Chapter 18
Image Data Mining Based on Wavelet Transform for Visualization of the Unique Characteristics of Image Data

Gebeyehu Belay Gebremeskel
Chongqing University, China

Zhou Shangbo
Chongqing University, China

Yi Chai
Chongqing University, China

Su Xu
Chongqing University, China

ABSTRACT

Mining techniques can play an important role in image decomposition, segmentation, classification and retrieval systems. As image data become more complex and growing at a fast pace, searching valuable information and knowledge implicit become more challenging than ever before. In this chapter, authors proposed a WT based DM techniques to optimize and characterize the unique feature of image retrieval, which is fundamental to optimize informative mathematical representation of image objects. Many software, including data exploratory tools such as DM packages contain fast and efficient programs that perform WT. Wavelets have quickly gained popularity among scientists and engineers, both in theoretical research and in applications. The authors discussed in details and introduced a novel method for image database analysis in different scenarios that foster the wide access of image data.

1. INTRODUCTION

Data Mining (DM) is a process of automatically extracting novel, useful, and understandable patterns from a large collection of data. Over the past decade, this area has become significant in many fields naming from the retailer -marketing to DNA -bioinformatics. DM techniques involve diverse dynamic and advanced tools, including wavelet to explore data sets as its nature and domain contexts (Jiawei, 2012). Wavelet theory could naturally play an important role in DM because it could provide data presentations that enable efficient and accurate mining process, which incorporated into the kernel for

DOI: 10.4018/978-1-5225-0983-7.ch018

many algorithms. Standard wavelet applications are mainly powerful for temporal-spatial data, which involves time-series, stream, and image data (Yu, Huang & Xia, 2012; Pimwadee, Aryya, George & Zhiyuan, 2011). It has been successfully applied to analyze large-scale image data using DM techniques. The approach is introducing a novel methodology how to reduce the amount of manual labor that usually comes with visualize and characterize big image data collections. With numeric and textual data, the techniques of extracting useful information from unstructured data have already been more or less established. However, with image-heavy datasets, processing methods such as object detection and text recognition are complex to be reliable and in most cases do not stand up to a comparison with a human doing the work.

Image mining is the process of searching and discovering valuable information and knowledge in large volumes of image data. It draws basic principles from Databases, Machine Learning (ML), Statistics, Pattern Recognition (PR) and 'soft' computing, which help to use Wavelet Transform (WT) based DM techniques enables a more efficient methods of the image data discrimination and analysis (Tao, Qi, Shenghuo & Mitsunori, 2002). However, image processing is one of those things people are still much better at than computers that use to know something new. Therefore, based on these two facts, we proposed WT based DM for visualization and characterization of the unique feature of image data. Besides DM algorithms, wavelet technique is growing importantly and having a lot of advantages that already exist numerous successful applications in image mining. The WT is syntheses of idea's computational methodologies are based orthonormal wavelet basis, which is fundamental to decompose, segments, extract and handling image data. WT based DM techniques, functions, or operators into different frequency components, the methodologies and image features' component with a resolution matched to its scale discussed in details.

The chapter is organized as first the introduction followed by section 2 about the related works, which focused on the facts and advancement of image DM in different approaches. It also includes image data managements and its attributes, segmentations, mining algorithms, distributed computing and others. In section 3, the wavelet technology, methodology and approaches are discussed. Section 4 discussed wavelet-based image annotations and measurements to visualize and characterize the detail and unique features of image data and mining techniques. In section 5, we present the summary of the chapter, which followed by the acknowledgment of the supporters of the chapter works and list of cited references.

2. A RATIONALE OF IMAGE DATA MINING

These days, image data are generated and collected in terabytes and petabytes, which are challenging to handle and analyze by traditional analytics tools. Such a huge amount of image data is generated in our daily life and each field, such as medical, satellite, all kinds of digital photographs images and others. These images involve a great number of useful and implicit information. The data types characterized by its volumes, variety, velocity … and complexity of image data that generated by the digital camera and video are unlabeled and difficult for users to discover. For example, manually annotating such a huge amount of image data is, time-consuming, laborious and prohibitively expensive (Xin, Lei, Xirong & Wei, 2007). To face the challenge, such as unlabeled online image resources, the proposed technique it is very important, which is capable and give clear insights to annotate and retrieve images based on their visual content. Therefore, WT-DM based image data synthesis is an automatic image processing