Nat. Volatiles & Essent. Oils, 2021; 8(4): 6276-6281

NVEO
Natural Volatiles &
Essential Oils

Image Feature Eradication Using Hadoop with GPU Accelerators

Dr. Sushil Kumar[1], Shaina[2*]

[1]Associate Professor, BBK DAV College for women, Amritsar

[2*]Assistant Professor, GNA University, Sri Hargobindgarh, Phagwara

Abstract—In this paper, analysis of distributed image processing framework is presented, so that it becomes very useful to beginners and to those who are new to distributed image processing field. The outcome of the study of distributed image processing with emphasis on benefits and application domains is described. Various research issues are examined and some future research directions are identified for distributed image processing. This study provides concise understanding of the distributed image processing area.

Keywords-Distributed Image processing; SEIP; GPU; Parallel processing; HADOOP

INTRODUCTION

These days, here exist several descriptions in the Cyberspace, and with the expansion of cloud computing and big data bids, several of those descriptions essential to be treated for unlike kinds of requests by means of detailed image dispensation procedures. Meanwhile, here already exist many varieties of image dispensation procedures and their differences, while new procedures are still developing. Subsequently, a current delinquency is that in what way the efficacy of massive image processing is enhanced and support the incorporation of existing executions of image processing procedures into the schemes. This paper recommends a distributed image processing scheme named SEIP and GUI, which is constructed on Hadoop, and pays extendable in node construction to funding various varieties of image processing procedures on distributed platforms with GPU accelerators. The scheme also uses a pipeline-based outline to accelerate enormous image file dispensation. A demonstration claim for image feature removal is calculated. The scheme is appraised in a small- scale Hadoop cluster with GPU accelerators, and the tentative results display the serviceability and productivity ofSEIP.In the period of big data, demands for massive file dispensation raise quickly, in which image information inhabits substantial amount, such as depictions entrenched in web pages, pictures unconfined in social system, images of possessions in shopping sites, and so on. Usually, these imageries essential to be treated for dissimilar types of submissions, like content-based image retrieval, image explanation and organization, and image content acknowledgment. Due to the dimensions of imageries and the involvedness of related procedures, it is essential to use dispersed systems with accelerators (e.g., GPU) to

process these massive images. There are four compensations of distributed schemes over remote processers: data sharing, which permits numerous workers or machineries to contact a common catalogue; device distribution, which consents numerous manipulators or apparatuses to segment their strategies; infrastructures, that is, all the apparatuses can interconnect with each other more simply than inaccessible processers; flexibility, i.e., a dispersed processer can feast the workload over the obtainable apparatuses in an effective method. Associated with the single node atmosphere, by retaining dispersed schemes, we can get augmented performance, amplified dependability, and improved elasticity On comparation skill of human and computer for dispensation of image, we find that a human intelligence can procedure imageries very reckless as of its skill to simplify and infer information from image data and large number of neurons employed in parallel, on the further hand, the electronic image dispensation has many confines associated to human intelligence. At current greatest of the applied image dispensation [1] the whole thing contract with enhancement, restoration, and segmentation of the image. While the image clarification and information implication still need human participation. However, due to progression in Artificial Intelligence [2], parallel calculation, and network-based calculation [3], the explanation of overhead is conceivable with brainy image dispensation with parallel and distributed systems. As more sophisticated submissions are required by workers, less dispensation time and faster response are required from submissions. With e-information explosion, the computational load has amplified manifold and possible solution is parallel and dispersed processing.

In meekest words, the parallelism is instantaneous dispensation by two or more dispensation units. The parallelism can be achieved by one of two main tactics: (1) by means of parallel processing hardware and (2) by means of distributed computing/processing scheme. In the previous method, a computer system having manifold computers is used to achieve calculation task, while in last method, number of machineries associated in network are used to achieve computation task. Both methods have their merits and demerits. Though, dependent upon the prerequisite of the request and accessible budget, the collection of construction is done.

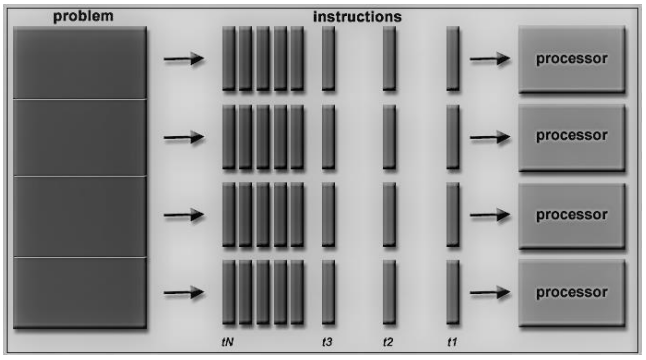

Fig.1. Parallel framework for massive image processing

The paper is structured as follows. Section II deliberates on aids of distributed image processing schemes. Section III defines the connected work. Section IV delivers the concept of SEIP system architecture.

I. BENEFITS OF DISTRIBUTED IMAGE PRCOESSINGSYSTEMS

1) Easy integration with the image analysis and image management systems.
2) Flexible, scalable and convenient to use.
3) Being a web-based service, no time consuming and cumbersome installation and configuration tasks.
4) No expenditure on outright purchase of software and hardware.
5) Anytime anywhere access because of the web-based architecture.
6) Ensure security and confidentiality of data.
7) Easy and convenient addition and modification of new algorithms by developers.
8) Image spooler utility, job queue facility and performance counter services for systematic and automated addition, routing and monitoring of jobs assigned to multiple connected image processing servers.
9) Facility for user to view job status for submittedjobs.
10) Status alerts to mobile devices to keep the user updated at alltimes.

RELATED WORK

In [4], a low-cost distributed computing approach for the post-processing of videos is described. A framework-based approach is used. Experimental results suggest that this approach lead to significant speedups with a reduction in operating costs by leveraging existing underutilized equipment.

A detailed report of results on an evaluation of performance of sequence of video frames is presented in [5], which remains a problem for video processing in distributed environments, and on parallel experiments using MapReduce on Hadoop.

A system for efficient image processing (SEIP) on distributed platform is build on hadoop in [6] using the pipeline-based framework that supports various image processing algorithms on distributed platforms with graphics processing unit (GPU) interface accelerators to accelerate massive image file processing. Experimental results on image datasets of 256 * 384, 400 * 400, 800* 800 pixels show the efficiency and usability of SEIP on distributed platforms.

A novel approach to develop distributed data driven storage systems is presented in [7]. The established technique permits accepting the squeezing, dismissal and security algorithms usage and procedure dismissal on the block level. This approach flexibly manages the redundancy, security permissions and resource consumption. In [8], key attainments within the GreenLand platform are tinted, concerning preparation, execution, and nursing the Grid procedures. Its expansion is founded on humble, but influential, notion of mathematical focused acyclic charts that are used in parallel and distributed performances over the Grid substructure. Due to the high complexity and size of input data satellite image processing requires high computing power. In order to be able to meet these requirements gProcess used the Grid execution platform. gProcess offered optimized execution and scheduling of multiple workflows so as to obtain the highest possiblethroughput.

A Hadoop image processing framework has been developed in [9] to abstract technical details of hadoopmapreduce platform. This framework addresses the problem of providing a system for computationally intensive data processing and distributed storage. It processes the large datasets easily and provides hadoop based library to support large scale image processing applications. The disadvantage is that the learning of hadoop requires alot of time and experience to develop useful applications.

In [10], a parallel processing based on integrated approach of Hadoop and CUDA for large scale image Processing is presented. It makes the usage of high dependability, scalability and fault tolerance competence of Hadoop system and high calculation power of CUDA for dispensation huge quantity of descriptions in highly well-organized manner. As Hadoop achieves well for information concentrated application over the use of HDFS and CUDA helps finest in case of calculate concentrated request, addition of both the framework delivers quicker implementation for image processing duty.

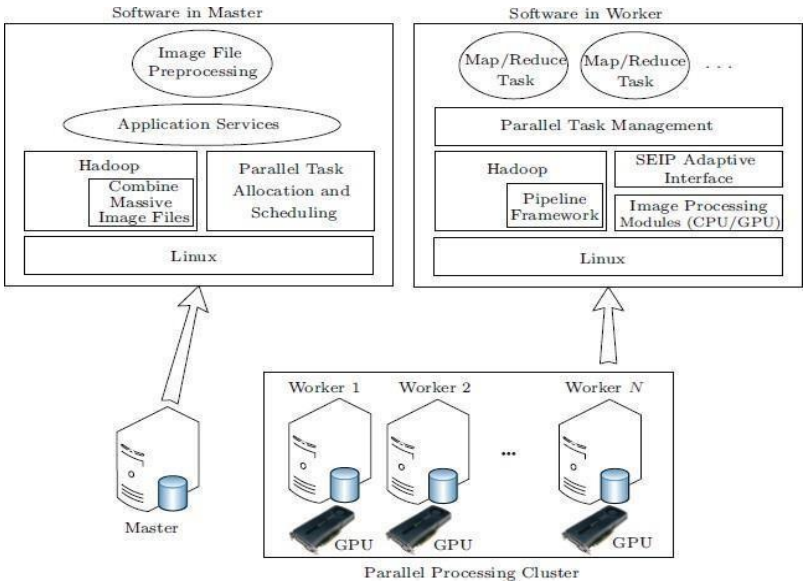

Fig.2. SEIP system architecture.

SEIP SYSTEM ARCHITECTURE

SEIP is built on Hadoop, and includes of single ace and dissimilar authorities in a cluster, and the ace and laborers are the same name node and data nodes of HDFS separately. The ace panels many authorities by apportioning projects to them, and laborers equipped with GPU and multi-center mainframes make huge image data formulating concurrently. Fig.2 establishes the outline manufacturing of SEIP, where components in oval are application-particular and should be revised or redid for a submission. The"pipelinestructure" module underpins pipeline management of image credentials for attendant/lessenprojects.

CONCLUSION AND FUTURESCOPE

A study of distributed image processing is presented in this paper. Various benefits of distributed image processing systems provided. A thorough survey has been done on various techniques and frameworks of distributed image processing framework. The application domains are also listed here. This study, which has considered different aspects, would be helpful to get concise understanding of the distributed image processing area. Further research study can be explored in specific hardware devices/processors used for image processing. In distributed image processing, analytical study on image processing and content-based image retrieval can provide details on algorithms, techniques, system architectures, etc. covering vast applicability of image processing operations. The outline will be improved incessantly, and our future effort will distillate on more flexible pipeline system with load-adjusting, amongst stages, as well as between CPUs-GPUs.

REFERENCES

[1] S. Haykin, Neural Networks: A Comprehensive Foundation. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.

[2] Coulouris, J. Dollimore, and T. Kindberg, Distributed Systems: Concepts and Design (4th Edition) (International Computer Science). Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.,2005.

[3] R. C. Gonzalez and R. E. Woods, Digital Image Processing (3rd Edition). Upper Saddle River, NJ, USA:Prentice-Hall, Inc., 2006

[4] Thompson, Craig, Christopher Brooks, and Jim Greer. "Distributed Image Processing for Automated Lecture Capture Post-Production." Multimedia, 2008. ISM 2008. Tenth IEEE International Symposium on. IEEE,2008.

[5] Yamamoto, Muneto, and Kunihiko Kaneko. "Parallel image database processing with mapreduce and performance evaluation in pseudo distributed mode." International Journal of Electronic Commerce Studies 3.2 (2012):211.

[6] Kokoulin, Andrey. "Methods for large image distributed processing and storage." EUROCON, 2013 IEEE. IEEE, 2013.

[7] Colceriu, Vlad, et al. "Mathematical Modeling of Distributed Image Processing Algorithms." International Journal of Advanced Computer Science and Applications(2013).

[8] Liu, Tao, et al. "SEIP: System for Efficient Image Processing on Distributed Platform." Journal of Computer Science and Technology 30.6 (2015):1215-1232.

[9] Vemula, Sridhar, and Christopher Crick. "Hadoop Image Processing Framework." 2015 IEEE International Congress on Big Data. IEEE,2015.

[10] Helly Patel, and Krunal Panchal. "IncopratingCuda In Hadoop Image Processsing Interface For DistributedImage Processing."Internation Journal Of Advance Research And Innovative Id

[11] eas In Education 2.4 (2016) :93-98.