Emotion Feature Selection from Physiological Signal Based on BPSO

Ruiqing Yang[1] **Guangyuan Liu**[2]

[1]School of Computer & Information Science, Southwest University, Chongqing 400715, P. R. China
[2]School of Electronic Information Engineering, Southwest University, Chongqing 400715, P. R. China

Abstract

In emotion recognition, many irrelevant and redundant features will affect recognition results, so feature selection is necessary. Aimed at emotion physiological signal feature selection, this paper proposed with improved discrete binary particle swarm optimization(BPSO) to increase the correct classification rate of emotion state. When recognizing four emotional states with nearest classifier by four physiological signals, the whole correct recognition rate is up to 85%. Experimental results demonstrate that the BPSO is an effective way to emotion physiological signals feature selection.

Keywords: Feature Selection, Binary Particle Swarm Optimization (BPSO), Physiological Signals, Emotion Recognition.

1. Introduction

Nowadays affective computing has become the hotspot of computer science. Through recording and analyzing physiological signals' features to recognize emotion has become an increasingly important field of research in affective computing and human computer interface [1]. The research of emotion recognition consists of facial expression, vocal, gesture, text, physiological signal recognition and so on. The research of emotion in physiological signal is the most difficult. Others are relatively easy, but can't detect the underlying emotional states. While analyzing physiological signals can detect the underlying emotion, it needs special equipments to record it [2].

In recent years, the research of emotion recognition is focused on facial expression, affective vocal, etc., while emotion recognition from physiological signals is beginning just now. It is necessary to select an optimal subset of features to recognize emotion and to minimize classification rate. Now, there are only some traditional methods used in feature selection. Picard et al. with SFFS [3], Fisher and SFFS-Fisher feature selection methods to classified physiological patterns for a set of eight emotions (including neutral), the correct recognition accuracy over 80% [4] [5]. Wagner et al. tested SFS, SBS, ANOVA feature reduction methods and a hybrid method of SFS and Fisher (SFS/Fisher), the recognition rate was raised from less than 80% up to 92.05% [6].

Feature selection is a problem of combinatorial optimization [7]. Its goal is to select a subset of d features from the given set of D ($D>d$) measurements, and without significantly degrading the performance of the recognition system. So we can use the method of solving optimal problem to resolve feature selection. Many algorithms have been proposed for feature selection, from simple algorithms like branch and bound algorithm, SFS, SBS and Plus-*l*-Minus-*r*, to more complex intelligent and heuristic algorithms such as simulated annealing, neural net pruning, Tabu search, genetic algorithm ,particle swarm optimization and so on. At present, we haven't find that computational intelligent methods were applied to resolve feature selection problem of emotion physiological signal.

The particle swarm optimization (PSO) was devised by Eberhart and Kennedy in 1995 [8]. It is an optimization algorithm inspired by social behavior of flocks of birds when they are searching for food, is one of the intelligent optimal algorithms. The algorithm is famous for its small population and parameters, easy to understand and realize. Firstly, the algorithm was applied to solve continuous optimal problem, and then Eberhart and Kennedy presented discrete binary particle swarm optimization (BPSO) [9], it was used to resolve the combinatorial optimization problem. Therefore, this paper proposed an improved discrete binary particle swarm optimization (BPSO) for emotion recognition tasks. The new algorithm was applied to select useful physiological signal features of emotion to increase the correct recognition rate. The results of the experiment show that BPSO algorithm is an effective method for physiological signal feature selection of emotion states.

2. Feature selection based on BPSO

2.1. Basic BPSO algorithm

The particle swarm optimization (PSO) algorithm is an optimization algorithm inspired by social behavior of flocks of birds when they are searching for food [8]. Particle swarm optimization shares many features with Genetic Algorithms and Evolutionary Programs. It uses a population of individuals, called particles, with an initial population distributed randomly over the search space. It searches for the optimal value of a function by updating the population through a number of generations. Each new population is generated from the old population with a set of simple rules that have stochastic elements. The basic PSO model consists of a swarm of m particles moving about in a D dimensional real value search space. Each particle, which is a potential global optimum of the function *f(x)* over a given domain *D*, is looked as a point in the *D* dimensional space and represented as $X_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$. Here subscript i means ith particle. Fitness value of all particles is evaluated by the fitness function to be optimized. And according to that value, the particle is updated to move towards the better area by the corresponding operators till the best point is found. In the iterative process, the position of each particle with its best fitness value, that is its local best, is remembered and denoted as $P_i = (p_{i1}, p_{i2}, \cdots, p_{iD})$. At the same time, the globe best, which is the position with the best fitness value of all particles, is also recorded as $P_g = (p_{g1}, p_{g2}, \cdots, p_{gD})$. Velocity, the rate of the position change for the ith particle is represented as $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$. At each times step, the velocity of all particles is adjusted as a sum of its local best value, globe best value and its present velocity, multiplied by the three constants *w*, *c1*, *c2* respectively, shown in Eq. (1). The position of each particle is also modified by adding its velocity to the current position [10], see Eq. (2).

$$V_{id} = w \times V_{id} + c1 \times rand() \times (p_{id} - x_{id}) + c2 \times rand() \times (p_{gd} - x_{id}) \tag{1}$$

$$x_{id} = x_{id} + V_{id} \tag{2}$$

where *c1* and *c2*, are two positive constant named as learning factors, normally set as *c1=c2=2*. To protect the particle fly out the search space, velocity V_{id} was constrained to the interval $V_{id} \in [-V\max, V\max]$, *Vmax* is an constant, set by user. w is an inertia parameter, was shown empirically to improve the performance of PSO when the value of *w* is reduced linearly from 0.9 to 0.4 over the number of generations of the search. Each particle updates its velocity according to local position and global position continuously, and then “fly” to the position of the best solution.

2.2. Binary particle swarm optimization algorithm

The binary PSO algorithm where the particles take the values of binary vectors of length *n* and the velocity defined the probability of bit x_{id} to take the value 1 reserved the updating formula of the velocity (see Eq.(1)) while velocity was constrained to the interval [0.0, 1.0] by a limiting transformation function $S(v)$.

$$S(x) = \frac{1}{1+e^{-x}} \tag{3}$$

Then the particle change its bit value by Eq. (4)

$$\begin{cases} x_{id} = 1, r < S(v_{id}) \\ x_{id} = 0, otherwise \end{cases} \tag{4}$$

where *r* is a random value in the range of (0,1).

2.3. Fitness function

In order to measure the performance of each particle, a pre- defined fitness function is applied to evaluate the fitness of each particle. In this paper, definition of the fitness function considered two elements: correct classification rate and the number of features was picked to up to the classification rate. If correct classification rates are same, little number of combinatorial features will be better. So the fitness function [11] definition as:

$$fitness = 10^5 \times (1 - Accuracy) + k \times Ones \tag{5}$$

Where *Accuracy* is the correct rate of classification, *Ones* is the number of picked feature, *k* is the balance parameter of correct rate and features number. The higher of *k* , the more important of features number. In this paper, the value of *k* is 0.5.

2.4. Description of binary particle swarm optimization

In BPSO algorithm, each particle can be represented as a feature vector, $X = (x_1, x_2, \cdots, x_D)$, $X_i \in \{0,1\}$, D is the number of original feature. If the *i*th bit value of *X* is 1, then the corresponding feature will be picked, otherwise, the feature will be discarded. Velocity represents the rate of the feature was picked.

The whole step of improved BPSO algorithm for feature selection are following:

- Step 1. set particle number of the swarm, number of iteration, threshold of error and iteration t;
- Step 2. initial velocity and position of the swarm, and calculate the fitness value of each particle according to Eq. (5), set it as the initial local optima and set the best value as the global optima; the initial value of V_i was set to 0;
- Step 3. evaluate the function of the swarm and update the local optima and the global optima;
- Step 4. update velocity and position of each particle;

- Step 5. calculate the hamming distance *d* between each particle and the global optima, if *d*<=*t*, then regenerate the particle;
- Step 6. stop the iterative if the terminal rule is satisfied, or go to Step 3.
- Step 7. output the global optima and corresponding particle, this particle is the best combinatorial features.

3. The results of feature selection based on BPSO

This section presents some of the results we obtained by applying the methods described above. Experiments were conducted with the dataset from the Multimedia Concepts and Applications of university of Augburg [12]. The dataset contains physiological data of four sensors: SC, EMG, RSP and ECG. 25 data sets of a single subject consecutively expressing four emotional states: Joy, Anger, Sadness, Pleasure, were collected. Calculate various statistical values, such as mean value, median value, minimum, maximum, standard deviation, approximation of first derivation, approximation of second derivation, pulse signal, amplitude signal, number of maxima divided by the total number of signal values and mean of the frequency spectrum in a given range, etc., as original features. 21 features were extracted from EMG and SC signals, 67 features were extracted from RSP signals, and 84 features were extracted from ECG signals. Overall 193 features were extracted from four physiological signals. These features were extracted from four different emotions corresponding physiological signals, so we should classify these data to four classes, each class has 25 samples, overall 100 samples.

In the experiment, the parameters were set as following: the swarm is 50 when use 193 features to test, and the learning factor is $c1=c2=2$, the value of w is reduced linearly from 0.9 to 0.4 over the number of generations of the search, the max velocity is set to 6. The feature set was selected by BPSO, nearest neighbor is applied to classify the emotion classes. The whole correct recognition rate is up to 85%. When we use 193 features to recognize four emotions, the correct rate is only 43%. We also tested with one physiological signal to recognize four emotions.

In the following tables, RR. represents correct recognition rate. The first column of the tables list the physiological signal used to classify four emotions. The second column to the fourth column list the results of max correct recognition rate, min correct recognition rate and average recognition rate. The average recognition rate is the average results of running 50 times.

Physiological signal	Max RR.	Min RR.	Average RR.
SC	47%	40%	45.88%
EMG	68%	66%	67.72%
RSP	73%	67%	69.86%
ECG	67%	57%	62.94%
SC,EMG,SP,ECG	85%	62%	70.10%

Table 1: Total correct recognition rate of four emotions.

Physiological signal	Max RR.	Min RR.	Average RR.
SC	47%	39%	45.56%
EMG	68%	66%	67.64%
RSP	73%	67%	69.76%
ECG	66%	57%	62.76%
SC,EMG, SP,ECG	76%	62%	66.00%

Table 2: Total correct recognition rate of four emotions (basic BPSO).

Table 1 is the results of BPSO based on Hamming distance. From the last column we can conclude that the recognition results of EMG and RSP are better when using single physiological signal to recognition four emotions, SC is the worst. The results in table 1 are more reasonable than table 2. But the results of all 193 features are still not satisfied. So we introduced mutation in basic BPSO. If the global optima is not change after several iteration, one dimension of each particle will be changed. The following table 3 shows the results.

Physiological signal	Max RR.	Min RR.	Average RR.
SC	47%	42%	45.94%
EMG	68%	65%	67.74%
RSP	73%	67%	70.20%
ECG	67%	60%	63.22%
SC,EMG, SP,ECG	86%	65%	79.10%

Table 3: The results of one dimension changed of each particle.

Number of features	4 Emotions	Arousal	Valence
193	86%	100%	86%

Table 4: Recognition results with sets of emotions.

From table 3 we can see that not only the four physiological signal results are improved but also the single signal. The average recognition rate of four physiological signals is remarkable, it is raised from 66% to 79.10%. We also tested discrimination between sets of emotion. We divided the emotions into groups of negative (anger/sadness) and positive (joy/pleasure) valence and into groups of high arousal (joy/anger) and low arousal (sadness/pleasure). It turned out that it was much easier to separate emotions along the arousal axis than along the valence axis. This result is accord with the result of Picard, Wagner and Andreas et al.

[13]. High and low arousal was recognized in about 100% of the cases. Negative and positive in only about 86% of the cases. Table 4 provides an overview of the results with the single sets of emotions.

4. Conclusions

When the feature selection from physiological signal was regarded as a combinatorial optimization problem, this paper adopted BPSO as feature selection method. The comparison of experimental results in part 3 show that BPSO with mutation is better. The recognition accuracy shows that features retrieved are acceptable. We can predict that BPSO with mutation will be more efficiency as it is used to feature selection problem. Future work will focus on collecting more physiological data and exploring the relationship between number of dimension change and the correct recognition rate.

Acknowledgement

This work is partially supported by Chongqing Science & Technology Commission Foundation of China (Grant NO. CSTC 2006BB2028). And we would like to thank Johannes Wagner at the Lab of Multimedia Concepts and Applications of university of Augburg for his help and permitting us to use their dataset.
Corresponding author: Guangyuan Liu

References

[1] L. Li and J.H. Chen, *Emotion Recognition Using Physiological Signals*, Springer-verlag, 2006.

[2] R.W. Picard. *Affective Computing*, MIT press, Cambridge, MA, 1997.

[3] P. Pudil, J. Novovicova and J. Kittler, Floating search methods in feature selection. *Pattern Reconition Letters*, 15:1119-1125, 1994.

[4] J.A. Healey, Wearable and Automotive Systems for Affect Recognition from Physiology, PhD thesis, MIT, Cambridge , 2000.

[5] R.W. Picard, E. Vyzas and J. Healey, Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 1175-1191, 2001.

[6] J. Wagner, J. Kim and E. André, From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification, *Prof. of the IEEE International Conference on Multimedia & Expo (ICME 2005)*, pp. 940-943, 2005.

[7] Z. Q. Bian, X. G. zhang, *Pattern Recognition*. Tsinghua University Press, Beijing, 2000.

[8] J. Kennedy and R.C. Eberhart, Particle swarm optimization, *Proceedings of 1995 IEEE International Conference on Neural Networks*, pp. 1942-1948 ,1995.

[9] J. Kennedy and R.C. Eberhart, A discrete binary version of the particle swarm algorithm, *Proc. of the 1997 Conf. on Systems, Man, and Cybernetics*, pp. 4104-4108, 1997.

[10] Y. Shi and R.C. Eberhart, A modified swarm optimizer, *Prof. of the IEEE International Conference of Evolutionary Computation*, pp. 69-73, 1998.

[11] L.Y. Qiao, X.Y. Peng and Y. Peng, BPSO-SVM Wrapper for Feature Subset Selection, *Acta Electronica Sinica*, 34:496-498, 2006.

[12] http://mm-werkstatt.informatik.uni-augsburg.de.

[13] A. Haag, S. Goronzy, P. Schaich and J. Williams, Emotion recognition using bio-sensors: First steps towards an automatic system. *LNCS*, 3068:36-48, 2004.