

Facial Movements Recognition Using Multichannel EMG Signals

Lin Shao

*Propaganda Department of the Party Committee, Information Center
Shan Dong University of Arts
Jinan, China
shaolin@sdca.edu.cn*

Abstract—With the development of intelligent wearable technology, the need for more effective and practical human-computer interaction becomes more and more urgent. The main purpose of this paper is to recognize rich facial movements with multi-channel EMG signals and try to apply them in human-computer interaction. In offline data analysis, we tested a variety of features and classification algorithms. In the two sets of facial motion recognition, the accuracy of 11 facial motion recognition can reach 98.3% and 98.8%, respectively. In the test based on the online typing system, an average accuracy of 83.9% was also achieved.

Keywords—human-computer interaction, machine learning, facial movements, surface electromyography

I. INTRODUCTION

In recent years, intelligent wearable technology has been gaining more and more attention from the academic and the business community around the world, which has been considered as the core and main form of the next generation of smart device [1][2]. Even so, there is still no human-computer interaction method that is particularly suitable for wearable devices.

Human computer interaction (HCI) technology using human physiological signals, such as EEG, EMG, EOG, ECG, has received worldwide attention in recent years[3]. This type of interaction does not require the involvement of hands and voice, and can automatically perceive user's intentions and status, which probably can help the users interact with the smart device more effectively, especially for wearable devices.

EEG signals have been widely used in human-computer interaction. Choi and Cichocki [4] developed a brain-controlled wheelchair system based on motor imagery, which can control wheelchair movement by imagining hand and foot movements.

Iturrate [5] et al. designed a brain-controlled wheelchair based on P300 visual evoked potentials. Yin [6] et al. developed a typing system based on steady state visually evoked potentials (SSVEP). However, due to the low signal-to-noise ratio of EEG signals, the richness and accuracy of interactive commands based on EEG are not enough to make them applicable in daily life.

Compared with EEG signals, EOG signals are easier to collect, and have higher signal-to-noise ratio and recognition rate. Ning [7] employed EOG to control wheelchair movement. Duguleana [8] studied the use of EOG signals to control robot motion. However, the limited number of instructions based on EOG makes it difficult to support complex human-computer interaction tasks. EMG signals have a high signal-to-noise ratio and a rich pattern of changes, which may provide us with a new human-computer interaction mode with rich commands, real-time and high accuracy. Pons [9] designed a robot which can identify six movements of hands by collecting the surface electromyogram signals and help the disabled accomplish many daily tasks. In this paper, we will focus on the identification of rich facial movements patterns through multi-channel EMG signals.

II. EXPERIMENTAL DESIGN

A. Volunteers and Acquisition Device

EMG data of facial movements is collected from 5 volunteers (4 males and 1 females, aged between 24-26) with Emotiv Epoc device. Emotiv Epoc has 14 active electrodes and 2 gyroscope (horizontal, vertical) channels available. The EMG sampling rate is 128Hz. The location and name of the electrodes are shown in Figure 1. The green points are the electrode positions.

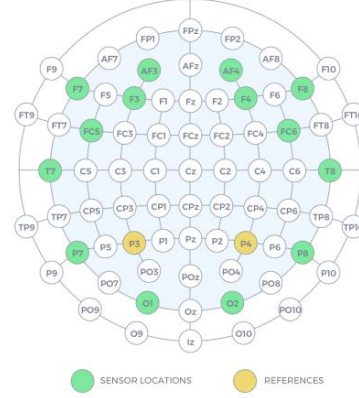


Fig. 1. Acquisition device and electrode distribution

B. Data Acquisition

All the participants signed the consent form before the experiments. We selected 11 facial movements which are suitable for human-computer interaction. Then subjects were required to complete the 11 facial movements with each movements 100 times. All facial movements are shown in Table I .

TABLE I. FACIAL MOVEMENTS

Number	Facial Movements
1	Single blink
2	Pick eyebrow
3	Bite
4	Glance to the left
5	Glance to the right
6	Look up
7	Look down
8	Head to left
9	Head to right
10	Head down
11	Head up

III. METHOD

A. Data Preprocessing

Band-stop filter is used to remove 50 Hz power frequency interference, and de-mean method is used. Through analyzing the experimental data of 5 subjects , we found that in the 11 facial movements, most movements can be completed in 1

second except ‘turn the head’with about 1.5 seconds. So we choose 1.5 seconds as the time length for signal segmentation. Furthermore, we find that the highly correlated channels were mainly concentrated in the forehead. So 8 EMG channels (AF3, F7, F3, FC5, FC67, F4, F8, AF4) and GyroX, GyroY were used in the following analysis.

B. Feature Extraction

Referring to [10-12], we extract time-domain, frequency-domain and time-frequency-domain features and the relationship features between channels. Through comparative experiments, we can find out the features of good performance. At last, 16 common and effective features, including 7 time-domain features, 6 frequency-domain features and 3 time-frequency-domain features are selected. Table II-IV give the formula of 16 feature extraction methods.

TABLE II. TIME-DOMAIN FEATURES

Feature Name	Feature Formula
Integrated EMG (IEMG)	$IEMG = \sum_{n=1}^N x_n $
Mean Absolute Value (MAV)	$MAV = \frac{1}{N} \sum_{n=1}^N x_n $
Segmental Energy (SE)	$SE_k = \sum_{i=k}^{k+\frac{1}{10}N} x_i$ $k \in \{1, \frac{N}{10}+1, \frac{2N}{10}+1, \dots, \frac{9N}{10}+1\}$
Modified Mean Absolute Value 1 (MMAV)	$MMAV = \frac{1}{N} \sum_{i=1}^N w_i x_i $ $w(i) = \begin{cases} 1, & 0.25N \leq i \leq 0.75N \\ 0.5, & otherwise \end{cases}$
Variance (VAR)	$VAR = \frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2$
Root Mean Square (RMS)	$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^N x_n^2}$

Feature Name	Feature Formula
Zero Crossing (ZC)	$ZC=ZC+1,$ when $\{x_i > 0 \text{ and } x_{i+1} < 0\}$ 或 $\{x_i < 0$ and $x_{i+1} > 0\}$

TABLE III. FREQUENCY-DOMAIN FEATURES

Feature Name	Feature Formula
Auto-Regressive coefficients (AR)	$x_k = \sum_{i=1}^N a_i x_{k-i} + e_k$
Frequency Median (FMD)	$F_{MD} = \frac{1}{2} \sum_{i=1}^M PSD_i$
Frequency Mean (FMN)	$F_{MN} = \frac{\sum_{i=1}^M f_i PSD_i}{\sum_{i=1}^M PSD_i}$ $f_i = (i * samplingRate) / (2 * M)$
Modified Frequency Median (MFMD)	$MFMD = \frac{1}{2} \sum_{j=1}^M A_j$
Modified Frequency Mean (MFMN)	$MFMN = \frac{\sum_{j=1}^M f_j A_j}{\sum_{j=1}^M A_j}$
Frequency Ratio (FR)	$FR_j = \frac{ F(\cdot) _{lowfreq}}{ F(\cdot) _{highfreq}}$

TABLE IV. TIME-FREQUENCY-DOMAIN FEATURES

Feature Name	Feature Formula
Short Time Fourier Transform (STFT)	$STFT_k(t, w) = \int W^*(\tau - t) x(\tau) e^{-jw\tau} d\tau$
Wavelet Transform (WT)	$W_x(a, b) = \int x(t) (\frac{1}{\sqrt{a}}) \Psi^*(\frac{t-b}{a}) dt$
Discrete Wavelet Transformation (DWT)	$C_{j,k}(f, \varphi_{j,k}) = 2^{-j/2} \sum_{n=-\infty}^{\infty} f(t) \varphi(2^{-j}t - k)$

C. Classifiers

Four classification methods are used to recognize 11 types of facial movements: random forest, k-nearest neighbor network (KNN), support vector machine (SVM) and decision tree.

IV. RESULT

A. Classification Methods

We use four classification methods: decision tree, random forest, KNN and SVM. Decision tree method improves the

performance of the model by adjusting the maximum number of splits and the learning rate. Random forest method improves the performance of the model by adjusting the maximum number of splits, number of learners and learning rate. KNN is employed with K-nearest neighbor based on Euclidean distance, K-nearest neighbor based on similarity and K-nearest neighbor based on weight, and the performance of the model is improved by adjusting the number of nearest neighbors. SVM improves the performance of the model by adjusting the type of kernel function. Kernel functions can be classified into three categories: linear kernels, polynomial kernels and Gaussian kernels. The results of parameter selection for each method are shown in Tables V to VIII.

TABLE V. DECISION TREE

	Maximum Number of Splits				
	100	300	500	800	1200
Accuracy(%)	90.3	90.3	90.3	90.3	90.3

TABLE VI. KNN

Accuracy(%)	Number of Nearest Neighbors					
	1	3	5	7	11	15
KNN based on euclidean distance	96.6	96.3	95.8	94.9	94.2	93.4
KNN based on similarity	96.8	96.3	96.5	95.9	95.3	95.0
KNN based on weight	96.6	96.6	96.1	95.7	94.9	93.8

TABLE VII. RANDOM FOREST

Accuracy(%)	Number of Learners			
	50	100	200	300
Maximum Number of Splits				
200	97.7	97.7	97.8	97.9
500	97.8	97.7	97.8	97.9
1000	98.2	97.8	98.0	98.0
1500	97.6	98.2	97.8	98.0

TABLE VIII. SVM PARAMETER SELECTION

	Type of Kernel Function		
	Linear kernels	Polynomial kernels	Gaussian kernels
Accuracy(%)	96.2	96.1	94.0

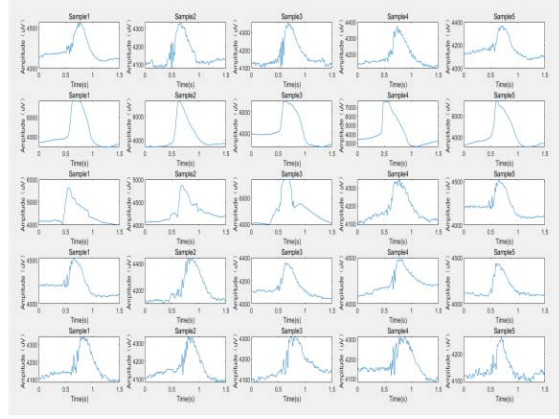
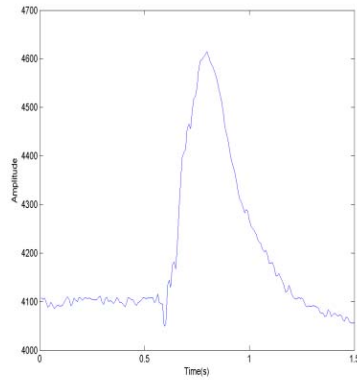


Fig. 2. The left picture shows a typical EMG signal of the FC5 channel with eyebrow picking. The right picture shows the EMG signal of the FC5 channel when with eyebrow picking for all five subjects.

B. Comparative Experiments

Eight comparative experiments were carried out. The best parameters obtained in the previous section were used for each classifier, and ten-fold cross validation was carried out. The results are shown in Table IX. (1) Raw data using all channels; (2) Raw data after channel selection with AF3, F7, F3, FC5, FC67, F4, F8, AF4, GyroX, GyroY; (3) Only time-domain features after channel selection; (4) Only frequency-domain features after channel selection; (5) Time-frequency-domain features after channel selection; (6) Time-domain, frequency-domain and correlation features between channels after channel selection; (7) Combine all kinds of features.

TABLE IX. COMPARATIVE EXPERIMENTS

Methods	Random Forest	KNN	SVM	Decision Tree
raw data+all channels	95.0	95.0	90.2	79.4
raw data+channel selection	95.5	95.3	89.8	77.0
time-domain features+channel selection	98.2	96.8	96.2	90.3
frequency-domain features+channel selection	93.8	83.6	92.1	86.7
time-frequency-domain features + channel selection	95.9	85.7	93.8	90.1
features between channels+channel selection	94.9	90.7	88.6	82.6
combine features+channel selection	98.3	97.2	96.8	92.3

C. Discussion of Experimental Results

From the above results, we can obtained:

- Random forest has the highest accuracy among all classification methods due to its powerful learning ability brought by integrated learning.
- The time-domain features have good performance best. The accuracy of SVM and KNN is higher than 96%, and that of random forest is 98.2%.
- The best performance is obtained when Combining all kinds of features for all the four methods.

V. MORE ROBUST SOLUTION

The above results show that facial movements can be accurately identified by EMG signals and so can be used in some simple human-computer interaction applications. We also design another system to obtain more robust interaction commands with facial EMG by recognizing the EMG patterns of deliberate single blink, deliberately two-blink, pick eyebrows, bite, long bite, left side bite, long left side bite, right side bite, long right side bite, left eye two-blink, right eye two-blink.

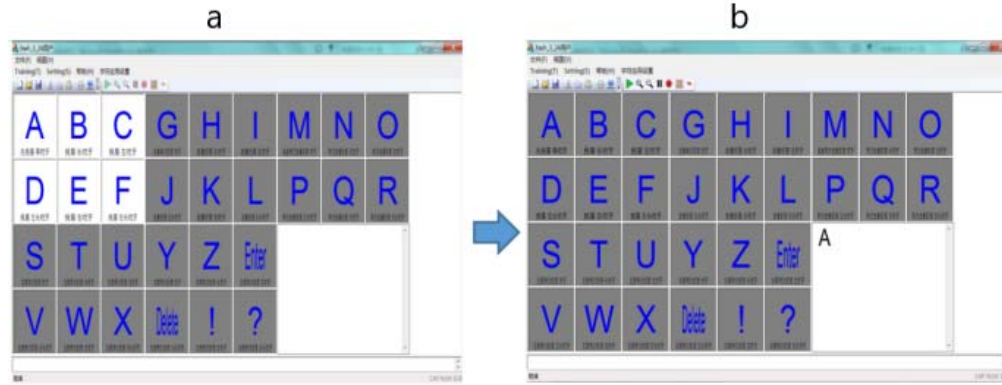


Fig. 3. User interface

TABLE X. COMMAND COMBINATION

	Bite	Long Bite	Left Side Bite ,	Long Left Side Bite	Right Side Bite	Long Right Side Bite
pick eyebrows	A	B	C	D	E	F
deliberate single blink	G	H	I	J	K	L
deliberately two-blink	M	N	O	P	Q	R
left eye two-blink	S	T	U	V	W	X
right eye two-blink	Y	Z	Enter	Delete	!	?

We use the similar experiment process to collect data from five subjects, and all kinds of features mentioned above and random forest as classifier. The results are shown in Table XI.

TABLE XI. RESULTS

Methods	Random Forest	KNN	SVM	Decision Tree
combine features+channel selection	98.8	96.5	97.9	92.7

VI. TYPING SYSTEM BASED ON FACIAL MOVEMENTS

From Table XI, we can find that the facial movements have high recognition accuracy, and these movements are rarely done involuntarily. A typing system is designed using the facial movements to verify the validity of our method.

A. System Structure

To implement a typing system, we need at least 26 commands (26 letters). But we don't have so much facial movement patterns, the commands are combined to provide enough interaction commands. The combination method is shown in Table X. For example, if we want to type the letter "A", we need to pick eyebrows first and then bite. In real-time systems, we need to segment the signals online before recognized the facial movements. Sliding window method [13]

is used to segment signals online. The segmented signals are extracted and then put into a random forest classifier to recognize and generate corresponding commands.

B. Typing System

Figure 3 shows the user interface of the typing system. There are 26 letters and "enter", "delete", "?", "!" on the interface. According to the combination method in Table X, we want to type the letter "A", first we need to pick eyebrows. Since the buttons A, B, C, D, E and F all need to do eyebrow picking first, after the system detects the eyebrow picking action, the buttons A, B, C, D, E and F are selected (Figure 3-a). After eyebrow picking, the user needs to start to bite immediately. After the system detects the biting action, the A letter will be typed in the white area (Figure 3-b). If no biting action is detected within 1.5 seconds, the selected button will be restored to the unchecked state.

C. System Test

A total of three subjects participated for testing, and each subject had four rounds. Subjects are asked to type out a sentence consisting of 30 characters. The test results are presented in Table XII. The average accuracy of the three subjects reached 83.9%. The three subjects typed at speeds of 2-2.5 seconds per character and could type 24-30 characters in one minute.

TABLE XII. TEST RESULTS

Accuracy(%)	Test 1	Test 2	Test 3	Test 4	Average
subject 1	66.7	83.3	76.7	86.7	78.4
subject 2	80.0	76.7	80.0	83.3	80.0
subject 3	86.7	93.3	96.7	96.7	93.4

VII. CONCLUSION

Several kinds of facial movements can be accurately recognized based on multi-channel EMG. This paper describes the methods in detail, including experiment design, feature extraction and classification. Finally, the validity and effectiveness of the method is illustrated by a typing system.

REFERENCES

- [1] Kumari, P., Mathew, L., Syal, P.: Increasing trend of wearables and multimodal interface for human activity monitoring: a review. *Biosens. Bioelectr.* 90, 298-307 (2017).
- [2] Chan, M., Esteve, D., Escriba, C., Campo, E.: A review of smart homes - present state and future challenges. *Comput. Meth. Prog. Biomed.* 91(1), 55-81 (2008).
- [3] Nam Y , Koo B , Cichocki A , et al. GOM-Face: GKP, EOG, and EMG-Based Multimodal Interface With Application to Humanoid Robot Control[J]. *IEEE Transactions on Biomedical Engineering*, 2014, 61(2):453-462.
- [4] Choi K and Cichocki A . Control of a Wheelchair by Motor Imagery in Real Time[C]// International Conference on Intelligent Data Engineering &Automated Learning. Springer-Verlag, 2008.
- [5] Iturrate I , Montesano L , and Minguez J . Synchronous EEG Brain-Actuated Wheelchair with Automated Navigation[C]// IEEE International Conference on Robotics & Automation. IEEE Press, 2009.
- [6] Yin E , Zhou Z , Jiang J , et al. A dynamically optimized SSVEP brain-computer interface (BCI) speller[J]. *IEEE transactions on bio-medical engineering*, 2014, 62(6).
- [7] Ning B , Li M , Liu T , et al. Human brain control of electric wheelchair with eye-blink electrooculogram signal.[J]. 2012.
- [8] Duguleana M and Mogan G . Using Eye Blinking for EOG-Based Robot Control[M]// The Stellar Populations of Galaxies. Springer Netherlands, 2010.
- [9] Pons J L , Rocon E , Ceres R , et al. The MANUS-HAND Dextrous Robotics Upper Limb Prosthesis: Mechanical and Manipulation Aspects[J]. *Autonomous Robots*, 2004, 16(2):143-163.
- [10] Phinyomark A , Phukpattaranont P , and Limsakul C . Feature reduction and selection for EMG signal classification[J]. *Expert Systems with Applications*, 2012, 39(8):7420-7431.
- [11] Puttasakul T , Sangworasil M , and Matsuura T . Feature extraction of wavelet transform coefficients for sEMG classification[C]// 2014 7th Biomedical Engineering International Conference (BMEiCON). IEEE, 2014.
- [12] Mahaphonchaikul K , Sueaseenak D , Pintavirooj C , et al. EMG Signal Feature Extraction Based on Wavelet Transform[C]// International Conference on Electrical Engineering/electronics Computer Telecommunications & Information Technology. IEEE, 2010.
- [13] Zhang J , Huang W , Zhao S , et al. Recognition of Voluntary Blink and Bite Base on Single Forehead EMG[J]. 2017.