

Improving Fuzzy Clustering Algorithm for Probability Density Functions and Applying in Image Recognition

Dinh Phamtoan¹, Ngoc Lethikim¹ Tai Vovan²

¹*Faculty of Engineering Van Lang University*

²*College of Natural Science, Can Tho University, Can Tho city, Vietnam*

Corresponding author: Tai Vovan (vvtai@ctu.edu.vn)

Abstract: This study introduces a measure called coefficient of within-cluster proximity (CWP) to evaluate the similarity of probability density functions (DFs) within clusters. After surveying the under and upper, and the computational problems of CWP, a fuzzy clustering algorithm for DFs is proposed. This algorithm can determine the suitable number of clusters and find the probability for each DF to belong to specific cluster. The convergence of the algorithm is considered in theory and illustrated by the numerical examples. The algorithm is applied to image recognition. The results show strong advantages of it in comparison to other algorithms. They also indicate the potential of the proposed approach in application to the data of different types.

Keywords: Automatic algorithm, density function, fuzzy cluster analysis, image recognition.

Mathematics Subject Classification: 62H30; 68T10

1 Introduction

Cluster analysis is intended to separate the initial data into sub-groups or “clusters” in such a way that the elements in the same clusters are more similar to each other than the elements form other clusters based on any criteria (Thao and Tai 2017). It is a crucial development of pattern recognition and an interesting topic of multidimensional statistics. Cluster analysis has a basis role in data mining as well as data analysis used in many areas. Therefore, it has been concerned by many statisticians (Chicco et al. 2003; Austin et al. 2005; Ciszak 2008).