

# Relative Confidence based Information Fusion for sEMG-based Pattern Recognition

Jinrong Li<sup>1,2</sup>, Yinfeng Fang<sup>2</sup>, Yong Ning<sup>1</sup>, Jing Jie<sup>1</sup>, Ping Tan<sup>1</sup>, Honghai Liu<sup>2</sup>

1. School of Automation and Electric Engineering, Zhejiang University of Science & Technology, Hangzhou, China

2. Intelligent Systems and Biomedical Robotics Group, School of Computing, University of Portsmouth, Portsmouth, UK

E-MAIL: [lijinrong\\_hz@163.com](mailto:lijinrong_hz@163.com), [yinfeng.fang@port.ac.uk](mailto:yinfeng.fang@port.ac.uk), [Ningyong0816@126.com](mailto:Ningyong0816@126.com), [jjing277@sohu.com](mailto:jjing277@sohu.com), [105011@zust.edu.cn](mailto:105011@zust.edu.cn), [Honghai.Liu@port.ac.uk](mailto:Honghai.Liu@port.ac.uk)

## Abstract:

This study aims to establish a decision fusion scheme that synthesizes the advantages of different classifiers and avoids uncertain decisions. Thus, relative confidence factors of each classifier were proposed to correct the classification decision made by each classifier, and the final classification result was derived by fusing all corrected decisions based on the combination of Dempster-Shafer's rule. The novel fusion scheme is evaluated in the scenario of sEMG-based hand movement identification, in which five classifiers are adopted. The experimental results demonstrated that the novel scheme can obtain higher classification accuracy and stability than the other methods.

## Key words:

sEMG; myoelectric; pattern recognition; evidence theory; information fusion

## 1. Introduction

Surface Electromyography (sEMG) signal, as one kind of electrophysiological signals, is extensively studied and applied in clinic and engineering[1,2]. Especially for the pattern recognition of hand movements, sEMG-signal has been revealed as a powerful tool due to its characteristics of analytical speed, high resolution, ease of collection and well security[3-4]. For the reason of noise and instability, however, the raw sEMG signal is rarely involved directly as a robust source of input data for pattern recognition systems. Instead of raw sEMG signals, the extracted features are usually taken as the input of pattern identification models. The existing features extraction methods mainly include time domain[5], frequency domain[6], time-frequency domain[7], and nonlinear characteristics[8] *etc.* Up to now, there're not yet any kind of features that have the ability to represent the whole mathematical characters of sEMG signals. Especially when more patterns need to be identified from sEMG signals, more kinds of features need to be extracted and fed to the pattern recognition model.

However, there're many noticeable risks faced by most classification algorithms as the dimension of input data increases, which is the so-called curse of dimensionality[9]. With adding of searching space dimension, the exponentially

increased algorithm complexity could lead to the collapse of computational speed. Moreover, because of some counter-intuitive geometrical properties in higher-dimension space, the algorithm may not be able to find the global optimal result in higher-dimension space, although it is easy in low-dimension space.

In order to improve the classification performance and avoid the curse of dimensionality, information fusion technology was adopted in this work. Dempster-Shafer (DS) theory (also known as evidence theory or Dempster-Shafer theory of evidence) [10] has been widely used in the field of pattern recognition since it was introduced in 1967, which is a framework that provides schemes to combine evidence from different sources and arrive at a degree of belief. In the applications of D-S theory, the main problem is how to define the basic belief assignment (BBA) for the identification frame based on some prior information. Moreover, a fact in pattern identification is that different classification results based on the same set of evidence may have different reliabilities and the reliability of each result should be taken into account during the fusion phase.

In this work, a novel evidence fusion method based on relative confidence is reported. The separability of each class was evaluated by a t-statistic created based on extracted features of training samples. Based on this, the relative confidence of a set of quadratic classifiers trained by various sets of features on each identified class was derived and used to correct the classification result for the purpose to take the advantage of each classifier and avoid its weakness effectively. At last Dempster's rule of combination was taken to fuse the corrected results produced by all classifiers to get the final identification output. In Section 4, it was demonstrated that the fusion method integrated with relative confidence factors can accurately identify the hand movements based on sEMG signals. Furthermore, a comparative study showed that the new fusion method could achieve better accuracy and robustness than four other classification methods which include three methods not using any fusion technologies and one just using the traditional D-S combination rules.

## 2. Relative confidence factors

Almost all machine learning methods will face to the curse of dimensionality as the dimension of input data increase. So, it's more reasonable to train a number of classifiers each based on a few of features rather than to train only one classifier based on all features. But because the input features of each classifier can't represent the mathematical characters of the original signal completely, the results of all classifiers need to be fused to obtain the final decision. It should be noticed that different classifiers also differ in their ability to indentify different classes, since they are trained based on different set of signal features, and thus this difference need to be taken into account during the fusion process.

Normally for pattern recognition issue, a general assumption is that the distribution of samples in each class is Gaussian. Therefore, the separability of two classes can be evaluated by the distance between their mean values. Usually, it's reasonable to judge two classes can be separated from each other with large probability when the distance between their population mean values is greater than two times of maximum variance, that is

$$|\mu_1 - \mu_2| > 2 \max(\sigma_1, \sigma_2) \quad (1)$$

Where  $\mu_1$  and  $\mu_2$  refers to the population mean values of the two classes,  $\sigma_1$  and  $\sigma_2$  refers to the variances of the two classes respectively. But we have no idea about the actual values of  $\mu_1$  and  $\mu_2$  just relying on a limited number of training samples. In this work, the probability  $P(|\mu_1 - \mu_2| > 2 \max(\sigma_1, \sigma_2))$  was adopted as an index to judge the separability between two classes, which can be described by Eq. (2), denoting  $\max(\sigma_1, \sigma_2)$  as  $\delta$

$$\begin{aligned} & p(|\mu_1 - \mu_2| > 2\delta) \\ & = p((\mu_1 - \mu_2) > 2\delta) + p((\mu_1 - \mu_2) < -2\delta) \end{aligned} \quad (2)$$

Assuming samples set  $\mathbf{X} = \{x_1, x_2, \dots, x_{n_1}\}$  and  $\mathbf{Y} = \{y_1, y_2, \dots, y_{n_2}\}$  come from two classes with Gaussian distribution, a *t-statistic* can be created as

$$t = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2), \quad (3)$$

where  $\bar{X}$  and  $\bar{Y}$  are the sample means of the two classes respectively, and  $S_w$  can be calculated by the following formula,

$$S_w = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}, \quad (4)$$

where  $S_1$  and  $S_2$  denote the sample variances of the two classes respectively.

Let  $f(t)$  represents the probability density function of t-distribution with the degree of freedom  $(n_1 + n_2 - 2)$ , for given values  $\alpha_1$  and  $\alpha_2$ , the interval probabilities of  $t$  are :

$$P(t < \alpha_1) = \int_{-\infty}^{\alpha_1} f(\tau) d\tau \quad (5)$$

$$P(t > \alpha_2) = \int_{\alpha_2}^{\infty} f(\tau) d\tau \quad (6)$$

Replace the variable  $t$  as its form defined in Eq. (3), the value of  $\alpha_1$  and  $\alpha_2$  can be resulted out as Eq. (7) based on our judgement standard, which was defined in Eq. (1)

$$\alpha_1 = \frac{\bar{X} - \bar{Y} - 2\delta}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \alpha_2 = \frac{\bar{X} - \bar{Y} + 2\delta}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (7)$$

As discussed above, the separable probability of two classes based on training samples, denoted as  $\rho_{1,2}$ , can be derived as bellow:

$$\begin{aligned} \rho_{1,2} & = P(|\mu_1 - \mu_2| > 2 \max(\sigma_1, \sigma_2)) \\ & = \int_{-\infty}^{\alpha_1} f(\tau) d\tau + \int_{\alpha_2}^{+\infty} f(\tau) d\tau \end{aligned} \quad (8)$$

If there are  $N$  classes need to be identified and the dimension of input data is  $M$ , the probability of the  $i^{th}$  class can be distinguished from all other classes based on each given feature can be expressed as a matrix  $\mathbf{\Omega}(i) \in \mathbf{R}^{M \times (N-1)}$ , the element of which, denoted as  $p(i)_{k,j}$ , represent the probability of the  $i^{th}$  class can be separated from the  $j^{th}$  class ( $j \neq i$ ) based on the  $k^{th}$  feature, i.e.  $\mathbf{\Omega}(i) = \{p(i)_{k,j}\}$ ,  $k=1, \dots, M; j=1, \dots, N$  and  $j \neq i$ , and

$$p(i)_{k,j} = P(|\mu_k(i) - \mu_k(j)| > 2 \max(\sigma_k(i), \sigma_k(j))) \quad (9)$$

Let  $\max_k(\mathbf{\Omega}(i))$  denote the vector composed by the column maximum in  $\mathbf{\Omega}(i)$ , which represent the maximal probability of  $Class_i$  can be separated from all other classes based on given features. Thus the minimum of this vector means the separable probability between the  $Class_i$  and the specific class which is most difficult to distinguish from. So, for  $N$  classes and  $M$  features, it's reasonable to take the single value  $SP(i)$ , which is defined as bellow, denote the separable probability of  $Class_i$ .

$$SP(i) = \min \left( \max_k(\mathbf{\Omega}(i)) \right) \quad (10)$$

After the definition of separable probability of each

class based on given features has been created, it will be much easier to define the formula for Relative Confidence Factor. If there are a total of  $K$  classifiers trained by different sets of features each, let vector  $\mathbf{V}(i)=[SP_1(i), \dots, SP_K(i)]$  denotes the separable probabilities of  $Class_i$  based on different sets of features, and then the relative confidence factor of the  $i^{th}$  classifier on  $Class_i$  was defined as,

$$Rbel_i(i) = SP_i(i) / \max(\mathbf{V}(i)), \quad (11)$$

which is the ratio of each separable probability of  $Class_i$  to the maximum of them, that means the value of relative confidence factor is between 0 and 1.

### 3. Material and Methods

#### 3.1 sEMG Data set and Preprocessing

The sEMG dataset(sEMG for Basic Hand movements Data Set) [11] analyzed in this paper includes electromyography data produced by hand movements of 6 subjects (2 females and 4 males), who were asked to repeat the six movements including: 1) Cylindrical (C): for holding cylindrical tools, 2) Hook (H): for supporting a heavy load, 3) Lateral (L): for holding thin, flat objects, 4) Palmar (P): for grasping with palm facing the object, 5) Spherical (S): for holding spherical tools, 6) Tip (T): for holding small tools. All of the movements were shown in Fig.1 and each movement was repeated 30 times lasting 6s for every time.



**FIGURE 1. Hand movements[11]: a) Cylindrical, b) Hook, c) Lateral, d) Palmar, e) Spherical, f) Tip**

Considering of wearing comfort and economic costs, only two surface EMG electrodes were placed on the main activity spots of the Flexor Capri Ulnaris and Extensor Capri Radialis to collect the raw sEMG signal at 500Hz. Therefore, for each subject a total of 180 6-second long 2-channel EMG signals were recorded.

A band-pass Butterworth filter with low cutoff frequency 15Hz and high cutoff frequency 500Hz was applied on the raw sEMG signal for the purpose to eliminate the noise coming from device hardware and testing environment. Then a 50-Hz Hampel filter was applied as well to eliminate the interference produced by AC power line.

The muscle contraction relating to a movement generally doesn't start at the beginning of each collected sEMG signal due to human reaction time and experimental conditions. So the sEMG active phases need to be extracted before further analysis. The sliding time-window approach[4] was applied to detect the active phase through checking if the value of IMEG(Integrated Electromyogram) exceed a predefined threshold.

#### 3.2 Feature extraction and Classifier training

Two sets of features, one was extracted in time domain and the other was extracted in frequency domain, were adopted as input data to train two classifiers respectively based on QDA(Quadratic Discriminant Analysis) in this work.

Three time domain sEMG features, including: mean absolute value (MAV), waveform length (WL), and logvariance(LogVar), were employed in the current study, which had demonstrated their efficacy in hand movements identification[5]. So six time domain features were extracted from the two sEMG channels.

The frequency domain features of sEMG are seldom used in sEMG-based movement recognition relative to time dominant features due to the non-stationary of sEMG and broadband characteristic noise. For this the empirical mode decomposition (EMD) method[12], which is designed to work well for nonstationary and nonlinear signal, was applied before frequency features extraction. EMD algorithm can decompose a signal into a series of narrow-band signals called intrinsic mode functions (IMFs). And the mean power frequency(MPF) of the first 3 IMFs decomposed from original sEMG signal were taken as frequency dominant features in this work. Therefore, the total number of frequency domain features is six which is the same as the dimension of time domain features..

For movement identifying from sEMG signals, we employed a quadratic discriminant analysis(QDA) classifier. Compared to linear discriminant analysis(LDA), QDA method could work better on some practical applications especially for lower-dimension input.

The extracted time domain features and frequency domain features were fed as input respectively to the classifiers and the actual kinds of movement were used as targets. Two classifiers, *Classifier\_TD*(taking time dominant

features as input) and *Classifier\_FD* (taking frequency dominant features as input), were trained by using data from individual subjects. And then the vector containing relative confidence factors of each classifier was calculated following the method discussed in Section 2.1.

### 3.3 Information fusing and movements identifying

The following steps, which are carried out sequentially, illustrate the process of movements recognition based on a given sEMG signal as well as the decision fusion scheme.

**Step 1:** Extract the time domain features and frequency domain features from the given sEMG signal.

**Step 2:** Feed the time domain features and frequency domain features to the two QDA classifiers: *Classifier\_TD* and *Classifier\_FD* respectively. The decision of QDA classifier is an N-dimension (N: the number of classes for classification, which equals to six in our experiment) vector describing the posterior probability distribution over all classes, which can be taken as the basic belief assignment that will be used in the following fusion step. So, let  $\mathbf{m}_i = [m_i(1), m_i(2), \dots, m_i(N)]^T, i = 1, 2$  denote apart the decision of the two classifiers, the element in which is the probability of each class, and  $\sum_{j=1}^N m_i(j) = 1$ .

**Step 3:** Calculate the relative confidence factors vectors for *Classifier\_TD* and *Classifier\_FD* by the method discussed in Section 2. Let  $\mathbf{rcof}_i = [rc_i(1), \dots, rc_i(N)], i = 1, 2$  denote the relative confidence factors vector calculated by time domain features and frequency domain features respectively.

**Step 4:** Correct the classification decision obtained in Step 2 based on the following equation,

$$\mathbf{m\_corr}_i(j) = m_i(j) \times rc_i(j) + \varepsilon/N, \quad (12)$$

where,  $\varepsilon$  is the residual belief defined as:

$$\varepsilon = 1 - \sum_{j=1}^N p_i(j) \times rc_i(j) \quad (13)$$

The vector  $\mathbf{m\_corr}_i$  expresses the corrected belief assignment over all classes, so the sum of it is one.

**Step 5:** Combine the corrected decisions of all classifiers based on D-S combination rules. Finally, the probability distribution over classes can be obtained under the combined effect of multiple sets of evidence, and then choose the class with maximum probability as the

classification result of the tested sample.

## 4. Results and Discussion

To illustrate the advantage of the proposed information fusion method based on relative confidence factors, the results of 5 classifiers, described as follow, were compared.

I. *Classifier\_TD*, trained based on time domain features.

II. *Classifier\_FD*, trained based on frequency domain features.

III. *Classifier\_All*, trained based on all features (time domain and frequency domain).

IV. *Classifier\_DS*, fusing decisions of *Classifier\_TD* and *Classifier\_FD* based on traditional D-S combination rules.

V. *Classifier\_RC*, fusing decisions of *Classifier\_TD* and *Classifier\_FD* based on relative confidence factors

For total of 30 sEMG signals produced by each gesture repeated by an individual subject, 2/3 of them were selected randomly as training samples and the remainder was used as testing samples to verify the performance of the classifier. The experiment was repeated 20 times for each subject.

Fig. 2 provided the average classification accuracy of the five classifiers during the 20 tests. It's shown in Fig.2 that *Classifier\_RC* using the fusion method proposed in this work demonstrated a better performance than the others in most tests (only except the 4<sup>th</sup>, 14<sup>th</sup> and 16<sup>th</sup> tests). *Classifier\_DS* using the traditional D-S combination method can achieve comparable accuracy as *Classifier\_RC* in some tests, but in more tests its performance is not as good as imagined, even worse than *Classifier\_TD*, that's because it doesn't take account of the classification confidence which is different for each classifier during the fusion procedure. Meanwhile, it can be seen that the robust of *Classifier\_RC* is the best among these classifiers. The reason lies in its ability to synthesize the advantages of different classifiers (*Classifier\_TD* and *Classifier\_FD*) and avoid the classification decision with large uncertainty, which makes the performance more stable when it faces to different set of training samples. On the other side *Classifier\_All* showed poor performance in almost all tests, just a little better than *Classifier\_FD*. Even it takes more comprehensive features (time domain and frequency domain) as input, but because QDA classifier has higher sensitivity to curse of dimension, the classification performance cannot achieve effective improvement.

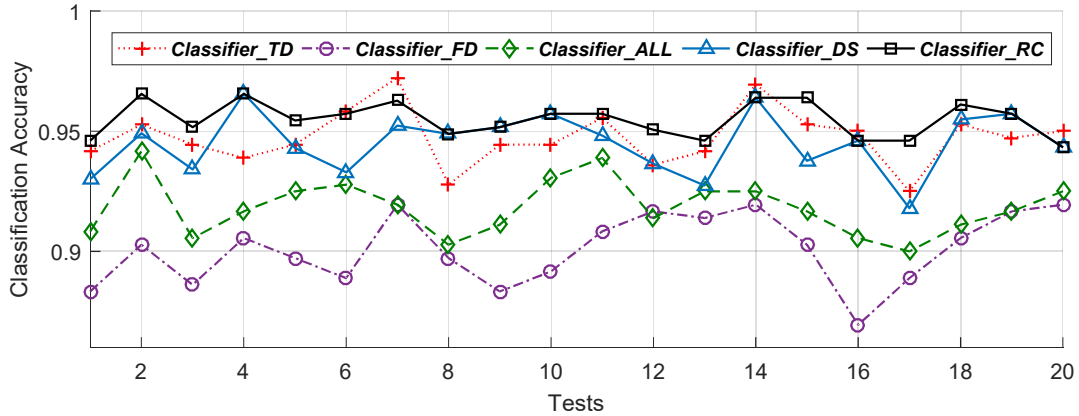


FIGURE 2. Comparison of classification accuracy of 5 classifiers

Tab.1 Classification accuracy of 5 classifiers

	Gestures					
	Cylindrical	Hook	Lateral	Palmar	Spherical	Tip
<i>Classifier_TD</i>	0.9623	0.9550	0.8923	0.9150	0.9583	0.9025
<i>Classifier_FD</i>	0.9350	0.9533	0.8667	0.8817	0.9783	0.8850
<i>Classifier_All</i>	0.9092	0.9242	0.8475	0.8892	0.9742	0.8608
<i>Classifier_DS</i>	0.9541	0.9678	0.8764	0.9047	0.9782	0.9023
<i>Classifier_RC</i>	0.9775	0.9733	0.9142	0.9375	0.9825	0.9200

Based on all classification results of the 20 tests, the accuracy of the 5 classifiers for the 6 gestures were listed in Tab. 1, from which, it could be found that *Classifier\_RC* could achieve better accuracy for each movement, which benefitted from the fusion of the results produced by *Classifier\_TD* and *Classifier\_FD*. But misclassification still exists for *Classifier\_RC* when the other two basic classifiers make mistakes simultaneously.

## 5. Conclusion

The concept of relative confidence factors of classifier over each class was proposed to evaluate the reliability of each decision produced by each classifier. A decision fusing scheme was introduced by using these relative confidence factors to correct the classification results (a probability distribution over all classes) of multiple classifiers. The use of relative confidence factors enables the fusion process to take the advantages of different classifiers and avoid the effects of uncertain decision. The performance of this novel method was verified by a sEMG-based pattern recognition experiment, in which two classifiers were trained by time domain features and frequency domain features respectively at first, and then the fusion method based on relative confidence factors was applied to fuse the decisions of these 2 classifiers. Compared with other 4 classifiers, the new method can obtain higher classification accuracy and

robustness.

## Acknowledgements

The authors would like to acknowledge the support from the China Scholarship Council (No.201608330411), Zhejiang Natural Science Foundation under Grant No. LY17C100001, EU Seventh Framework Programme (FP7)-ICT under Grant No. 611391, Natural Science Foundation of China under Grant No. 51677171, 51575412, 51575338, and 51575407, and Research Project of State Key Lab of Digital Manufacturing Equipment & Technology of China under Grant No. DMETKF2017003.

## References:

- [1] Berni KC, Dibai-Filho AV, Pires PF, Rodrigues-Bigaton D. Accuracy of the surface electromyography RMS processing for the diagnosis of myogenous temporomandibular disorder. *J. Electromyogr Kinesiol.* 2015, 25(4):596-602.
- [2] Ferrari D, Kuriki HU, Silva CR, Alves N, Mícolis de Azevedo F. Diagnostic accuracy of the electromyography parameters associated with anterior knee pain in the diagnosis of patellofemoral pain syndrome. *Arch Phys Med Rehabil.* 2014, 95(8):1521-1526.

- [3] Fang Y, Hettiarachchi N, Zhou D, Liu H. Multi-modal sensing techniques for interfacing hand prostheses: a review. *IEEE Sens. J.* 2015, 15(11): 6065–6076.
- [4] Krasoulis A, Kyranou I, Erden MS, Nazarpour K, Vijayakumar S. Improved prosthetic hand control with concurrent use of myoelectric and inertial measurements. *J Neural Eng.* 2017, 14(7):71-84
- [5] Hargrove LJ, Englehart K, Hudgins B. A comparison of surface and intramuscular myoelectric signal classification. *IEEE Trans Biomed Eng.* 2007, 54(5):847–53.
- [6] Hyonyoung H, Sungho J. Simg. Pattern Classification Using Bayeeian Model. *35th Annual International Conference of the IEEE EMBS.* 2013, 7: 6647-6650.
- [7] Duan F, Dai L, Chang W, Chen Z, Zhu C, Li W. sEMG-based identification of hand motion commands using wavelet neural network combined with discrete wavelet transform. *IEEE Trans Ind Electron.* 2016, 63(3):1923–1934.
- [8] Englehart K, Hudgin B. A Robust, Real-time Control Scheme for Multifunction Myoelectric Control. *IEEE Trans. on Bio. Eng.* 2003, 50(7): 848-854.
- [9] Keogh E, Mueen A. “Curse of dimensionality,” in *Encyclopedia of Machine Learning.* Springer. 2010, 257–258.
- [10] Klawonn F, Schwecke E. On the axiomatic justification of Dempster’s rule combination. *International Journal of Intelligent Systems.* 1992 7(5):469–478
- [11] Sapsanis C, Georgoulas G, Tzes A. Emg based classification of basic hand movements based on time-frequency features. In *Control & Automation (MED), 21st Mediterranean Conference.* 2013: 716–722.
- [12] Wu, Huang NE. Ensemble empirical mode decomposition: A noiseassisted data analysis method. *Advances in Adaptive Data Analysis.* 2009, 1(1): 1–41.