



PCA based clustering for brain tumor segmentation of T1w MRI images



Irem Ersöz Kaya^a, Ayça Çakmak Pehlivanlı^b, Emine Gezmez Sekizkardeş^c, Turgay Ibrikci^{c,*}

^aMersin University, Software Eng. Dept. 33440 Tarsus, Mersin, Turkey

^bMimar Sinan Fine Arts University, Department of Statistics, Istanbul, Turkey

^cCukurova University, Electrical-Electronics Eng. Dept. 01330, Adana, Turkey

ARTICLE INFO

Article history:

Received 1 May 2016

Revised 10 October 2016

Accepted 23 November 2016

Keywords:

Dimension reduction

PCA algorithms

Clustering

k-means

Fuzzy C-Means

ABSTRACT

Background and objective: Medical images are huge collections of information that are difficult to store and process consuming extensive computing time. Therefore, the reduction techniques are commonly used as a data pre-processing step to make the image data less complex so that a high-dimensional data can be identified by an appropriate low-dimensional representation. PCA is one of the most popular multivariate methods for data reduction. This paper is focused on T1-weighted MRI images clustering for brain tumor segmentation with dimension reduction by different common Principle Component Analysis (PCA) algorithms. Our primary aim is to present a comparison between different variations of PCA algorithms on MRIs for two cluster methods.

Methods: Five most common PCA algorithms; namely the conventional PCA, Probabilistic Principal Component Analysis (PPCA), Expectation Maximization Based Principal Component Analysis (EM-PCA), Generalize Hebbian Algorithm (GHA), and Adaptive Principal Component Extraction (APEX) were applied to reduce dimensionality in advance of two clustering algorithms, K-Means and Fuzzy C-Means. In the study, the T1-weighted MRI images of the human brain with brain tumor were used for clustering. In addition to the original size of 512 lines and 512 pixels per line, three more different sizes, 256 × 256, 128 × 128 and 64 × 64, were included in the study to examine their effect on the methods.

Results: The obtained results were compared in terms of both the reconstruction errors and the Euclidean distance errors among the clustered images containing the same number of principle components.

Conclusion: According to the findings, the PPCA obtained the best results among all others. Furthermore, the EM-PCA and the PPCA assisted K-Means algorithm to accomplish the best clustering performance in the majority as well as achieving significant results with both clustering algorithms for all size of T1w MRI images.

© 2016 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Brain tumor is a mass or a growth of abnormal cells in brain tissue. There are two types of tumors, cancerous and non-cancerous. Researches show that over a third of brain tumors are cancerous, and brain tumors rank among the top of cancer related deaths. Therefore, early diagnosis and treatment are crucial for survival [1].

The detailed aspects of brain tumors can be displayed by imaging modalities such as X-Ray, Ultrasonography, Computed Tomography (CT), and Magnetic Resonance Imaging (MRI), enabling clin-

ical doctors to understand the texture of brain tumors and determine the type of therapy. The location and the extent of the tumor considering its posture in compliance with the surroundings are the determining factors in the therapy decision.

Since the capacity to obtain brain tumor images has outstripped the ability to analyze and segment these images manually, the studies on computational segmentation of brain tumor have been inevitably motivated. In recent years, brain tumor segmentation has become one of the most challenging tasks in medical image analysis.

Clustering is ideally used for image segmentation because of its ability to find out the complex relationships hidden in large unlabeled data sets based on a heuristic search for interesting and reasonable features [2–4]. Image clustering is simply mapping the image into clusters such that the set of clusters presents the same information about the image as the entire image-set collection. The

* Corresponding author.

E-mail addresses: iremer@mersin.edu.tr (I.E. Kaya), ayca.pehlivanli@msgsu.edu.tr (A.Ç. Pehlivanlı), eminegezmez@yahoo.com (E.G. Sekizkardeş), ibrikci@cu.edu.tr, ibrikci@cukurova.edu.tr (T. Ibrikci).