EXTREME COMPRESSION FOR NOVELTY DETECTION IN BANDWIDTH CONSTRAINED ENVIRONMENTS

By

NAGESWAR CHATURVEDULA

Bachelor of Technology

Jawaharlal Nehru Technological University

Hyderabad, Telangana, India

2012

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
MAY, 2015

EXTREME COMPRESSION FOR NOVELTY DETECTION IN BANDWIDTH CONSTRAINED ENVIRONMENTS

Thesis Approved:

Dr. Christopher Crick

Thesis Adviser

Dr. Johnson Thomas

Dr. David Cline

Name: NAGESWAR CHATURVEDULA

Date of Degree: MAY, 2015

Title of Study: EXTREME COMPRESSION FOR NOVELTY DETECTION IN BANDWIDTH CONSTRAINED ENVIRONMENTS

Major Field: COMPUTER SCIENCE

Many applications have been launched based on images. More than two hundred thousand images are uploaded to Facebook alone every minute. Accordingly, memory and bandwidth requirements for uploading images keep increasing. In this digital era, image storage and transmission of images over a channel have become more frequent. In addition, many environments exist that have limited memory and low bandwidth capacity. Applications like the Mars Curiosity Rover have extreme constraints on their memory and bandwidth. However, applications of this sort require a great deal of image processing and transmission. The problem is to look for novelty in unmanned places. To detect novelty, we compare all the images with each other that are captured by the machine. Comparison between extremely large images, such as those taken by scientific instruments on remote planets, takes a great deal of processing time, and transmitting those images back to Earth-based facilities is an even greater challenge. To avoid this problem, we propose a solution of extreme compression that forms a signature, solving both storage and transmission problems. These signatures are further used to detect novelty which yields almost same comparison accuracy as before compression. The whole idea reduces the similar data transfers by concentrating on novel data and also on available bandwidth.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

Image processing has been increasing rapidly after introducing image applications like Instagram, Twitter, Facebook, Picasa etc. Storage for these huge numbers of images has also been increasing accordingly. Every second millions of images of different sizes pass through networks. High definition images and videos are a priority these days. High definition images are assumed to occupy large amounts of space in a memory. More importantly, in case of image similarity testing, high definition images tend to use large amounts of time for comparison. In environments like space missions, robots navigate and learn under extreme constraints of memory and bandwidth for data transfer. These systems deal with multiple images to navigate along a path, to detect some novelty or to send the images when there is a climatic change etc. An example of one such space mission is the Curiosity Rover. This kind of robotic navigation system uses images to navigate and also to find novelty in an unmanned environment. Due to the problem of limited memory, however, robotic machines must assess the value of their old data to determine what to save and what to ignore. Hence, we introduce Image compression in research work to improve the performance of the system without losing data.

1.1 Image Signature

The objective of "image signatures" in this research is to reduce the irrelevance and redundancy of the image data in order to be able to store or transmit data in an efficient form. The best image quality at a given bit-rate (or compression rate) is the main goal of most image compression algorithms; however, there are other important properties of image compression. The quality of a compression method often is measured by the peak signal-to-noise ratio. This measures the amount of noise introduced through a lossy compression of the image; however, the subjective judgment of the viewer also is regarded as an important measure, perhaps the most important. In this work, we aim to generate a signature and hence we compress images extremely. We don't look for image clarity or definition to be visible for human in an understandable format. But we still have a minimal form of an image and using that minimal form we process the image using different methods for similarity measures. We have seen algorithms like OASIS [11] to find similar images among many. We use the same algorithm with a very small optimization and extract similarity data. In this case, we do not consider similarity, but rather non-similar data, with the goal of performing novelty detection.

1.2 Image Comparison

Image comparison is done for several measures such as image similarity, Image equivalence etc. One of the basic approaches to compare two images is pixel by pixel for both of the images. But it takes a lot of time to give the result for high resolution images. The main purpose of image compression is to make comparisons faster in any technique. As far as Image similarity is concerned, comparisons among image signatures yield faster results. In the idea, a new image that is ready for comparison is processed with signature and compared with every other image in memory.

1.3 Bandwidth Management

There are many low bandwidth channels in this world which are experiencing congestion. To avoid this congestion, we need to manage the data that has to be sent over the network. This research work basically focuses on robotic vehicles which are trying to work efficiently in unmanned environments. If the robots are left alone in such environments, they have to maintain some contact with the operators using channels and to solve the purpose of its presence there, they need to send the data in regular time intervals to the operators. If the robot is not associated with any kind of machine learning technique, the operator should expect lots of unnecessary data being transmitted over the available bandwidth channels. So, the concept of novel data detection is associated with image signatures to process and send over the network.

Section III and Section IV give us a clear design and experimental analysis about the current idea.

CHAPTER 2

Related Work

The present research is proposed to movable robotic environments. Fields such as Space explorations, automated navigation systems can be fitted with this machine learning approaches to make an advancement in Artificial Intelligence and Technology. The following are the some of the research works that are in progress and that have made progress in recent times to the fields that this research work is into.

Fu-Tai An et al. [6] has discussed about dynamic bandwidth allocation in Ethernet Based Passive Optical Networks (E-PON). The idea discussed in the research work is purely for optimizing network efficiency in available bandwidth.The authors have proposed a hybrid slot-size/rate media access protocol (MAC) to ensure that the data with highest priority will have short delays and also packet delay variation as minimum. Using this protocol, the elastic buffer size which is used to mitigate delay variations and delays in the network can be reduced. This protocol is capable of assigning very large sizes for the slots when there is low priority traffic. Eventually, the results shown in the work proves that there is a very high raise in the network throughput when offered load is high. A queue length is used as a weighting factor just to guarantee fairness to low priority traffic. The results were quite convincing about the packet delays and delay variation parameters when compared to present conventional slot-size. These kind of concepts are highly useful in low bandwidth environments which also helps in prioritizing the data to be sent.

Suet-Peng Yong et al. [9] has come up with the similar concept of finding the novel data. The only difference between the current idea and this research work is that detection algorithms are conducted on videos in Youtube. The authors have taken about three videos which show a wildlife scenes such as tigers and lions chasing zebras and buffaloes. The research work mentioned that it is quite important for a machine to detect novel data in the whole video and extract it to show what exactly is the video about. The experiments involved different resolution videos and thousands of frames in each of them. The program could detect a bunch of frames such as a tiger chasing , a lion attacking the buffalo , a zebra defending an attack from a tiger etc. Prior to the experiments, the program was trained with objects like grass, lion,sky, zebra etc. The program ultimately generated a single class file which helped itself detecting the key frames in which the objects were involved. This is similar to what a video catalog like youtube shows when entered for a query to search. Video thumbnails are the best example for this.

Chia-Yuan et al. [5] discusses about human fall detection through multiple cameras from different corners of the room. This works on the novel detection of human fall in emergency situations. The main purpose of this work is to visualize aged people activity at home when they are left unattended. This can give us home monitoring just by live video feed to be watched from anywhere.Similarly,Vineet S.Chaoji et al. [12] also describes about novelty detection by masking in video frames in a parking lot surveillance.The system detects a car entering into a parking lot, a person walking near the lot and examines them as a novel event.The same data is used by the system itself to get trained. Later, similar events of people walking, cars moving are not recorded as novel data.

NASA [3] has successfully tested a data transmission using laser beams. The laser system is called Optical Payload for Lasercomm Science (OPALS) which was shipped by Space X agency to the International Space Station (ISS). When the ISS enters the horizon, it maintains contact with ground stations for approximately 100 seconds. In this time interval, a video of about 1 Gigabyte was transferred through laser beam for about 148 seconds giving us the transfer speed of 6 megabytes per second. NASA claimed that the same video would have taken about 10 minutes when transmitted through present bandwidth channel. But we have a very high time constraint for this communication which is a considerable factor.

Jonathan J.Hull et al. [7] proposed a detection mechanism between two image based documents. Both documents are divided into fixed grids and fixed grids are imposed on the document images. Pass codes in the compressed data were used as features A feature vector is determined from pass codes inside the grid. These featured vectors are used to compare a bunch of documents and obtain similar documents to input images.

CHAPTER 3

Background

On July 6,1997 , Mars Pathfinder's Sojourner Rover rolled onto Mars' surface. It was the first robotic rover to land and to Mars. Every rover is created to find and get information from locations where a human is not present. A perfect example is Mars Curiosity Rover. Till date it is able to dig into mars' surfaces , record sunsets , examine the temperatures etc. Communications have been quite difficult between operators on earth and the rover. It takes help of orbiters of mars to transfer any bulk of data which is for very limited time in a day. We can face similar environments on earth too. Desert areas, Forest areas etc., can be a couple of examples for such environments. We cannot expect these type of environments with high speed bandwidths. Also, we should take care of how a rover works in unmanned environments. Consider a situation of a rover taking images and sending them to its operator. The purpose of this could be knowing about a location without human presence. This purpose can be solved by taking images and sending them to the operators. Now, we have bandwidth problems, less connectivity and duplication of data. This duplication occurs if we just give the instruction to the machine to take pictures and send. If a machine is taught to detect which a novel data and which is not, we can avoid duplicate data transfers which ultimately helps in efficient utilization of bandwidth.

Image comparisons can be used in many environments. Facebook uses image comparisons to detect faces in images. Google searches for images in millions of web pages when asked for a query. [11] In the year 2010, a research work named "Online Algorithm for Scalable Image Similarity" (OASIS) proposed an idea about image similarity and their rankings to consider what images are similar and what is not. In a web search dialog box, as we enter a query as "dogs", the machine takes a reference dog image and compares using all the images in the prioritized web pages. More importantly, the comparison was taking very less time to identify the similar images to the query. The OASIS approach was basically designed to perform better in large datasets. The experiments conducted in that approach involved about 30,000 images. Later, to check the algorithm's efficiency, the authors chose a dataset for similarity measures that contained 2.3 million images which lasted only 2 days for processing. This whole idea has been upgraded comfortably to overcome some of the drawbacks. A detailed explanation for this idea is elaborated in technical approach section.

Histograms [10] are the simplest form of representing an image. Each and every image editor at-least has a histogram box these days. Applications like Adobe Photoshop, Picasa, Auto-Desk Maya etc., are capable constructing operations directly on histograms of image directly.It occupies very less memory than other compression method/ signature in recent times.Histograms can also be optimized with other operations like normalization, equalization etc.,

Histogram normalization [1] is generally used on the images where there are lot of distractions. It is usually built based on the maximum value of the array of pixel intensities. Normalized histogram does not change its shape compared to original histogram of an image. All the values of the histogram are brought to the required range by performing basic mathematical calculations.For example, when we need a range 50-130 to changed to 0 to 255, we subtract 50 from minimum and maximum value and multiply with 255/130 for each value in the array.

Histogram equalization [4] is used to increase global contrast of numerous images. This method more efficiently works when the usable data is in close contrast values. It is a straight forward technique and an operator for inverse function in mathematics. The effects produced after equalizing an image can go unusual but this method is highly efficient in satellite , x-ray , thermal imageries etc. Also, images with very low color depth cannot have this method because it may give undesirable results. For example, it might produce better results when applied on 16 - bit color space than on 8-bit color space. This method basically distributes higher intensity values through out the image when applied on higher color images. So this equalization can actually bring down the whole image into certain intensity which solves the problem when a couple of images differ with small contrast values.

CHAPTER 4

Design and Methodology

In this we describe briefly about the sequence of generating signatures, comparing the signatures to detect novelty and also give information about the bandwidth efficiency after novelty detected. we consider some of the drawbacks in some popular platforms and discuss the idea including solutions. We take a real time example to explain about the constrained environments wherever this particular idea is applicable.

Lot of applications these days involve huge data transfers, memory, detection mechanisms etc. All those applications can be seen in rovers, drones and robots. Drones from Washington D.C alone are capable of flying for several hours just to aim and kill a terrorist some where in Afghanistan. This kind of operations general happen only by live video streaming from drone's camera. Another example that suits the best for the idea is Mars Curiosity Rover. This rover communicates with Mars orbiters for 8 minutes of a day. The communication channel has a bandwidth as low as 32 kilobits per second from the rover to earth operating stations. This rover has a memory of about 8 Gigabytes. In this environment, we have a memory constraint and a bandwidth constraint. Also, if we check the raw images taken by the rover, on an average, 60% of the data look similar to human vision. This brings up the problem of similar data transfer in a low bandwidth environment where a machine cannot even decide what to send and what not to. So, to teach the machine for the novel data, we have come up with a unique idea where any intelligent vehicle such as rover, drone etc., can take images and store their signatures which help the machine to decide the novel data and send using the bandwidth efficiently.

The first technique that we embed the design with is scaling down the images. Image signature actually creates an impression that a human cannot visualize it in an understandable view. But here , we scale down the image to an extremely low percentage which helps in consuming very less memory and also we can have a better view of images than some signature techniques. For the sake of experiment and analysis, we take five different signature sizes to test which signature size is capable of finding better novel data. On each signature size generated, we use an optimized algorithm of OASIS[11] where the algorithm is capable of finding similarity metrics when a bunch of images are given as an input. This algorithm is named as Pixel Sampling in this research work. The output of Pixel Sampling after taking each and every scaled images as input decides the efficiency of each signature size. Fig.4.2 shows an outline of what sizes are images are scaled down to. The efficiency can be seen clearly in experimental analysis section for a couple of datasets.

The Second technique we add to the design is generating a histogram of an image. Histogram occupies very less space than any other heavy techniques or algorithms to form an image signature. The signature size takes maximum of 2 Kilobytes of memory for any image. Shortly, it is graph plot to show intensities of pixels from 0 to 255 on the X-axis and count of those intensities of pixels on Y-axis. More detailed explanation can be found in technical approach section.

Histograms can be enhanced by normalizing them and also by equalizing them to bring all the values to certain level of almost equal intensities. These two methods are also used on histograms to test the images that differ with certain intensity type such as brightness , darkness etc. Apart from generating five results for each signature size by image de-resolution, Histogram has its own comparison result for every couple of images and besides normalization and equalization of histograms have their results respectively. The whole outline of the design can be seen in Fig4.1, Fig.4.2 and Fig.4.3

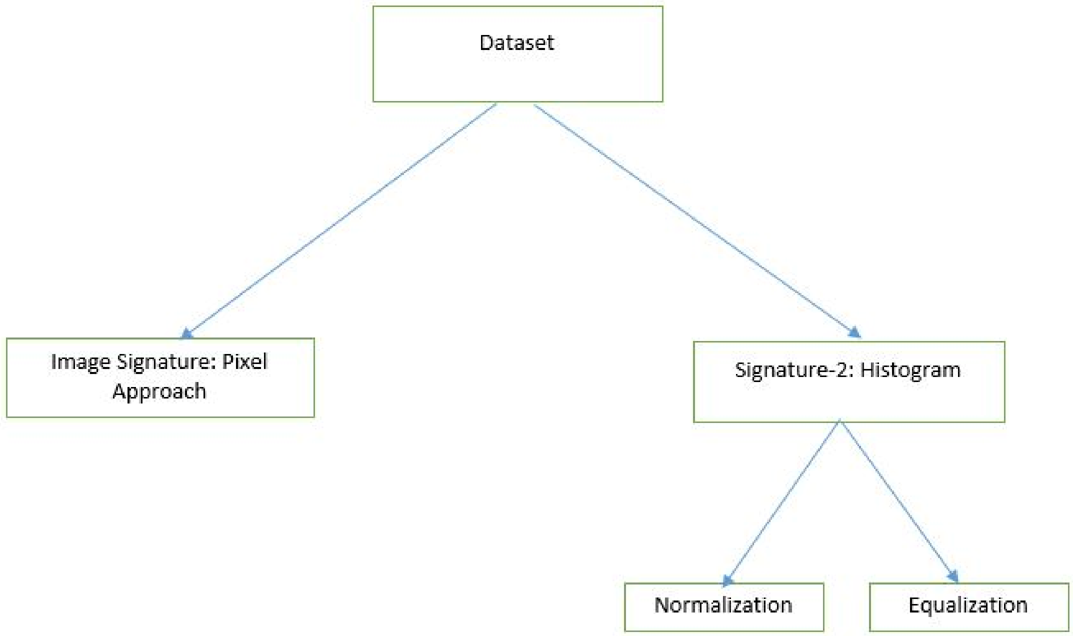

Figure 4.1: Image Signature Generation Outline

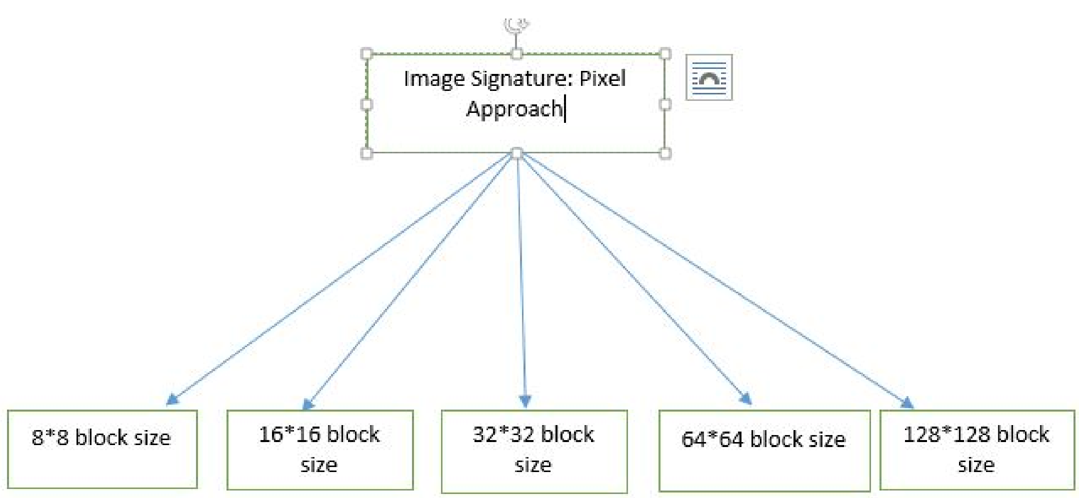

Figure 4.2: Images to scale down to test with Pixel Sampling

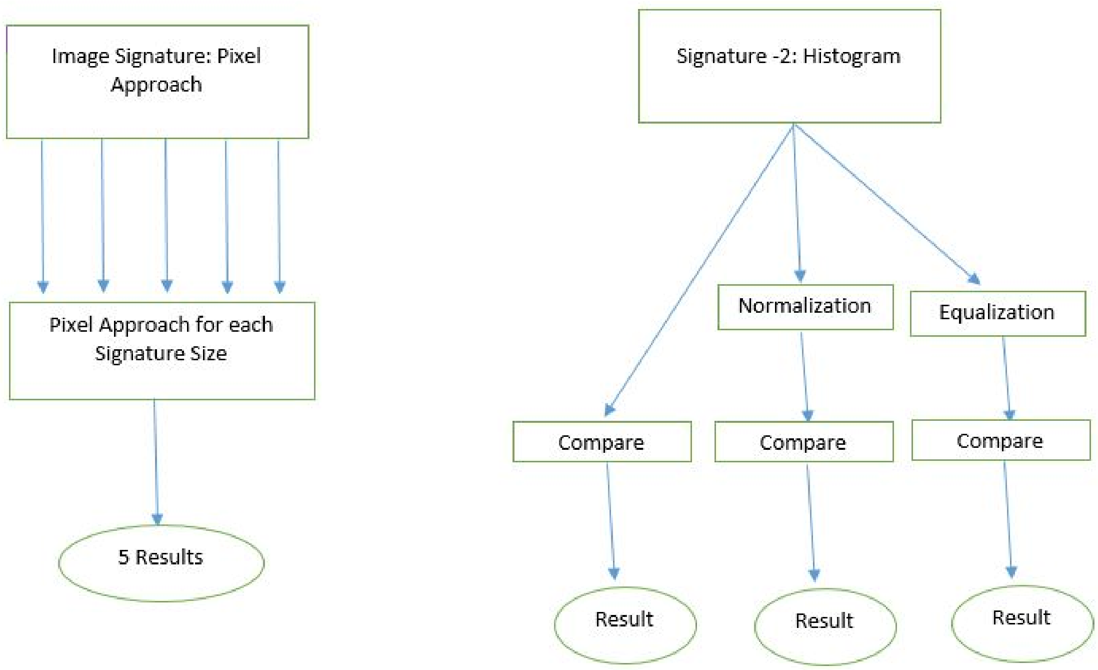

Figure 4.3: Methods to Generate Results

CHAPTER 5

Technical Approach

Faster and Efficient computation is our priority. To achieve this, images are first compressed and then compared because the smaller are the images, the faster is the comparison. The whole idea is designed and developed in complete Java Platform.

5.1 Pixel Approach

Since comparison of each and every pixel of two images take lot of time, we follow Pixel Sampling. In this approach, we only compare a block of pixels of both images starting from left top corner of each image. We take the center pixel of the block as a point to draw a circle with some random radius inside the block. We select only a bunch of points on the circle in which the neighboring points are equidistant to each other. It is recommended to select only a few points , lets say 8,16 or 24 etc. Because, selecting too many pixel points on the circle takes more time for comparison. We perform binary test for the selected points on the circle and append the binary values to give a binary sequence. This binary sequence of both images are compared to test whether the block is exactly similar or not. For Binary test, we take threshold value which is actually the average of each block images' center pixels. If the selected value is above the threshold limit, the binary value for that particular pixel is given as '1' or if the selected pixel value is less than or equal to threshold limit, the binary value for that pixel is given as '0'. So, if we select eight values on the circle and we perform binary test on those pixel points, we get 8-bit binary sequence of first block of one image. Similarly, we perform the same test on the first block of second image.

If the first blocks of both images are similar to each other , the comparison continues with second block of pixels of both images and so on. So, instead of comparing all the pixels of one image to all the pixels of another, we are reducing the number of pixel comparison which can still give the best result estimated.

5.2 Histogram Analysis

The idea is basically a research platform for many techniques to be tested which upgrades the performance to detect novelty in images. One of the most popular and very simple technique is getting a signature of an image in the form of histogram. Any huge image can be represented in the form of histogram which is actually signature. We basically convert each image into gray scale image and get the RGB value for each pixel until the control reaches the respective height and width of an image. That way we have an array of values indexing from 0 to 255.This indexing remains same for every image. It can practically be seen that comparison of count of similar pixels is more simple than comparing each and every pixel for a couple of images at a time.

5.3 Histogram Normalization

Normalization is just another form of histogram. This technique also takes an input of gray scale image. we collect each and every pixel of the image and gather the value of each pixel in red channel , value of each pixel in green channel and also value of each pixel in blue channel. All these values are counted in a three dimensional array. The whole histogram values are adjusted according to the maximum value in the three dimensional array. It can be explained more like pulling a normal histogram from both ends which reduces all the other values in the histogram. This technique is really helpful to maintain consistency among dynamic set of data for avoiding mental distraction.

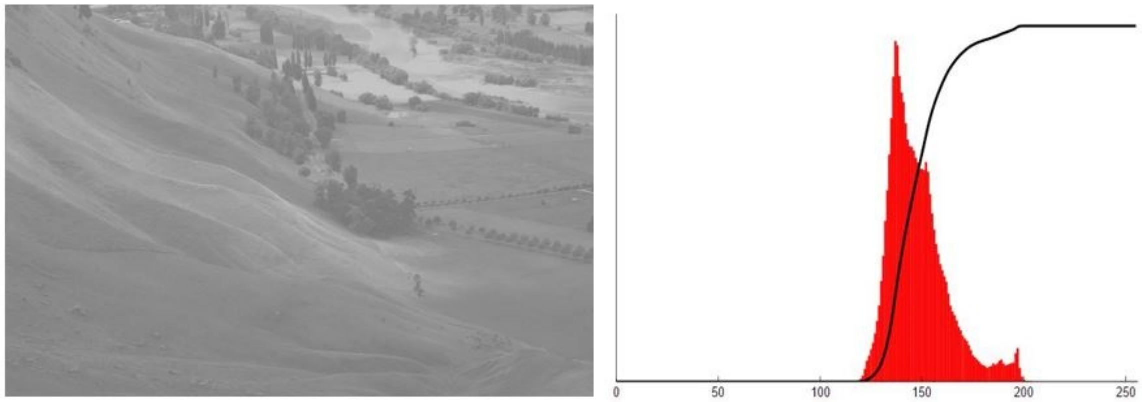

Figure 5.1: Non-equalized Image and Its Histogram

5.4 Histogram Equalization

Histogram equalization is another process to enhance a histogram. This process distributes the high intensity values among other intensities equally. This method produces undesirable results when applied on low color images. For example, it works good on 16-bit gray scale images than on 8-bit gray scale images. The color distribution in 8-bit gray scale images still lowers the overall intensity which affects the image so much. The main difference between histogram normalization and histogram equalization is that histogram equalization changes the shape of histogram where as histogram normalization doesn't. This method uses cumulative distribution function (cdf) on histogram of an image and then cdf is scaled and normalized to the range 0-255. These scaled values then substituted in an equalized formula including the extremities. Fig.5.1 and Fig. 5.2 show the exact difference between a non-equalized image and an equalized image.

This idea can be implemented in different platforms such as Communications in Space Networks, Remotely located machines, Robots that navigate with learning, Military missions for enemy activities and search operations.

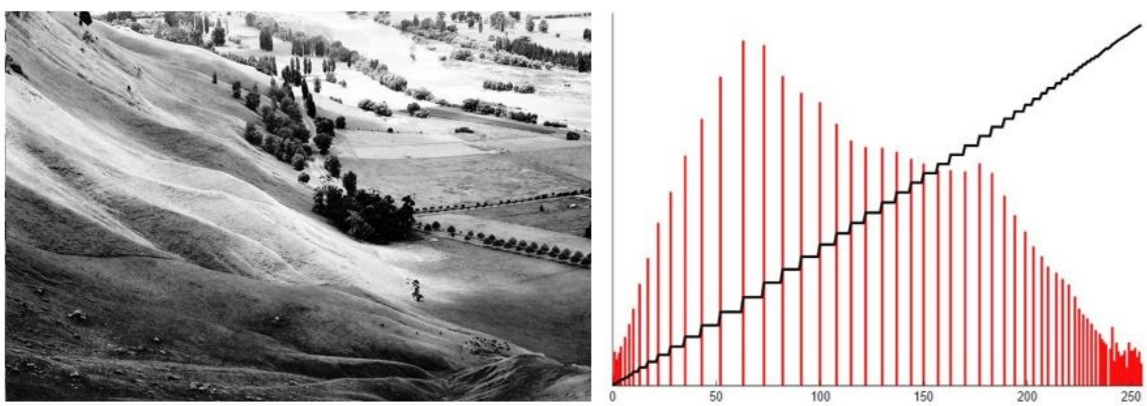

Figure 5.2: Equalized Image and Its Histogram

CHAPTER 6

Experimental Results and Analysis

We have performed experimental tests using the proposed idea on a couple of datasets. First dataset collected from Flickr image gallery and the other one taken by NASA's Mars Curiosity Rover. The second dataset is quite important for us because the whole idea is proposed for the environments similar to Curiosity Rover. Results and their analysis are described in the subsections below.

6.1 Experimental Setup

Experiments were conducted on 252 node Linux based Cluster. This cluster consists of 252 standard compute nodes, each with dual Intel Xeon E5-2620 Sandy Bridge hex core 2.0 GHz CPUs, with 32 GB of 1333 MHz RAM. The cluster also contains two fat nodes each with 256 GB RAM and an NVIDIA Tesla C2075 card. The aggregate peak speed is 48.8 TFLOPs, with 3048 cores, 8576 GB of RAM.

For the whole experimental analysis we have results of two total datasets. One dataset is collected from Flickr high definition images and the other dataset has been collected from NASA's Mars Jet Propulsion Laboratory [2] website. The website could provide very recent image dataset that has been collected directly from the Mars Curiosity Rover.Fig.6.2, Fig.6.4, Fig.6.6 are results that are generated from the Mars image dataset.

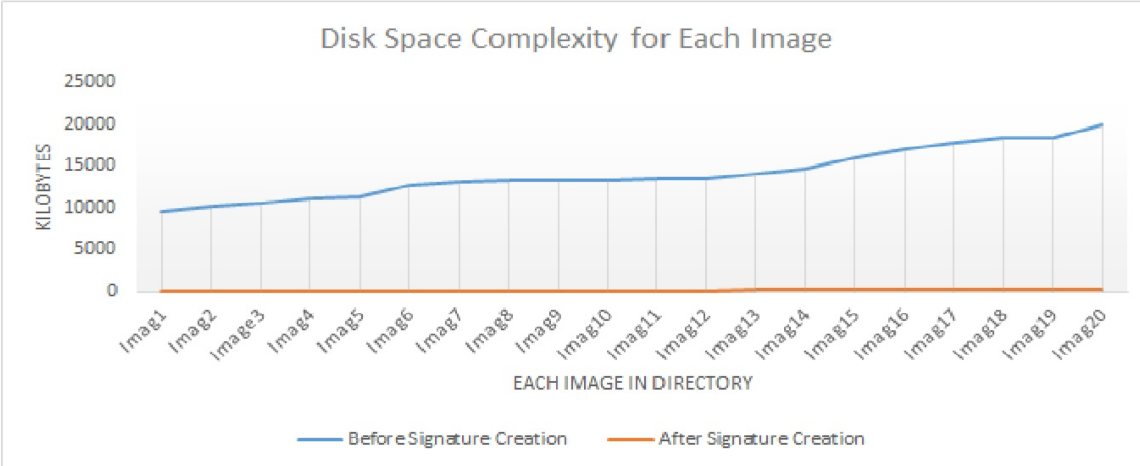

Figure 6.1: Disk Space Complexity for Each Image in FLICKR Dataset

6.2 Results

6.2.1 Space Complexity

Image signatures occupy very less disk space. We have collected an image dataset from Flickr with extremely high resolution(at least 4K pixels each). The dataset contained about 20 images each of at least 10 Mega Bytes of Memory. Each of these images are then extremely scaled down to create signatures for detection of novelty.Each high resolution image is divided into blocks and using the Pixel Approach algorithm, we have collected the center pixel intensity and respected pixel intensities on the circumference of the circle leaving a radius of 2 from the center pixel . Fig.?? shows the exact comparison of memory occupancy before the signature creation and after the signature creation. Y-axis in the chart shows the memory occupancy in Kilo Bytes and X-axis shows two different situations.

Fig.6.1 shows the signature size and Original image size comparison for each image individually. The minimum size of an original image is 10123 kilo bytes where as the maximum size of an image in dataset is 20362 Kilo Bytes. Similarly, the minimum signature size of the dataset image is as low as 20 kilo bytes and maximum signature size is 215 kilo bytes.

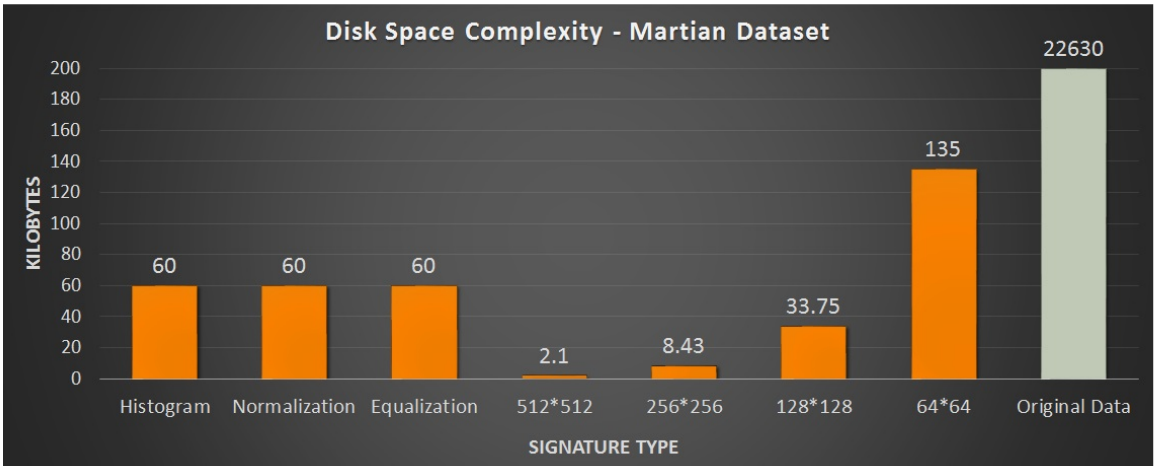

Figure 6.2: Disk Space Complexity in MARTIAN Dataset

6.2.2 Comparison Efficiency

In this idea, Comparison is done to find out the novelty in any two images. Comparisons take longer time when performed on high resolution images. When this particular system is tested for efficiency before creating signatures, the processing time for comparing just two high resolution images took more than one minute to produce a result. High processing times due to overload of huge amounts of data cannot be acceptable in small robots when sent to unmanned environments. Our idea is to make sure that the comparison efficiency remains same when tested on original images or on bunch of signatures. In this experiment, we have created different sizes of signatures for the same dataset. The dataset is scaled down to different signature sizes using 64*64 , 128*128, 256*256 and 512*512 block sizes in the Pixel Sampling. Binary comparison test is then applied on each block signature to test the comparison efficiency.

Fig.6.3 gives us a complete picture of each signature size and its comparison efficiency. As we can see, 128*128 of the block size is least efficient situation for novelty detection. But, image blocks with 64*64 of the dimensions has some details left. So, there has been an improvement in efficiency from 66% to 83%. As we scan through, the efficiency remained same for 128*128 and 64*64 of dimensions of image blocks. This proves that we can stick to 128*128 of original image block which gives us better results than less efficient situations. The result for 32*32 signature dimensions also shows that it is equally efficient with greater dimensions. So, to reduce the time and space complexity for signature creations and comparisons, we can consider scenario of image block size to 128*128.

Fig.6.4 shows the efficiency of second dataset from Mars Curiosity Rover. The experiments are conducted using different signature sizes and different signature types. Image Signatures with Pixel Sampling, histograms of images, normalizing the histograms and equalizing the histograms are simple types used for image signature generation. Using Pixel Approach, the block size of images was restricted to four sizes such as 64*64 , 128*128, 256*256 and 512*512. Besides we processed normal histogram data, normalized form of histogram and equalized form of histogram for every image to compare it with every other image. The graph (Fig.6.4) shows the percentage of efficiency on the Y-axis and signature types on X-axis. 512*512 block dimensions could not give better result compared to other sizes. we found that the dataset lost more of its definition or the minimum information required for the Pixel Sampling algorithm to generate better results when compared to 128*128 block dimensions.

6.2.3 Bandwidth Efficiency

One of the best examples for low bandwidth environments is the data communication between Mars Curiosity Rover and Earth station. The Rover can transmit up to 32 Kbits/sec and it has to rely on mars' orbiters when a bulk of data has to be sent. In this kind of environments, replicated data transfers doesn't really create a very good impression. Since, we have performed tests to find the novel data, we can directly choose not to send any data that is similar to novel data. This idea can reduce

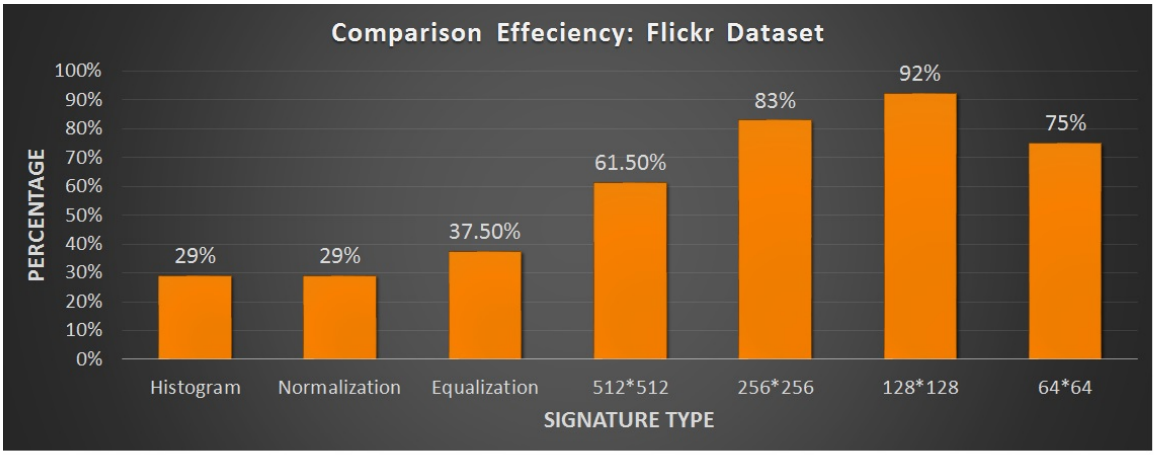

Figure 6.3: Comparison Efficiency for FLICKR Dataset

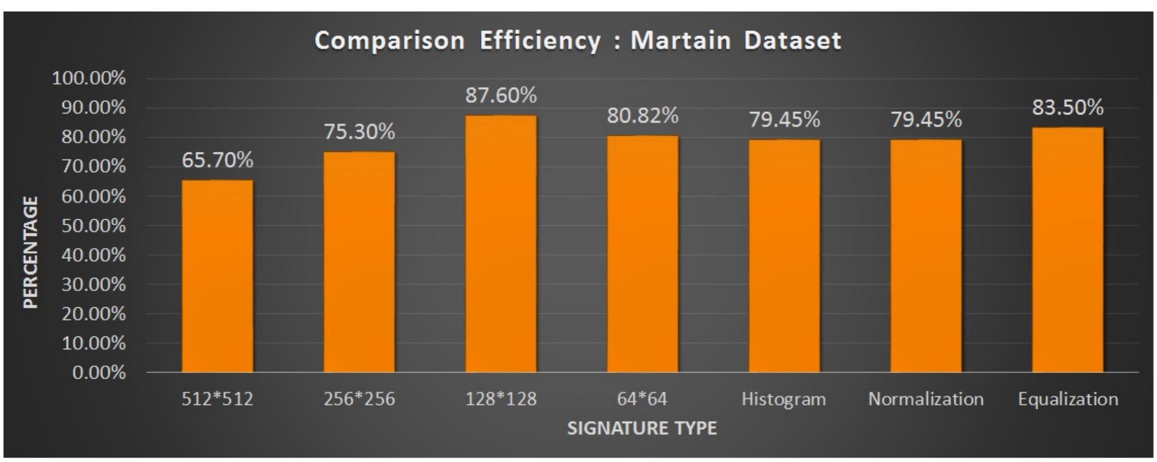

Figure 6.4: Comparison Efficiency for MARTIAN Dataset

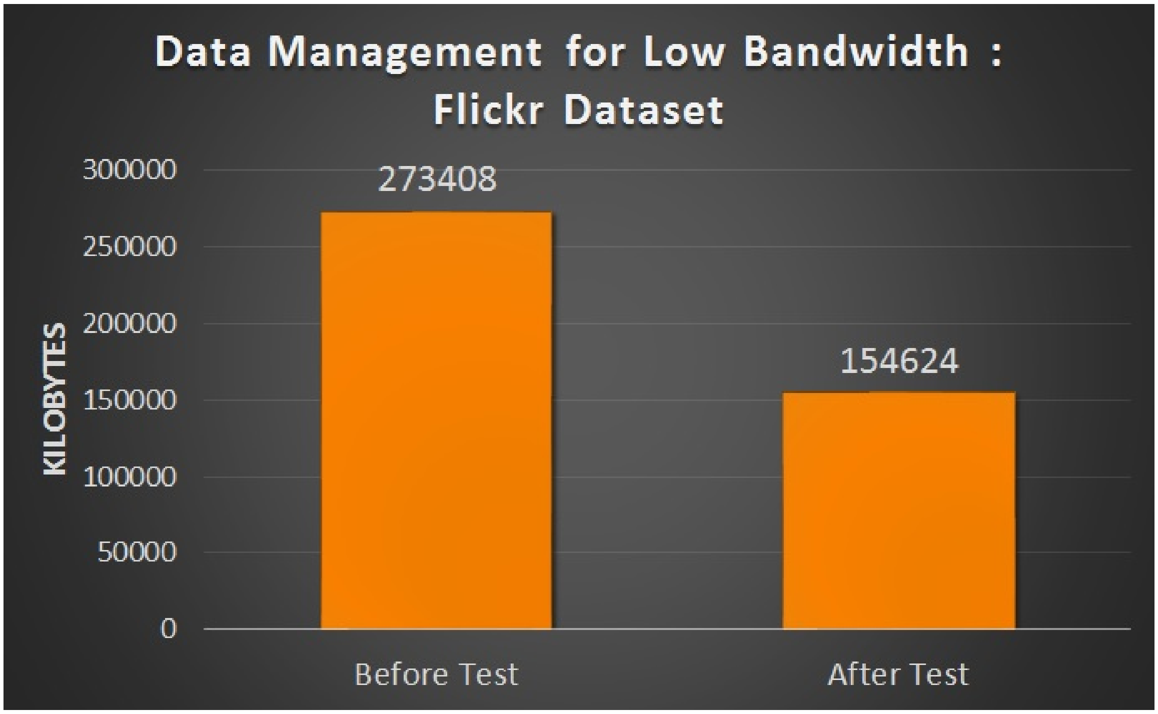

Figure 6.5: Bandwidth Efficiency for FLICKR Dataset

lot of disk space and can also make data transfers more efficient in low bandwidth environments.

Fig.6.5 shows the data transmission in Megabytes before the test and after the test. The test meaning only the novel data is collected and sent via available bandwidth and rest of the similar data is given a necessary action. In this example, this idea could save 111 Megabytes of data from being sent over the low bandwidth network. The similar data is removed from the disk by just saving their signatures in the memory and occupancy of signatures is about 45 Megabytes of memory in the disk.

Fig.6.6 shows the data transmission of Dataset-2 before the test and after the test. This Dataset contained images of Martian landscapes, all of equal dimensions taken from one of the cameras of the rover. The total dataset occupied 22630.40 Kilobytes of memory.But after the test, we could find the data that has to be transfered was reduced to 9000.96 Kilobytes. The left over data is actually novel and had 13629.44 Kilobytes of similar images which is no really useful.

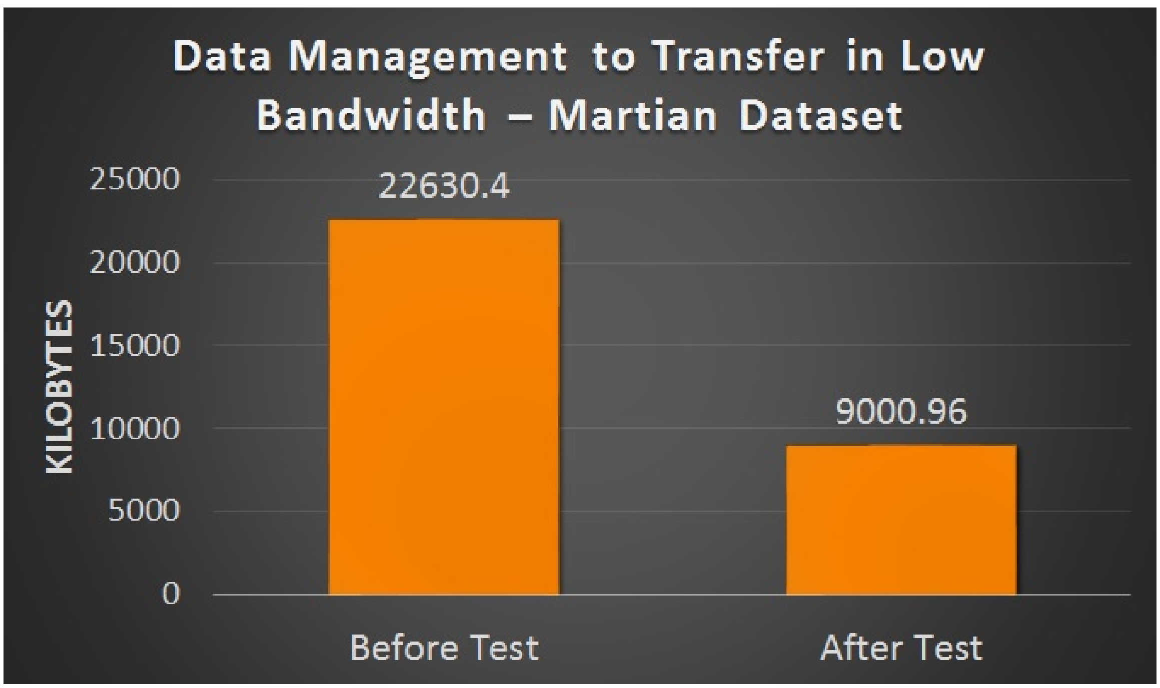

Figure 6.6: Bandwidth Efficiency for MARTIAN Dataset

CHAPTER 7

Conclusion and Future Work

In this study, we proposed an idea of extreme compression of images which forms signatures. These signatures are very much helpful to find the novelty by reducing the time and space complexity.Considering the results of main dataset(dataset-2), we found that simple technique as normalization works more efficient than any other techniques used in the experiments. But when we look for deep scanning and efficiency, Pixel Sampling works better. OASIS [11] was used to find similar images for a query and show them but here we optimized the OASIS approach to find the similar data and ignore it to find the novelty. This idea can be a platform for so many other techniques like Hough Transforms [13], SIFT algorithm [8] etc., to make the comparison results more accurate.

We can also optimize this work by using it in video platform such as surveillance cameras where we can just save a key frame of the anything that is unusual for the camera. We can use image comparisons like Pixel Sampling to see how similar the key-frames are and of course similarity can be detected just by signatures.

CHAPTER 8

Acknowledgment

We acknowledge the faculty, Cowboy - High Performance Computing Cluster (HPCC) management at Oklahoma State University for supporting the research and providing resources that have contributed the results demonstrated in this paper.

REFERENCES

[1] Acharya and Ray. Image processing: Principles and applications. In *Wiley - Interscience*, 2005.

[2] National Aeronautics and Space Administration. http://mars.jpl.nasa.gov/. 2015.

[3] Sebastian Anthony. Nasa installs space laser on the iss, uses it to transmit high-speed data back to earth. June 9, 2014.

[4] N Bassiou and C. Kotropoulos. Color image histogram equalization by absolute discounting back-off. In *Computer Vision and Image Understanding*, volume 107, pages 108–122, July- August 2007.

[5] Gang Miaou Pei Hsu Sung Chia-Yuan, Huang Shaou. A customiszed human fall detection detection system using omni-camera iimage and personal information. 2011.

[6] Kyeong Soo Kim Ian M. White Fu-Tai An, Yu-Li Hsueh and Leonid G. Kazovsky. A new dynamic bandwidth allocation protocol with quality of service in ethernetbased passive optical networks. 2014.

[7] Jonathan J.Hull. Document image similarity and equivalence detection. June 1997.

[8] Tony Lindeberg. Scale invariant feature transform. 2012.

[9] Jeremiah D. Deng Suet-Peng Yong and Martin K. Purvis. Wildlife video key-frame extraction based on novelty detection in semantic context. 2013.

[10] Ed Sutton. Histograms and the zone system. In *Illustrated Photography*.

[11] Gal Chechik Varun Sharma Uri Shalit, Samy Bengio. Large scale online learning of image similarity through ranking. 2010.

[12] Roger S. Gaborski Vishal S. Vainganka r. Vineet S.Chaoji, Ankur M.Teredesai. Venus: A system for novelty detection in video streams with learning. June 2013.

[13] Wikipedia. Hough transform. 30 March 2015.

VITA

NAGESWAR CHATURVEDULA

Candidate for the Degree of

MASTER OF SCIENCE

Thesis: EXTREME COMPRESSION FOR NOVELTY DETECTION IN BANDWIDTH CONSTRAINED ENVIRONMENTS

Major Field: COMPUTER SCIENCE

Biographical:

Education: Completed the requirements for the Masters degree with a major in Computer Science at Oklahoma State University in MAY, 2015.