E-LEARNING IMAGE COMPRESSION USING K-MEANS-BTC IN PUBLIC CLOUD

Sivakumar, R.D.[1], Dr. Ruba Soundar, K.[2]
[1]Ph.D. Research Scholar, Department of Computer Science,
Bharathiar University, Coimbatore, India
[2]Professor and Head, Department of CSE, P.S.R. Engineering College, Sivakasi, India.

Abstract

In today's environment, everything will be expressed in the image. The image play an important role in the digital world to convey the message mainly in e-learning. While using the images in e-learning, the size should be low. Hence, the image compression is the necessary one for storing and transmission in the cloud. The aim of this paper is to develop an algorithm for reducing the size without affecting the quality of the image and increase the transmission speed. In the first phase of our algorithm, the vector quantization technique in K-Means algorithm the original image is transformed into semi-compressed image. During the second phase of our algorithm the semi-compressed image is further compressed by using Block Truncation Coding method which converts the semi-compressed image into time domain. The proposed algorithm is an efficient one which acquires the less space and also the speed is increased on cloud by showing the experimental results and comparison of the image.

Keywords: E-Learning, Cloud, K-Means, BTC, Compression

1. INTRODUCTION

Image Compression is an essential procedure in the modern world where there is a requirement of storing large amount of data using as much less storage space as possible. Image compression can be achieved using various algorithms which may be divided into lossless and lossy categories. In lossless compression scheme, the reconstructed image after compression is numerically identical to the original image whereas in lossy compression techniques the compression ratio is high. Removal of redundant data is a very important part of image compression [12][23].

Data Mining technology is used to give the user an ability to extract meaningful patterns from large database. Information retrieval systems have made large quantities of textual data available. Extraction of meaningful patterns from this date is very difficult. Current tools for mining structured data are inappropriate for free text and present architecture for extracting patterns that hold across multiple documents. Data mining technology has created a new opportunity for exploiting the information from the databases. Patterns in the data, such as associations among similar items purchases, enables target marketing to focus on what things the customers are likely to purchase. Data mining, the extraction of hidden predictive information from large databases, is a powerful, new technology with great potential to help many information storehouse. Data mining tools calculate upcoming trends and activities, allowing company to make practical, knowledge-driven decisions [3].

In this paper, the algorithms and techniques for clustering and block truncation coding has been presented. The performance measurement and complexities of algorithms have also been analysed. The combination of K-Means clustering and BTC yields good result.

1.1 E-Learning

E-learning refers to the use of electronic media and Information and Communication Technologies (ICT) in education. E-learning is broadly inclusive of all forms of educational technology in learning and teaching. E-learning is inclusive of, and is broadly synonymous with

multimedia learning. E-learning can occur in or out of the classroom. It can be self-paced, asynchronous learning may be instructor-led, synchronous learning. E-learning includes numerous types of media that deliver text, audio, images, animation and streaming video and includes technology applications and processes such as audio or video tape, satellite TV, CD-ROM, and computer based learning, as well as local intranet/extranet and web based learning. ICT, whether free-standing or based on either local networks or the Internet in networked learning underlay many e-learning processes [2].

1.2 Cloud Computing

A cloud can be defined as a place for users to create or store files. The cloud computing is a collection of server delivering resources that can be accessed remotely via the Internet in real-time. These servers are housed in a bunker like structure called a Data Centre and data, software applications are not housed on the computer; they are on a service's cloud of web servers (often virtual servers) usually accessed by users through the Internet using a browser like chrome or IE [8].

Deployment Models

The NIST definition for the four deployment models are as follows:

- ***Public Cloud*** – The public cloud infrastructure is available for public use alternatively for a large industry group and is owned by an organization selling cloud services.
- ***Private Cloud*** – The private cloud infrastructure is operated for the exclusive use of an organization. It may be either on or off premises.
- ***Hybrid Cloud*** – A hybrid cloud combines multiple clouds where those clouds retain their unique identities but are bound together as a unit.
- ***Community Cloud*** – A community cloud is the one where the cloud has been organized to serve a common function or purpose [5].

1.3Block Truncation Coding

Block truncation coding is well known technique for image compression, It (BTC) divides the original image into small sub blocks of size n x n pixels and after the division of image, it reduces the number of gray levels within each block, reduction of gray level is performed by a quantizer. Threshold reconstruction values are calculated for each block and a bitmap of the block is obtained for that values. We replace all the pixels whose values are greater than or equal (less than) to the threshold by a 1(0), in this bit map. Then for each segment (group of 1s and 0s) of the bitmap, reconstruction value is calculated [14].

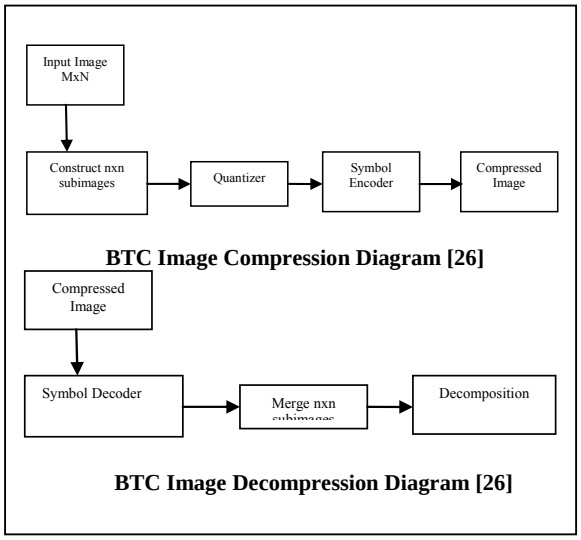


Figure1: Diagram for Block Truncation Coding

Merits of BTC

1. It is very fast encoding method.
2. BTC requires low memory space and less complexity.
3. Implementation is easy.
4. It is low transmission errors [3][26].

BTC Algorithm

Step1. Input images divide into several non overlapping block of size n x n. For example 4 x 4 or 8x8 and so on.

Step 2. Compute the two statistical values x (mean) and σ (standard deviation). Both values are calculated for each block of Image by using following equations.

A) $x=1/n \sum_{i=1}^{n} x_i$

B) $\sigma = \sqrt{1/n \sum_{i=1}^{n} (x_i - \bar{x}_i)}$

Here x_i represent the i^{th} pixel value of the image block and n is represents total number of pixel in particular block

Step3. The x and σ are termed as a quantizers of this technique.

The x is the threshold value of presenting bit plan and obtained by comparing each pixel value of image with defined threshold value.

$$\begin{cases} 1 & x_i \geq \bar{x} \\ 0 & x_i \geq \bar{x} \end{cases}$$

Here block of pixel is represents as a B "1" to represent a pixel whose gray level is greater "0" to represent a pixel whose gray level is less By this process each block reduce the bit plan.

Step 4. In decoding phase encoded block are reconstruct by replacing 1 as H (high intensity) and 0 as a L (low intensity) which are given by as follow.

$$H = \bar{x} + \sigma\sqrt{p/q}$$

$$L = \bar{x} + \sigma\sqrt{q/p}$$

Where p and q are the number of 0"s and 1"s in the compressed bit plane respectively [3] [26].

1.4Clustering

Clustering is the process of grouping similar objects into groups. This is extensively used in image compression. The clustering center of each partial area is acting as a representative of the corresponding type. It is the basis of pattern recognition [10].

1.4.1 K-Means Clustering

The K-Means Clustering algorithm is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. [4]. The K-means algorithm is a extensively used VQ technique known for its efficiency and speed. This is an iterative clustering algorithm that generates a codebook which is a set of cluster centroids from the training data using an appropriate distance function suitable for the given application [10].

K-Means Clustering Algorithm

The various steps of the standard K-Means clustering algorithm is as follows :

1. The number of clusters is first initialized and accordingly the initial cluster centers are randomly selected.
2. A new partition is then generated by assigning each data to the cluster that has the closest centroid.
3. When all objects have been assigned, the positions of the K centroids are recalculated.
4.Steps 2 and 3 are repeated until the centroids no longer move any cluster [4].

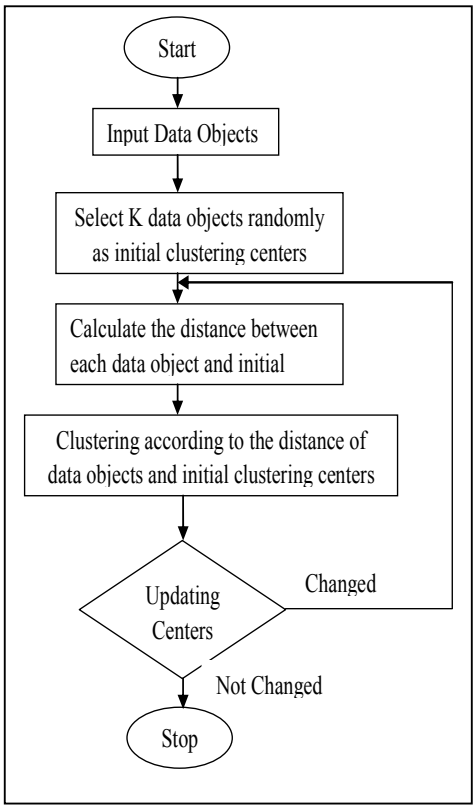


Figure2: Block Diagram for K-Means Clustering Algorithm

The paper is structured as follows: Section 2 describes the outline of the related works, Section 3 point out the proposed method, Section 4 presents the image quality measure which includes PSNR and Compression Ratio, Section 5 provides the experimental results, comparison and discussion of the proposed method and Section 6 concludes with the future enhancements.

2. RELATED WORKS

Somasundaram,K. and Rani,M.M.S., have published a research article [15] entitled "*Novel K-Means Algorithm for Compressing Images*" . In this paper, the author proposed the possibility of application of statistical parameters for choosing the initial seeds for K-means algorithm. The selection of initial seeds depends on the statistical features of input data set [22].

The novelty in their approach is the judicious selection of initial seeds based on variance, mean, median and mode parameters [19][22]. Considering mode value of each dimension of the data adds uniqueness.

Malekar,N.C. and Sedamkar,R.R. have published a research article [7] entitled "*Novel K-Means Clustering Approach for Compressing Hyperspectral Image*". In this paper, the author pointed out to compress the hyperspectral image by using k-means clustering. The compression ratio is improved by using k-means clustering approach by clustering pixels into classes based on pixel spectral similarity with other class members [18][27].

Rani,M.M.S. and Chitra,P., have published a research article [11] entitled "*Region of Interest Based Compression of Medical Images Using Vector Quantization*". In this paper, the author proposed method mainly focuses on compressing medical images with different codebook sizes for region of interest and non region of interest using vector quantization method. It achieves high compression ratio without compromising the quality of reconstructed ROI image

Kumar,S. and Nancy have published a research article [3] entitled *"K-Mean Evaluation in Weka Tool and Modifying it using Standard Score Method"*. In this paper, author proposed the modified approach of K-Means clustering and algorithm has been designed. The entire data will be normalized using standard score method which is also called z score and then cluster will be formed using Euclidean distance. The fast clustering process will reduce the system resources and provides the efficient technique to generate the clusters.

Jain,A. and Chawla, S., have published [5] a research article entitled "*E-learning in Cloud"*. In this paper, Cloud based e-learning will enable people to build their learning around their specific needs with the cloud allowing relevant, tailored content to be created for the users without any dependence on IT to update delivery platforms accordingly.

Tambe, M., *et.al.,* have published [12] a research article entitled "*An E-learning System with Image Compression"*. In this paper, focused to provide an e-learning platform, mainly deals with the provision of storing the hand written notes in digitized form and also storing the images, Diagrams, Audio-video lectures etc., on Cloud. The data will be stored by reducing its size which would be beneficial as it acquires less space on cloud, thus data manipulation speed will be increased. The paper is achieved to use the various algorithms like Thresholding, Grayscaling.

Mohammed, D. and Abou-Chadi, F., proposed [9] a *"Image Compression using Block Truncation Coding"*. In this paper, two algorithms were selected namely, the original Block Truncation Coding (BTC) and Absolute Moment Block Truncation Coding (AMBTC) and a comparative study was performed [21]. Both of two techniques rely on applying divided image into non overlapping blocks. They differ in the way of selecting the quantization level in order to remove redundancy [20].

Gupta,P., Bansal, V., and Purohit, G.N., have published [6] a research article entitled *"Block Truncation Coding for Image Compression Technique"*. In this paper investigated of gray-scale images by using Block Truncation Coding Technique. The technique is to achieved better quality image reproduction. It is comparative study of BTC technique and proposed a method that is based on Block Truncation algorithm for image compression [26].

Vimala, S., Uma, P. and Abidha,B., have published [17] a research article entitled "*Improved Adaptive Block Truncation Coding for Image Compression"*. In this paper proposed a method called the Improved Adaptive Block Truncation Coding based on Adaptive Block Truncation Coding. The feature of inter-pixel redundancy is exploited to reduce the bit-rate further by retaining the quality of the reconstructed images [24].

Shashikumar, S., Parakale, A. and Mahavir, B.M., have published [13] a research article entitled *"Image Compression using Absolute Moment Block Truncation Coding"*, In this paper, Absolute Block Truncation Coding technique is used for image compression. AMBTC algorithm is a lossy fixed length compression method that uses a Q level quantizer to quantize a given region of the image. This technique rely on applied divided image into non overlapping blocks [20].

Almrabet, M.M., Zerek, A.R., Chaoui,A. and Akash, A.A., have published [1] a research article entitled *"Image Compression using*

Block Truncation Coding". In this paper, the Block Truncation Coding algorithm uses a two-levels (one bit) nonparametric quantizes that adapts to local properties of the image. The quantize that shows great promise is one which preserves the local sample moments. The quantize is compared with standard (minimum mean-square error and mean absolute error) one bit quantizares.

3. PROPOSED METHOD

Algorithm: K-Means-BTC

Input : Image, K (number of clusters)

Output :Compressed Image

Phase I:

Step 1 : Randomly place k points into the space represented by the objects that are being clustered. These points represent initial group centroids.

Step 2 : Assign each object to the group that has the closest centroid.

Step 3 : When all objects have been assigned, recalculate the positions of the k centrodids.

Step 4 : Repeat steps 2 and 3 until the stopping criteria is met.

Phase II:

Step 5 : The semi-compressed image split into many non overlapping block of size n x n.

Step 6 : Calculate the two statistical values of mean and standard deviation for every block of the Image

Step7. These statistical methods are known as a quantizers. Here x is the threshold value. It is attained by evaluating the every pixel value.

Step 8. After that the encoded blocks are restructured by swapping one as high intensity(H) and zero as a low intensity(L).

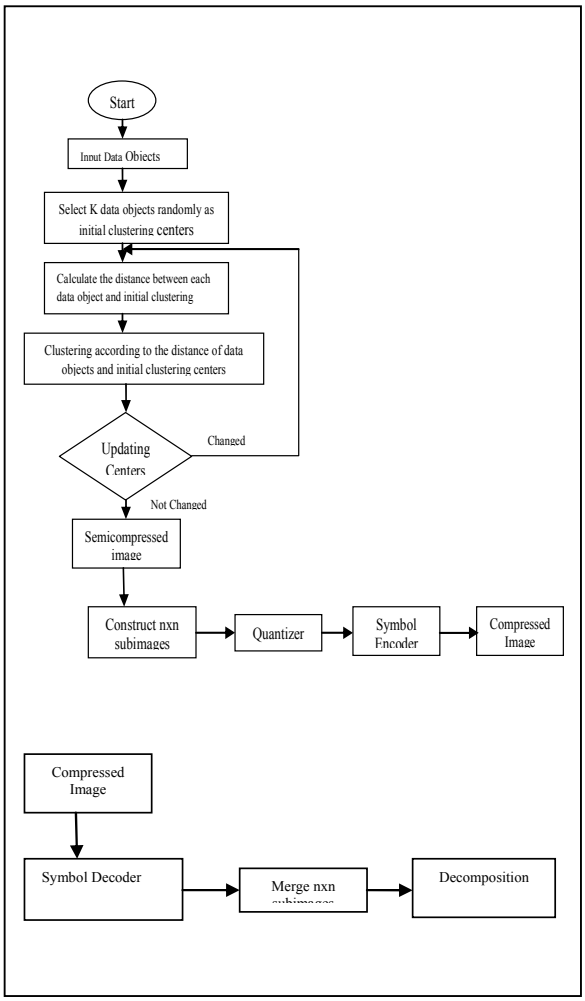


Figure3 : Block Diagram for K-Means Clustering Algorithm and Block Truncation Coding

4. IMAGE QUALITY MEASURE

There are various parameters present which are used to measure the performance of different compression algorithm. Some examples of the performance parameters of image compression are given below: [14]

4.1. Peak Signal to Noise Ratio (PSNR)

PSNR is an important parameter for image compression. It is measurement of the peak error present between the compressed image and original image. For better quality of image PSNR should be as high as possible. [14]

$$PSNR = 10log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$
$$= 20log_{10}\left(\frac{MAX_I}{MSE}\right)$$
$$= 20log_{10}(MAX_I) - 10log_{10}(MSE)$$

4.2. Compression Ratio

Compression Ratio (CR) is the ratio of size of compressed image to the size of original image [25]. Compression ratio should be as high as possible to achieve better compression [14].

$$\text{Compression ratio} = \frac{\text{uncompressed size}}{\text{compressed size}}$$

4.3. Mean Square Error

Mean Square Error (MSE) is cumulative difference between the original image and compressed image. MSE should be as minimum as possible for better quality of image [14]

5.EXPERIMENTAL RESULTS

Performance of the K-Means, BTC and K-Means-BTC has been evaluated for a set of standard test e-learning images, viz., 'Monalisa, 'Leena' and 'Consumer Club e-book. All the images are of size 256x256. K-Means-BTC is compared with Block Truncation Coding. Table 1 shows the reconstructed images of K-Means, BTC and K-Means-BTC.

Table1 : Original Images and Reconstructed Images

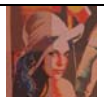
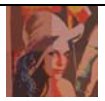
Original Images	K-Means	BTC	Proposed Method K-Means-BTC
			
Fig. 1. Monalisa Image			
			
Fig. 2. Leena Image			
			
Fig. 3. Consumer Club E-Book Image			

Table 2 : Comparative Study among K-Means, BTC and K-Means-BTC Methods

Image Name and Size	Methods	Compression Ratio	PSNR Value
Monalisa and Size – 11 KB	K-Means	1.1	30.98
	BTC	1.38	21.08
	K-Means-BTC	1.57	20.89
Leena and Size – 769 KB	K-Means	23.16	29.94
	BTC	19.23	21.76
	K-Means-BTC	24.03	21.47
Consumer Club E-Book and Size – 16 KB	K-Means	1.6	24.05
	BTC	1.6	14.62
	K-Means-BTC	1.78	15.19

Table2 shows the compression ratio and PSNR value for the images Monolisa, Leena and Consumer Club is compared with K-Means, BTC and K-Means-BTC. The performance is measured based on two parameters PSNR and Compression Ratio. From Table 2, it is seen that performance of the method K-Means-BTC is better than BTC algorithm on the basis of the two parameter measures.

6.CONCLUSION

In this paper, the algorithms that are dealing the image compressions are compared and analyzed. The algorithm compressed the image by using BTC from the centroids of K-Means. Then the compression ratio and PSNR value of KBTC are calculated with the K-means and BTC algorithm. This algorithm compressed the e-learning images and occupies less space for storage and low bandwidth while transmitting over the public cloud. In future, to make the algorithm more efficient the Genetic Algorithm will be applied at the beginning of K-MEANS-BTC to obtain the optimal centroids.

REFERENCES

[1] Almrabet, M.M., Zerek, A.R., Chaoui,A. and Akash, A.A., "Image Compression using Block Truncation Coding", International Journal of Sciences and Techniques of Automatic Control & Computer Engineering", Volume 3 : Issue 2, 2009.

[2] Duggal,S. Ali,N., and Sharma. M.K., "Designing Issues for E-learning Modules in Cloud Platform", International Conference of Advance Research and Innovation (ICARI-2015)", 2015.
[3] Gupta,P., Bansal, V., and Purohit, G.N., "Block Truncation Coding for Image Compression Technique", International Journal of Emerging Research in Management & Technology,, Volume 4 : Issue 4, 2015.
[4] Hartono., Ongko,E., and Abdullah, D., "Determining a Cluster Centroid of K-Means Clustering using Genetic Algorithm", International Journal of Computer Science and Software Engineering (IJCSSE)", Volume 4 : Issue 6, 2016.
[5] Jain,A. and Chawla, S., "E-learning in the Cloud", International Journal of Latest Research in Science and Technology", Volume 2 : Issue 1, 2013.
[6] Kumar, S. and Nancy, "K-Mean Evaluation in Weka Tool and Modifying It using Standard Score Method", International Journal on Recent and Innovation Trends in Computing and Communication, Volume 2: Issue 9, 2014.
[7] Malekar,N.C. and Sedamkar,R.R., "Novel K-Means Clustering Approach for Compressing Hyperspectral Image", International Journal of Advancements in Research & Technology, Volume 3 : Issue 1, 2014.
[8] Manro,S. Singh,J., and Joshi,A.S., "Managing e-learning using Clouds: A Cost-effective boon in 21st Century", International Journal of Computers & Distributed Systems", Volume 1: Issue 1, 2012.
[9] Mohammed, D. and Abou-Chadi, F (Senior Member, IEEE)., "Image Compression using Block Truncation Coding", Cyber Journals: Multidisciplinary Journals in Science and Technology", Journal of Selected Areas in Telecommunications (JSAT), 2011.
[10] Rani,M.S.M., "A Genetic Algorithm Based K-Means Algorithm For Compressing Images", International Journal of Engineering and Computer Science, Volume 4: Issue 9, 2015.
[11] Rani,M.M.S. and Chitra,P., "Region of Interest Based Compression of Medical Images Using Vector Quantization", International Journal of Computational Science and Information Technology(IJCSITY), Volume 4 : Issue 1, 2016.
[12] Tambe, M. Maniyar,A., Patil, R. and Uttekar, K., "An E-learning System with Image Compression", (IJCSIT) International Journal of Computer Science and Information Technologies", Volume 5 : Issue 6, 2014.
[13] Shashikumar, S., Parakale, A. and Mahavir, B.M., "Image Compression using Absolute Moment Block Truncation Coding", International Journal of advanced Scientific and technical research, Volume 3 : Issue 3, 2013.
[14] Singh,A.P., Potnis,A. and Kumar,A., "A Review on Latest Techniques of Image Compression", International Research Journal of Engineering and Technology (IRJET), Volume 3 : Issue 7, 2016.
[15] Somasundaram,K. and Rani,M.M.S., "Novel K-means Alogirthm for Compressing Images", International Journal of Computer Applications, Volume 18 : Issue 8, 2011.
[16] Utkarsh,A. and Chakraborty, C., "Image Compression: A Comparative Study Between ANN and Traditional Approach", International Journal OF Computer Applications, Volume 139, 2016.
[17] Vimala, S., Uma, P. and Abidha,B., "Improved Adaptive Block Truncation Coding for Image Compression", International Journal of Computer Applications, Volume 19 : Issue 7, 2011.
[18] http://www.ijoart.org/research-paper-publishing_january-2014.shtml
[19]http://www.ijser.org/researchpaper/Novel-K-Means-Clustering-Approach-for-Compressing-Hyperspectral.pdf
[20]http://www.ijarcsee.org/index.php/IJARCSEE/article/view/275
[21]http://ijarcsee.org/index.php/IJARCSEE/article/download/275/243
[22]http://www.ijser.org/paper/Novel-K-Means-Clustering-Approach-for-Compressing-Hyperspectral.html
[23]http://www.ijcaonline.org/research/volume139/number4/utkarsh-2016-ijca-909145.pdf
[24]http://connection.ebscohost.com/c/articles/74273248/novel-spatial-domain-lossless-image compression-scheme
[25] https://www.researchgate.net/publication/260290173_Comparative_study_of_Block_Truncation_Coding_with_EBTC[26]https://www.ermt.net/docs/papers/Volume_4/4_April2015/V4N4-148.pdf
[27]http://www.ijoart.org/docs/Novel-K-Means-Clustering-Approach-for-Compressing.pdf