WristPress: Hand Gesture Classification with two-array Wrist-Mounted pressure sensors

Yufei Zhang[1], Bin Liu[1], Zhiqiang Liu[1], Jinyang Huang[1], Rui Sun[1]

Abstract— **This paper presents a hand gesture recognition system WristPress based on only the pressure sensors, which can reflect the different pressure changes of different hand gestures. Two arrays of force sensitive resistors (FSRs) are arranged around the wrist to capture the pressure fluctuation with the subtle muscle and tendon movements of different gestures, which can help to identify similar gestures for achieving more functions. For distinguishing more gestures with similar muscle and tendon movements, the temporal features and the spatial features of pressures are selected and designed to characterize the relation of every tiny pressure changes corresponding to the muscle and tendon movements at different positions around the wrist. In the WristPress system, 24 kinds of one-gestures, which cover not only the finger movements but also rotations around the wrist and forearm, are classified with an overall 10-fold cross validation classification accuracy of 97.40%. In addition, the WristPress prototype is non-obtrusive with a small size, and is well suited to existing wearable device forms, such as smart watches and a bracelet that are already mounted on the wrist. Our study shows that the temporal features and the spatial features of these pressures can reflect the the correlation between different pressure sensors can improve the accuracy of the hand gesture classification, and the kNN classifier has the best classification accuracy performance 97.40% with a low time complexity.**

I. INTRODUCTION

Hand gesture recognition has been developed as an intelligent, natural and convenient way of human-computer interaction (HCI), which can help the disable people communicate better with others [1]. With increasing use of wearables, the smart watches and other wrist-mounted wearables show many opportunities to detect interactions with the hand on the same side as a wearable [2]. However, a lot of current techniques can be impractical due to some factors, such as requirements for additional sensors mounted independently of the wearable, many kinds of sensors or inability to sense enough types of hand gestures [3].

For improving the usability of the hand gesture system, the gesture interface needs to satisfy the following criteria [4]: First, the good user experience should be guaranteed, thus the disable people can use the interface without physical discomfort or embarrassment. Second, the interface should be easily accessible and don't invade the disable people or other people's privacy in any scenario. Third, enough kinds of hand gestures need to be classified with a high classification accuracy, thus the disable people can utilize the interface for the desired application. To satisfy the above criteria, several kinds of hand gesture classification systems have been studied in the literature: camera based systems [5], [6] , Electromyography (EMG) and Inertial Measurement Unit (IMU) based systems [7], novel approaches [8], [9], pressure and IMU based systems [3], [4]. However, the heavy camera based systems cannot be always carried by the users, and the privacy requirements limit the vision based approach in some scenario. In the EMG and IMU based systems, using the EMGs will introduce many electrodes around the forearm or the wrists, the precise positions of which are closely related to the accuracy of the hand gesture recognition and can be only determined by the specialists [7]. Compared with the above approaches, the pressure and IMU based schemes are more appropriate for the hand gesture recognition, which utilize the off-the-shelf wrist-mounted pressure sensors to distinguish the different muscle and tendon movements around the wrist. In [4], an array of FSRs worn around the wrist were used to distinguish only six kinds of finger pinch gestures with the accuracy of above 80%.

However, with the continuous enrichment of wearable devices, more and more interactive gestures are needed to correspond to different functions. For example, more than a dozen functions have been integrated into smartwatches, such as playing music, answering calls, detecting heart rate and so on. Traditional touch interfaces are hindered by big-finger occlusion on smartwatches [10], and gesture-based interaction can solve the occlusion problems. Furthermore, not only the finger pinch gestures but also some other hand and wrist gestures should be able to be recognized for more useful applications [3]. And considering a better user experience, one-hand interactive gestures are preferred. However, due to the high similarity of one-hand gestures, it's difficult to recognize a lot of gestures only by one-array pressure sensors.

In this paper, we design a hand gesture recognition system with two-array and wrist-mounted pressure sensors to classify a variety of one-hand gestures with a high classification accuracy. The key contributions of this paper mainly are two-fold:

Firstly, we present a simple and low-cost gesture input system that can measure the pressure distribution around the wrist due to the muscle and tendon movements. Then an one-hand gesture set is collected for the validation of the hand gesture recognition, which including a variety of similar gestures. Secondly, with fully consideration of the various elastic strength problem and the similar muscle and tendon

[1]Yufei Zhang, Bin Liu, Zhiqiang Liu, Jinyang Huang and Rui Sun are with the Department of electronic engineering and Information Science, University of Science and Technology of China; Key Laboratory of Electromagnetic Space Information, Chinese Academy of Sciences, Hefei 230027, China, (e-mail: zyf001@mail.ustc.edu.cn; flowice@ustc.edu.cn; lzhq28@mail.ustc.edu.cn; huangjy@mail.ustc.edu.cn, ruisun@mail.ustc.edu.cn).

978-1-5386-7477-2/19/$31.00 ©2019 IEEE

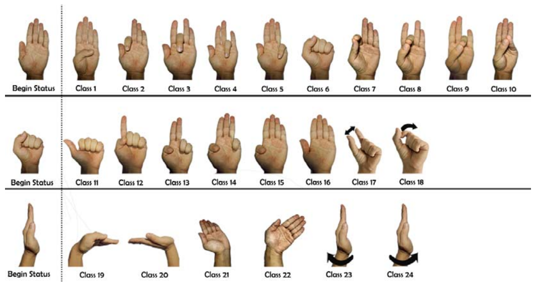

Fig. 1. The hand gesture set contains 18 finger gestures and 6 wrist gestures. The first 10 finger gestures from class 1 to class 10 are changed from the static relax status of the hand. And the following 8 finger gestures from class 11 to class 18 are changed from the static fist status.

movements, the spatial and temporal properties of two-array pressure sensors are designed as the main features, which can commendably reflect the different subtle muscle and tendon movements of similar gestures.

II. SYSTEM PROTOTYPE

A. Hand Gesture Set

In order to measure the usability of our model and the ability to distinguish similar gestures, we designed the gesture sets as shown in Fig. 1, which contains some common gestures and some gestures which are difficult to distinguish. In this paper, we group gestures into two classes: 1) Finger Gestures: these gestures involve movements of the fingers while the wrist keeps steady; 2) Wrist Gestures: the hand movements rotate the whole hand around the wrist joint while the fingers do not flex or extend and they keep steady. And in the gesture set, several finger gestures are very similar and easy to be confused with each other. Thus it is a great challenge to classify these similar gestures with a high classification accuracy.

B. Hardware and Software

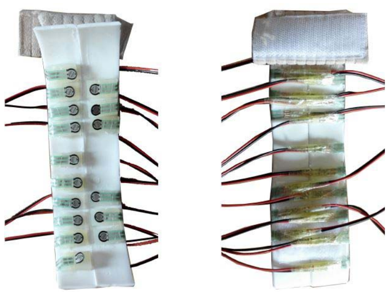

Fig. 2. WristPress prototype. Details of the two arrays of force sensitive resistors (FSRs). Left shows the front of the prototype and right give the details of the opposite.

The WristPress prototype are given and shown in Fig. 2, in which 16 FSRs in two arrays are placed on the independent silicone pads. These silicone pads are soft, and the independence of each other can better collect precise pressures in each position around the wrist. These silicone pads are hold in place by a custom velcro and elastic wrist strap, and the degree of tightness between the pressure sensors and the wrist can be adjusted by the strap. Then, a MSP430FR5969 microcontroller is used to collect 16 channel pressure sensors (FSR400) with 120Hz sample frequency of each sensors.

III. HAND GESTURE RECOGNITION APPROACH

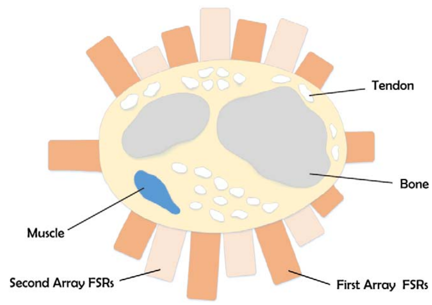

Fig. 3. Cross section of the wrist, the two arrays of FSRs are shown in their locations with heights proportional to values of the FSRs.

A. Problems

As for these similar finger or wrist gestures as shown in Fig. 1, how to distinguish them from each other is a difficult problem. To better solve the problem, two sub-problems are taken seriously and solved, which are given as follows,

a) Various Elastic Strength: For different users, the preferred strengths of the elastic wrist cannot be always kept constant. Even for the same user, the elastic strength may be not the same. Thus the absolute value of pressure values cannot be directly used to the classifier.

b) Similar muscle and Tendon Movements: Based on the biomechanical study of flexor tendons and the finger movements, the pressure values reflects indirectly the level of the muscle and tendon movements. Some muscles or tendons have the similar reflection for some different fingers' flexion and extension, which makes the classification more difficult as shown in Fig. 3.

B. Feature Extraction

To classify similar hand gestures, the proper features should be carefully designed to cope with the above sub-problems. Firstly, to solve the various elastic strength problem, normalization method can be applied to the various pressure values for the effective features. Then we find that when different one or several fingers flex, the muscle or tendon movements in some moment may be similar to each other, however, the temporal properties of the whole finger gesture process are different. In addition, due to the different distribution of tendons in the corresponding region, the change rules between the first-array FSRs and the second-array FSRs are not the same between different hand gestures. Especially for the similar gestures, two-array pressure sensors can get more useful information to distinguish them. Therefore, the correlation relationship of the two-array FSRs can be used to construct effective features for the hand gesture recognition. In summary, the spatial and temporal properties of the two-array FSRs can be taken into consideration while we design these following features.

1) Temporal Features: Firstly, the statistic feature $\frac{\mu_i}{\sigma_i}$ can be used to measure the overall change level of ith channel of pressure data, where $\mu_i = \frac{\sum_{n=1}^{N} x_i(n)}{N}$ is the mean value of the gesture pressure data $x_i(n)$ of channel i, $\sigma_i = \sqrt{\frac{\sum_{n=1}^{N}(x_i(n)-\mu_i)^2}{N}}$ and $i \in \{1, 2, 3, ..., 16\}$ means the channel index. Then, the first order difference $\Delta x_i(n) = x_i(n+1) - x_i(n)$ of the origin pressure data of each gestures x_i can reflect the change of the pressure and directly show the change speed of the muscle and tendon movements. With considering the effect of uncertain noise, the mean of the first order difference over one channel of the gesture sample data $\frac{\sum_{n=1}^{N-1} \Delta x_i(n)}{N-1}$ is selected as a important temporal feature to indicate what kind of trend of change type the gesture's pressure data is, where N is the number of samples in one channel of gestures.

2) Spatial Features: In this paper, these two-array FSRs sensor the pressure values in each position around the wrist, which can capture each subtle muscle and tendon movements. The correlation relationships of every two FSRs can construct the spatial properties of these two positions around the wrist, which can be expressed as $\frac{\sum_{n=1}^{N} x_i(n) \cdot x_j(n)}{N}$, where $i, j \in \{1, 2, 3, ..., 16\}$ and $i \neq j$. Finally, not only the spatial features of one cross section but also the correlation of two cross sections of the wrist are obtained to construct the 3-D spatial structure of the wrist when the finger or wrist flex and muscles or tendons move.

C. Classifiers

In this paper, six kinds of classical classification methods are introduced as the classifiers to distinguish each gesture. They are the Support Vector Machine (SVM) with a linear kernal, the multinomial Naive Bayes (NB), the k-NearestNeighbor (kNN), the Logistic Regression (LR), the Random Forest (RF) and the Decision Tree (DT). And a Python module *Scikit-learn* as the machine learning library is adopted to quickly realize these classifiers [11]. After the feature extraction, both of the temporal features and the spatial features are used to train these classifiers. In order to avoid overfitting, 10-fold cross-validation is used to train these classifiers. The classification accuracy (CA) used in this paper is the ratio of the number of correctly classified gestures to the number of the whole testing gestures.

IV. Results

In this experiment, 10 participants (9 male, 1 female) are recruited to acquire the hand gestures to build the hand gesture set, and the wrist circumferences are in the range of 13.3cm and 18.2cm. Total 24 kinds of hand gestures include 18 finger gestures and 6 wrist gestures are captured, and each gesture is sampled 20 times for 2 second interval for each participant. Finally, 480 gesture samples are collected for one participant, and total of 4800 gesture samples are in the hand set in which each gesture owns 200 samples. To ensure that the participants make the right gestures, we teach each participant how to make these gestures for 10 minutes. The participants do the gestures with the left hand, while the right hand records the beginning and ending of gestures.

These classifiers run in a laptop (Operate system: macOs Sierra v10.12.1, Processor: 1.6GHz Intel Core i5, Graphics: Intel HD graphics 600 1536MB). Firstly, we check whether the design of two-array FSRs can improve the classifier performances. As shown in Fig. 4, we only use the data of first-array or the second-array FSRs to train and test these six classifiers, and then we compare the classifier performances with that using the two-array FSRs. We can see that the classifiers with first-array FSRs generally outperform that with the second-array FSRs, and the classifiers with both two-array FSRs have the best performances. It is because that the first-array FSRs contain more FSRs and are closer to the wrist for capturing more obvious muscle and tendon movements. Further on, comparing with one array FSRs,

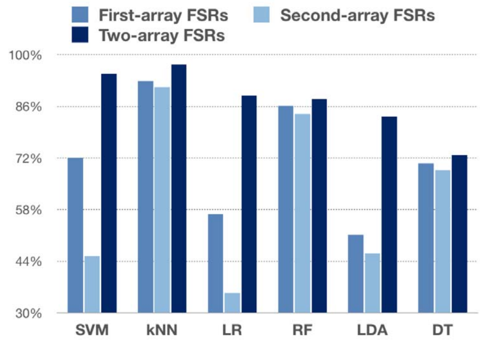

Fig. 4. Classifier performances with using only the first-array FSRs, only second-array FSRs and the two-array FSRs.

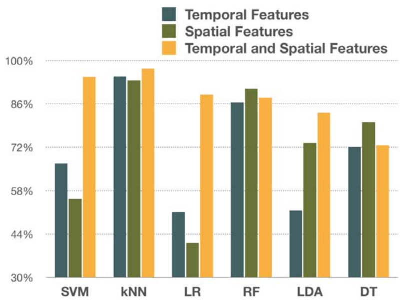

Fig. 5. Classifier performances with using only the temporal features, only the spatial features and both of the temporal and spatial features.

TABLE I
PERFORMANCES OF DIFFERENT CLASSIFIERS IN DETAILS

Classifiers	Classification accuracy	Cost time (s)
SVM	94.75%	8.16
kNN	**97.40%**	**3.94**
LR	88.98%	15.27
RF	87.98%	2.72
LDA	83.19%	0.53
DT	72.73%	11.98

the two-array FSRs can get more useful information and construct more details of the 3-D spatial features when the tendon and muscle move. Then, we compare the classifier performances with different feature sets as shown in Fig. 5. We can find that the SVM, kNN, LR and LDA classifiers have the best performances with both the temporal and spatial features, while the performances of the RF and DT classifiers with both the temporal and spatial features are between that with the temporal features and that with the spatial features. Thus the temporal features and the spatial features plays different roles in different classifiers. And the kNN classifier with both the temporal and spatial features can achieve the highest performance over all other classifiers.

To measure which classifier has the best choice for the hand gesture recognition, we compare the classification accuracy and cost time performances of each classifier applied to the hand gesture set. As shown in Tab. I, we can see that the kNN classifier can obtain the largest classification accuracy 97.40%. And the implementation complexity of the kNN classifier is very low and it can be easily applied to a wearable device in the future work. In addition, the confusion matrix using the kNN classifier is given in Fig. 6, we can find that the misclassification occurred most often in the $9th$, $14th$ and $15th$ gestures. 7% of the $9th$ gesture (ring pinch) are judged as the $8th$ gesture (middle pinch). It is likely to be because the muscle and tendon movements of the middle and the ring finger are too similar when we make the middle pinch and the ring pinch. As for the $14th$ and $15th$ gestures, 6% of the $14th$ gesture are wrongly predicted as the $15th$ gesture, while 10% of the $15th$ gesture are wrongly divided into the $14th$ gesture. As shown in Fig. 1, compared with the $14th$ gesture relaxing three finger (index, middle, ring), the pinky in the $15th$ gesture is simultaneously relaxed with these three fingers. Probably the muscle and tendon movements of the pinky are covered by that of these three fingers, thus some samples of $14th$ and $15th$ gestures are mixed with each other. And we can see that the wrist gestures can achieve high classification accuracies, all of which are larger than 98%. It means the wrist gestures are more easily identified than the finger gestures, because both of the muscle and tendon movements are much intense compared with the subtle muscle or tendon movements of finger gestures.

V. Conclusion

In this paper, we design a WristPress prototype, a wrist-mounted one-hand gesture recognition device with two-array FSR sensors. A hand gesture set which consists of 18 finger gestures and 6 wrist gestures is collected, in which some gestures are very similar. By capturing two-array pressure values around the wrist and using the designed features, the minor differences in tendon movement of similar gestures can be detected. Finally, six kinds of classical classifiers are used to the hand gesture set with the temporal and spatial features. The highest classification accuracy of 97.40% can be obtained using the kNN classifier.

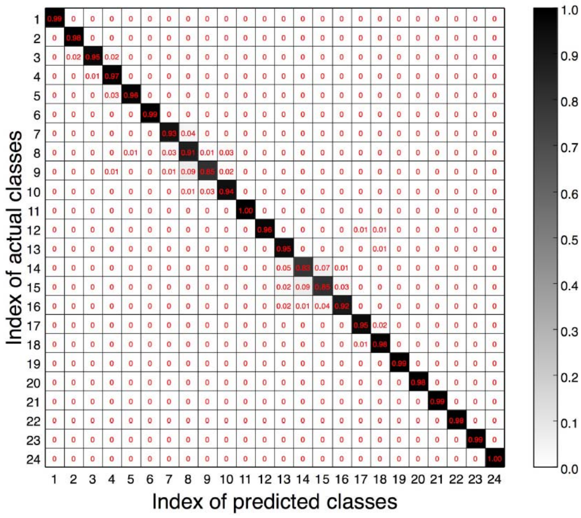

Fig. 6. Confusion matrix across all hand gesture data using kNN classifier, the y label is the index of the actual classes, and the x label is the index of the predicted classes. The number in $(i, j)th$ small box presents the ratio of the predict class j with the actual class i to the total number samples of the class i.

References

[1] S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 1–54, Nov. 2012.

[2] F. Kerber, P. Lessel, and A. Krüger, *Same-side Hand Interactions with Arm-placed Devices Using EMG*. New York, USA: ACM, Apr. 2015.

[3] J. McIntosh, C. McNeill, M. Fraser, F. Kerber, M. Löchtefeld, and A. Krüger, "EMPress: Practical Hand Gesture Classification with Wrist-Mounted EMG and Pressure Sensing ," in *the 2016 CHI Conference*. New York, USA: ACM Press, 2016, pp. 2332–2342.

[4] A. Dementyev and J. A. Paradiso, "WristFlex: Low-Power Gesture Input with Wrist-Worn Pressure Sensors ," in *the 27th annual ACM symposium*. New York, USA: ACM Press, 2014, pp. 161–166.

[5] D. Way and J. Paradiso, "A usability user study concerning free-hand microgesture and wrist-worn sensors," in *2014 11th International Conference on Wearable and Implantable Body Sensor Networks*. IEEE, 2014, pp. 138–142.

[6] D. Kim, O. Hilliges, S. Izadi, A. D. Butler, J. Chen, I. Oikonomidis, and P. Olivier, *Digits: freehand 3D interactions anywhere using a wrist-worn gloveless sensor*, ser. freehand 3D interactions anywhere using a wrist-worn gloveless sensor. New York, USA: ACM, Oct. 2012.

[7] X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, and J. Yang, "A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 41, no. 6, pp. 1064–1076, Sept. 2016.

[8] Y. Su, M. H. Fisher, A. Wolczowski, G. D. Bell, D. J. Burn, and R. X. Gao, "Towards an EMG-Controlled Prosthetic Hand Using a 3-D Electromagnetic Positioning System," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 1, pp. 178–186, Sept. 2016.

[9] G. Laput, R. Xiao, X. A. Chen, S. E. Hudson, and C. Harrison, "Tomo: Wearable, Low-Cost, Electrical Impedance Tomography for Hand Gesture Recognition ," in *the 27th annual ACM symposium*. New York, USA: ACM Press, 2014, pp. 389–394.

[10] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen, "Orbits: Gaze interaction for smart watches using smooth pursuit eye movements," in *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. ACM, 2015, pp. 457–466.

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011.