IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 13, 2019, accepted May 9, 2019, date of publication May 14, 2019, date of current version May 28, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2916480

Three-Dimensional Vibration-Based Terrain Classification for Mobile Robots

CHENGCHAO BAI, (Student Member, IEEE), JIFENG GUO, (Member, IEEE), AND HONGXING ZHENG

School of Astronautics, Harbin Institute of Technology, Harbin 150001, China

Corresponding author: Jifeng Guo (guojifeng@hit.edu.cn)

This work was supported by the Manned Space Advance Research Fund under Grant 060101.

ABSTRACT Extraterrestrial celestial patrol missions have introduced very strict requirements for the performance of rovers, due to their high cost. Vision-based or Lidar-based environment sensing technology has matured. However, due to its perceptual characteristics, it is impossible to predict the traversability of the terrain completely, and it lacks the judgment of the physical properties of the terrain, such as the degree of hardness and softness. Due to the spectrum of risks that the rover is facing, a wide range of detection processes is required. This research paper proposes a terrain classification approach based on 3-D vibrations induced in the rover structure by the wheel-terrain interaction. Initially, the acceleration information of the three directions is obtained by using the Inertial measurement unit of the rover. Then, the characteristics of the vibrations of the known terrain are learned. The Fast Fourier Transformation (FFT) is used to transform the labeled three-axis vibration vectors into a frequency domain. Then the training feature vectors are obtained through normalization. Taking into account the characteristics of the environment, an improved back propagation (BP) neural network is used to get the mapping relationships between the vibrations and the terrain types. Finally, classification testing has been conducted on five kinds of environments, including concrete, grassland, sand, gravel, and mixed. After 20 times random testing experiments, the classification accuracy has proven to be in the range 88.99%–100%, which verified the validity and the robustness of the algorithm and laid a foundation for the subsequent identification of terrain characteristic.

INDEX TERMS BP artificial neural network, rover, terrain classification, vibration.

I. INTRODUCTION

Intelligence is the development trend of mobile robots in the future. High-efficiency perception of the environment can be achieved through the fusion of multi-sensors. Especially for terrain recognition, it will determine the success or failure of the task execution of mobile robots. As a kind of intelligent mobile robots, planetary rovers will provide effective support for planetary exploration.

Extraterrestrial celestial patrol mission is very difficult and costly. It has high safety requirements, autonomy, and very low fault tolerance. Due to the unpredictability of the environment, adding cognition to the planetary exploration rover so that it can understand the surrounding environment of the premise, is particularly crucial [1]. With the advancement of control technology [2], [3], the rover can operate autonomously. In the face of more and more abundant detection tasks, accurate terrain classification is vital.

The associate editor coordinating the review of this manuscript and approving it for publication was Bora Onat.

At present, relying on sensing devices such as Vision, Lidar and so on, the rover information and the relative position data of the environmental obstacle can be well obtained. At the same time, the features are processed and analyzed to get more accurate terrain classification results [4]–[6].

However, vision is susceptible to changes in lighting image distortion and other factors. At the same time, the existing sensing mode cannot effectively realize the physical properties of the flat surface (e.g. the type of material, the degree of hardness). As a result, there are unforeseen risks in the detection task. For example, the rover may fall inside when it passes through a loose and flat place. Faced with similar non-geometric hazards, identification of the detected terrain is significant in order to perform robust high-precision classification [7].

According to different sensing modes, the terrain classification can be divided in two main categories. The first is the contact classification, such as vibration and touch; the second is the non-contact classification, such as vision, laser radar, sound and so on. Among them, the non-contact

type has high discrimination capability on objects like large stones and steep slopes of the terrain. This enables the rover to avoid the most obvious dangers [8]. In Wuhan University, He *et al.* [9], proposed a hierarchical classification approach. The Conditional Random Field (CRF) and the Bayesian Network (BN) were employed to incorporate prior knowledge, in order to facilitate SAR image classification. Manduchi *et al.* [10] proposed a new obstacle detection and terrain classification method for autonomous offline navigation. Initially, binocular ranging was used to detect obstacles. Then, a color-based classification system was designed to mark the identified obstacles. Finally, the algorithm was used to analyze the radar data and to achieve effective classification of the terrain and the obstacles. Lalonde *et al.* [11] from Carnegie Mellon University, successfully divided Lidar data in three categories using local 3D point cloud statistics. The ''scatter'' class characterized porous volumes, such as grasses and canopies; the ''linear'' class characterized elongated objects such as lines and tree ridges; and the ''surfaces'' one was related to solid targets such as the ground. Valada *et al.* [12] used the sound features of wheel interaction to perform successful classification-identification. They proposed a learning method based on Convolutional Neural Networks. The robustness of this algorithm was verified as follows: 1) It outperformed the traditional audio classification methods; 2) It maintained its robustness under varying Gaussian white noise and fine-tuning noise enhancement; 3) Though low-quality data loggers were used in a very noisy environment, the algorithm had still reliable performance. After that, A. Valada and Burgard also applied the deep long-short term memory (LSTM) based recurrent model to the proprioceptive terrain classification [13]. Zhao *et al.*, also proposed a terrain classification method based on the sound and vibration data of the wheel-terrain interaction. Compared with the traditional handcrafted domain-specific features, a two-step selection method of optimal feature subsets that combines ReliefF and mRMR algorithms was proposed. Finally, the effective classification of different data was performed by the combination of multiple classifiers [14].

The contact terrain classification is mainly based on vibration and touch. It is analyzing the vibration and the acceleration signals caused by the interaction between the rover wheel and the surface of the celestial [15]. Iagnemma *et al.* [16] of the Massachusetts Institute of Technology proposed an online terrain parameter estimation method, based on simplified forms of classical terramechanics equations in 2002. This research introduced for the first time, the visual and the vibration based rover terrain classification and recognition method. It mainly realized the following: a) measuring the sinking of the planet rover wheels with a vision-based approach b) online estimation of the terrain parameters using tactile-based approach and c) classification of the terrain using the vibration feedback based method [17]. Based on [16], Brooks *et al.* [18] proposed a more complete vibration-based terrain classification method in 2005. The offline learning training of the classifier, used the vibration data of the mark number. At the same time, linear online classification and identification of the terrain was performed [19]. Weiss *et al.* 2006, from University of Tübingen, Germany [20], used Support Vector Machines (SVM) to perform vibration-based terrain classification. They employed a Radial Basis function kernel and a feature extraction method. Initially, unprocessed acceleration data were introduced. Totally 8 features were calculated, and normalized to form a feature vector. In 2007, Weiss *et al.* proposed a technique for measuring vibration acceleration in different directions, in order to improve classification performance. Through FFT transformation, each feature was normalized and processed using SVM. The results have shown that when the y-direction acceleration was the feature vector, the classification accuracy rate was higher than that the case of the z-direction acceleration feature vector. Moreover, the highest accuracy was obtained when the three were combined [21]. Collins and Coyle [22] proposed a response-based terrain input classification method. Speed dependence was eliminated or reduced compared to existing vibration methods. This approach, used the AGV vibration transfer function to map the vibration output to the terrain input, and validated it in the simulation, using surface contours from real terrain. Tick et al. [23] proposed a multi-level terrain classifier, based on angular velocity in 2012. The innovation lied in the use of acceleration and angular velocity measurements to characterize features in all basic directions. They also employed sequential forward floating feature selection for feature screening, and linear Bayesian normal classifiers for classification. Different feature sets were generated according to different speed conditions, and the classifier was also switched according to different speeds.

In addition, with the continuous development of artificial intelligence, more and more intelligent algorithms, such as CNN and unsupervised leaning, are used for terrain sensing [24]–[26]. Zeltner [27] used a deep convolutional neural network to implement a vision-based terrain classification. Park *et al.* [28] proposed a new classification network framework based on LSTM unit and ensemble learning. They provide a good research direction for the subsequent output of terrain classification results from vibration data [29]–[32].

Combined with the above analysis, an improved terrain classification method based on three-dimensional vibration information is proposed. The main contribution of this paper is that the effective mapping of vibration features to terrain classification is realized based on improved BP neural network for the first time. Meanwhile, a selective method for terrain classification is added. Compared with the previous methods such as SVM, the improved BP structure is simpler, and it does not need artificial design features. Besides, it is easy to implement in practical applications, and has stronger adaptability to unknown environments. The remainder of this paper is organized as follows.

In Section 2, we give a detailed description of terrain classification methods, including the elaboration of algorithm framework, feature extraction based on FFT, and the design

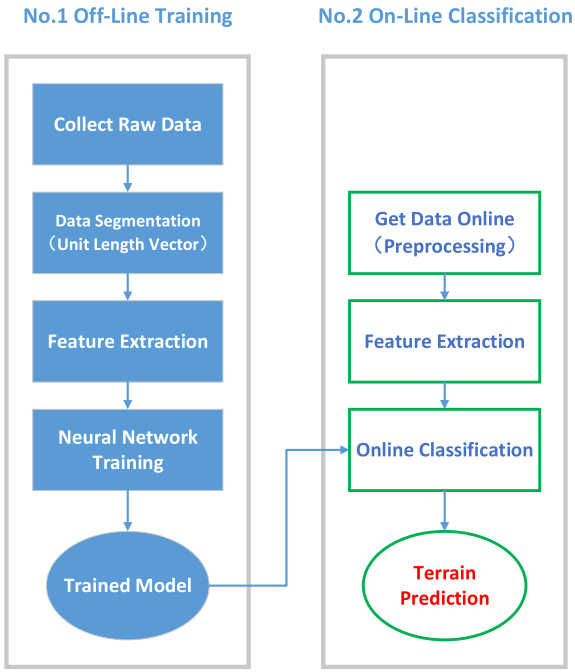

FIGURE 1. Schematic diagram of algorithm flow framework.

of improved BP neural network. The Section 3 compares and analyzes the rationality and correctness of algorithms based on actual test environment. In the end, the conclusion is presented in Section 4.

II. TERRAIN CLASSIFICATION METHOD

A. METHOD OVERVIEW

The entire classification process is divided in two phases: 1) off-line training; 2) online classification. The model should be trained properly, in order to obtain a higher resolution for different terrains, in the online classification process. As shown in figure 1, the 3-D vibration raw data collected by the sensor is first initialized and divided to form a vector with a duration of 1 second; Then, the frequency domain is transformed by the Fast Fourier Transformation (FFT), and the eigenvectors for training are obtained by normalization; Finally, a BP neural network is developed to perform parameter learning and a network model that can be used online is obtained.

B. FEATURE EXTRACTION

In the training process, it is always necessary to learn the vibration signal characteristics of the known terrain type. For this reason, the experimental platform needs to traverse different surfaces multiple times, to collect vibration information. The performed experiment used three single-axis vibration sensing units with an operating frequency of 100 Hz. In order to facilitate the processing of the data, the collected vibration data was segmented. Each segment corresponded to one second travel of the unmanned platform, so that a vector of 1×100 size could be generated. Finally each vector was marked as its corresponding terrain type.

Then, the original vibration signal was converted to the frequency domain. First, the raw data processing was standardized, and each vibration vector was normalized to a vector with a mean value of zero and a standard deviation of one. Then a Fast Fourier Transformation of 100 points was performed. Thus, the original time-domain data of the experimental platform was transformed in the frequency domain. This transformation has proven capable to determine the difference between various terrains. After applying the FFT to all vectors, they were normalized in the closed interval [0, 1]. Standardization prevented high-level data from prevailing in later training. In this way, after the original data was normalized, each index was in the same order of magnitude, which was suitable for a comprehensive comparative evaluation.

As it was mentioned in the introduction, the current vibration-based terrain classification methods are mainly based on separate measurements in the vertical direction (z). The reason is that the terrain changes have a greater impact on the upper and lower directions. Based on this, the paper adds the measurements of the horizontal data, that is, the front-back (x) and left-right (y) directions, which increases the dimension of the data. This approach aims to find the best representation pattern between the multi-dimensional data and the terrain categories.

At the same time, this paper also introduces a simple and effective three-axis information classification method. In each terrain testing phase, the vibration data of the three axes, namely $(x_{1:100})$, $(y_{1:100})$, and $(z_{1:100})$ are obtained first. Then, the normalized signal is transformed by the FFT to obtain the $F(x)_{1:100}$, $F(y)_{1:100}$, $F(z)_{1:100}$. Finally, each signal is normalized in the closed interval [0, 1]. Next, the transformed three-axis signal is connected as a feature vector. At this time, the dimension of the feature vector is 1×300. Finally, the Artificial Neural Network (ANN) is trained on these feature vectors.

This paper collects vibration data for 2 minutes for different terrain. According to the above segmentation method, 120 sets of data are collected for each type of terrain, so a total of 600 sets of data are collected. In the training process, 500 sets of data are selected randomly, and the remaining 100 sets of data are used as test samples.

C. BP ARTIFICIAL NEURAL NETWORK TRAINING

The BP ANN is a multi-layer feedforward one [33]. It performs forward signal and reverse error transmissions. In forward propagation, the input signal is processed from the input layer, through the hidden layer, to the output one. The neurons' state of each layer can only affect the state of the next layer of neurons. If the error in the output layer is not acceptable, the signal is redirected backwards, and the network weights and the threshold are adjusted according to the prediction error. In this way, the predictive output of the BP ANN is continuously approaching the expected one. The architecture of the BP ANN is shown in the following figure 2.

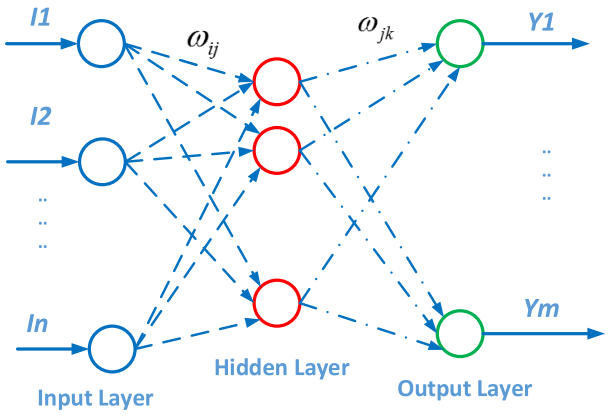

FIGURE 2. Architecture of a BP artificial neural network.

In the figure above, $I_1, I_2, \cdots, I_n$ is the input vector of the BP ANN. Here it is the preprocessed eigenvector. $Y_1, Y_2, \cdots, Y_m$ is the predicted vector of the BP ANN. As it can be seen from the above figure, the BP ANN can be regarded as a nonlinear function. The input and the predicted value are the independent and dependent variables of the function, respectively. When the number of input nodes is n and the number of output nodes is m, the ANN expresses the mapping function relationship from n independent variables to m dependent ones.

A BP neural network must be trained before the performance of online classification. It is characterized by associative memory and it retains prediction ability through training. The training process includes the following steps:

1) Network initialization. Determine the number of the nodes (neurons) for the input layer, for the hidden layer, and for the output layer, denoted as n, l, m respectively, according to the system input sequence (I, Y). Initialize the connection weights ω_{jk}, ω_{ij} between the neurons of the input, hidden and output layers. Initialize the hidden layer's threshold a, and the threshold b of the output layer, and also do the same for the given learning rate and the neuron excitation function.

2) Implicit layer output calculation. The hidden layer output H is calculated based on the following: the input variables I, the connection weights ω_{ij} of the input and implicit layers and the threshold α of the hidden layer.

$$H_j = f(\sum_{i=1}^{n} \omega_{ij} I_i - a_j) \quad j = 1, 2, \cdots, l \tag{1}$$

In equation 1, l is the number of the hidden layer nodes; and f is the hidden layer excitation function (see function 2).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

3) Output layer. The BP ANN prediction output O is calculated based on the hidden layer's output H, on the connection weight ω_{jk} and on the threshold b.

$$O_k = \sum_{j=1}^{l} \omega_{jk} H_j - b_k \quad k = 1, 2, \cdots, m \tag{3}$$

4) Error calculation. The network prediction error e is calculated based on the ANN's prediction output O and on the expected output Y.

$$e_k = \sqrt{\frac{1}{m} \sum (O_k - Y_k)^2} \; k = 1, 2, \cdots, m \tag{4}$$

5) Weight update. Update the ANN's connection weights ω_{ij}, ω_{jk} according to the obtained prediction error e.

$$\begin{aligned} \omega_{ij} &= \omega_{ij} + \eta H_j (1 - H_j) I_i \sum_{k=1}^{m} \omega_{jk} e_k \\ &\quad (i = 1, 2, \cdots, n; j = 1, 2, \cdots, l) \\ \omega_{jk} &= \omega_{jk} + \eta H_j e_k \\ &\quad (j = 1, 2, \cdots, l; k = 1, 2, \cdots, m) \end{aligned} \tag{5}$$

where η is the learning rate

6) Threshold update. The node's thresholds a, b is updated based on the ANN's prediction error e.

$$\begin{aligned} a_j &= a_j + \eta H_j (1 - H_j) \sum_{k=1}^{m} \omega_{jk} e_k \quad j = 1, 2, \cdots, l \\ b_k &= b_k + e_k \quad k = 1, 2, \cdots, m \end{aligned} \tag{6}$$

7) It is determined whether the algorithm's iterations should stop. If they should not stop, return to step 2.

For classification of five different terrain types, a three-layer BP neural network is designed in this paper. The input layer consists of 300 neurons, corresponding to the input eigenvectors. Hidden layer contains 25 neurons. And there are 5 neurons in the output layer, which correspond to five different terrain types. The activation function of the hidden layer chooses sigmoid function.

After the training is completed, the neural network can enter the online classification stage. The experimental platform traverses the unknown terrain and it collects vibration signals. Once per second, the acceleration signals emitted in three directions are combined to generate a 1×300 vector $(x_{1:100}, y_{1:100}, z_{1:100})$. Normalize each component, then transform the three components of the vector with FFT to get $(F(x)_{1:100}, F(y)_{1:100}, F(z)_{1:100})$. They should be normalized separately in order to get the testing vector. The trained ANN is used to classify the testing vectors and it returns an estimation of the terrain type.

The BP ANN employs the gradient correction learning algorithm, for the update of both the weights and the threshold. The weights and the threshold are corrected from the negative gradient direction of the network prediction error. The accumulation of the previous experience is not considered, and the learning process converges slowly. An additional momentum method can be used to solve the problem. The weight learning function of the additional momentum is the following:

$$\omega(k) = \omega(k-1) + \eta \Delta\omega(k) + \alpha \left[\omega(k-1) - \omega(k-2)\right] \tag{7}$$

where $\omega(k)$, $\omega(k-1)$, $\omega(k-2)$ are the weights at time $k, k-1, k-2$; α is the learning momentum. The parameters of this paper are set to $\eta = 0.1, \alpha = 0.05$.

According to the network prediction error, the training error can be calculated by the following formula

$$E = \frac{1}{train\_size} \sum_{train\_size} \sqrt{\frac{1}{m} \sum_{k=1}^{m} e_k^2} \tag{8}$$

Corresponding to this experiment, $m = 5$, $train\_size = 500$.

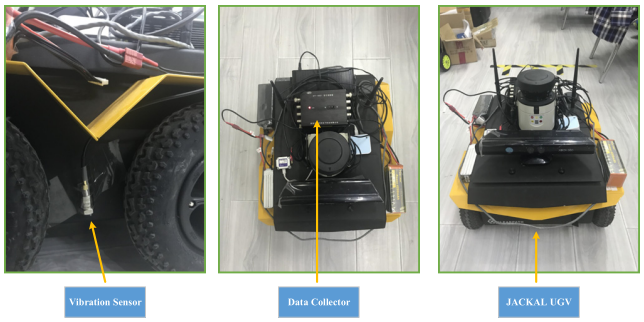

FIGURE 3. Jackal unmanned ground vehicle and sensors.

III. EXPERIMENTAL RESULTS

A. PLATFORM OVERVIEW

In order to verify the correctness of the method described above and the feasibility of its application in the actual environment, we have used the Jackal unmanned ground vehicle [34] as the carrier. It was tested in five different terrains (flat concrete, grass, sand, gravel, grass and stone mixed) with measuring equipment such as vision (Kinect), laser radar (SICK LMS 100), vibration sensor (AFT601D), and IMU (LPMS-USBAL2). This paper focuses on the vibrations' testing. As shown in Figure 4, the comparison of vibration data from different terrains is given. Subsequent research will compare and analyze the relationship between different sensors and their potential fusion values.

B. RESULTS AND ANALYSIS

First, the original signals of five terrain types have been collected. The processed data that have been found suitable for the classification, have been obtained by the above method as shown in Figure 5. It has been found that the original signal in the time domain has small discrimination, and the feature becomes relatively obvious after transforming it into the frequency domain. This can be beneficial to the training of the subsequent neural network and it can greatly improve the classification accuracy.

Moreover, the five types of terrain (flat concrete, grass, sand, gravel, mixed) have been represented by the numbers 1, 2, 3, 4, and 5 respectively. Figure 6 shows a comparison between the predicted versus the actual terrain categories in the testing process.

The abscissa indicates the number of test samples, and the ordinate corresponds to the above five topographies respectively. Red snowflakes represent the predicted value, while triangles and circles represent the true value of different topographies. Combining with Figure 7, it can be seen that

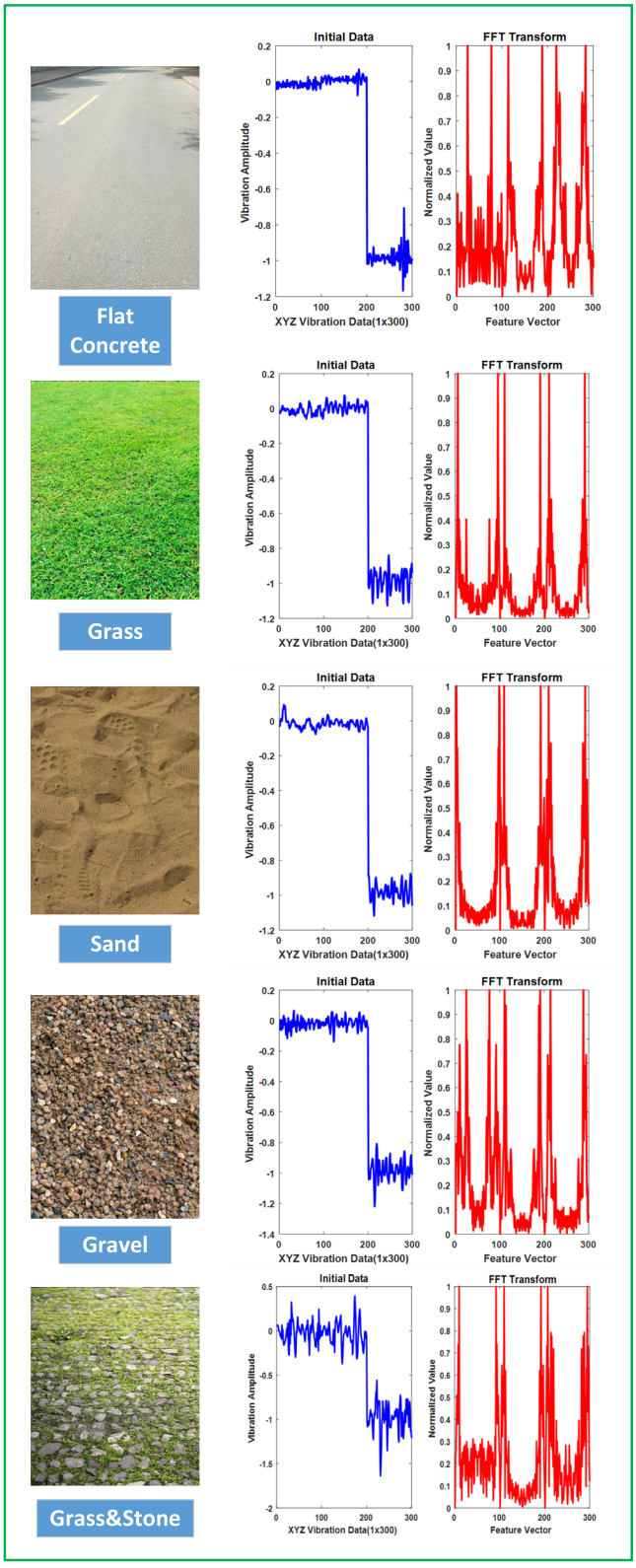

FIGURE 4. Raw data and preprocessed data from different terrain.

there are two prediction errors in 100 test samples, so the accuracy of this experiment is 98%. In addition, the error values of this two misclassification are 1 and 3, which can

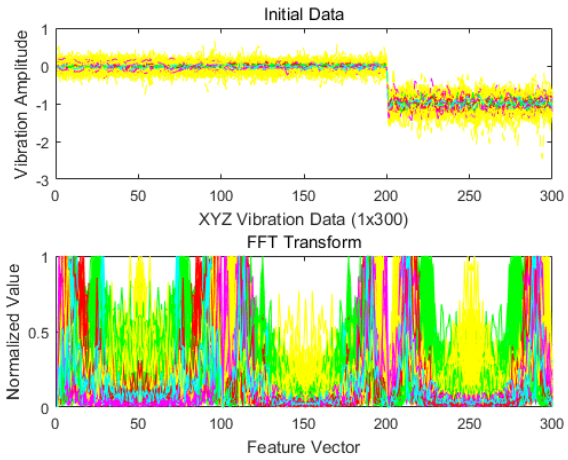

FIGURE 5. The original signal and the FFT processed one.

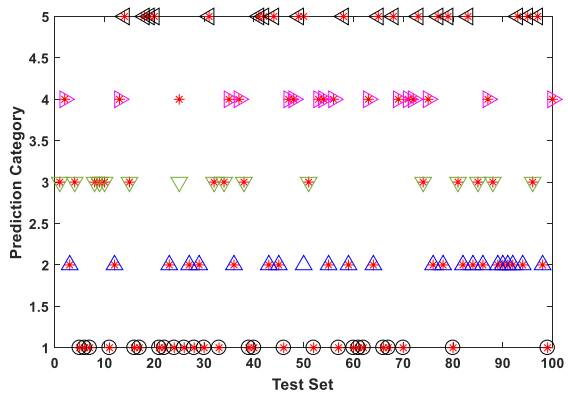

FIGURE 6. Comparison of predicted terrain and actual terrain.

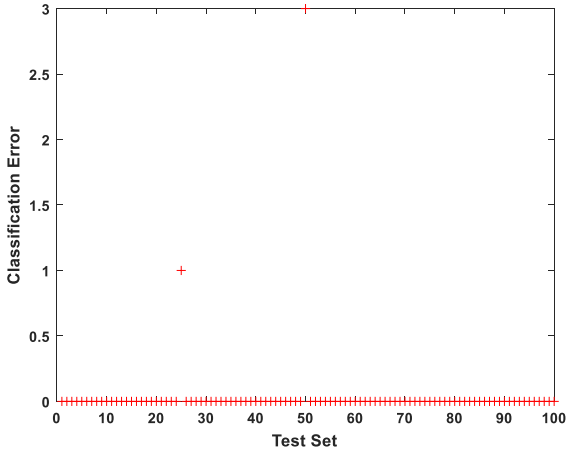

FIGURE 7. Classification error results of BP neural network.

be found in Figure 6. That is terrain 3 is misclassified into terrain 4, so its error is 1, terrain 2 is misclassified into terrain 5, so its error is 3. By this way, the mismatched association between terrain can be defined, which can provide reference for improving classification accuracy. According to this experiment, sand and gravel, as well as grass and mixed, are easy to confuse. There are two reasons for this. One is that

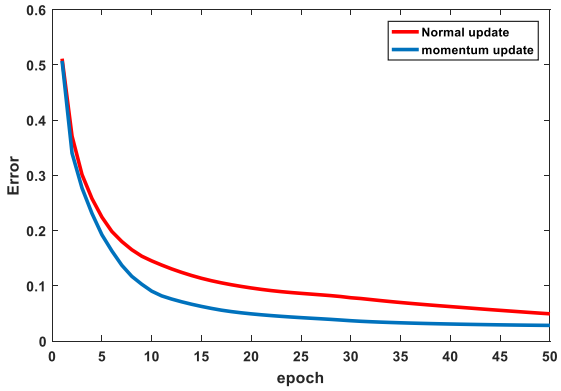

FIGURE 8. Convergence curve of training error.

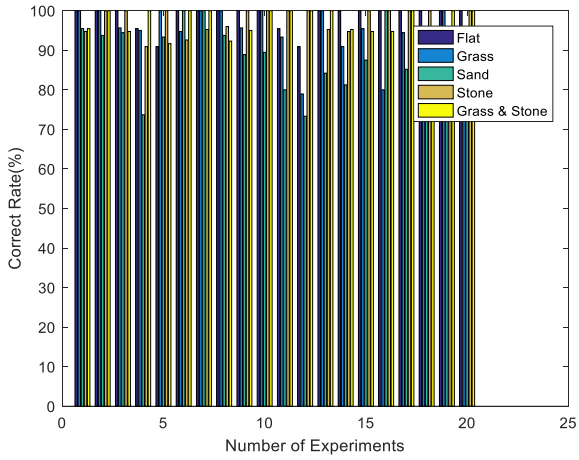

FIGURE 9. 20 times experimental comparison results.

the distinction between the features of the selected terrain is not large, and the other is that the training has a certain randomness.

The experimental results show that the flat concrete terrain has a correct rate of 100%, the grass terrain has a correct rate of 95.65%, the sand terrain has a correct rate of 93.75%, the stone terrain has a correct rate of 100%, and the mixed terrain has a correct rate of 100%. At the same time, according to Figure 8, we compare the declining trend of the error between the improved BP neural network and the non-improved BP neural network. It can be seen that although the final error is close and the improved accuracy is limited, the training speed of BP network is accelerated by the method of additional momentum when the weight is updated, which shows that the proposed method has a certain improvement in training convergence.

Considering the randomness of a single experiment, the experiment used a multi-validation model, and a total of 20 classification verifications were performed. The results are shown in the Figure 9. From the final results, the classification accuracy of five types of terrain is 98.64%, 94.81%, 88.99%, 97.71%, 97.07% respectively. The results show that the classification accuracy of flat concrete, grass, stone and mixed terrain is better than that of sand terrain.

At the same time, according to Figure 9, the classification accuracy of sand terrain is only 73.33% in the 13th experiment. Combining with the recognition accuracy of the whole sand terrain, it can be seen that the low accuracy may be caused by the randomness of single training. The random selection of training samples may contain a small amount of sand characteristic data, which makes the learning accuracy inadequate. In addition, sand terrain is less distinguishable from other types of terrain. Through the analysis of each prediction result, it can be seen that the probability of sand being misclassified into stone is higher, which is also the reason of its low accuracy. Besides, the choice of platform and the speed of operation will have an impact on the final accuracy, which will be analyzed in depth in the follow-up study.

IV. CONCLUSION

On the one hand, terrain classification and recognition technologies can make up for the lack of visual perception on physical properties. On the other hand, it can provide a new navigation means in the case of sensor failure. Due to the fact that no additional infrastructure is required the overall cost is not increased. At the same time, the verification of the method has proven its contribution towards the increase of rover's reliability. Both its potential value and its significance are huge. From the experimental results, it can see that there are higher recognition rates for five different types of terrain materials. Moreover, the accuracy rate of the verifications (which were performed 20 times in a random mode) is maintained above 88.99%. In the future research, the unmanned platform will be equipped with Vision, Lidar, IMU, and sound sensors to extend the experiment. The advantages of vibration perception can be given by comparison, and the fusion between multiple sensors can also be explored.

REFERENCES

[1] O. Patil, N. Shivtarkar, S. Sengupta, and J. Kolap, "Review on planetary rovers Lunokhod to curiosity and beyond," *Res. Rev., J. Space Sci. Technol.*, vol. 7, no. 1, pp. 15–23, 2018.

[2] C. Li and Z. Qu, "Distributed finite-time consensus of nonlinear systems under switching topologies," *Automatica*, vol. 50, no. 6, pp. 1626–1631, 2014.

[3] C. Li, Z. Qu, and M. A. Weitnauer, "Distributed extremum seeking and formation control for nonholonomic mobile network," *Syst. Control Lett.*, vol. 75, pp. 27–34, Jan. 2015.

[4] C. Wong, E. Yang, X.-T. Yan, and D. Gu, "Adaptive and intelligent navigation of autonomous planetary rovers—A survey," in *Proc. NASA/ESA Conf. Adapt. Hardw. Syst.*, Jul. 2017, pp. 237–244.

[5] S. Zhang, S. Liu, Y. Ma, C. Qi, H. Ma, and H. Yang, "Self calibration of the stereo vision system of the Chang'e-3 lunar rover based on the bundle block adjustment," *ISPRS J. Photogramm. Remote Sens.*, vol. 128, pp. 287–297, Jun. 2017.

[6] H. Wu *et al.*, "Optimum pipeline for visual terrain classification using improved bag of visual words and fusion methods," *J. Sensors*, vol. 2017, Mar. 2017, Art. no. 8513949.

[7] B. Wilcox, "Non-geometric hazard detection for a mars microrover," in *Proc. Conf. Intell. Robot. Field, Factory, Service, Space*, vol. 2. Washington, DC, USA: NASA, 1994, pp. 675–684.

[8] L. Ojeda, J. Borenstein, G. Witus, and R. Karlsen, "Terrain characterization and classification with a mobile robot," *J. Field Robot.*, vol. 23, no. 2, pp. 103–122, 2006.

[9] C. He, X. Liu, D. Feng, B. Shi, B. Luo, and M. Liao, "Hierarchical terrain classification based on multilayer Bayesian network and conditional random field," *Remote Sens.*, vol. 9, no. 1, p. 96, 2017.

[10] R. Manduchi, A. Castano, A. Talukder, and L. Matthies, "Obstacle detection and Terrain classification for autonomous off-road navigation," *Auton. Robots*, vol. 18, no. 1, pp. 81–102, 2005.

[11] J. F. Lalonde, N. Vandapel, D. F. Huber, and M. Hebert, "Natural Terrain classification using three-dimensional ladar data for ground robot mobility," *J. Field Robot.*, vol. 23, no. 10, pp. 839–861, 2006.

[12] A. Valada, L. Spinello, and W. Burgard, "Deep feature learning for acoustics-based Terrain classification," in *Robotics Research* (Springer Proceedings in Advanced Robotics), A. Bicchi and W. Burgard, Eds. Cham, Switzerland: Springer, 2018.

[13] A. Valada and W. Burgard, "Deep spatiotemporal models for robust proprioceptive Terrain classification," *Int. J. Robot. Res.*, vol. 36, nos. 13–14, pp. 1521–1539, 2017.

[14] K. Zhao, M. Dong, and L. Gu, "A new Terrain classification framework using proprioceptive sensors for mobile robots," *Math. Problems Eng.*, vol. 2017, Sep. 2017, Art. no. 3938502.

[15] Q. Li, K. Xue, H. Xu, W. Pan, and T. Wang, "Vibration-based Terrain classification for mobile robots using support vector machine," *Robot*, vol. 34, no. 6, pp. 660–667, 2012.

[16] K. Iagnemma, H. Shibly, and S. Dubowsky, "On-line Terrain parameter estimation for planetary rovers," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2002, pp. 3142–3147.

[17] K. Legnemma, C. A. Brooks, and S. Dubowsky, "Visual, tactile, and vibration-based Terrain analysis for planetary rovers," in *Proc. IEEE Aerosp. Conf.*, Piscataway, NJ, USA, Mar. 2004, pp. 841–848.

[18] C. A. Brooks, K. Iagnemma, and S. Dubowsky, "Vibration-based Terrain analysis for mobile robots," in *Proc. IEEE Int. Conf. Robot. Automat.*, Piscataway, NJ, USA, Apr. 2005, pp. 3415–3420.

[19] C. A. Brooks and K. Iagnemma, "Vibration-based Terrain classification for planetary exploration rovers," *IEEE Trans. Robot.*, vol. 21, no. 6, pp. 1185–1191, Dec. 2005.

[20] C. Weiss, H. Frohlich, and A. Zell, "Vibration-based Terrain classification using support vector machines," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Piscataway, NJ, USA, Oct. 2006, pp. 4429–4434.

[21] C. Weiss, M. Stark, and A. Zell, "SVMs for vibration-based Terrain classification," in *Autonome Mobile Systeme*. Berlin, Germany: Springer, pp. 1–7, 2007.

[22] E. G. Collins and E. J. Coyle, "Vibration-based Terrain classification using surface profile input frequency responses," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2008, pp. 3276–3283.

[23] D. Tick, T. Rahman, C. Busso, and N. Gans, "Indoor robotic Terrain classification via angular velocity based hierarchical classifier selection," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2012, pp. 3594–3600.

[24] S. Luo, J. Bimbo, R. Dahiya, and H. Liu, "Robotic tactile perception of object properties: A review," *Mechatronics*, vol. 48, pp. 54–67, Dec. 2017.

[25] P. Kozlowski and K. Walas, "Deep neural networks for Terrain recognition task," in *Proc. Baltic URSI Symp.*, May 2018, pp. 283–286.

[26] R. D. Rosenfeld *et al.*, "Unsupervised surface classification to enhance the control performance of a UGV," in *Proc. Syst. Inf. Eng. Design Symp.*, Apr. 2018, pp. 225–230.

[27] F. Zeltner, "Autonomous Terrain classification through unsupervised learning," M.S. thesis, Dept. Comput. Sci., Elect. Space Eng., Luleå Univ. Technol., Luleå, Sweden, 2016.

[28] J. Park, K. Min, H. Kim, W. Lee, G. Cho, and K. Huh, "Road surface classification using a deep ensemble network with sensor feature selection," *Sensors*, vol. 18, no. 12, p. 4342, 2018.

[29] C. Chorley, C. Melhuish, T. Pipe, and J. Rossiter, "Development of a tactile sensor based on biologically inspired edge encoding," in *Proc. IEEE Int. Conf. Adv. Robot. (ICAR)*, Jun. 2009, pp. 1–6.

[30] M. Madry, L. Bo, D. Kragic, and D. Fox, "ST-HMP: Unsupervised spatio-temporal feature learning for tactile data," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May/Jun. 2014, pp. 2262–2269.

[31] L. Cao, R. Kotagiri, F. Sun, H. Li, W. Huang, and Z. M. M. Aye, "Efficient spatio-temporal tactile object recognition with randomized tiling convolutional networks in a hierarchical fusion strategy," in *Proc. AAAI Nat. Conf. Artif. Intell. (AAAI)*, 2016, pp. 3337–3345.

[32] Y. Gao, L. A. Hendricks, K. J. Kuchenbecker, and T. Darrell, "Deep learning for tactile understanding from visual and haptic data," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2016, pp. 536–543.

[33] X. Wang, F. Shi, L. Yu, and Y. Li, *43 Case Analysis of MATLAB Neural Network*. Beijing, China: Beihang Univ. Press, 2011.

[34] *Clearpath*. Accessed: Oct. 2018. [Online]. Available: https://www.clearpathrobotics.com/

CHENGCHAO BAI was born in Zhangjiakou, China, in 1990. He received the B.S. degree in aerospace engineering from the Harbin Institute of Technology, China, in 2013, where he is currently pursuing the Ph.D. degree in aerospace engineering. He has been taking successive postgraduate and doctoral programs of study for doctoral degree, since 2013.

His research interests include intelligent sensing, SLAM, terrain classification, state estimation, motion planning, and optimization. He is also interested in robot intelligence and its application.

JIFENG GUO was born in Xi'an, China, in 1977. He received the B.S., M.S., and Ph.D. degrees from the Harbin Institute of Technology, China, in 2001, 2004, and 2007, respectively, all in aerospace engineering.

From 2004 to 2007, he served as a Lecturer and an Associate Professor with the Harbin Institute of Technology. Since 2015, he has been a Professor with the School of Astronautics, Harbin Institute of Technology. He has authored two books, more than 100 articles, and more than 30 inventions, and holds ten patents. His research interests include intelligent sensing, autonomous planning, on-orbit service, and collaborative control. He is a member of the editor board of the *Journal of Unmanned Systems Technology*.

HONGXING ZHENG was born in Qitaihe, China, in 1990. He received the B.S. degree from the Harbin University of Science and Technology, China, and the M.S. degree from the Harbin Institute of Technology, China, where he is currently pursuing the Ph.D. degree.

His current interests include autonomous planning, decision making, and task assignment.