Review of Various Techniques for Medical Image Compression

Harpreet Kaur
Department of CSE
Chandigarh University
Gharuan, Punjab

Rupinder Kaur
Assistant Professor
Department of CSE
Chandigarh University
Gharuan, Punjab

Navdeep Kumar
Assistant Professor
Department of CSE
Chandigarh University
Gharuan, Punjab

ABSTRACT
Image Compression generally refers to reducing the size of an image for the purpose of minimizing storage space as well as reducing the transfer time when transmitted over the network. Compression is very useful in medical imaging as a large quantity of storage is needed for storing medical images which can further be delivered for diagnosis. An appropriate technique for compression is needed for saving storage capacity as well as network bandwidth. It is also necessary that the valuable information should not be lost after compression of an image. In this paper various techniques of lossless image compression for medical images have been reviewed. The evaluation of performance is based on the parameter compression ratio.

Keywords
Lossless compression; medical images; compression ratio

1. INTRODUCTION
Medical imaging, which generates digital photograph of internal body structure, grants a powerful tool for analysis, hospitalization along with surgery, therefore turns to be an important portion of current healthcare delivery. This is a rapidly expanding field with the highest source of data and collection of modalities for example, computerised tomography (CT), positron emission tomography, ultrasound images, X-ray images, X-Ray mammography (MG), magnetic resonance imaging and various others [1]. Digital images generally accommodate great amount of superfluous data, and hence they are normally compressed to get rid of redundancy as well as lessen the space for storage and the network bandwidth [2]. In case the loss of visual quality is reasonable the modern compression techniques generate huge compression rates. On the other hand, in ROI (region of interest), there are many cases in which insufficiency in diagnostically significant regions is not affordable by the physicians. Hence it is necessary to have an approach that has high rate for compression along with superior quality of the ROI. The normal idea is to save quality in diagnostically vital areas while applying lossy encoding to the remaining regions [3].

Because of the need for precise diagnosis, lossless image compression is the imperative approach in lessening the unwanted size without losing accuracy of the images, granting the images to be fully restored at the time of decompression. To cut down or lessen the storage capacity of memory without interfering with the image quality and to transfer the image rapidly over the network with least transmission error, many algorithms for image compression have been introduced. There are commonly two kinds of image compression techniques: Lossy and Lossless [4]. Lossy compression is associated with some information loss that enables it to achieve much greater compression on the account of preciseness. It produces compression artifacts, when the data from the compressed image is recompiled using lossy method. JPEG is the example of lossy compression technique. In this method, the image reconstructed subsequent to compression is numerically similar to original image. An exact reproduction of the actual image can be obtained with it. To compress medical images lossless method is generally used. Compression ratio is the ratio between original file size and compressed file size and is calculated as.

$$CR = \frac{Input\,image}{Compressed\,image} \qquad ...(1)$$

The paper is divided into the various sections as follows. Section 2 provides a brief overview of image compression. Section 3 discusses the techniques for image compression. Section 4 provides an overview of the previous work related to image compression. Section 5 discusses the results based on compression ratio related to different techniques that have been studied and section 6 concludes the paper.

2. IMAGE COMPRESSION
Image compression means saving the image in the form of a stream of bits that is possibly small and to project the decoded image on the monitor that is as exact as required. Now take an illustration of an encoder and decoder as depicted in diagram. The image is transformed into a sequence of binary data known as the bit stream when it is received by the encoder. The resultant image is the output of the decoder when decoder obtains the encoded bit sequence. Image compression is that when the overall data quantity of the input image is greater than the overall data quantity of the received bit stream. The complete compression flow is depicted in Fig 1.

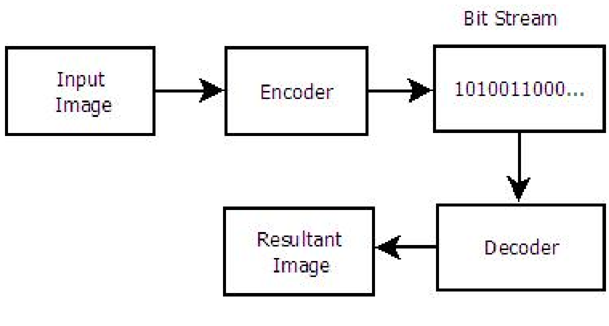

Fig 1: Basic flow of Image compression

Distortion: It is the difference between actual and the compressed image. It is given using Mean Square Error measured in dB.

$$MSE = \frac{\sum_{M,N} \left[I_1(m,n) - I_2(m,n) \right]^2}{M * N} \quad \ldots(2)$$

Where I_1 is the actual image, I_2 is the compressed image and (m, n) is used for rows and columns.

Fidelity or Quality: It refers to the similarity between the Actual and compressed image. It is calculated using Peak Signal to Noise ratio.

$$PSNR = 10 \log_{10} \frac{255}{MSE} \quad \ldots(3)$$

3. COMPRESSION TECHNIQUES FOR IMAGE

Normally there are two kinds of techniques which are used for compression of digital images named as: lossy and lossless compression techniques depicted in Fig 2.

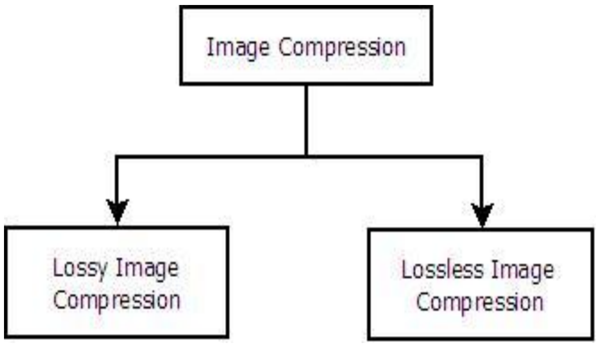

Fig 2: Image compression Classification

3.1 Lossless Image Compression

The compression in which the image remains similar as the original image after decompression is called as Lossless compression. This compression technique most presumably exploits factual redundancy to state information more accurately without any significant loss in data [4]. As described before, lossless techniques are chosen for technological drawing, satellite imaging as well as medical imaging and many more. In lossless compression some methods are entropy based and rest are dictionary based. Run-length encoding, Huffman encoding, Arithmetic coding are Entropy based methods and dictionary based technique is LZW which will be later discussed one by one.

3.1.1 Run-length simple encoding (RLE)

It is an extremely simple type for compression of an image where runs of information are saved as the single information value as well as count, in spite of original run. It is utilized for sequential information and is useful for repeated information. This method changes sequence of similar symbol (pixel), known as runs [5]. The run length code used for gray scale images is expressed by a series {Vi, Ri} where Vi denotes the pixel intensity and Ri denotes the various consecutive pixels having the intensity value Vi. This is very important for data which contains plenty of such runs such as, basic graphic images like icons, line drawings and animations. This method is not significant for files which do not have several runs because it could elevate the size of the file. Lossless compression is performed by Run-length encoding. Fax machines also use this technique for compression.

3.1.2 CALIC

It stands for Context-based, Adaptive, Lossless Image Codec. This coder was given by Wu and Memon. It helps to obtain greater compression ratios generally for the continuous-tone images as compared to other techniques that have comparatively lower execution time as well as space complexities. It lays main emphasis on image information modeling. An exclusive characteristic of this method is the using of a huge amount of modeling contexts in order to constraint a nonlinear predictor as well as adapt predictor to different source statistics [2]. The lesser time as well as space complexity is attributed to effective method for quantizing the modeling contexts. It is a method of sequential coding which encodes as well as decodes in the raster scan arrangement with a unique pass throughout the images. CALIC basically works in two types of modes known as binary as well as continuous-tone. Though more elaborative than various other lossless image coders, it is computationally easy, using mostly integer arithmetic and also easy logic. Both algorithms of encoding and decoding are appropriate for parallel implementation as well as pipelined hardware implementation also by sustaining sequential buildup.

3.1.3 Huffman encoding

Entropy encoding technique that is useful for lossless compression is called Huffman coding. Huffman developed this method which is nowadays utilized as a “back-end” for other methods of compression [6]. The term denotes the usefulness of code table of variable-length required for the encoding of source symbol in which the code table of variable length is derived in the particular manner that relies on the calculated probability of appearance for every probable value of source symbol. Pixels of an image are taken as symbols. Symbols that occur very commonly are given a lesser number of bits, where as those symbols that appear less intermittently are given a comparatively greater number of bits. It is prefix code [7] which depicts that code of a symbol does not possess to be a prefix of code of some other symbol.

3.1.4 Lempel-Ziv-Welch (LZW)

Abraham Lempel, Jacob Ziv and Terry Welch produced a universal algorithm for lossless compression called LZW. Welch published it in 1984 as a revised execution of the LZ78 technique published in 1978 by Lempel as well as Ziv. It is dictionary dependent coding which can be static as well as dynamic. In a static process of dictionary coding, the dictionary is rigid during the process of encoding as well as decoding but in dynamic process of dictionary coding; the used dictionary is restructured on fly. This technique is easy to execute, and possesses a potential for extremely high performance inside the hardware implementations. The GIF format of image uses LZW compression which was the technique of the famously utilized UNIX file compression application. It became the initial widely used compression algorithm for universal images on computers.

3.1.5 Prediction-based Algorithm

A major element of the research process has emphasized on a particular kind of compression method, basically called as lossless DPCM or predictive coding [2]. In case the prediction is practically precise, the allocation of prediction error is determined near to zero and has considerably lesser zero-order entropy as compared to the actual image. The present JPEG standard makes use of a predictive method in the lossless mode. This gives eight predictors for the process of selection. In answer to the need for proposing a new lossless compression method, nine proposals were submitted to ISO. Amongst the first round estimation, it is evident that proposals based on transform-coding do not give as good ratio for compression as compared to algorithms that are proposed on the basis of predictive methods. The three generally proposed predictors are: GAP, ALCM and MED.

3.2 Lossy Image Compression

The name itself clarifies that lossy compression is related to the loss of some data. The image that is compressed is identical to the actual uncompressed image, but is not same like the previous because during compression some of the information regarding the image is lost. This method is normally suited for images where minimum loss of data is recommendable to attain a significant decrease in the bit rate. JPEG is the most popular technique which is used for lossy compression. This compression technique results in a greater compression rate as compared to lossless compression. Methods of lossy compression are scalar quantization and vector quantization.

3.2.1 Scalar Quantization

It is a common category of quantization. This quantization, generally represented by Y=Q (x) and is a procedure of utilizing a quantization function denoted by Q in order to map a 1-D (scalar) value x of the input to any scalar resultant value Y. This quantization is as basic as well as instinctive to use as rounding off the high-precision integers to the closest integer, or the closest multiple of another unit of a precision.

3.2.2 Vector Quantization (VQ)

This is a traditional quantization method of signal processing that permits the development of the probability density functions through the allocation of prototype vectors. This was firstly utilized for compression of images [8]. It basically works by separating a series of points known as vectors into sets having almost the equal number of points nearest to them. The vector quantization has a powerful property known as density matching, particularly for recognizing the density of huge and higher-dimensional information. Because information points are expressed with the index that is related to the closest centroid, the infrequent data have high error and the frequently appearing data possess low error. It is the reason for Vector Quantization being suitable for lossy compression of data. It can be utilized for lossy data rectification as well as density estimation.

4. RELATED WORK

Xiwen Zhao et al. [9] built up an effective lossless method for image compression depending upon structure prediction where an image was divided into structure and non-structure areas. The structure areas were encoded using structure prediction and the non-structure areas were encoded using an already existing technique called Context-based, Adaptive, Lossless Image Codec (CALIC). The broad experimental results showed that the proposed method was exceptionally proficient in lossless compression of images, particularly for images with specified structure components. The disadvantage of the proposed technique was its complexity because of costly structure prediction and the classification threshold was selected manually so it might not be optimal with respect to coding efficiency.

Suresh Yerva et al. [10] gave a way to deal with lossless image compression inside spatial domain for continuously-toned images utilizing a novel idea of image folding. The technique proposed utilized the property of redundancy and adjacent neighbor for prediction. This method of data folding provided an easy approach for compression, which gave better compression efficiency as well as has lesser computational complexity when compared with the standard SPIHT method for lossless compression. The proposed algorithm worked better in case of smooth images and comparatively provided better results on text images rather than the photographic images. The technique faced the flaw of greater magnitude during the calculation of difference data due to greater distance in pixels at higher levels.

Yi-Fei Tan et al. [11] provided a technique for image compression that could be used for lossless as well as lossy compression. This technique included a threshold value, diverse compression ratios could be obtained by altering the threshold value. In case the threshold value was zero, lossless compression could be carried out. Lossy compression was carried out when the value of threshold acquired positive values. The proposed technique permitted the determination of image quality during the compression procedure. This gave more adaptability to user in picking the proper type of compression, as well as meeting preferable ratio of compression by altering the threshold values corresponding to required image quality.

Arif Sameh Arif et al. [12] presented a technique for lossless image compression of those parts that act as target in Fluoroscopy images, computing the region of interest. Here the targeted parts were pharynx and esophagus. The hybridization of two algorithms that are Run Length as well as Huffman coding was used in order to enhance the efficiency of compression. The results of proposed technique concluded that it increased the compression ratio by 300% which was better than the existing method. It also helped in achieving benefits regarding storage and transmission of images over the network.

J. Papitha et al. [13] addressed many previous techniques for DICOM image compression and analyzed the accuracy of the algorithms with the help of visual inspection method. The results were also verified with respect to compressed size of image, compression ratio, Peak Signal to Noise Ratio. At last, the optimal compression method resulted from this investigation had been described with regard to the visual perception as well as performance metrics. Based on the comparative study it was concluded that the LDT technique was better than other techniques as it resulted in less MSE value, greater value of PSNR, higher MSSIM as well as visual inspection.

Srikanth S. et al. [14] implemented the SPIHT and EZW algorithms along with Huffman encoding by utilizing various wavelet families and comparing the PSNRs as well as bit rates of these wavelet families. This algorithm was carried out on various images and showed that the result of the proposed methodology were better as they provided good compression ratio and high image quality when compared with the traditional methods. The preprocessing as well as advanced SPIHT methods provided better quality recreation at decoder along with lesser computational complexity when compared with existing methods. The simulation results concluded that the proposed algorithm yielded challenging value of PSNR at less bit rate.

Victor Sanchez et al. [15] introduced enhancements to HEVC intra coding procedure for lossless method of compressing the grayscale medical images that were characterized by their greater quantity of edges. In particular, an alternative angular as well as planar prediction modes was proposed, which relied on differential pulse code modulation with an improved range of directionalities. The DPCM decoding method that was implemented maintained the coding format of HEVC. Evaluation results proved that the proposed differential pulse code modulation (DPCM) modes effectively calculate the huge quantity of edges in images accomplishing bit-rate savings up to 15%.

Atef Masmoudi et al. [16] proposed an effective adaptive arithmetic coding to be used in lossless compression that relied on finite mixture model. Hence for every block of image, the proposed method described a combination of distributions generated from its consequent blocks. The parameters for mixture have been evaluated using EM algorithm based on maximum likelihood. The experimental results concluded that the compression efficiency by using the proposed technique was much better as compared to usual arithmetic encoders and JPEG-LS, by an increase of 21% and 31%, respectively.

Krishan Gupta et al. [17] introduced a new method termed as KMK used for the compression of an image.The compression by using KMK Technique was more as the repetition of number of pixels increase. This technique helped in compressing the image in vertical, horizontal or diagonal directions. The main advantage of using this technique was that it could provide better compression when compared with some other methods. The disadvantage that occurred was that in case pixel of the image repeated slightly than its destined position then the proposed method was not much appropriate. The proposed technique also improved the size of image as well as reduced storage of data without any particular loss in image.

Arif Sameh Arif et al. [18] proposed a proficient method for compression of fluoroscopic images by using lossless technique. The main role of this paper was to extract the regions of interest by using appropriate shapes in the fluoroscopic images then hybridization of Huffman encoding and Run-length algorithm was used for compressing the extracted Region, to improve compression ratio. The proposed technique concentrated on automatically selecting the most suitable shape amongst a selection of existing shapes to constitute 2 ROI regions along with least redundant areas in the extracted region. The results of proposed technique showed an improvement in the compression ratio by approximately 400% when the comparison is made with existing techniques.

5. RESULTS

The given table provides a comparative study of the results obtained by using different techniques of lossless compression on the basis of compression ratio. The greater the compression ratio, the more the image is compressed.

Table 1: Comparison of Compression Ratio based on different techniques.

Techniques	Compression Ratio (%)
DPCM [15]	15.14
CALIC [16]	28.27
Huffman encoding [7]	54.29
Run-length encoding [7]	100.06
R. Huff -Seg [12]	300
Auto Shape-R.Huff [18]	400

6. CONCLUSION

Image compression refers to compressing the significant bits of an image such that the quality of the image does not get affected. In this paper various techniques for lossless compression of medical images along with a brief overview of lossy compression technique has been presented. Lossless image compression is being widely used in many applications as it leads to minimum loss of data. It is an efficient technique to be used in hospitals for diagnosis of several disorders. Auto Shape-R.Huff method for lossless compression is much efficient as it provides greater compression ratio.

7. REFERENCES

[1] Xin Bai , Jesse S. Jin & Dagan Feng, "*Segmentation-based multilayer diagnosis lossless medical image compression*", Pan-Sydney Area Workshop on Visual Information Processing, pp. 9-14, 2004.

[2] Ming Yang & Nikolaos Bourbakis, "*An overview of lossless digital image compression techniques*" IEEE 48th Midwest Symposium Circuits and Systems, pp. 1099-1102, 2005.

[3] Charalampos Doukas & Ilias Maglogiannis, "*Region of interest coding techniques for medical image compression*", IEEE Engineering in Medicine and Biology Magazine, Vol. 26, No 5, pp. 29- 35, 2007.

[4] Mohammad Kabir Hossain, Shams M Imam, Khondker Shajadul Hasan & William Perrizo, "*A lossless image compression technique using generic peano pattern mask tree*", IEEE 11th International Conference on Computer and Information Technology ICCIT, pp. 317-322, 2008.

[5] Mohamed Y. Eltabakh, Wing-Kai Hon, Rahul Shah, Walid G. Aref & Jeffrey S. Vitter, "*The SBC-tree: An index for run-length compressed sequences*", ACM In Proceedings of the 11th International Conference on Extending Database Technology: Advances In Database Technology, pp. 523-534, 2008.

[6] Micha Stabno & Robert Wrembel, "*RLH: Bitmap compression technique based on run-length and Huffman encoding*", Information Systems, Vol. 34, No. 4, pp. 400-414, 2009.

[7] Qiusha Min, & Robert J.T. Sadleir, "*A hybrid lossless compression scheme for efficient delivery of medical image data over the internet*", IEEE Second International Conference on Computer Modeling and Simulation, Vol. 1, pp. 319-323, 2010.

[8] Jau-Ji Shen & Hsiu-Chuan Huang, "*An adaptive image compression method based on vector quantization*", First International Conference on Pervasive Computing Signal Processing and Applications (PCSPA), pp. 377-381, 2010.

[9] Xiwen Zhao & Zhihai He, "*Local structure learning and prediction for efficient lossless image compression*", IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 1286-1289, 2010.

[10] Suresh Yerva, Smita Nair & Krishnan Kutty , "*Lossless image compression based on data folding*", IEEE International Conference on Recent Trends in Information Technology (ICRTIT), pp. 999-1004, 2011.

[11] Yi-Fei Tan & Wooi-Nee Tan, "*Image compression technique utilizing reference points coding with threshold values*", IEEE International Conference on Audio, Language and Image Processing (ICALIP), pp. 74-77, 2012.

[12] Arif Sameh Arif, Rajasvaran Logeswaran, Sarina Mansor & Hezeru Abdul Karim, "*Segmentation and compression of pharynx and esophagus fluoroscopic images*", IEEE International Conference on Signal and Image Processing Applications (ICSIPA), pp. 220-225, 2013.

[13] J. Papitha, G. Merlin Nancy & D. Nedumaran, "*Compression techniques on MR image- A comparative study*", IEEE International Conference on Communications and Signal Processing (ICCSP), pp. 367-371, 2013.

[14] Srikanth S. & Sukadev Meher, "*Compression efficiency for combining different embedded image compression techniques with Huffman encoding*", IEEE International Conference on Communications and Signal Processing (ICCSP), pp. 816-820, 2013.

[15] Victor Sanchez & Joan Bartrina-Rapesta, *"Lossless compression of medical images based on HEVC intra coding*", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) , pp. 6622-6626, 2014.

[16] Atef Masmoudi, Afif Masmoudi& William Puech, "*An efficient adaptive arithmetic coding for block-based lossless image compression using mixture models*", IEEE International Conference on Image Processing (ICIP), pp. 5646-5650, 2014.

[17] Krishan Gupta, Mukesh Sharma & Kirti Saneja, "*A new lossless KMK technique for image compression*", IEEE International Conference on Optimization Reliabilty and Information Technology (ICROIT), pp. 405-408, 2014.

[18] Arif Sameh Arif, SarinaMansor, Rajasvaran Logeswaran & Hezerul Abdul Karim, "*Auto-shape Lossless Compression of Pharynx and Esophagus Fluoroscopic Images*", Springer Science Business Media New York, 2015.