International Journal of Engineering & Technology, 7 (3.34) (2018) 855-859

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Multimedia Big Data Processing Techniques: a Survey

Chaithra[1], Roshan Fernandes[2]

[1,2]Department of Computer Science and Engineering, NMAM Institute of Technology
Nitte, India
**Corresponding author Email: [1]chaithrasa02@gmail.com, [2]roshan_nmamit@nitte.edu.in*

Abstract

Massive amount of multimedia information are used in day to day life, especially in social media such as Facebook, Whatsapp, Instagram etc. Along with this, most of the real time applications, produce large scale video data which has to be continuously monitored, analysed and processed, so that easier interpretation of required patterns can be done, and real time applications such as detecting traffic flow, predicting weather can be eased out. In this paper, the various existing work done on multimedia big data processing using Hadoop have been discussed. A review on Hadoop-MapReduce system along with HDFS and how multimedia processing has been done in a parallel and distributed manner using it, have also been examined and discussed. In addition to that, the availability and benefits of Real time big data processing tools, methods, and framework such as Apache Spark is been highlighted, and need for the discovery of much more better real time multimedia big data processing framework have been talked about**.**

Keywords*: Hadoop, MapReduce, HDFS, Apache Hadoop, Apache Spark*

1. Introduction

1.1 Big Data

Big Data is a recent technology which has become a very important aspect in many organizations, which mainly plays a role of collecting large scale specific and useful information about several problems such as marketing, financial, transportation etc. Different domains that use Big Data widely are research, science, medical, crime detection, business and many more. It mainly helps in business for analyzing the trends and also helps in forecasting, and has various real time applications such as detecting traffic flow, predicting weather etc. It is also used in processing large scale images and videos. For example, CCTV cameras present in garment shops generates large scale video data continuously, which can be used for automatic detection of violation of specific rules.

The collection and hence processing of Big Data can be categorized most of the times into batch processing. Processing the data and producing the results within few seconds is supported by Real-time processing, and hence is very quick. This mainly requires collecting the input continuously, process it and then produce the output. Batch processing collects the data in a certain period of time, and processing takes place once the data is stored in data stores.

1.2. Hadoop a Distributed Framework

Hadoop plays a major role in analysing and storing huge amount of data especially unstructured data in distributed manner. This is a framework which supports processing the data by parallelizing the work, for which it requires distributing the data to work nodes, and hence forming a Hadoop cluster which can be extended easily. Its main component is the distributed file system called HDFS which concentrates on storing and sharing, and also helps in retrieving huge files. Since the files are distributed to several work nodes, the reliability and availability of the system is increased. After taking place the role of distribution in Hadoop, the analysing part of the data and processing part can be taken place in MapReducer.

Large scale video processing can be easily done using hadoop system which mainly involves analysing and then processing huge amount of images and videos, so that interpretations of semantic patterns can be eased out. Big video processing in real time applications may require analysing of large amount of video footage which will be generated continuously. Hence visual big data analytics helps to process such data in a quick manner. HadoopMapReduce technology in coordination with several algorithms for video processing helped in applications which require detecting objects in large scale video data.

Hadoop image processing is also a popular recent technology. Some of the image processing applications, which are mainly designed for small scale purpose and for performing computations locally, are not effective for handling the large scale multimedia data. Hence, Hadoop image processing framework is been developed, which supports massive image processing by providing Hadoop-based library. A MapReduce framework used along with hadoop helps in easier processing of large scale images.

1.3 Hadoop-Mapreduce System

MapReduce is a programming system initially created to encourage the processing of immense amount of data specifically in parallel manner, and in a fault tolerant and reliable way, on clusters. When MapReduce is combined with Hadoop Distributed File System (HDFS), it will be easier to process massive amount of data. HDFS-MapReduce system is usually alluded to as Hadoop. The essential unit of data, utilized as a part of MapReduce is a (Key, Value) pair. Any form of data, may be

structured or unstructured has to be converted to this form, before giving it to MapReduce part..

2. Literature Survey

Raju, Kishan, Harikrishna, Abhilash [1] have introduced specific methods which help in handling massive amount of video data using a framework called Hadoop. HadoopMapReduce Technology in coordination with several video processing algorithms mainly supported detection of objects in massive video data. Similarly Husain, Meena, Akash, Harish, Shiddu and Sonal [2] developed a distributed system using Apache Hadoop framework. Their intention was to count the occurrences or the frequency of each textual word from large scale video data or video frames. The handling of large scale images is exceptionally simultaneous, batch-wise processing of text content reading and tallying the event of each word by utilizing MapReduce system was actualized. Since in both the cases, the implementation is done on Hadoop clusters, there is a large improvement in running times because of parallel work when brought in comparison to the serial work on single node.

Dinesh, Vishnu and Krishna [3] developed a method for detection of vehicle riders (specifically to bikes) automatically, who violate particular rules by not wearing helmets. For this purpose, framework has been developed called visual big data analytics, mainly for analysing and processing large scale images and videos. Here the experiment was taken using 5-fold cross validation, so that for each combination of representation and model, the performance will be validated. The results shows that best performance and accuracy of 95% and above is achieved, by using HOG descriptors. Along with that, the false alarm rate is less than 1%. Main advantage of using this framework is that, it has the ability to work in real time and return results in real time.

Similarly, Toropov, Liangyan, Shanghang, Satwik, Jose [4] developed a method to detect or inspect traffic flow. The input taken was from low quality, low frame rate city camera (video). Major stages involved background subtraction, scene geometry, car detection and car counting. In both cases, in order to improve the accuracy they have captured many representative features together, later they have taken the support of HOG features along with color histogram. Since Earth Mover's Distance is more robust, it is used for measuring similarity of histogram.

Basil, Ramesh, Abraham and Vikas [5] introduced a Framework for analysing large surgical videos, which was named by Surgical Video Analysis Framework using Hadoop. The parallel processing of large scale images that is present in the HIB bundle is taken up by Mappers, in addition to data they record information of instrument frame that is identified. For local image processing mappers make use of three feature extraction methods, namely SIFT which is Scale Invariant Feature Transform, Haralick Texture Descriptor with SVM, and Speeded Up Robust Features with SVM. The latter two methods has shown improved accuracy over SIFT. With the increasing amount of images, generation of HIB bundle is linear. HIB bundle consisted of more than 7000 images, which was added just in 72 seconds, which is hence an advantage towards using Hadoop multimode cluster.

Jose and Molto [6] deals with visual data from many streaming multimedia Television and also radio stations, which uses Hadoop framework to record, process, and visualize the data. In order to automate the visualization process for stages such as Automatic Speech Recognition, R framework was created. Optimizing large number of Hadoop nodes was accentuated, and Enhanced performance is achieved by using techniques namely diarization. The shown parallel work in MapReduce Framework helped in lessening processing time.

Gabriel and Paulo [7] eased out the agricultural soil analysis by introducing a Big Data platform, mainly from Computed Tomography (CT) images. Big data environment along with MapReduce can process more prominent number of recreated CT images (2D and 3D) compared to that of sequential recreation or reconstruction, in light of the fact that Hadoop provides distributed and parallel calculation.

Saman and Mehdi [8] used Big Data Spark/ PySpark stage, in order to outline, create and effectively test another pipeline for medical imaging data (specifically fMRI), specifically on single node. This permitted to peruse and stack imaging data, change over the data to Resilient Distributed Datasets, keeping in mind the end goal to control and perform in-memory data processing in parallel, and change over the definite outcomes to imaging format. These methods improved the performance and were four times faster in terms of processing time, when contrasted with the previous work in Python. It demonstrated easy usability and also storage results in various data structures. The method also supports scalability, because the pipeline can be extended to have multiple nodes, and also supports huge data examination and hence improves performance and processing time compared to that of single node.

Roshan, Deepu, Chandrashekhar and Swarup [9] used Hadoop for data processing and Hbase for storing output, and worked with MapReduce framework for parallel processing, specifically of remote sensing satellite data. Since the workload can be distributed over multiple clusters of Hadoop, the performance and speed up can be easily achieved and improved significantly. Different image processing algorithms namely, Image Registration, Gaussian Filter and image mosaicking have been used in order to analyze the performance of framework in obtained results.

Honghua, Ming, Li Li, Ye Tian, Yanfeng [10] concentrates on main technology of storing and processing for continuously observable video data, in view of big data and data mining values, and also outlined the big data analysis framework for surveillance video. The examination on real-time or ongoing intelligence analysis advances the video observation from manual inspecting to productively progress ahead of time, post examination, acknowledgement of insightful data examination, determining. The results showed that cloud computing technology in view of big data processing platform gives a capable help to powerful information stockpiling and supports processing in parallel manner.

Bharat, Pavan and Girish [11] made use of incremental background model learning in order to outline a viable online multi-linear subspace learning algorithm. This modelling in co-ordination with proper post processing steps, is valuable to detect or identify erroneous events, consequently the foreground or closer view in the video. Different available datasets have been used to test the algorithm effectiveness Swapnil, Abhishek, Praveen [12] gave an answer to one of the image processing application which helps in examining and analyzing images from industrial modern application namely, OLED –Organic-Light-Emitting-Diode Printing for OLED center detection. In order to parallelize the work, Hadoop as well as HIPI (Hadoop Image Processing Interface) is been utilized. It also improves the performance and processing time significantly by processing images in parallel manner, using mapper task.

Alberto, Ruben, Ruben [13] proposed a continuous ongoing Big data architecture, mainly to gather, keep up and dissect massive amount of images which are in consideration with the issue of detection of glasses on human faces automatically. In order to get the space around the eyes Face normalization algorithm is utilized. One of major operator used for classification is the Local Binary Pattern. Initially, to detect four sided box of the glasses, SVM'S are used for classifying. To get the correct faces, non-maximum suppression is applied.

Xiaomeng, Huadong, Haitao, Yi Tang, Yue Kou [14] extended Hadoop by developing Hadoop Video Processing Interface (HVPI), so that it helps to deal with video analytic applications. It helps out the developers by giving simple touse r/w interface so that they can rapidly fabricate vast scale video analytic application in view of Hadoop. In order to port existing video analytic

applications written in c/c++ into Hadoop platform, native programming interface is been provided. Similarly, Chungmo, Daecheol, Minwook, Cheolgi, Euiseong [15] supported video processing in parallel manner, in cloud environment by producing an extensible framework for video processing on Apache Hadoop. Image processing is employed to OpenCV, and Video coder is employed to FFmpeg, The system provides scalability of 75% in 8-core environment with two quad core systems.

A Manju, P Valarmathie [16] supported to associate video resources by creating a model using Semantic Link Network model. For search query users can choose the attributes. This is mainly possible by fuzzy reasoning as well as knowledge extraction process. Due to this, video examination or analysis can be dynamically managed, as well as the defined fuzzy rules helped in making the decisions, and hence this helps in efficient event extraction.

Andrey and Timofei [17] presented a MapReduce Image Processing framework (MIPr), where the image processing is done in the distributed manner. It also helps by providing libraries as well as high level API for image processing. MIPr provides parallel and more simplified method for image processing. Similarly, Sridhar and Christopher [18] provided a significant support for large scale image processing, by providing a Hadoop based library, by developing a Hadoop Image Processing Framework. The system gives another level of straightforwardness and effortlessness for making massive amount of image processing applications over HadoopMapReduce framework.

Tongke Fan [19], brought up with Canopy+ K-means. In large scale data, understanding is quite complex and efficiency of clustering is a problem. So clustering algorithm, where the storage ability of Hadoop cluster is fully utilized by MapReduce model. The operation speed has found a great improvement. In hadoop environment, Canopy K-means as well as K-means shows a very good speed-up ratio.

Mazhar, Awais, Anand, Seugmin [20] deals with real time applications, for handling VoIP flows, in order to control high speed traffic easily. This is done by implementing in a parallel manner, using Hadoop ecosystem along with the Spark. Evaluation results show the effectiveness and efficiency and they could prove that the system's accuracy is more than other techniques by making comparative study

Kyoungsoo, Jaemin, Jongtae, Yeonwoo, Jaesoo [21] developed a scheduling scheme, which played a great role in decreasing the deadline missing of jobs. For this the deadlines will be given while processing massive amount of multimedia data, say videos and images by utilizing MapReduce approach. Results proved that, the method decreased the time of completion by 13.8%, and deadline success ratio is improvised by 11% on an average, when bought in contrast with other existing scheme

David, Michael and Pavel [22] tried to speed up the multimedia feature extraction process. Grid Computing, Apache Storm, Apache Spark and Apache Hadoop were the four different approaches that were compared. All approaches tried to speed-up the process significantly when brought in contrast with sequential approach. In that Apache Hadoop showed that it was beneficial when large scale massive data need to be stored into HDFS and then processed, which hence supports batch processing approach and improves the processing time. After that, Apache Spark was initially as one of the distributed structure that can beat Hadoop. A deliberation called RDD predominantly upholds the Spark speed, which enables the execution calculations in memory. Extensive variety of workloads such as cluster applications, iterative calculations and streaming, are greatly handled by Apache Spark. This method adds simplicity, and makes it easier to join various processing situations.

Large scale data is easily handled by Hadoop due to its distributied nature, but for high speed real time processing, these tools are not suitable and are not efficient. So Mazhar, Hojae, Awais, Anand, Gwanggil [23] developed a real time Big Data architecture for which proved to be efficient, in analyzing and processing large scale real time multimedia data. They retained the distributed nature benefit from hadoop, but along with that they also added GPU (Graphic Processing Unit) and Spark to get the full benefit. In addition to that, a different implementation of MapReduce, and Apache Spark streaming is employed. By this, they were able to record the real time information from remote areas, and then divide the work among the various Hadoop data nodes, utilizing the Hadoop Distributed File System associated with Apache Spark, and also utilizing GPU's which works with parallel, distributed nature of Hadoop ecosystem. The results proved that framework developed, is more proficient mainly for real time cases compared to other existing methods.

Sl. No	**Author**	**Data Used For Proc-essing**	**Tool, Technique/ Algorithm used**	**Merits**	**Demeri-ts**
1	U. Raju, N.Kishan, HarikrishnaKotteAbhilash [1]	Video	Backgrou-nd Subtraction Hadoop, Mapreduce, Object Detection	Improvem-ent in the running time	No methods to support real-time process-ing
2	Moula, Meena S, Akash K, Harish, Shiddu, SonalShetty [2]	Video	Mapreduce, Apache Hadoop, Framework Job-Tracker, Tasktracker HDFS	Improved computational speed	Memory shortage problem, requires high end machine configu-ration
3	Dinesh , Vishnu and Krishna Mohan [3]	Video	Backgrou-ndSubtractionHistogram Of Oriented Gradients (HOG) , Scale Invariant Feature Transform (SIFT) And Local Binary Patterns (LBP)	Best performan-ce, better accuracy	Not able to detect static objects and also other moving objects like human, cars etc.
4	Evgeny, Liangyan, Shanghang, Satwik, F.Moura [4]	Low quali-ty, low frame rate city video	Backgrou-nd SubtractionScene Geometry, Car Detection, And Car Counting	More accuracy	Poor performance in case of closer parts of image
5	Basil, Ramesh , Dr. Abraham Varghese,Dr.Vikas [5]	High-definiti-on video	Hadoop Image Processing Interface, SIFT, SURF with SVM, Haralick Texture Descriptor with SVM	More accuracy	Accura-cy of system using SIFT is very poor, accuracy using SURF and SVM is not that reliable
6	Jose Herrera, German Molto[6]	Video	HadoopFrameworkR Framework DiarizationOptimizati-on, Mapreduce	Enhanced performan-ce	With large block size parallelism is worse and may require more storage
7	Gabriel M. Alves, Paulo Cruvinel[7]	CT Images	Hadoop Framework	Improved analysis	Not able to handle situationlike semi-arid,

			Mapreduce		crop rotation systems etc.
8	SamanSarraf, Mehdi Ostadhash-em [8]	Medi-calimag-ing data	Big Data Spark /Pyspark Platform, Mllib, Spark SQL	Ease of use, in-memory data processing, improved processing time	No parallel comput-ation develop-ed
9	Roshan, Deepu, MaruthiChandrase-khar, Shanti Swarup [9]	Satell-ite Image data	MapreduceFrameworkHbase, Image Registration, Image Segmentation,Gaussian Filter	7X speed could be achieved on large scale image data processing	requires high end machine configu-ration
10	Honghua Ming Fang, Ye Tian, Yanfeng, Li Li [10]	Video	Intelligent Analysis,Big Data Acquisition TechnologyMapreduce, HDFS	Large-scale parallel processing	Need of deep mining for complex abstract-ions
11	Venkatesh,Pavan Reddy K Girish [11]	Video	Backgrou-ndModeling, Event Detection	Improvised incremental technique	Unable to detect objects for videos with moving backgr-ound
12	SwapnilArsh, AbhishekBhat, Praveen Kumar [12]	Image	Hadoop, HIPI	Significant performan-ce improvement	Splitting of data is crucial to complet-ely utilize the power of parallel process-ing
13	Alberto Fernandez Ruben Casado, Ruben Usamentia-ga [13]	Images	Local Binary Pattern, Support Vector Machine Classifiers, Real-Time Big Data Architectu-re	Works in real time	Need of deep mining for complex abstract-ions represe-ntation
14	Xiaomeng, Huadong , Haitao , Yue Kou, Yi Tang [14]	Video	Hadoop Video Processing Interface HVPI, Distributed Video Processing, JNI, Hadoop Steaming, Hadoop Pipes	Scalable and efficient	Takes extra time to transfer other data blocks to machine which the mapper locates.
15	Chungmo, Daecheol, MinwookCheolgi, Euiseong [15]	Video	FFmpeg, Apache Hadoop, OpenCV, MapReduce	Very fast, Very stable	Hyper threadi-ng was found to be not very effectiveperform-ance loss due to I/O overhead
16	A.Manju, P. Valarmath-ie [16]	Video	Big data, Association Relation, Semantic Link Network, Fuzzy Reasoning,.	Efficient, Dynamical-ly managed	Event extracti-on may lose some correct results
17	AndreySozykin, TimofeiEpanchint-sev [17]	Image	Image processing, mapreduce, hadoop, distributed computing, MIPr framework	Scalability, fast processing	Process-ing large scale image requires more time
18	Sridhar Vemula, Christopher Crick [18]	Image	HDFS, HadoopMapReduce Hadoop Image Processing Framework	All data resides in HDFS, Highly Parallelized.	No support to multim-edia, no much support to paralleli-sm
19	Tongke Fan [19]	Image	Canopy + K-means algorithm, Hadoop, MapReduce Model	Improved speed-up ratio	Comparisons done: k-means is slower than canopy+ k-means algorit-hm.
20	M. MazharRathore, Awais , Anand, Seungmin Rho [20]	Voice over IP (VoIP) and multi-media content	Hadoop, HDFS, VoIP flow detection algorithm	Higher accuracy, efficient high speed traffic process	No much machine learning and deep learning used in order to classify network traffic
21	KyoungsooJaemin, Jongtae, Yeonwoo, Jaesoo [21]	Large multi-media data	MapReduce framework, scheduling scheme	Less time, good success ratio	Slight increase in mean execute-on time of jobs
22	David Mera, Michael and Pavel [22]	Vide-os and imag-es	Apache Hadoop, Apache Spark, Grid Computing	Improved speed-up ratio, parallel work	Compar-isons done: *Apache hadoop*-overhead costs *Apache spark*-cannot be used where dataset process-ed only once
23	Mazhar, Hojae , Awais, Anand, Gwanggil [23]	Real time City traffic vide-os	Hadoop, GPU, Apache Spark, MapReduce	Real time, efficient	CPU based MapRe-duce works slower than GPU based implem-entation.

The above table summarizes various existing work on multimedia processing using hadoop, the input taken, the techniques used, and their merits and demerits.

3. Conclusions

In this paper, the various existing work done on multimedia big data processing using Hadoop has been discussed. A review on Hadoop-MapReduce system along with HDFS and how multimedia processing has been done in a parallel and distributed manner using it is also examined and discussed. It provides a greater storage capacity and performance. The most prominent advantage of Hadoop is it can be expanded easily and is highly scalable. Hadoop is open source and has distributed nature, and has benefit of flexibility and economy. Although large scale data is easily handled by Hadoop due to its distributed nature, for high speed real time processing, these tools are not suitable and are not efficient.

The real-time big video data processing architecture along with intelligent analysis, and real-time multimedia data processing framework such as Apache Spark developed in recent days is highly beneficial to handle massive amount of continuously generating video data, especially from CCTV Cameras in Retail stores, banking location etc. But for utilizing the capacity of big data architecture properly, we must explore more and more in the future, especially about the real-time processing methods, algorithms and so on. Adding on, for high level extraction and making use of complex abstractions as data representations, Deep Learning can be coordinated and included, by learning through hierarchical manner.

References

[1] U. S. N. Raju, N. K. Varma, H. Pariveda and K. A. Reddy, "Query Object Detection in Big Video Data on Hadoop Framework," 2015 IEEE International Conference on Multimedia Big Data, Beijing, 2015, pp. 284-285.

[2] M. Husain, Meena S M, A. K. Sabarad, H. Hebballi, S. M. Nagaralli and S. Shetty, "Counting occurrences of textual words in lecture video frames using Apache Hadoop Framework," 2015 IEEE International Advance Computing Conference (IACC), Banglore, 2015, pp. 1144-1147.

[3] D. Singh, C. Vishnu and C. K. Mohan, "Visual Big Data Analytics for Traffic Monitoring in Smart City," 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, 2016, pp. 886-891.

[4] E. Toropov, L. Gui, S. Zhang, S. Kottur and J. M. F. Moura, "Traffic flow from a low frame rate city camera," 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, 2015, pp. 3802-3806.

[5] B. C. Sunny, Ramesh R, A. Varghese and V. Vazhayil, "Map-Reduce based framework for instrument detection in large-scale surgical videos," 2015 International Conference on Control Communication & Computing India (ICCC), Trivandrum, 2015, pp. 606-611.

[6] Jose Herrera, German Molto, "Detecting Events in Streaming Multimedia with Big Data Techniques", 2016 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing.

[7] G. M. Alves and P. E. Cruvinel, "Big Data Environment for Agricultural Soil Analysis from CT Digital Images," 2016 IEEE Tenth International Conference on Semantic Computing (ICSC), Laguna Hills, CA, 2016, pp. 429-4.

[8] S. Sarraf and M. Ostadhashem, "Big data application in functional magnetic resonance imaging using apache spark," 2016 Future Technologies Conference (FTC), San Francisco, CA, 2016, pp. 281-284.

[9] R. Rajak, D. Raveendran, M. C. Bh and S. S. Medasani, "High Resolution Satellite Image Processing Using Hadoop Framework," 2015 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), Bangalore, 2015, pp. 16-21.

[10] H. Xu, M. Fang, L. Li, Y. Tian and Y. Li, "The value of data mining for surveillance video in big data era," 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(, Beijing, 2017, pp. 202-206.

[11] B. Venkitesh, P. K. R. K and M. G. Chandra, "Sequential Multilinear Subspace Based Event Detection in Large Video Data Sequences," 2015 IEEE 22nd International Conference on High Performance Computing Workshops, Bangalore, 2015, pp. 48-51.

[12] S. Arsh, A. Bhatt and P. Kumar, "Distributed image processing using Hadoop and HIPI," 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, 2016, pp. 2673-2676.

[13] Fernández, R. Casado and R. Usamentiaga, "A Real-Time Big Data Architecture for Glasses Detection Using Computer Vision Techniques," 2015 3rd International Conference on Future Internet of Things and Cloud, Rome, 2015, pp. 591-596..

[14] X. Zhao, H. Ma, H. Zhang, Y. Tang and Y. Kou, "HVPI: Extending Hadoop to Support Video Analytic Applications," 2015 IEEE 8th International Conference on Cloud Computing, New York City, NY, 2015, pp. 789-796.

[15] C. Ryu, D. Lee, M. Jang, C. Kim and E. Seo, "Extensible Video Processing Framework in Apache Hadoop," 2013 IEEE 5th International Conference on Cloud Computing Technology and Science, Bristol, 2013, pp. 305-310.

[16] Manju and P. Valarmathie, "Organizing multimedia big data using semantic based video content extraction technique," 2015 International Conference on Soft-Computing and Networks Security (ICSNS), Coimbatore, 2015, pp. 1-4.

[17] Sozykin and T. Epanchintsev, "MIPr - a framework for distributed image processing using Hadoop," 2015 9th International Conference on Application of Information and Communication Technologies (AICT), Rostov on Don, 2015, pp. 35-39.

[18] S. Vemula and C. Crick, "Hadoop Image Processing Framework," 2015 IEEE International Congress on Big Data, New York, NY, 2015, pp. 506-513.

[19] Tongke Fan, "Research and implementation of user clustering based on MapReduce in multimedia big data", Multimedia Tools and Applications, Springer, 2017

[20] M. MazharRathore, Awais , Anand, Seungmin Rho, "Exploiting encrypted and tunneled multimedia calls in high-speed big data environment", Multimedia Tools and Applications, 2018, Volume 77, Number 4, Page 4959,Springer.

[21] Kyoungsoo, Jaemin, Jongtae, Yeonwoo, Jaesoo, "An efficient MapReduce scheduling scheme for processing large multimedia data", Multimedia Tools and Applications, 2017, Volume 76, Number 16, Page 17273 , Springer.

[22] [22] David Mera, Michael and Pavel, "Speeding up the multimedia featureextraction: a comparative study on the big data approach", Multimedia Tools and Applications, 2017, Volume 76, Number 5, Page 7497 , Springer.

[23] [23]M. MazharRathore, Hojae Son, Awais Ahmad, Anand Paul, GwanggilJeon, "Real-Time Big Data Stream Processing Using GPU with Spark Over Hadoop Ecosystem", International Journal of Parallel Programming, 2017, Page 1,Springer.