



# Efficient ant colony optimization for image feature selection



Bolun Chen<sup>a</sup>, Ling Chen<sup>b,c,\*</sup>, Yixin Chen<sup>d</sup>

<sup>a</sup> College of Information Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China

<sup>b</sup> Department of Computer Science, Yangzhou University, Yangzhou, China

<sup>c</sup> State Key Lab of Novel Software Tech, Nanjing University, Nanjing, China

<sup>d</sup> Department of Computer Science and Engineering, Washington University in St Louis, USA

## ARTICLE INFO

### Article history:

Received 2 January 2012

Received in revised form

3 October 2012

Accepted 29 October 2012

Available online 6 November 2012

### Keywords:

Ant colony optimization

Dimensionality reduction

Feature selection

Image classification

## ABSTRACT

Feature selection (FS) is an important task which can significantly affect the performance of image classification and recognition. In this paper, we present a feature selection algorithm based on ant colony optimization (ACO). For  $n$  features, existing ACO-based feature selection methods need to traverse a complete graph with  $O(n^2)$  edges. However, we propose a novel algorithm in which the artificial ants traverse on a directed graph with only  $O(2n)$  arcs. The algorithm incorporates the classification performance and feature set size into the heuristic guidance, and selects a feature set with small size and high classification accuracy. We perform extensive experiments on two large image databases and 15 non-image datasets to show that our proposed algorithm can obtain higher processing speed as well as better classification accuracy using a smaller feature set than other existing methods.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Reduction of pattern dimensionality via feature extraction is one of the most important tasks for pattern recognition and classification. Feature selection has considerable importance in areas such as bioinformatics [1–3], signal processing [4–8], image processing [9–11], text categorization [12], data mining [13], pattern recognition [14–18], medical diagnosis [19] and remote sensor image recognition [20]. The goal of feature selection is to choose a subset of available features by eliminating unnecessary features. To extract as much information as possible from a given image set while using the smallest number of features, we should eliminate the features with little or no predictive information, and ignore the redundant features that are strongly correlated [21,22]. As a

result, a large amount of computation time can be saved. The selected subset of features used to represent such classification function influences several aspects of image classification, including the time required to learn a classification function, the accuracy of the learned classification algorithm, the time-space cost associated with the features, and the number of samples required for training.

Ant colony optimization (ACO) is an evolution simulation algorithm proposed by Dorigo et al. [43–45]. Inspired by the behaviors of the real ant colony, they recognized the similarities between the ants' food-hunting activities and traveling salesman problem (TSP), and successfully solved the TSPs using the same principle that the ants have used to find the shortest route to food source via communication and cooperation. ACO has been successfully used for system fault detecting, job-shop scheduling, frequency assignment, network load balancing, graph coloring, robotics and other combinatorial optimization problems [45].

ACO is a multi-agent system where communications among artificial ants result in a positive feedback

\* Corresponding author at: Department of Computer Science, Yangzhou University, Yangzhou, China. Tel.: +86 514 87870 026; fax: +86 514 878 87937.

E-mail addresses: [Chenbolun1986@gmail.com](mailto:Chenbolun1986@gmail.com) (B. Chen), [lchen@yzu.edu.cn](mailto:lchen@yzu.edu.cn) (L. Chen), [chen@cse.wustl.edu](mailto:chen@cse.wustl.edu) (Y. Chen).

behavior to guide the ant colony to converge to the optimal solution. Pheromone information, which simulates the chemical substance the real ants lay on the route they passed, is assigned to the edges of the graph. Artificial ants are used in ACO to travel in the graph to search for optimal paths according to the pheromone and problem-specific local heuristics information. The pheromone on each edge is evaporated at a certain rate at each iteration. It is also updated according to the quality of the paths containing this edge. Artificial ants are usually associated with a list that records their previous actions, and they may apply some additional operations such as local search, crossover and mutation, to improve the quality of the results obtained. Compared with GA, ACO has some advantages such as allowing positive feedback, distributed computing, and constructive greedy heuristic search.

In recent years, some ACO based methods for feature selecting are reported. Basiri et al. [1] used ACO to select features for predicting post-synaptic activity in proteins. Nemati et al. [6] proposed an ACO based algorithm to reduce features size in automatic speaker verification. Aghdam et al. [12] proposed an ACO based feature selecting algorithm for text categorization. Kanan et al. [15] also presented an improved ACO algorithm for feature selection in face recognition. Their algorithm can select the optimal feature subset in terms of shortest feature length and the best performance of classifier. Vieira et al. [37] proposed an algorithm for feature selection based on two cooperative ant colonies, which minimizes two objectives: the number of features and the classification error. Two pheromone matrices and two different heuristics are used for these objectives. Subbotin [39] presented a modified ACO algorithm for feature selection. Akarsu et al. [46] proposed a hybrid modeling approach based on ACO for clustering and feature selection.

In some methods for feature selection, ACO is also combined with other methods to improve the quality of feature selection and classification. Huang [38] presented a hybrid ACO-based classifier model that combines ACO and support vector machines (SVM) to improve classification accuracy with a small and appropriate feature subset. To simultaneously optimize the feature subset and the SVM kernel parameters, the feature importance and the pheromones are used to determine the transition probability; the classification accuracy and the weight vector of the feature provided by the SVM classifier are both considered to update the pheromone. Based on ACO and rough sets, Chen et al. [47] proposed an approach to feature selection which adopts mutual information based feature significance as heuristic information. In their algorithm the artificial ants start their tour from the feature core, which changes the complete graph to a smaller one. Sivagaminathan et al. [48] presented a hybrid method based on ACO and artificial neural networks to address feature selection.

In most ACO-based feature selection methods, the nodes in the graph are used to represent the features. Ants traverse on the graph to look for a path containing part of the nodes which indicate the features selected.

For  $n$  features, most ACO-based feature selection methods use a complete graph with  $O(n^2)$  edges. This means that  $O(n^2)$  pheromone and heuristic information will be computed and stored.

In this paper, we present an ACO-based feature selection algorithm, ACOFS, to reduce the memory requirement and computation time. In this algorithm, the artificial ants traverse on a directed graph with only  $2n$  arcs. The algorithm uses classifier performance and feature set size to guide search, and optimizes the feature set in terms of its size and classifier performance. Experimental results on large image and non-image datasets show that the proposed algorithm has superior performance. Compared with other existing algorithms, our algorithm can obtain higher processing speed and classification accuracy with a smaller feature set.

The rest of the paper is organized as follows. Section 2 reviews related work on feature selection. Section 3 presents the proposed ACO based algorithm, ACOFS, for feature selection. Section 4 describes the implementation details of ACOFS. Section 5 shows and analyzes the experimental results obtained by ACOFS and compares the performance with other similar methods. Section 6 draws conclusions.

## 2. Related work on feature selection

Methods for feature selection can generally be classified into two main categories: model-free methods and model-based methods [22]. Model-free methods do not assume a particular model and only use statistical tests on the data to select features. In the model-based methods, models with different sets of features are compared, and the model that minimizes the output error is selected. In this case, the feature selection is actually a problem of global combinatorial optimization in machine learning, which reduces the number of features, removes irrelevant, noisy and redundant data, and results in acceptable classification accuracy.

Feature selection algorithms can also be divided into two categories according to the way they process and evaluate features: feature ranking methods and subset selection methods. Feature ranking methods rank the features by a metric and eliminate all features that do not achieve an adequate score. Subset selection methods search the set of possible features for the optimal subset. To find the optimal feature subset, the entire search space contains all the possible subsets of features. For  $n$  features, the size of the search space is  $2^n$ . The problem of finding the optimal feature subset is NP-hard. In the simplest way, the best subset can be found by evaluating all the possible subsets by exhaustive search [22] which needs to enumerate all the subsets of features. Although this method guarantees an optimal solution, finding the optimal subset is time consuming and impractical even for a moderate sized feature sets. For most practical problems, an optimal solution can only be guaranteed if a monotonic criterion for evaluating features is applied. But this assumption rarely holds in real-world applications [23]. Decision tree search methods, with proper branch conditions, can limit the search space to the best

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات