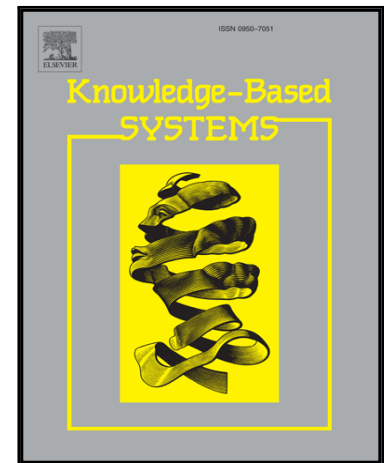


Sparse Learning based Fuzzy C-means Clustering

Jing Gu , Licheng Jiao , Shuyuan Yang , Jiaqi Zhao

PII: S0950-7051(16)30491-9
DOI: [10.1016/j.knosys.2016.12.006](https://doi.org/10.1016/j.knosys.2016.12.006)
Reference: KNOSYS 3757



To appear in: *Knowledge-Based Systems*

Received date: 4 July 2016
Revised date: 1 December 2016

Please cite this article as: Jing Gu , Licheng Jiao , Shuyuan Yang , Jiaqi Zhao , Sparse Learning based Fuzzy C-means Clustering, *Knowledge-Based Systems* (2016), doi: [10.1016/j.knosys.2016.12.006](https://doi.org/10.1016/j.knosys.2016.12.006)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Sparse Learning based Fuzzy C-means Clustering

Jing Gu*, Licheng Jiao, Shuyuan Yang, Jiaqi Zhao

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, International Research Center for Intelligent Perception and Computation, Joint International Research Laboratory of Intelligent Perception and Computation, Xidian University, Xi'an 710071, China

Abstract: Recently sparse representation (SR) based clustering has attracted a growing interests in the field of image processing and pattern recognition. Since the SR technology has favorable category distinguishing ability, we introduce it into the fuzzy clustering in this paper, and propose a new clustering algorithm, called sparse learning based fuzzy c-means (SL_FCM). Firstly, to reduce the computation complexity of the SR based FCM method, most energy of discriminant feature obtained by solving a SR model is reserved and the remainder is discarded. By this way, some redundant information (i.e. the correlation among samples of different classes) in the discriminant feature can be also removed, which can improve the clustering quality. Furthermore, to further enhance the clustering performance, the position information of valid values in discriminant feature is also used to re-define the distance between sample and clustering center in SL_FCM. The weighted distance in SL_FCM can enhance the similarity of the samples from the same class and the difference of the samples of different classes, thus to improve the clustering result. In addition, as the dimension of stored discriminant feature of each sample is different, we use set operations to formulate the distance and cluster center in SL_FCM. The comparisons on several datasets and images demonstrate that SL_FCM performs better than other state-of-art methods with higher accuracy, while keeps low spatial and computational complexity, especially for the large scale dataset and image.

Keywords: Sparse representation, fuzzy c-means, clustering, set operations.

1. Introduction

Data clustering is a key and fundamental topic in data analysis [1] and data mining [2]. Actually, clustering is to divide a dataset into several meaningful groups. The samples in a group are more similar in a specific metric than the others in different groups. There are two main clustering strategies: hard clustering

* Corresponding author at: Electronic Engineering Institute, Xidian University, Xi'an 710071, China. Tel.: +86 029 88202661.

E-mail address: xuer6126@126.com { Jing Gu }.

Download English Version:

<https://daneshyari.com/en/article/4946404>

Download Persian Version:

<https://daneshyari.com/article/4946404>

[Daneshyari.com](https://daneshyari.com)