Telematics and Informatics 34 (2017) 299–310

Contents lists available at ScienceDirect

Telematics and Informatics

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

Clustering-based compression connected to cloud databases in telemedicine and long-term care applications

Wei-Yen Hsu

Department of Information Management, National Chung Cheng University, 168 University Rd. Sec. 1, Minhsiung Township, Chiayi County 621, Taiwan

ARTICLE INFO

Article history:
Received 2 March 2016
Received in revised form 11 May 2016
Accepted 15 May 2016
Available online 26 May 2016

Keywords:
Clustering-based compression
Telemedicine
Long-term care
Hopfield neural network

ABSTRACT

A novel clustering-based compression work connected to cloud databases is proposed for the applications of telemedicine and long-term care in this study, where the goal is to enhance information transfer rate and storage capacity to further improve communication between medical staffs and patients in long-term care and telemedicine. The proposed system mainly involves three-dimensional histogram competitive Hopfield neural network (CHNN) clustering, regionalization, and modified block truncation coding (BTC). Three-dimensional histogram CHNN clustering and regionalization are proposed to achieve better clustering accuracy within three-dimensional spaces and simultaneously overcome the problems of fluctuating initial values of clustering. Modified BTC is also proposed to analyze clustering regions with different compression rates according to their importance in order to greatly preserve important image feature information under the condition of smaller image sizes. The experimental results indicate that the proposed system is adaptive and performs better than several previous methods. It is also suggested being suitable for the applications of telemedicine and long-term care connected to cloud databases.

1. Introduction

Since the beginning of tele-aided (e.g. telemedicine) and medical tele-technology (e.g. tele-long-term care) systems, the process and uses of tele-data and tele-information have been limited to some well-understood external electrical and electromagnetic devices, such as personal computers, mobile phones, tablets, bluetooth instruments, network devices, and servers. However, the potential of creative technologies has become increasingly obvious, and widely various concepts and applications are in the process of development in the past decades (Doukas et al., 2011; Robinovitch et al., 2013). Currently, cloud databases, electronic medical records, machine learning, and long-term care are the most popular new issues and technologies for telemedicine and tele-care (Piette et al., 2015; Yang et al., 2015). Although commercial applications of the clustering-based compression systems are still limited and problematic, the researchers currently involved in telemedicine and long-term care have recognized that the transfer and storage of medical important data or information in cloud database environments using clustering-based compression approaches are significantly important (Livi et al., 2013).

In recent decades, clustering-based compression applied to digital images or data has become increasingly important in telemedicine and long-term care environment, since they are used more and more in healthcare-related and medical-related applications (Castiglione et al., 2015). A wide variety of digital imaging technologies may simultaneously produce a large amount of data or images for each patient. Hence, the process and storage of them are considered as a kind of problem of big data. For example, three-dimensional medical volume data are very large in size, and they are not suitable for storage

E-mail addresses: shenswy@gmail.com, shenswy@mis.ccu.edu.tw

http://dx.doi.org/10.1016/j.tele.2016.05.010
0736-5853/

with relatively slow network access. Accordingly, the efficient transmission and storage of such images, mainly in cloud databases, requires advanced compression techniques (Yue et al., 2013), such as clustering-based compression. More specifically, it is easy to note that the compression is essential for efficient transmission and storage of big data to remote cloud databases. In addition, the visual quality of these images in cloud databases is required to be high to guarantee correct assessments, analyses and diagnoses. Therefore, the clustering-based compression system connected to cloud databases should achieve the data availability anywhere and anytime to provide the benefits for both the patients and medical staffs. That is, the system can provide medical staffs with convenient aids to access resources easily and efficiently from cloud databases. In general, it is important to emphasize that the realization of clustering-based compression through cloud databases produces several challenges, such as data transmission, storage and visualization. Hence, the technical issues of data use in a cloud-database environment must be addressed by accommodating the users with the needs of telemedicine and long-term care. The aim is to provide the necessary infrastructure for allowing special users to use images or data over cloud databases from anywhere through different devices.

2. Objective

Clustering-based compression connected to cloud databases is a novel technology and concept to quickly transferring important information from one place to others by clustering and then compressing data before transmission. Further, it might be a highly feasible approach for clustering-based compression to assist medical staffs in telemedicine and long-term care. Compression based on the clustering of information and data associated with hospital and commercial real or cloud databases have grown rapidly in the last decade. Clustering-based compression analysis connected to cloud databases is based on clustering and compression using the state-of-the-art or novel algorithms. For this, the conventional compression approach can be enhanced by means of several ways. These include, preprocessing or an appropriate selection of the best parameters is known to improve the performance (Livi et al., 2013; Tan and Sun, 2014; Tsekouras and Tsolakis, 2013). However, the data and information are generally non-stationary and chaotic. This suggests that it is necessary to provide an adaptive and good-performance system for the applications of telemedicine and long-term care when it is connected to the cloud databases.

3. State of the art

In general, clustering-based compression procedure involves two main processes, image clustering and compression. Image clustering is an important issue because it greatly affects the effectiveness and efficacy of compression rates. Image clustering based on unsupervised neural-network or genetic-algorithm-based approaches usually partitions pixels or regions with common image properties into the same class. The neural-network and genetic-algorithm-based clustering algorithm is an unsupervised and automated technology extensively used for image clustering (Saprikis, 2013; Sim et al., 2014; Hsu, 2015a). A widely variety of methodologies have been presented for the clustering of images (Hancer et al., 2012; Huang et al., 2015; Lin et al., 2015).

An artificial-bee-colony-based image clustering approach (Hancer et al., 2012) is present to find the clusters of images, in which the number of clusters is specified. Image clustering is an approach of image analysis that supports high-level description of image content for image understanding where the aim is to find the relationships between the images and clusters. However, the clustering results are easily influenced by the noise of images and the initial parameters of the artificial bee colony approach. A neighborhood intuitionistic fuzzy c-means clustering algorithm (Huang et al., 2015) is proposed with a genetic algorithm. This clustering algorithm technology can retain the advantages of an intuitionistic fuzzy c-means clustering algorithm to maximize benefits and reduce noise/outlier influences through neighborhood membership. Furthermore, the genetic algorithms are used simultaneously to select the optimal parameters of the proposed clustering algorithm. It has been successfully applied to the clustering of different regions of magnetic resonance imaging and computerized tomography scanning. However, there would be greatly different for the results of clustering if the parameters of the intuitionistic fuzzy c-means clustering algorithm and genetic algorithm are not carefully selected. A functional principal component analysis (FPCA) and randomized sparse clustering algorithm (Lin et al., 2015) is present for medical image analysis. The FPCA is extended from one dimension to two dimensions to fully capture the space variation of image the signals. A randomized algorithm for accurate feature selection is then present in image clustering analysis. The image signals contain a large number of redundant features which provide no additional information for clustering analysis. The widely used methods for removing the irrelevant features are sparse clustering algorithms using a lasso-type penalty to select the features. It is applied to both the liver and kidney cancer histology image data from the TCGA database. However, the results of clustering are usually restricted to the quality of the acquired image signals or data. In the present study, we propose a three-dimensional histogram competitive Hopfield neural network (CHNN) clustering to cluster a medical image into several homogeneous regions via histogram and CHNN. In addition to more efficient computation, it can also achieve more precise classification.

For the compression process, vector quantization (VQ) is usually adopted to perform codebook design from training vectors (Gersho and Grey, 2012; Hsu, 2015b; Yang et al., 2011). Its purpose is to create a codebook for which the average distortion caused by the approximation of a training vector and a codevector in the codebook is minimized. It can be considered