EEG Feature Extraction Based on Rough Set

Zhendong Mu

Institute of Information Technology, Jiangxi University of Technology, Nanchang, Jiangxi 330029, China

Keywords: Feature extraction; Rough set; Electroencephalograph (EEG) signal,

Abstract. Rough set theory has been successfully utilized in a variety of information mining fields. Since the EEG found that many researchers to study the biological signals in electrical brain research, the EEG feature extraction algorithm is a major research content. In this paper, rough set theory is taken as a research tool, EEG feature extraction, through 30 subjects collected EEG analysis, collect the EEG applied to the identification, the results show that rough set of feature extraction method can achieve identification.

Introduction

Up to now, some achievements have been obtained by using the rough set theory to analyze temporal data. Ostroff S. put forward a real-time temporal logic frame, which indicates temporal series using event variables. Berndt J. D. detected the mode of time series using dynamic programming. Bazan Jan G. et al. made analysis of market data using rough set and dynamic reduction and succeeded. Golan R. made analysis of Canada stock data and put forward the thinking for converting time series to a traditional information system. Anders T. B. brought forward a time series information system and the concept to implement the system, and formalized the thinking for converting time series to an information system. When rough set was used to process temporal data, previous study mainly paid attention to the time series features of mined temporal data, i.e. a strict time sequence is kept between objects. Time series information is divided into two types: the time series without real-time restriction and the time series with real-time restriction. The former can be considered as a time string arranged according to time, and the time interval (sampling rate) between events is constant; while the latter must consider the time intervals between events, which differs from the former one greatly.

With the development of biological information technology, complex human body information can be known by collecting various kinds of signals from human body. The knowledge of such information helps to treat many complex diseases and realize the development of many human body imitating techniques. Many of such signals are the ones of time series, such as electrocardio, electromyogram and EEG signals, etc. Complex time series include many physiological signals. The feature extraction of these signals is one of the main subjects in biological information study. Such feature extraction is often done by introducing a signal analysis method to biological identification technology, such as Fourier transformation, wavelet analysis, etc. However, physiological signals are special signals. Therefore, how to extract features from signals themselves is a difficulty baffling many study team.

In this paper, rough set is used for feature extraction, and features used in the identification study showed that characteristic can be used to represent individual characteristics, identification of the results of more than 90%.

Rough set theory

In 1982, Professor Z. Pawlak et al. in Poland Warsaw University of Technology came up with the concept of rough set, and up to now a perfect rough set theory system has been formed. In this system, the information in the real world can be generally indicated by a piece of information. Each line in the information is called an instance (entity, object), whose nature is reflected by assigning some variables. The main component of an information table knowledge expression system is a set of studied objects. The knowledge of such objects is described by assigning their attributes (features) and attribute values (characteristic values). Generally, an information table knowledge system can be expressed as S = (U, R, V, f), where

U= the set of objects;

R= $C \cup D$; Subsets C and D are called a conditional attribute set and a resulting attribute set respectively;

V= $\bigcup_{r \in R} V_i$, the set of attribute values; Vi refers to the attribute value range of Attribute $r \in R$, i.e. the range of r.

f- f: U x R --> V is an information function, which assigns the attribute value of each object x in U.

For each attribute subset B⊆ R, we define an indiscernibility binary relation (indiscernibility relation) IND (B), i.e.

IND (B) $= \left\{ (x, y) \mid (x, y) \in U^2, \forall b \in B(b(x) = b(y)) \right\}$.

Definition 1: Assume $X \subseteq U$. When X can be described exactly by the attribute subset, B (i.e. the merge of indiscernibility set on U determined according to attribute subset B),   X can be defined by B. Otherwise X cannot be defined by B. The definable set of B is also called the exact set of B, and the indefinable set the inexact set or rough set of B (rough set for short).

Definition 2: For each definition X (subset of instances) and indiscernibility relation B, both the maximum definable set included in X and the minimum definable set including X can be determined according to B. The former is called the lower approximation set of X (indicated by $B_-(x)$), and the latter the upper approximation set of X (indicated by $B^-(x)$).

Definition 3: Given knowledge system S= (U, R, V, f). For each subset $X \subseteq U$ and indiscernibility relation B, the upper and lower approximation sets of x can be expressed by the basic definition of B respectively as follows:

$B_-(x) = \bigcup \left\{ Y_i \mid (Y_i \in UIND(B) \wedge Y_i \subseteq X \right\}$;

$$B_{-}(x) = \bigcup \{ Y_i \mid (Y_i \in UIND(B) \land Y_i \cap X \neq \phi \};$$

Where, U|IND（B）$= \{ X \mid (X \subseteq U \land \forall x \forall y \forall b(b(x) = b(y))) \}$, is the division of U by the indiscernibility relation B, that is to say, the basic set B of discourse domain U.

Results

The EEG data used in this paper is the EEG acquisition of college students of the BCI Laboratory of Jiangxi Institute of Technology. EEG acquisition is the use of 40-Neuroscan amplifier, which are obtained by scan 4.3 software, the reference electroded by way of the right mastoid as reference electrode, using 1000Hz sampling rate, band acquisition using 200Hz low-pass, high-pass 0.05Hz and 50Hz notch. Each experiment are collected continuously for 5 minutes, each subject every 4 experiment.

30 subjects, each subject to select 200 EEG, a total of 6000 EEG, EEG mix these 30 subjects, then using the proposed method of identification, for which 6000 Brain electric signal identification, the results are shown in Fig. 1
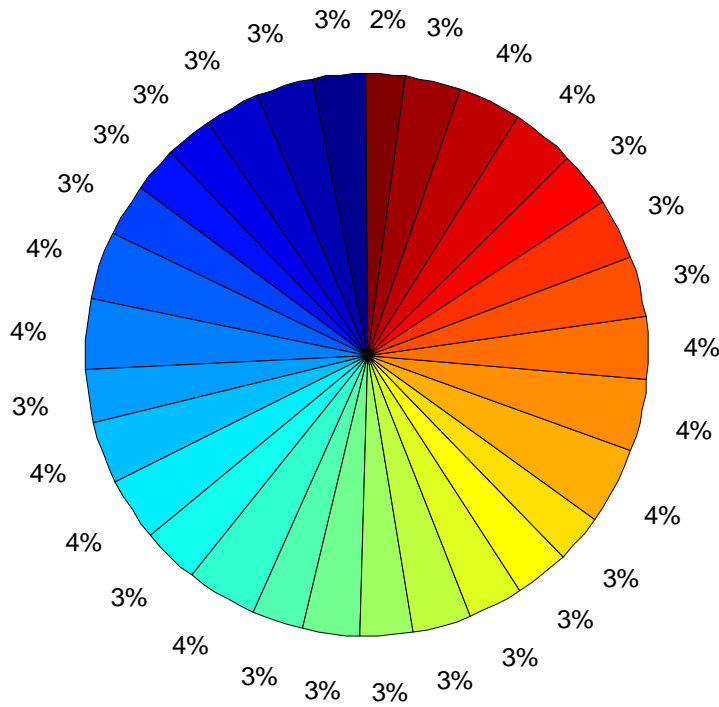

Fig. 1   6000 EEG and subject probability of matching

Fig. 1 shows the 6000 EEG matching relationship, the data shows that this matching relationship proportion of total brain points is between 2-4%, indicating that the number of brain signals for each subject identified in 2000- between 4000, compared with the 3000 EEG each subject, each subject is to identify the proportion of 70% or more. Description accurately extracts the characteristics of the subject.

6000 EEG is incorrectly identified as the number of EEG other people is 436, accounting for 7.2% of all EEG, which is called misidentification, it refers to not judge their own EEG signal.

Acknowledgements

References

[1]Marcel S, Millán J del R. Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation[J]. IEEE Trans Pattern Anal Mach Intell, 2007, 29(4): 743-752.

[2] Mukherjee A, Dey G, Dey M, et al. Web-Based Intelligent EEG Signal Authentication and Tamper Detection System for Secure Telemonitoring[M]//Brain-Computer Interfaces. Springer International Publishing, 2015: 295-312.

[3] Pham T, Ma W, Tran D, et al. Multi-factor EEG-based user authentication[C]//Neural Networks (IJCNN), 2014 International Joint Conference on. IEEE, 2014: 4029-4034.

[4] Al-Hudhud G, Alzamel M A, Alattas E, et al. Using brain signals patterns for biometric identity verification systems[J]. Computers in Human Behavior, 2014, 31: 224-229.

[5] La Rocca D, Campisi P, Scarano G. Stable EEG Features for Biometric Recognition in Resting State Conditions[M]//Biomedical Engineering Systems and Technologies. Springer Berlin Heidelberg, 2014: 313-330.

[6]Peng Xiaohu, Wang Guofeng, Wei Jinghan. Face recognition specificity ERP study brain mechanisms [J] Aerospace and Biomedical Engineering, 2005, 17 (6): 438-443.

[7]Zhang Yan, hole Fanchang, Guo YingImplicit or explicit memory: study of female faces attractive memory preferences ERP [J] ACTA: Natural Sciences, 2012, 48 (1): 160-168.

[8]Keenan J P, McCutcheon B, Freund S, et al. Left hand advantage in a self-face recognition task[J]. Neuropsychologia, 1999, 37(12): 1421-1425.

[9]Palaniappan R, Danilo P. Biometrics from Brain Electrical Activity: A Machine Learning Approach[J]. IEEE Trans Pattern Anal Mach Intell, 2007, 29(4): 738-742.

[10]Gupta C N, Palaniappan R, Swaminathan S. Novel analysis technique for a brain biometric system[J]. Int J Med Eng Inf, 2008, 1(2): 266–273.