Int J Comput Vis
DOI 10.1007/s11263-013-0680-6

Active Image Clustering with Pairwise Constraints from Humans

Arijit Biswas · David Jacobs

Received: 25 February 2013 / Accepted: 23 November 2013

Abstract We propose a method of clustering images that combines algorithmic and human input. An algorithm provides us with pairwise image similarities. We then actively obtain selected, more accurate pairwise similarities from humans. A novel method is developed to choose the most useful pairs to show a person, obtaining constraints that improve clustering. In a clustering assignment, elements in each data pair are either in the same cluster or in different clusters. We simulate inverting these pairwise relations and see how that affects the overall clustering. We choose a pair that maximizes the expected change in the clustering. The proposed algorithm has high time complexity, so we also propose a version of this algorithm that is much faster and exactly replicates our original algorithm. We further improve run-time by adding two heuristics, and show that these do not significantly impact the effectiveness of our method. We have run experiments in three different domains, namely leaf, face and scene images, and show that the proposed method improves clustering performance significantly.

Keywords Clustering · Active learning · Human in the loop · Pairwise constraints · Image labeling

1 Introduction

Clustering, or *unsupervised learning*, is a critical part of the analysis of data. There has been a huge volume of work on clustering (Jain 2010), producing many interesting and effective algorithms. However, all clustering algorithms depend on some method of computing a distance between items to be clustered that reflects their similarity. For most tasks, automatically computed distances provide useful information about similarity, but still produce significant errors. This leads even the best clustering algorithms to produce clusters that do not contain objects from the same class.

We therefore propose a new clustering method that brings a human into the loop. In many tasks, experts, or even naive humans, can provide very accurate answers to the question of whether two objects belong in the same cluster. In spite of this accuracy, it is not practical to expect people to cluster thousands of objects into meaningful groups. Our goal, therefore is to meld human and automatic resources by directing valuable human attention to those judgments that are most critical to improving clusterings produced by automatic means.

We illustrate the value of this approach with examples of clustering in surveillance videos and plant images.

- There are many applications for clustering faces or actions in surveillance videos. This could allow, for example, an analyst to determine whether the same person has visited a number of locations, or to find different people who have performed similar actions. Images from videos have variations in pose, illumination and resolution that make automatic analysis extremely challenging, so that automatic clustering will be quite error-prone. But a person can readily look at two face images or actions and tell if they are similar.
- There has been a great deal of interest recently in obtaining large, labeled image sets for plant species identification (Kumar et al. 2012). Classifiers that can identify species require large sets of leaf images, labeled with their species. Accurately labeling such images requires

A. Biswas (✉) · D. Jacobs
Computer Science Department, University of Maryland,
College Park, MD 20742, USA
e-mail: arijitbiswas87@gmail.com

D. Jacobs
e-mail: djacobs@umiacs.umd.edu

Published online: 10 December 2013