Signal Processing 109 (2015) 172–181

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Automatic image annotation using feature selection based on improving quantum particle swarm optimization

Cong Jin [a], Shu-Wei Jin [b]

[a] School of Computer Science, Central China Normal University, Wuhan 430079, PR China
[b] Département de Physique, École Normale Supérieure, 24, rue Lhomond, 75231 Paris Cedex 5, France

ARTICLE INFO

Article history:
Received 6 September 2014
Received in revised form
24 October 2014
Accepted 25 October 2014
Available online 20 November 2014

Keywords:
Automatic image annotation
Visual feature selection
Optimization algorithm
Improvement operation
Ensemble stratagem

ABSTRACT

Automatic image annotation (AIA) is a task of assigning one or more semantic concepts to a given image and a promising way to achieve more effective image retrieval and analysis. It is a typical classification problem. Due to the semantic gap between low-level visual features and high-level image semantic, the performances of many existing image annotation algorithms are not satisfactory. This paper presents a novel AIA scheme based on improved quantum particle swarm optimization (IQPSO) algorithm for visual features selection (VFS) and an ensemble stratagem based on boosting technique to improve performance of image annotation approach. To maintain the population diversity, the measure method of population diversity and improvement operation are proposed. To achieve better performance of AIA scheme, the measure of population diversity is as a control condition of VFS process. The classification result of an ensemble classifier is as the final annotation result rather than individual classifier. The experimental results confirm that the proposed AIA scheme is very effectiveness. When using proposed AIA scheme over three image datasets respectively, the annotation results are satisfactory.

1. Introduction

In the past decade, the content-based image retrieval (CBIR) has been largely explored. In CBIR, an image has no direct relevance with high-level semantic content, but only with low-level visual features such as textures, colors and shapes, etc. [1,2]. However, the users often prefer retrieving images using high-level semantic content. The gap between high-level semantic contents and low-level visual features is the major difficulty that hinders further development of CBIR [2,3]. Image annotation techniques are used to bridge the gap [3,4]. Because it is very difficult that a large image dataset is annotated manually, AIA is an important task.

AIA is a process of automatically labeling image with set predefined keywords to represent image semantic [3–6]. For classifying correlation between keywords and visual features, classification algorithm is adopted to automatically annotate unlabeled images. However, the contribution of each visual feature is not the same when computing the similarity measure and AIA scheme must solve a basic problem in which features are more appropriate than the others in order to express the concept of the current query, whose essence is the VFS problem.

Feature selection techniques have been widely used in various areas, e.g., natural language processing [7], intrusion detection [8], video processing [9], online group [10], bioinformatics [11] and so on. For AIA, VFS is expected to improve its performance [12,13], particularly in situations of large image dataset. Even if not to improve the accuracy of image annotation, reducing number of features still has many advantages, e.g. reducing the number of measurements required, shortening training and execution times, and improving classifier compactness, transparency, and interpretability etc. The existing

E-mail address: jincong@mail.ccnu.edu.cn (C. Jin).

http://dx.doi.org/10.1016/j.sigpro.2014.10.031

feature selection approaches [14–16] deal with feature sequentially one by one, which does not seem appropriate for large image databases.

Currently, a lot of image annotation schemes are based on image segmentation [17], which is an important method to extract the targets of the image, and these targets can well describe the content of the image. However, there are some disadvantages for image segmentation, such as segmentation results fragile and erroneous etc. In this paper, we process directly the image rather than the first segment image. The main contributions of this work are the following:

(1) In this work, the measure method of population diversity is proposed to maintain the population diversity and is also a control condition of VFS process. The proposed VFS based on QPSO is possible to simplify the representation of the image.
(2) To achieve better performance of AIA scheme, an improvement operation of QPSO is proposed. This operation can avoid premature convergence.
(3) An ensemble stratagem based on boosting technique is used for AIA scheme, i.e., an ensemble annotation result rather than individual annotation result is as the final annotation result.

The rest of this paper is organized as follows. Section 2 introduces relate work. Section 3 describes the standard QPSO and its improvement method; they will be used in feature selection study. Proposed VFS approach is described in Section 4. Section 5 introduces the image annotation scheme including the visual feature vector, image annotation approach and ensemble strategies of classifiers. Section 6 presents the experimental studies including the experimental image datasets, evaluation of annotation accuracy, parameters initialization, experimental results and comparison. Finally, conclusions are given in Section 7.

2. Related work

For AIA, one of the easiest and simplest is template matching approach [18–21]. The semantic template to be learned from a collection of sample images is used to bridge the gap between low-level visual features and high-level semantic concepts [18]. Currently, the most common image annotation approach has two types, i.e., classification-based and probabilistic modeling-based approaches.

In first type, image annotation can be viewed as a classification problem [18–21], which can be solved by using a classifier. For annotating an image without caption, first, represent image into a low-level features vector. Then, classify the image into a category. Finally, propagate the semantic of the corresponding category to the image. The unlabeled image may be automatically annotated.

In second type, probabilistic model [22,23] attempts to infer the joint probabilities between images and semantic concepts. Images given in the class can be regarded as instances of stochastic process that characterizes the class. Then, statistical models, such as Markov, Gaussian, and Bayes etc. are trained and images are classified based on probability computation.

An effective image annotation approach should be able to deal with a large number of images, allowing users to query for images of interest efficiently and effectively. Usually, an image is represented by different visual features, if the dimensions of the visual feature vector is too high, it is difficult to achieve the desired performance even for medium size databases. On the other hand, not all the visual features are useful to image annotation and the irrelevant and/or redundant visual features should be removed. How to remove these useless visual features is the main challenge for bridging the semantic gap. Currently, some image annotation approaches based on feature selection technique have been proposed [18,24–27]. In [24], Ferecatu et al. used linear principal component analysis (PCA) to reduce dimension or clean up noise. However, if PCA is used to reduce dimension, the new visual features will be linear combination of original features, and it will be hard to explain the meaning of results and determine important original visual features. In [25], Lin et al. described the application of the information gain and a pixel density filter (PDfilter) for selecting the most suitable visual features, however, authors only used color feature, did not use other visual features such as texture and shape etc. This indicates their technology has some limitations. In [26], Wang et al. developed a mechanism to automatically assign different weights to different visual features in order to identify the most important visual features and discard irrelevant visual features. Although this method had good performance when used in image annotation, it is not suitable to evaluate visual feature weights one by one. Optimization model had also been used to VFS for the image annotation. In [18], an image annotation technique based on VFS was proposed, and which used a bi-coded chromosome genetic algorithm simultaneously to select optimal class-pair feature subsets and corresponding optimal weight subsets. In [27], Li et al. proposed dynamic Adaboost learning with VFS based on parallel genetic algorithm for image annotation. Although these VFS methods using the optimization models had shown their prominent performance, the disadvantage of these techniques was that they were only used on the 2000 images database.

3. QPSO and its improvement

3.1. Standard QPSO

Standard quantum particle swarm optimization (QPSO) [28] was proposed to overcome premature convergence, and it is simple and easy to understand. It is initialized with a population of random potential solutions and the algorithm searches for satisfying some performance. The potential solutions, called particles, are flown through a multidimensional search space. Each particle has a position represented by a vector X. A swarm of particles moves through a d-dimensional problem space. P_{best} and P_{global} are the best previously visited position of the particle and the best value of all particle position values respectively, where $X(t)=(X_1(t), X_2(t), \ldots, X_d(t))$.

In this paper, the fitness function of particle X is defined by the annotation accuracy, denoted by $F(X)$.