Distributed Approach to Process Satellite Image Edge Detection on Hadoop Using Artificial Bee Colony

Tapan Sharma, Amity University, Uttar Pradesh, India
Vinod Shokeen, Amity University, Uttar Pradesh, India
Sunil Mathur, Maharaja Agrasen Institute of Technology, Rohini, Delhi, India

ABSTRACT

The remote sensing domain has witnessed tremendous growth in the past decade, due to advancement in technology. In order to store and process such a large amount of data, a platform like Hadoop is leveraged. This article proposes a MapReduce (MR) approach to perform edge detection of satellite images using a nature-inspired algorithm Artificial Bee Colony (ABC). Edge detection is one of the significant steps in the field of image processing and is being used for object detection in the image. The article also compares two edge detection approaches on Hadoop with respect to scalability parameters such as scaleup and speedup. The experiment makes use of Amazon AWS Elastic MapReduce cluster to run MR jobs. It focuses on traditional edge detection algorithms like Canny Edge (CE) and the proposed MR based Artificial Bee Colony approach. It observes that for five images, the scaleup value of CE is 1.1 whereas, for MR-ABC, it is 1.2. Similarly, speedup values come out to be 1.02 and 1.04, respectively. The algorithm proposed by authors in this article scales comparatively better when compared to Canny Edge.

KEYWORDS
Artificial Bee Colony, Bio-Inspired, Canny Edge, Edge Detection, Image Processing, MapReduce, SequenceFile, Swarm Intelligence

INTRODUCTION

Edge detection is one of the oldest and basic components in the field of image processing. It is the main step required for feature extraction, object detection, and segmentation (Sergio et al., 1987). Edge detection is also a memory-intensive job and becomes crucial to work on better algorithms while focussing on large images. In this experiment, the authors have run a MapReduce program to perform edge detection techniques on satellite images. MapReduce is the distributed processing framework which is designed to process a large set of data in a distributed manner (http://hadoop.apache.org/). This processing engine works upon Hadoop Distributed File System (HDFS). Bigdata and analytics are nowadays are crucial components of any business (Krimpmann & Stühmeier, 2017). The information needs to be processed at a faster rate without compromising on scalability and accuracy of the results. As remote sensing imagery is also witnessing huge growth, it is important to explore novel approaches to process large amount of data. Given there may be different mechanisms to process images on Hadoop, the authors have compared different approaches.

DOI: 10.4018/IJSSMET.2020040105
There are many algorithms for edge detection but in this study, the authors have chosen two algorithms with different behavior to run on Hadoop. The first is a gradient operator based Canny Edge algorithm (Canny, 1986), while the other is a bio-inspired Artificial Bee Colony (Karaboga, 2005). These approaches have been designed to be executed in an optimized manner by leveraging the big data platform.

There are various gradient operators used for edge detection such as Robert Operator, Sobel Operator, Prewitt Operator, Laplace of Gaussian (LoG) and Canny Edge. Robert Operator is the most simple and quick operator to detect edges and uses 2*2 convolution kernels. Sobel Operator convolves with the original image to calculate approximations of the derivatives. It uses two 3*3 masks for calculation. Prewitt Operator’s mask values are different than Sobel, however, performs in a similar way. LoG is one of the most advanced techniques for edge detection. It is based on the calculation of zero crossings in the second derivative of an image. Canny Edge algorithm is more complex and advanced than the other algorithms. At the same time, it produces a much better output by lowering the error rate and reducing false positives.

The gradient is calculated for each image coordinate. The pixel is considered an edge if the gradient value exceeds the threshold amount. The intensity of edges is relatively higher than nearby pixels. Canny Edge is considered the most complex among all but produces the best output (Juneja & Sandhu, 2009). This paper compared the qualities of all the gradient operators and found that the effect of noise is the least in Canny Edge. Again, Katiyar & Arun (2014) showed that Canny Edge performs best in object extraction. The algorithm produces very few false edges as compared to all others. Sobel operator took the least time. The studies have also derived new Mathematical morphological edge detection practices having a quality similar to existing techniques (Kaur & Aggarwal, 2011). The work of Maini and Aggarwal (2009) compared and provided advantages and disadvantages of these algorithms. It was shown that although Canny Edge Detection is computationally expensive it performs better than other operators, under noisy conditions.

Artificial Bee Colony (ABC) is a nature-inspired metaheuristic algorithm, the type of stochastic algorithm that brings random-ness to the solution (Bianchi et al., 2009). Such algorithms are different than deterministic algorithm where the solution is generally fixed for a particular problem. Each time the algorithm is executed it produces the same result. On the other hand, stochastic or metaheuristic algorithms produce different paths to the solution and are not repeatable. The objective of such an algorithm is to have an efficient algorithm that should work most of the time and at the same time produce a good quality result. Bio-inspiring or nature-inspiring algorithms have a wide range of application. One of the recent studies proposed cuckoo search optimization approach for community detection in social media and compared it with other Swarm algorithms (Babers & Hassanien, 2017). Similarly, another study has reviewed different algorithms like Ant Colony Optimization, Bacteria Foraging, and many others being used for antenna designs in a wireless system (Dey & Ashour, 2016).

ABC algorithms can be applied for edge detection and classification in satellite images. Researchers applied ABC to achieve the unsupervised classification of remote sensing images. They chose this algorithm due to its reduced memory requirements and rapid convergence (Deriche & Fizazi, 2015). This paper compares the performances of ABC with Particle Swarm Optimization and other classification techniques. The comparison indicated that ABC can be used for multivariate clustering purpose (Karaboga & Ozturk, 2009). In another study conducted by Jayanth et al. (2015), ABC is employed to classify satellite data. The results of this study showed that ABC outperforms traditional classification algorithms such as Maximum Likelihood and Support Vector Machines.

Yigitbasi and Baykan (2013) proposed edge detection of greyscale images using ABC. The result is compared with operator-based algorithms such as Robert, Prewitt, Sobel and Canny Edge. They used Hamming Distance to compare two images. As per research by Goel et al. (2015) metaheuristic algorithms Cuckoo Search (CS) and ABC worked efficiently in classifying landcover features in satellite images using metrics such as the Kappa Coefficient.
Automatic Partitioning of Large Scale Simulation in Grid Computing for Run Time Reduction
[www.igi-global.com/chapter/automatic-partitioning-large-scale-simulation/64162?camid=4v1a](www.igi-global.com/chapter/automatic-partitioning-large-scale-simulation/64162?camid=4v1a)

Securing Data Storage By Extending Role-Based Access Control
[www.igi-global.com/article/securing-data-storage-by-extending-role-based-access-control/105508?camid=4v1a](www.igi-global.com/article/securing-data-storage-by-extending-role-based-access-control/105508?camid=4v1a)