An efficient optimized independent component analysis method based on genetic algorithm

Liangmin Li1, Guangrui Wen2, Jingyan Ren3, Xiaoni Dong4, Lin Liang5
1Key Laboratory of Automotive Transportation Safety Enhancement Technology of the Ministry of Communication, School of Automobile, Chang'an University, Xi'an 710064, China
2, 4, 5School of Mechanical Engineering, Xi'an Jiao Tong University, Xi'an 710049, China
3Grid Electric Power Science & Technology Co., Ltd., Xi'an 710075, China
2, 5Corresponding authors
E-mail: 1leelm@chd.edu.cn, 2grwen@mail.xjtu.edu.cn, 3ren.j.y@163.com, 4birdy_dong@163.com, 5lianglin@mail.xjtu.edu.cn
(Received 30 May 2013; accepted 5 November 2013)

Abstract. Three simulation experiments are designed to evaluate and compare the performance of three common independent component analysis implementation algorithms – FastICA, JADE, and extended-Infomax. Experiment results show that the above three algorithms can’t separate the mixtures of super-Gaussian and sub-Gaussian precisely, and FastICA fails in recovering weak source signals from mixed signals. In this case an independent component analysis algorithm, which applies genetic algorithm to minimize the difference between joint probability and product of marginal probabilities of separated signals, is proposed. The computation procedure, especially the fitness evaluation when signals are in discrete form, is discussed in detail. The validity of the proposed algorithm is proved by simulation tests. Moreover, the results indicate that the proposed algorithm outperforms the above three common algorithms significantly. Finally the proposed algorithm is applied to separate the mixture of rolling bearing sound signal and electromotor signal, and the results are satisfied.

Keywords: independent component analysis, FastICA, JADE, extended-Infomax, genetic algorithm, rolling bearing.

1. Introduction

Blind source separation (BSS) [1] is one of the research focuses of signal processing area, its purpose is to recover source signals from mixed signals, without any prior knowledge (or with very little information) about the source signals. There are two main methods for BSS: principal component analysis (PCA) [2] and independent component analysis (ICA) [3]. PCA is based on the second-order statistics. Its main purpose is to eliminate the correlation between signals, so it is mainly used to compress the dimension of data. ICA considers the higher-order statistics of data, output signals of ICA are mutually independent. Owing to this property, ICA has been widely used in many areas such as audio processing [4], image processing [5] and vibration signal processing [6].

Algorithms for ICA include FastICA [7], JADE [8], Infomax [9] and many others. FastICA is an iterative algorithm maximizing non-Gaussianity as a measure of statistical independence. As FastICA algorithm has the advantage of quick convergence and free from setting iteration step length, it has become the most popular algorithm for ICA. But the update formula of FastICA includes activation functions that depend on the distribution of source signals, which is unknown for practical problems. Infomax algorithm proposed by Bell and Sejnowski is another widely used ICA implementation algorithm. This algorithm has the same drawback as FastICA – activation function selection problem. This disadvantage led to the appearance of an extension version – extended-Infomax algorithm [10], which is able blindly to separate mixed signals with sub- and super-Gaussian source distributions, and has been proved to outperform traditional Infomax algorithm. In 1999 another widely used ICA algorithm, JADE algorithm, appeared. It’s a fourth order technique, involving the diagonalization of cumulant matrices. Lots of literatures proved the success of these algorithms in solving BSS problems. Note that these are mostly normal BSS