Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 52 (2018) 882–890

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Research on feature importance evaluation of wireless signal recognition based on decision tree algorithm in cognitive computing

Lin Li [a], Juzhen Wang [b,*]

[a] State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System (CEMEE), Luoyang, Henan 471003, China
[b] Zhongxing Telecommunication Equipment Corporation (ZTE Corporation), 55 Keji Road South, Nanshan District, Shenzhen 518057, China

Received 29 July 2018; received in revised form 31 August 2018; accepted 10 September 2018
Available online 20 September 2018

Abstract

Cognitive computing is an important method in the field of wireless signal processing, analysis and recognition. How to select features to complete the cognitive computing quickly and effectively is an important role in real application. In this paper, three kinds of features are extracted from six communication signals: power spectrum entropy, singular spectrum entropy and wavelet energy entropy. And the importance of the features is evaluated. Box-diagram and recognition rate are used for the evaluation of single feature. The visual boundaries of feature classification are used to evaluate two features. Meanwhile, the confusion matrix and the visualization model of decision tree are given for more detailed evaluation. The evaluation results show that the combination of power spectrum entropy and singular spectrum entropy can get the best recognition performance.

Keywords: Cognitive computing; Modulation recognition; Feature evaluation

1. Introduction

Nowadays, the rapid development of wireless communication technology increases the complexity of electromagnetic environment (Wu, Li, & Lin, 2017; Sun et al., 2018). Modulation recognition as an intermediate part between signal detection and signal demodulation has been paid more attention. Various recognition algorithms are gradually proposed (Sastry and Shekar, 2016; Walenczykowska and Kawalec, 2016). The earliest automatic modulation recognition (Weaver et al., 1969) was proposed in 1969 by C. Weaver et al. Generally speaking, the methods of automatic modulation recognition have two categories: one is the maximum likelihood method based on hypothesis testing; and the other is the pattern recognition method based on feature extraction. An automatic modulation recognition system based on feature extraction usually consists of three parts: data preprocessing, feature extraction and classifier designing.

Over the past several decades, numerous studies on feature parameters had been conducted, such as instantaneous frequency (Ma, Wang, Lin, & Jin, 2018) and instantaneous phase (Subbarao, Noorbasha, Thati, & Professor, 2015), feature extraction based on signal transform domain (Thakur, Madan, & Madan, 2015), fractal feature (Shi, Dou, Lin, & Li, 2018; Wang et al., 2017) and deep features based on deep learning (Zhou et al., 2018), etc. With the development of information theory, information entropy was gradually used as signal feature Li and Guo, 2015.

* Corresponding author.
E-mail address: wjzheu@163.com (J. Wang).

https://doi.org/10.1016/j.cogsys.2018.09.007
1389-0417/

Entropy could be used to calculate the information amount. The larger amount of information entropy, the greater of the signal uncertainty. Great progresses have been made in the application of entropy features. N. Arunkumar picked the entropy features to distinguish focal EEG and non-focal EEG (Liu, Guan, & Lin, 2017). Jingchao Li applied Shannon Entropy, Singular Spectrum Entropy, and Singular Spectrum Exponential Entropy as recognition features for the study of radar emitter identification in low SNR environment (Li and Ying, 2014).

As for the classifier designing, a lot of researches had been finished, including Artificial Neural Network(ANN) (Wang and Guo, 2017), K-Nearest Neighbor classifier (KNN) (Lin, Zhu, Zheng, Dou, & Zhou, 2017), Grey Relationship Classifier (Li, 2015; Lin, Li, Yin, & Dou, 2018), Support Vector Machine (Ali and Yangyu, 2017), and semi-supervised learning with GANs (Tu, Lin, Wang, & Kim, 2018). What's more, the decision tree algorithm was another choice for classifier. Decision Tree was a fast and efficient algorithm for classification and prediction in data mining (Batra and Agrawal, 2018). It converted the complicated decision-making processes into rules or judgments, which were easy to understand.

The classic decision tree algorithms mainly included ID3, C4.5 and CART. Both of ID3 and C4.5 were proposed by J. Ross Quinlan (Quinlan, 1983). In the literature (Guo et al., 2017), ID3 and C4.5 were proposed to assemble a feature detection model for detecting the illegal using of mobile devices. Comparative analysis of decision tree algorithms: ID3, C4. 5 and random forest were taken in paper (Sathyadevan and Nair, 2015). Recogniton and Regression tree (CART) (Nivedha and Sairam, 2015) was another decision tree algorithm, which was proposed by L. Breiman et al. In the CART classification, the Gini Index was used to select the best data-partitioning feature (Zheng, Saxena, Mishra, & Sangaiah, 2018). As is well known, the Gini Index was similar to the information entropy (Shi et al., 2018). There were many applications of CART algorithm, Leszek Rutkowski applied the CART algorithm to classify the data stream. Decision tree induction was used to analysis indoor climate in paper (Geronazzo, Brager, & Manu, 2018).

As an important branch of machine learning, cognitive computing has become an important method for cognitive radio (Lin, Wang, Wang, & Dou, 2016; Yang, Ping, Sun, & Aghvami, 2017), especially in signal recognition (Tang, Tu, Zhang, & Lin, 2018; Wang, Guo, Dou, & Lin, 2018). Cognitive computing requires computers to quickly perform signal recognition and learning ability closer to the human brain (Jia, Gu, Guo, Xiang, & Zhang, 2016). Therefore, an effective feature importance evaluation method needs to be found. In this paper, how to select the proper feature has been studied. Three kinds of information entropy have been proposed as signal feature, including power spectrum entropy, singular spectrum entropy and wavelet energy entropy. Box-diagram and recognition rate are proposed for the evaluation of signal feature. The visual boundaries of feature classification are proposed for the evaluation of two features. Finally, the best feature can be selected for the best recognition performance.

2. Evaluation method

In this paper, Power Spectrum Entropy(PE), Singular Spectrum Entropy(SE) and Wavelet Energy Entropy(WE) are evaluated in cognitive computing. Six different kinds of signals are used for example, including 2ASK, 4ASK, 2FSK, 4FSK, BPSK and QPSK. The work flow is shown in Fig. 1.

Firstly, entropy features are extracted from six different signals. In order to ensure the accuracy of the evaluation results, all feature combinations are discussed.

Features are evaluated in their own combination, and the best one is selected respectively. For single features, box-diagrams are used to evaluate the feature. And recognition rate is used to verify the analysis. For two features, the visual boundaries of feature classification are proposed to evaluate the feature. Confusion matrix is used to verify the analysis. At last, the best three feature combinations are evaluated by the recognition rate based on decision tree classifier. According to the evaluation result, the best feature combination can be determined.

3. Basic theory

3.1. Entropy features

Information entropy can be used to measure the uncertainty and complexity of modulation signal. Therefore,

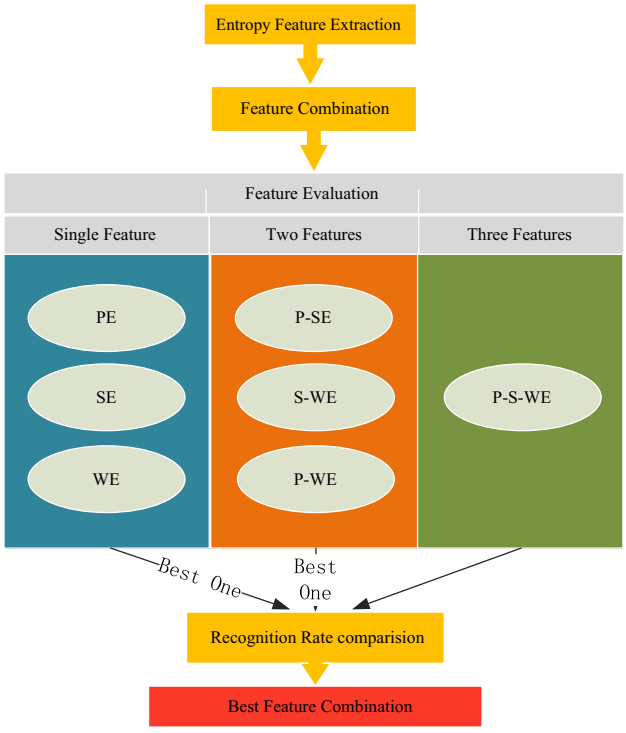

Fig. 1. Method diagram.