, China

Correspondence should be addressed to Zhiyuan Shen; shenzy@nuaa.edu.cn

Received 18 October 2019; Revised 5 January 2020; Accepted 18 January 2020; Published 12 February 2020

Academic Editor: Alberto Campagnolo

Copyright © 2020 Zhiyuan Shen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As air traffic volume increases, the air traffic controller (ATC) fatigue has become a major cause for air traffic accidents. However, the conventional fatigue-detecting methods based on speech are neither effective nor accurate because the speech signals are nonlinear and complicated. In this paper, an ATC fatigue-detecting method based on fractal dimension (FD) is proposed. Firstly, a special speech database of ATC radiotelephony communications is constructed. These radiotelephony communications are obtained from Air Traffic Management Shandong Bureau of China. Then, speech signals implement a wavelet decomposition and FD calculation. The calculation result shows the significant difference among the FD of the speech signal before and after fatigue. Furthermore, a novel fatigue feature of the ATC based on the FD of speech is built. A series of experiments are conducted to detect the ATC fatigue with the fatigue feature comparison process and a support vector machine (SVM). The results show that the accuracy in detecting ATC fatigue based on FD was 92.82%, which are higher than the state-of-the-art methods. The research provides a theoretical guidance for Air Traffic Management Authority on detecting ATC's fatigue, while it may provide reference for the fatigue assessment in other professional fields of civil aviation.

1. Introduction

The rapid development of civil aviation results in continuing increases in the volume of air traffic in China. Despite the rapid growth in the number of flights, the growth rate of the number of ATCs (air traffic controllers) is relatively slow [1]. The associated increasing working pressures are making ATCs more vulnerable to fatigue. This situation has led to frequent air traffic accidents in recent years [2]. The accurately detecting method of ATC fatigued state has gradually attracted the attention of experts and scholars in the field of civil aviation.

The fatigue-detecting methods for ATCs can be divided into subjective and objective methods [3]. A popular subjective method is the fatigue scale. The fatigued state is detected according to the score on the scale [4]. For example, Chalder et al. reported the Chalder fatigue scale [5]. Subjects were asked to fill out the scale before and after work. Although this method is easy to implement, it cannot detect the

fatigued state either rapidly or accurately. Therefore, the objective method has received a considerable amount of research interest of researchers. A popular objective method uses instruments, equipment, and other auxiliary tools to determine the fatigued state. This kind of method records changes in certain indicators of human physiology, biochemistry, behavior, or other characteristics. Objective methods can be also subdivided into contact and noncontact detection methods. This classification is based on whether the detection tool needs to be in contact with the human body when attempting to detect fatigue. Contact detection methods mainly detect the fatigued state by recording changes in physiological indicators of the tested person (such as the EEG, ECG, or heart rate) [6–8]. Some European researchers have proposed obtaining the fatigued state of ATCs by analysing five physiological indices. These indices involve skin voltage, skin conductivity, skin blood flow, body temperature, and instantaneous heart rhythm [9]. Chen et al. reported the subjective symptoms and physiological

measures of fatigue in air traffic controllers [3]. This study was carried out on 102 ATCs in Taiwan. The tests collected the physiological information of flicker fusion threshold, thumb/index finger strength, and systolic and diastolic pressure. The process of detecting requires physical contact with the tested person. Although this detection method is highly accurate, its applicability is poor. Physical contact potentially interferes with the work they are performing. That also makes it difficult to implement.

Noncontact detection methods mainly employ facial- or vocal-feature-based detection method. The facial-feature-based detection method involves collecting images of the face. These images are used to detect the fatigued state based on facial-features (such as movements of the eyes and mouth) [10, 11]. Nie et al. found the significant difference among the indicators before and after experiment the experimental subjects undergo fatigue [12]. These indicators include eye closure time, percentage of eyelid closure (PERCLOS) value, frequency, and number of blinks. The fatigue of experimental subjects is obvious when eye closure time is 3.5 s/min, PERCLOS value 6%, and blink frequency 0.4 times/s. Di Stasi et al. improved the accuracy of fatigue test results by studying how the characteristics of eye movement reflect a fatigued state [13]. Here, 12 subjects were asked to perform a prolonged and demanding visual search task under fixation conditions. A major disadvantage of this method is that the image acquisition equipment needs to remain in front of the face of the subject while collecting facial images. This situation induces a certain psychological pressure. In contrast, the vocal-feature-based detection method is currently receiving a considerable amount of research interest due to its high accuracy and applicability [14]. Krajewski et al. introduced a general framework for detecting accident prone fatigue states based on prosody, articulation, and speech quality-related speech characteristics [15]. The advantage of this measurement method is that the speech data used for detecting fatigue can be easily obtained without needing sensors or calibration. Krajewski et al. subsequently proposed a framework for detecting a fatigued pattern that combines speech nonlinear dynamics analysis with a machine-learning classification algorithm [16]. The utilization of speech nonlinearity greatly improves the accuracy of fatigue detection. Deng et al. proposed an internal model-based neural network control for unknown nonaffine discrete-time multi-input multi-output (MIMO) processes in the nonlinear state space form under model mismatch and disturbances [17]. Palo et al. proposed a time-frequency source feature for emotional speech classification [18]. The feature spans four dimensions. This method improves the detection speed while maintaining the detection accuracy.

The remainder of this paper is organized as follows. In Section 2, a time-frequency vocal source feature is introduced. Then, the relationships between radiotelephony communication and the chaos in a speech signal are reported. In Section 3, a revised feature called the speech wavelet fractal feature (SWFF) is proposed. The process of constructing this feature is described in detail. Then, a novel fatigue-detection method for ATCs based on SWFF is

proposed. In Section 4, a series of experiments are conducted. In Section 5, conclusions and future research directions are presented.

2. Chaos in Radiotelephony Communication

2.1. Radiotelephony Communication. The radiotelephony communication recorded by an ATC-speech recording system in each time period contains all the voices of ATCs and aircrews. Therefore, it is necessary to extract the speech data of the ATCs firstly. According to the characteristics of radiotelephony communication, semantic recognition is used to achieve this goal. Radiotelephony communication refers to a method of issuing and executing instructions between ATCs and aircrews. This communication has the following characteristics. When the ATC talks to an aircrew, they hold down the button on the headset cable. They release the button after the end of the call when the aircrew first establishes contact with the ATCs. The communication structure adopted by the ATC is aircraft call sign + control-unit code + content, while that of the aircrew is control-unit code + aircraft call sign + content. After the first call, the call structure of the ATCs is aircraft call sign + content and that of the aircrew is content + aircraft call sign.

This revealed that the ATC always states the aircraft call sign first when they reply to the aircrew. The aircrew uses the aircraft call sign as the closing remark at all times except when first establishing contact with the ATC. Therefore, the aircraft call sign should be the beginning of the speech transmission from the ATC. The aircraft call sign at the time of receiving is used as the sign to end the call.

These features make it relatively easy to timely distinguish the voice of ATCs. An audio editing software (GoldWave) was used to intercept the voice of ATCs from the radiotelephony communication at different times. In addition, fatigue is more likely to make lead ATCs to be frequent pauses, hesitation, or even speak incorrect instructions when they issue control instructions. This is also consistent with the known physiological mechanisms of fatigue. Therefore, in this paper, when there is a problem with the instructions, ATC is fatigue.

2.2. Hurst Vocal Source Feature. Vocal source features extracted from speech signals contain important information about the distribution of harmonic [19]. In addition, the characteristics of the excitation source affect the spectrum envelope of short-term speech. Because these different sounds have different characteristics, vocal features have previously been studied for automatic detection of emotion in speech [20]. The Hurst Vocal Source Feature (HVSF) introduced in this paper is a time-frequency feature used in a speaker recognition and verification system. This feature consists of a vector containing the Hurst index. It is closely related to the excitation source. The Hurst index ($0 < H < 1$) indicates the time correlation or scale of speech signal $\{X(i)\}_{i=1}^N$. Its autocorrelation coefficient function (ACF) decreases gradually in the following form:

$$\rho(k) \sim H(2H-1)k^{2(H-2)}, \quad k \longrightarrow \infty. \quad (1)$$

The value of H can be associated with the spectral characteristics of $\{X(i)\}_{i=1}^N$. The proposed HVSF can represent the emotional state of the person speaking. In the process of extracting the HVSF, time-frequency multi-resolution analysis captures the high-order correlation of speech samples. This correlation is also found when source features are extracted from linear prediction residuals. Therefore, HVSF is closely related to the characteristics of the excitation source. This relationship can be utilized for recognizing emotions [21].

The process of HVSF extraction based on wavelet analysis [18] is as follows:

Step 1: speech signals are decomposed into approximate coefficients ($a(j, k)$) and detail coefficients ($d(j, k)$) using the discrete wavelet transform. The j is the decomposition scale ($j = 1, 2, \dots, J$). The k is the coefficient index of each scale.

Step 2: for each scale j , variance $\sigma_j^2 = (1/n_j) \sum_k d(j, k)^2$ is derived from the detail coefficient, where n_j is the number of possible coefficient values of each scale.

Step 3: weighted linear regression is used to obtain the relative slope of $(j, \mathcal{Y}_j = \log_2(\sigma_j^2))$. The value of H is obtained as $H = ((1\alpha)/2)$.

Step 4: the HVSF is composed of $j+1$ values in $H[H_0, H_1, \dots, H_j]$. Component H_0 is calculated from the original speech signal. Other values ($H_1, \dots, H_j$) are obtained by repeating steps 1 to 3 for each j detail coefficient sequence. The solution process of HVSF for $j=3$ (i.e., $[H_0, H_1, H_2, H_3]$) is shown in Figure 1.

Figure 2 depicts the distribution of the H values of speech signals corresponding to fatigued and normal states. The speech signals here were randomly selected from the speech instructions of ATCs. The fatigued and normal states are distinguished by whether the ATCs can continue to perform their duties normally (i.e., whether the speech given by the ATC has the problems mentioned in Part 2.1). The subjective feelings of the ATC and the Chalder fatigue scale are also considered. In this project, the Daubechies wavelet filter was used to decompose the discrete wavelet transform. It can be seen that there is no obvious difference in the distribution of the H index between the fatigued and normal states.

This problem is due to three limitations of applying the HVSF to detect fatigue in ATCs. First, the H index cannot adequately indicate the changes in chaos of radiotelephony communication. Second, unlike the speech data in the Berlin Database of Emotional Speech (EMO database) or other databases of emotional voices, the fatigue detection of ATCs is based on the speech characteristics of different voices (different speech contents). The radio communication itself is also distinctive and has been discussed above. The speech data in EMO database or other databases are derived from the speech produced by the same person in different emotional states in which the semantic content is the same. This is obviously not practical for the present application. Third, the speech data of ATCs are contaminated by noise

due to influences from equipment and environment associated with the data recording process. A revised vocal feature for detecting ATC fatigue is proposed as follows.

2.3. Chaos in a Speech Signal. The H can be used as an index to judge whether time-series data follow a random walk or a biased random walk [22]. Because this index cannot adequately indicate the changes in chaos of radiotelephony communication, this paper illustrates how the chaos in a speech signal changes the presence of fatigue from a different aspect. Based on Takens' embedding theorem, the chaotic nonlinear dynamic model of speech signals is reconstructed using a delay phase diagram method [23]. The process is in the phase space of the discrete time series of a speech signal. The model describes the phase-space topology of the strange attractor of speech [24]. When reconstructing the speech sequence $\{X(i)\}_{i=1}^N$ sampled in discrete time, the vector point set $P(i)$ in m -dimensional space with delay τ is obtained:

$$P(i) = \{X(i), X(i + \tau), \dots, X[i + (m-1)\tau]\}_{i=1}^{N-(m-1)\tau}. \quad (2)$$

The velocity of the airflow when speaking decreases when a person is in a state of fatigue. The friction and viscous force of the airflow increase due to the softening and cooling of the vocal-duct wall. This physiological change will reduce the energy of the airflow turbulence in the vocal-duct boundary layer [16]. Turbulence forms the basis of a chaotic speech signal. Any change in the turbulence directly affects the chaotic characteristics of a speech signal. Figures 3–6 depict the phase-space trajectories of different speech states when the same person utters the same speech. The right and left subgraphs of all four graphs show the phase-space trajectories of speech signals produced in the normal and fatigued states, respectively. Four words and numbers (i.e., speed, height, 182, and 134375) that are highly representative of the content of radiotelephony communication. The figures clearly indicate that the degree of speech fluctuation is significantly lower in a fatigued state than in a nonfatigued state. Therefore, the changes of chaos in speech can be quantitatively evaluated by using a reconstruction model of speech signals.

3. Speech Wavelet Fractal Feature

3.1. Fractal Dimension and Fatigue. The chaos in speech signal is related to fractal theory. The trajectory change of a nonlinear system in the process of chaotic evolution has some universality. Aerodynamics shows that the generation of a speech signal is a nonlinear process. Furthermore, the production of a speech sound (in particular, breathing sounds such as frictional and explosive sounds) involves the generation of eddies in the boundary layer of the vocal tract that eventually turns into turbulence, which has been proved to be a kind of chaos. This qualitative result forms the basis of applying chaos and fractal theory to the analysis of speech signals. Fractal is a complex system whose complexity can be described by noninteger dimension called the fractal dimension (FD).

FIGURE 1: Extraction processing for the HVSF assuming that the multidimensional estimator based on wavelet analysis has three steps of decomposition. “HC” represents the calculation method of the H index described above.

FIGURE 2: Distribution of the H values of speech samples in two emotional states. It can be seen that there is no obvious difference in the distribution of the H index between the fatigued and normal states.

Fractal dimension is the main parameter to describe fractal. The FD indicates the complexity of a fractal set. It is not a simple extension of Euclidean dimension, instead has many new connotations [25]. Generally, in Euclidean geometry, a line or curve is one dimensional. A plane or sphere is two dimensional. A geometry with length, width, and height is three dimensional. However, the complex fractal (such as coastline, koch curve, and shelpensky sponge) cannot be described the dimension of the integer. The FD has broken through the limit of the integer dimension of a general topological set. The importance of the FD is that it can be defined by data and calculated approximately experimentally. It is related to H as follows [26]:

$$H = 2 - D, \quad (3)$$

where D is the FD. On this basis, a revised fatigue feature of the ATC is proposed.

The formula for calculating the fractal dimension is as follows:

$$D = \lim_{\varepsilon \rightarrow 0} \left[\frac{\log N(\varepsilon)}{\log(1/\varepsilon)} \right], \quad (4)$$

where ε is the side length of a small cube and $N(\varepsilon)$ is the number needed to cover the measured geometry with the small cube. The formula is to determine the fractal dimension by covering the measured geometry with a small cube. For random fractal, different approximate methods can be used to calculate it, and some appropriate methods can also be used to measure it.

In order to obtain qualitative information about dynamic systems, it is often necessary to have sufficient information about the state evolution. However, in many practical engineering applications it is only possible for data acquisition equipment to obtain 1-D vectors containing system information, namely, time-series data. Therefore, it is necessary to extract the qualitative information of the system from the experimental time series, in which reconstructing the phase space is the first step to detecting weak signals. The FD of time series in this paper is calculated directly in the time domain, which not only simplifies the computational complexity but also achieves the same effect as phase-space reconstruction. A time series $\{X(i)\}_{i=1}^N$ with length N is set up. There are k new time series X_k^m that are obtained by reconstructing the time series with a delay method. The new time series X_k^m has the following form:

$$X_k^m: X(m), X(m+k), X(m+2k), \dots, X\left(m + \text{int}\left(\frac{N-m}{k}\right)k\right), \quad m = 1, 2, \dots, k. \quad (5)$$

The curve length $L_m(k)$ of each X_k^m can be calculated using

$$L_m(k) = \frac{1}{k} \left[\left(\sum_{i=1}^{\text{int}((N-m)/k)} |X(m+ik) - X(m+(i-1)k)| \right) \times \frac{N-1}{\text{int}((N-m)/k)k} \right]. \quad (6)$$

FIGURE 3: Phase trajectory when saying “speed.” Here, dimension is 3. Delay τ is 6. This parameter setting is the same below. The right and left subgraphs of all four graphs show the phase-space trajectories of speech signals produced in the normal and fatigued states, respectively. The figures clearly indicate that the degree of speech fluctuation is significantly lower in a fatigued state than in a nonfatigued state.

FIGURE 4: Phase trajectory when saying “height.”

The length of the total sequence can be approximated as the average of the length of the sequence curve generated by k delays:

$$L(k) = \frac{1}{k} \sum_{m=1}^k L_m(k). \quad (7)$$

For different values of k , a set of curve data related to k and $L(k)$ can be obtained. Curve $\text{lb}(L(k)) \sim \text{lb}(1/k)$ can be

drawn out. If it is a straight line, the relationship between k and $L(k)$ is as follows:

$$L(k) \sim k^{-\text{FD}}. \quad (8)$$

Linear fitting is used to obtain the straight line:

$$\text{lb}(L(k)) = D \times \text{lb}\left(\frac{1}{k}\right) + C. \quad (9)$$

FIGURE 5: Phase trajectory when saying “182.”

FIGURE 6: Phase trajectory when saying “134375.”

A method for determining $k_{\max}$ is proposed, which involves changing the value of $k_{\max}$ from $k_{\max} = 1$ when the abovementioned method is used to calculate the FD of a time series. When the value of FD no longer clearly changes, the corresponding $k_{\max}$ value is the most suitable for calculating the FD of this kind of time series by using the abovementioned algorithm. The specific calculation method of FD is shown in Figure 7.

The FD is calculated for the speech data which is used to calculate the HVSF feature in Section 2, as shown in Figure 8.

It is not difficult to see from the graph that the FD is obviously smaller for speech produced in a fatigued state than in a normal state. Furthermore, analysis shows that this situation applies to different ATCs in two states. Table 1 lists the FDs for some marked voice instructions from different ATCs recorded during the same time period (07:00 to 10:00 on April 26, 2018).

3.2. Speech Wavelet Fractal Feature. The speech data of an ATC are contaminated by noise due to the influence of

FIGURE 7: The specific calculation method of FD.

FIGURE 8: Comparison of speech FDs under different conditions. It is not difficult to see from the graph that the FD is obviously smaller for speech produced in a fatigued state than in a normal state.

data-acquisition equipment and the environment. Considering this problem, an improved vocal feature of ATC fatigue is proposed. The noise in radiotelephony communication contains more energy at low frequency. In this paper, wavelet decomposition is used to extract the detail coefficients of the ATC speech signal to reduce the influence of noise. In wavelet decomposition, it is very important to choose the appropriate decomposition scale and wavelet basis function. The decomposition scale (j) is closely related to the frequency range of speech signals and the frequency distribution of wavelet decomposition. The

TABLE 1: The FDs for some marked voice instructions. Instead of giving the ATCs' real name in the form, they are replaced by letters from A to N.

Voice instruction	ATC	Sex	Position	FD	State
1	M	M	Area	1.773	Fatigue
2	D	M	Tower	1.792	Fatigue
3	C	M	Tower	1.829	Fatigue
4	L	M	Area	1.849	Fatigue
5	B	M	Tower	1.862	Normal
6	I	M	Approach	1.863	Fatigue
7	N	M	Area	1.876	Normal
8	E	M	Area	1.882	Normal
9	H	M	Approach	1.887	Fatigue
10	J	M	Approach	1.890	Normal
11	K	M	Approach	1.898	Normal
12	F	M	Approach	1.900	Normal
13	A	F	Tower	1.905	Normal
14	G	M	Approach	1.927	Normal

frequency distribution of speech signals on each scale after wavelet decomposition is shown in Figure 9.

If the signal is decomposed in four layers, the frequency range of the fourth-level low-frequency coefficients is 0–500 Hz. If the signal is further decomposed into five layers, the frequency range of the fifth-layer low-frequency signal is 0–250 Hz. The energy and information in a speech signal is generally present between 300 and 3400 Hz. Therefore, it is meaningless to decompose the speech signal in the fifth level. Therefore, for a speech signal with a sampling frequency of 8 kHz, a wavelet decomposition scale (j) of 4 can be chosen.

Different wavelet basis functions will have different impacts on noise reduction. Generally, the wavelet bases that can produce the most coefficients near to zero are chosen. When wavelet transform is applied to signals, the selected wavelet bases are better able to satisfy the properties of symmetry or antisymmetry, compact support, and orthogonality simultaneously. The main properties of common wavelet bases are listed in Table 2. It is not difficult to see that the Daubechies wavelet is highly consistent with the abovementioned requirements. Therefore, the Daubechies wavelet was chosen as the wavelet basis function.

When the wavelet decomposition scale and wavelet basis are determined, the speech signal can be decomposed by using the wavelet transform. Then, the detail coefficients can be extracted. The FD of the detail coefficient d_i of each layer is also calculated:

$$D(d_i) = \text{FD}[d_i, k_{\max}(d_i)], \quad (10)$$

where FD is the FD calculation method described in this paper, for $i = 1, 2, 3, 4$, $k_{\max}(d_i) = 10$, and $D(d_i)$ represents the FD of the detail coefficients of d_i layer. The FD comparison of the detail coefficients of each layer is shown in Figure 10. Finally, the SWFF of the ATC fatigued speech is built, which is composed of the FD of the speech signal and its detail coefficients. In the following formula, $D(x)$ represents the fractal dimension of the speech signal:

$$\text{SWF} = [D(x), D(d_1), \dots, D(d_4)]. \quad (11)$$

FIGURE 9: Frequency distribution of the speech signal on different scales after wavelet decomposition. High-frequency coefficient is the detail coefficient and the approximate coefficient.

TABLE 2: Main properties of common wavelet bases.

Wavelet basis	Orthogonal	Biorthogonal	Compact support	CWT	DWT	Symmetry
Harr	Yes	Yes	Yes	Yes	Yes	Yes
Daubechies	Yes	Yes	Yes	Yes	Yes	Approximate
Biorthogonal	No	Yes	Yes	Yes	Yes	No
Coiflets	Yes	Yes	Yes	Yes	Yes	Approximate
Symlets	Yes	Yes	Yes	Yes	Yes	Approximate
Morlet	No	No	No	Yes	No	Yes
Mexican hat	No	No	No	Yes	No	Yes
Discrete Meyer	Yes	Yes	No	Yes	Yes	Yes

(a)

(b)

FIGURE 10: Continued.

FIGURE 10: The FD comparison of the detail coefficients of each layer of the speech signal in two states. Blue line indicates normal, and red line indicates fatigue: the (a) first layer, (b) second layer, (c) third layer, and (d) fourth layer.

A novel fatigue-detection method for ATCs can be proposed this moment. This method takes SWFF as fatigue-detection feature. Then, a support vector machine (SVM) is used to detect fatigue, as shown in Figure 11. The first step of the method is to construct a speech database of control instructions corresponding to the individual ATC. The speech signals in the database are marked as normal or fatigued. The second step is to decompose the voices in the database using four-layer wavelet decomposition. Then, the detail coefficients of each layer are extracted. The third step is to calculate the FD of the sequence in the second step according to the method described in this paper. The FD of the sequence is used to obtain the SWFF of the ATC. Finally, an SVM is applied to detect ATC fatigue. The implementation process of the method is reported in detail by experiment.

4. Method and Experiment

4.1. ATC Speech Database. The fatigue experienced by ATCs mainly results from a poor working environment and inadequate rest due to an excessive workload. The factors affecting fatigue can be roughly divided into aspects such as personality characteristics, the available facilities and equipment, the operating environment, duty scheduling, and organizational management.

In order to address the aims of the present study, each speech signal was numbered according to certain rules, as presented in Table 3. These rules are worked out by taking factors (such as working time, age, position, and the individual feelings of ATCs) into consideration. Finally, an ATC speech database is constructed. This database could support future research into the fatigue of ATCs. The numbering scheme used consists mainly of numbers and English letters that indicate certain factors.

FIGURE 11: Fatigue-detection method for ATCs.

4.2. Support Vector Machine Settings. The SWFF of 696 speech signals in the ATC speech database were selected. Then, a SVM was used to detect fatigue status. The selected speech signals were recorded from 02:00 to 17:00 on April 26, 2018. They were sampled at 8000 Hz. In all samples, the number of negative samples is less than that of positive samples. Therefore, in the simulation experiment, we selected all the fatigue samples, and then select positive samples with the same number of negative samples according to the time sequence of speech samples being recorded. These voices

TABLE 3: Numbering rules of the speech database constructed in this study.

Number position	Expression	Explanation
1	Control categories	R: area control, A: approach control, T: tower control
2	ATC's rank	5: Level 5, 4: Level 4, 3: Level 3, 2: Level 2, 1: Level 1, 0: Trainee
3-10	Time (UTC)	Numbers 3 to 6: starting time of work, Numbers 7 to 10: ending time of work
11	Sex	F: female, M: male
12-13	Age	Arabic number representation of real age
14-15	Order	Nn:N is a digital indicator and n is an Arabic numeral, indicating the n th instruction issued by a ATC during working
16-17	Status	14th and 15th: voice command; 1: error, 2: ambiguity, 3: hesitation or pause, 4: fatigue

Note. (1) The Arabic numerals immediately following N are treated as a numbering number regardless of the number of digits. (2) Number position 14-15 indicates the speech state. When the speech is normal, there number position 16 is vacant.

contained speech data of ATCs of different genders, ages, and positions. These voices comprised 348 in a normal condition and 348 in a fatigued condition. The 174 voices in a normal condition and 174 voices in a fatigued condition were assigned to the training set. The remaining 348 speech samples were assigned to the test set. It should be noted that each speech sample was different in terms of the semantic content and signal length. SVM is a popular method in machine learning. This method is "robust." A few support vectors determine the final result and are not sensitive to outliers, which eliminates a large number of redundant samples. Based on these advantages, the process of training and detecting sample data using an SVM is shown in Figure 12.

$\{x_j, y_j\}_{j=1}^N$ is a given set, where $x_j \in R^n$ is the j th input vector and $y_j \in R$ is the corresponding output. The overall model by weighted LS-SVM is formulated as [27]

$$\hat{y}_v = \frac{\sum_{i=1}^C P_{iv} \hat{y}_{iv}}{\sum_{i=1}^C P_{iv}}. \quad (12)$$

The i th weight coefficient of x_v is calculated by

$$P_{iv}(x_{iv}) = A_{iv}^1(x_{iv}^1), A_{iv}^2(x_{iv}^2), \dots, A_{iv}^n(x_{iv}^n), \quad (13)$$

$$A_{ij}^t = \exp\left(-\left(\frac{x_{ij}^t - \theta_i^t}{\beta_i^t}\right)^2\right), \quad \theta_i^t = z_i^t,$$

$$\beta_i^t = \lambda \sqrt{\frac{\sum_{j=1}^N (\mu_{ij})^m (x_{ij}^t - \theta_i^t)^2}{\sum_{j=1}^N (\mu_{ij})^m}},$$

where θ_i^t is the t th component of the center θ_i , β_i^t is the t th component of the width β_i , and λ is the parameter to control the overlapped ratio. The reconstructive set in the i th subspace can be expressed by

$$\Delta i = \left\{ (x_{ij}, y_{ij}) \mid \theta_i^t - \beta_i^t \leq x_{ij}^t \leq \theta_i^t + \beta_i^t, y_{ij} \right\}. \quad (14)$$

A_{ij}^t is the membership grade, $t = 1, 2, \dots, n$. Weighted LS-SVM employs fuzzy c-means clustering to decide the number of rules, which is based on the following formula:

FIGURE 12: The fatigue-detection process for ATCs.

$$\min J_m(\mu_{ij}, z_i) = \sum_{i=1}^C \sum_{j=1}^N (\mu_{ij})^m \|x_j - z_i^2\|, \quad (15)$$

$$\sum_{i=1}^C \mu_{ij} = 1; \quad 0 < \sum_{j=1}^N \mu_{ij} < N,$$

where $m \in (1, \infty)$ is a fuzzy exponent, μ_{ij} denotes the degree that x_j belongs to the rule i , $\mu_{ij} \in U$, and z_i is the i th cluster center. The novel LS-SVM considers general errors that include noises of input variables and output variables as empirical errors [28].

Considering that the radial basis function (RBF), namely, the Gauss kernel function, has better antijamming ability for noise in data, the classical robust RBF is chosen as the kernel function of the SVM in the ATC fatigue-detection method proposed in this paper. The RBF kernel in this research is the same as the activation function used by Mu and Zhang [29]. The mathematical model of kernel function is as follows:

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0, \quad (16)$$

where γ is the parameters of kernel function.

A K-fold crossvalidation (K-CV) method is used to obtain the values of the penalty factor $c = 9.7656 \times 10^{-4}$ and Gamma parameter $\gamma = 0.5$. In the experiment, the 696 original data were divided into $K = 6$ groups (generally average). Each subset data is used as a verification set, and the rest $K-1$ subset data is used as a training set so that K models can be obtained. The average of classification accuracy of the final verification set of these K models is used as the performance index of the classifier under this K-CV. K is set to 6. Except where stated otherwise below, the above parameters were used for the SVM in the experiment experiments.

FIGURE 13: Test results for support vector machines with HVSF and RBF kernels. The red asterisk in the figure indicates the true category of speech samples, while the blue circle indicates the predicted category. The two markers will overlap when the predicted and real categories are the same, otherwise they will be separated. Vertical coordinates of 0 and 1 indicate that the speech sample has been labelled as normal and fatigue, respectively.

FIGURE 14: Test results for support vector machines with SWFF and RBF kernels.

TABLE 4: The results for different fatigue testing methods (348 test samples).

Fatigue features	Classifier	Kernel function	Detection rate	
			Fatigue (%)	Accuracy (%)
pH	SVM	RBF	81.61	63.84
	SVM	Polynomial	95.42	85.63
SWFF	BP	RBF	90.80	85.21
	SVM	RBF	96.55	92.82

4.3. Result and Analysis. In order to fully demonstrate the detection accuracy of the proposed method, different combinations of fatigue features and classifiers were simulated in this study. First, the HVSF was used as the fatigue feature. The SVM was used as the classifier to detect the fatigued state of ATCs. The test results are shown in Figure 13. The kernel function of SVM is RBF kernel function. Subsequently, the SWFF is used to replace the HVSF as the fatigue feature to be detected. The experiment results in Figure 14 show that the fatigued-state detection results based on SWFF characteristics are superior in terms of accuracy than those based on HVSF. When the predicted and real categories are the same, the detection results can be considered as correct. Detection accuracy is the ratio of the number of correct detection to the number of samples in the detection set. Detection accuracy A_{det} is calculated as follows:

$$A_{\text{det}} = \frac{N_{\text{corr}}}{N_{\text{det}}}, \quad (17)$$

where N_{corr} is the number of correct detection and N_{det} is the number of samples in the detection set.

Second, considering that the use of different kernel functions for the SVM may affect the detection results, the detection results were also analysed when the fatigue characteristics were SWFF and the kernel functions of the SVM differed, as listed in Table 4. That table also gives the parameter settings, fatigue-detection rate, and total detection accuracy of different methods. It is clear from the table that when the kernel function is a polynomial, the accuracy of the test results were reduced to 85.63%. The results obtained when the classifier for detecting a fatigued state was changed to a backpropagation (BP) neural network are also presented in the table. The accuracy when using the BP neural network as a classifier was 8% lower than when using the SVM.

The test results for the 348 speech instructions of the ATCs indicate that the best performance in detecting a fatigued state was obtained when using SWFF. The classifier used here for detection was an SVM, whose kernel function is RBF, with a wavelet decomposition scale (j) of 4 and a db1 wavelet basis function. This method produced a high detection accuracy of 92.82%. In particular, the fatigue-detection rate of the proposed method is 96.55%. This is very important for aviation security. In short, the technology for detecting ATC fatigue proposed in this study is superior to other advanced fatigue-detection technologies.

5. Conclusion

In this paper, the radiotelephony communications of ATCs have been analysed. The chaos in radiotelephony communications has also been discussed. The chaos can be used to accurately judge the fatigued state of ATCs. The phase-space trajectories are significantly more complex of a fatigued state than that of a normal state. This study has introduced the HVSF for fatigue detecting. Because of the specific characteristics of radiotelephony communications, the HVSF cannot represent the fatigued state of ATCs well. There is no significant difference in the HVSF between fatigued and normal speech signals. Therefore, a revised vocal feature of ATCs called SWFF was proposed based on FD. This feature shows a great change in speech signal when an ATC is fatigued. The FD is obviously smaller for speech in a fatigued state than in a normal state. Furthermore, analysis shows that this situation applies to different ATCs in two states.

A fatigued speech database of ATC has been constructed. The file name of each voice in the database represents the state information related to that voice. This database could support future research related to the fatigue of ATCs. A method for detecting ATCs in a fatigued state is proposed based on SVM technology and the SWFF of ATCs. This method is robust for noise contamination. The experiment results obtained for different fatigued-state detection methods demonstrate the superiority of the proposed method. The accuracy of the proposed method was at 92.82%. That is higher than the accuracy of the other fatigued-state detection methods analysed. In particular, the fatigue-detection rate of the proposed method is 96.55%. This is very important for aviation security. The research provides a theoretical guidance for Air Traffic Management Authority on detecting ATC fatigue, while it may provide reference for fatigue assessment in other professional fields of civil aviation.

Data Availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to their containing information that could compromise the privacy of Air Traffic Management Shandong Bureau of China.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Zhiyuan Shen, Guozhuang Pan, and Yonggang Yan are equal contributors.

Acknowledgments

The authors thank the Air Traffic Management Bureau of Civil Aviation Administration of China for the financial support by the Research on Monitoring and Management of Fatigue State in Air Traffic Controller project and also thank

the Shandong Air Traffic Management Sub-Bureau of Civil Aviation Administration of China for their help in the supplying speech data of air traffic controllers.

References

- [1] Civil Aviation Administration of China (CAAC), *Statistical Bulletin of Civil Aviation Industry Development*, Civil Aviation Administration of China (CAAC), Beijing, China, 2018.
- [2] J. Huibin and L. Yawei, *Evaluation of Detection Performance of Critical Flicker Fusion Frequency and Perclos on Air Traffic Controller Fatigue*, China Science paper, Beijing, China, 2018.
- [3] M.-L. Chen, S.-Y. Lu, and I.-F. Mao, "Subjective symptoms and physiological measures of fatigue in air traffic controllers," *International Journal of Industrial Ergonomics*, vol. 70, pp. 1–8, 2019.
- [4] V. Riethmeister, U. Bültmann, M. Gordijn, S. Brouwer, and M. D. Boer, "Investigating daily fatigue scores during two-week offshore day shifts," *Applied Ergonomics*, vol. 71, pp. 87–94, 2018.
- [5] T. Chalder, G. Berelowitz, T. Pawlikowska et al., "Development of a fatigue scale," *Journal of Psychosomatic Research*, vol. 37, no. 2, pp. 147–153, 1993.
- [6] S. Arnau, T. Möckel, G. Rinkenauer, and E. Wascher, "The interconnection of mental fatigue and aging: an eeg study," *International Journal of Psychophysiology*, vol. 117, pp. 17–25, 2017.
- [7] H. Shitong, J. Li, P. Zhang, and W. Zhang, "Detection of mental fatigue state with wearable ecg devices," *International Journal of Medical Informatics*, vol. 119, pp. 39–46, 2018.
- [8] H. Mansikka, P. Simola, K. Virtanen, D. Harris, and L. Oksama, "Fighter pilots' heart rate, heart rate variation and performance during instrument approaches," *Ergonomics*, vol. 59, no. 10, pp. 1344–1352, 2016.
- [9] G. Costa, "Evaluation of workload in air traffic controllers," *Ergonomics*, vol. 36, no. 9, pp. 1111–1120, 1993.
- [10] K. Naitoh, W. Karwowski, and A. Murata, *A Method for Predicting the Risk of Virtual Crashes in a Simulated Driving Task Using Behavioural and Subjective Drowsiness Measures*, The official publication of the Ergonomics Research Society, London, Francis, 2017.
- [11] M. M. Ibrahim, J. J. Soraghan, L. Petropoulakis, and G. Di Caterina, "Yawn analysis with mouth occlusion detection," *Biomedical Signal Processing and Control*, vol. 18, pp. 360–369, 2015.
- [12] B. Nie, X. Huang, Y. Chen, A. Li, R. Zhang, and J. Huang, "Experimental study on visual detection for fatigue of fixed-position staff," *Applied Ergonomics*, vol. 65, pp. 1–11, 2017.
- [13] L. L. Di Stasi, M. B. Mccamy, A. Catena, S. L. Macknik, J. J. Cañas, and S. Martinez-Conde, "Microsaccade and drift dynamics reflect mental fatigue," *European Journal of Neuroscience*, vol. 38, no. 3, pp. 2389–2398, 2013.
- [14] M.-J. Caraty and C. Montacié, "Vocal fatigue induced by prolonged oral reading: analysis and detection," *Computer Speech & Language*, vol. 28, no. 2, pp. 453–466, 2014.
- [15] J. Krajewski, R. Wieland, and A. Batliner, *An Acoustic Framework for Detecting Fatigue in Speech Based Human-Computer-Interaction*, Springer, Berlin, Germany, 2008.
- [16] J. Krajewski, A. Batliner, and M. Golz, "Acoustic sleepiness detection: framework and validation of a speech-adapted pattern recognition approach," *Behavior Research Methods*, vol. 41, no. 3, pp. 795–804, 2009.
- [17] H. Deng, Z. Xu, and H.-X. Li, "A novel neural internal model control for multi-input multi-output nonlinear discrete-time processes," *Journal of Process Control*, vol. 19, no. 8, pp. 1392–1400, 2009.
- [18] H. K. Palo, M. N. Mohanty, and M. Chandra, "Efficient feature combination techniques for emotional speech classification," *International Journal of Speech Technology*, vol. 19, no. 1, pp. 135–150, 2016.
- [19] N. Wang, P. C. Ching, N. Zheng, and T. Lee, "Robust speaker recognition using denoised vocal source and vocal tract features," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 1, pp. 196–205, 2011.
- [20] M. E. Ayadi, M. S. Kamel, and F. Karray, "Survey on speech emotion recognition: features, classification schemes, and databases," *Pattern Recognition*, vol. 44, no. 3, pp. 572–587, 2011.
- [21] G. Zhou, J. H. L. Hansen, and J. F. Kaiser, "Nonlinear feature based classification of speech under stress," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 3, pp. 201–216, 2001.
- [22] V. V. Nishawala and M. Ostoj-Starzewski, "Acceleration waves on random fields with fractal and hurst effects," *Wave Motion*, vol. 74, 2017.
- [23] B. Maie, L. Jaanus, S. Anna, and H. Hiie, "Spectral asymmetry and higuchi's fractal dimension measures of depression electroencephalogram," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 251638, 8 pages, 2013.
- [24] R. J. Povinelli, M. T. Johnson, A. C. Lindgren, F. M. Roberts, and J. Jinjin Ye, "Statistical models of reconstructed phase spaces for signal classification," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2178–2186, 2006.
- [25] B. P. Matteo, B. Gennaro, P. Tessa, C. Ron, B. Marco, and C. Corrado, "Relationship between isometric muscle force and fractal dimension of surface electromyogram," *BioMed Research International*, vol. 2018, Article ID 5373846, 9 pages, 2018.
- [26] M. D. Ruiz-Medina, E. Porcu, and R. Fernandez-Pascual, "The dagum and auxiliary covariance families: towards reconciling two-parameter models that separate fractal dimension and the hurst effect," *Probabilistic Engineering Mechanics*, vol. 26, no. 2, pp. 259–268, 2011.
- [27] C. Y. Sun, C. X. Mu, and X. M. Li, "A weighted ls-svm approach for the identification of a class of nonlinear inverse systems," *Science in China Series F:Information Sciences*, vol. 52, pp. 770–779, 2009.
- [28] C. Mu, C. Sun, and X. Yu, "Internal model control based on a novel least square support vector machines for mimo nonlinear discrete systems," *Neural Computing and Applications*, vol. 20, no. 8, pp. 1159–1166, 2011.
- [29] C. Mu and Y. Zhang, "Learning-based robust tracking control of quadrotor with time-varying and coupling uncertainties," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, pp. 259–273, 2020.