A MEMORY-ASSISTED LOSSLESS COMPRESSION ALGORITHM FOR MEDICAL IMAGES

Zhinoos Razavi Hesabi[1], *Mohsen Sardari*[2], *Ahmad Beirami*[2], *Faramarz Fekri*[2], *Mohamed Deriche*[3], *and Antonio Navarro*[1]

[1]Instituto de Telecomunicacoes/DETI- Universidade de Aveiro, Campus Universitrio Santiago, Aveiro, Portugal
[2]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[3]Department of Electrical Engineering, King Fahd University of Petroleum & Minerals, Saudi Arabia
Email: {zrazavi, navarro}@av.it.pt, {mohsen.sardari, beirami, fekri}@ece.gatech.edu, mderiche@kfupm.edu.sa

ABSTRACT

Rapid growth of emerging medical applications such as e-health and tele-medicine requires fast, low cost, and often lossless access to massive amount of medical images and data over bandlimited channels. In this paper, we first show that significant amount of correlation and redundancy exist across different medical images. Such a correlation can be utilized to achieve better compression, and consequently less storage and less communication overhead on the network. We propose a novel memory-assisted compression technique, as a learning-based universal coding, which can be used to complement any existing algorithm to further eliminate redundancies across images. The approach is motivated by the fact that, often in medical applications, massive amount of correlated images from the same family are available as training data for learning the dependencies and deriving appropriate reference models. Such models can then be used for compression of any new image from the same family. In particular, Principal Component Analysis (PCA) is applied on a set of images from training data to form the required reference models. The proposed memory-assisted compression allows each image to be processed independently of other images, and hence allows individual image access and transmission. Experimental results on Xray images show that the proposed algorithm achieves 20% improvement over and above traditional lossless image compression methods reported in the literature.

1. INTRODUCTION

A large amount of digital images is produced by hospitals and medical centers and stored in databases everyday. This vast amount of data translates into huge storage requirements. Further, in telemedicine and teleradiology applications, these images need to be transmitted over the network, consuming high bandwidth capacity. Therefore, compression is a necessity for handling this vast amount of images. More importantly, in computer assisted diagnosis, *loss* of any part of the information contained in the data can be detrimental. Therefore, *lossless* image compression is the method of choice for medical images. Most of the literature on lossless compression for medical images merely considers compression algorithms that are based on spatial redundancy removal within an individual image so as to achieve better compression performance. Hence, opportunities in using cross-image redundancies to further compress data are often overlooked!

The goal of this work is to develop practical algorithms that can first exploit both intra- and inter-image redundancies for a dictionary of medical images. Secondly, we require that the algorithms allow individual (random) access to images. In other words, when a retrieval request for an individual image is received, the proposed algorithm should be able to retrieve the requested image without decompressing the whole database. We argue that the compression problem discussed above can effectively be solved using the recently developed "memory-assisted compression" technique developed by the authors [1]. As such, we discuss the adaptation of well-known compression algorithms in the literature to this concept of memory-assisted compression framework, which can be formulated as a two phase compression scheme. The first phase is learning, in which the algorithm runs over a subset of images in the database, and extracts the commonalities shared among all the images. Such information is stored and used in the next phase; the memory-assisted compression. We should note that, in several applications, the subset of the images in the first phase is readily available. As one example, in telemedicine applications, where medical images are taken every day and transmitted over the network, the images from the previous day, for example, can serve as the memory for the next day, and so on.

Our focus here is mostly on the second phase, i.e., the memory-assisted compression. In the second phase, for compression of every image in the database, the common information stored in the memory is used to eliminate the inter-image redundancy and only the residual is fed to traditional lossless compression algorithms. We reiterate here that the proposed two-phase structure enables individual access to all the images without the need to decompress the whole database and at the same time, all the dependencies present among the images are used for efficient compression.

2. BACKGROUND AND RELATED WORK

As previously described, the use of correlation information across images has been very limited in lossless image compression. Most of the previous work only focuses on the compression of a single image regardless of the redundancy that is present in the set of images to which this specific image belongs. In this section, we review some of published research on lossless image compression based on individual medical images.

MacMahon *et al.* [2] proposed a form of adaptive block

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

Cosine Transform coding, in which considerable compression of digital radiographs with minimal degradation of image quality is allowed. Their results obtained for chest radiological images showed a compression ratio as high as 25. Ekstrand [3] presented a lossless compression algorithm based on Context Tree Weighting (CTW). The algorithm performs optimally in terms of redundancy for a wide range of data sources including medical gray scale images. The results show enhanced performance compared to JPEG, JBIG, and CALIC. Asraf *et al.* [4] proposed a novel hybrid lossy and lossless compression method using neural networks, vector quantization, and Huffman coding. The method was tested on CT (Computerized Tomography) images achieving a compression ratio of 5 to 10. A lossless medical image compression method was presented by Kil *et al.* [5]. The method was based on redundancy analysis and segmentation of image into Variable Block Size (VBS) in order to extract similarities and smoothness of blocks. It was reported that the technique outperforms Huffman, JPEG-LL and lossless JPEG2000 by 10-40%. Ghrare *et al.* [6] introduced a lossless image coding algorithm based on pixel redundancy reduction and using 2 matrices of Gray Scale and Binary. The algorithm achieved a maximum compression ratio of 4. Miaou *et al.* [7] proposed an image compression technique which combines JPEG-LS and interframe coding with motion vectors showing a better than JPEG-LS alone. They tested their algorithm with six capsule endoscope image sequences and improved the average compression by 13.3% and 26.3% over JPEG-LS and JPEG-2000, respectively. We emphasize here that our focus is on removing inter-image redundancy using memory-assisted compression. Again, the proposed algorithm can be used as an add-on to any existing image compression technique. In our experiments, we chose to work with the JPEG-LS, CALIC, CTW, and bzip2 algorithms.

3. PROBLEM SETUP AND PROPOSED APPROACH

The main rationale behind memory-assisted compression is learning source statistics at some intermediate entities, then leveraging the memorized context to reduce redundancy of the universal compression of finite length sequences. To the best of our knowledge, this is the first attempt in using cross-image correlation in medical image compression. Indeed, we propose to apply the concept of memorization with medical image sequences. The basic problem setup in a telemedicine application is displayed in Fig. 1. The source contains a set of correlated medical images (eg. Chest X-ray images) at the server node S (e.g., the central hospital) that need to be encoded and transmitted to the destination node (e.g. remote hospital) D through the network. We further assume that hospital D has already memorized a database of previously transmitted images in its database which is also shared with S. In the absence of *memorization*, traditional compression techniques can still be applied for transmitting the sequence of images from S to D. Here, we argue that the communication overhead to send the images from S to D can substantially be reduced if memorized context is available to the encoder and the decoder. In this paper, we present two different scenarios to compare results of traditional lossless image compression algorithms with our proposed memory-assisted compression algorithms.

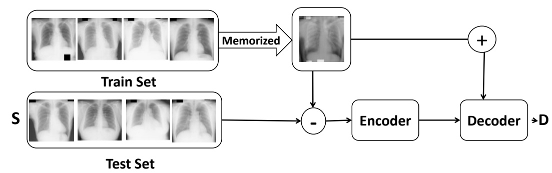

Fig. 1. The proposed memory-assisted compression algorithm.

First, we just apply traditional compression algorithms, from state-of-the-art, on a set of medical images. In this scenario, redundancy is only considered within a single image for encoding and decoding it. The problem is that every single image is encoded without considering other images in the same set leading to low compression ratio and additional overhead. Next, we apply our memory-assisted compression algorithm using the simple PCA algorithm [8] within the same set of images. In this case, the encoder and decoder have access to the memorized context for compression of new unknown images. We show that we can obtain a significant improvement in compression ratio for lossless medical images over state-of-the-art algorithms used in the literature.

4. THE PROPOSED MEMORY-ASSISTED COMPRESSION ALGORITHM

In this section, we describe the proposed algorithm in further details. Consider a basic scenario in which a set of X-ray images is available at node S, and node D requests one of the images, as shown in Fig. 1. This scenario can be an abstraction of a transmission problem or a storage and retrieval problem. Our benchmark is the case in which each image is compressed individually and sent to D. Then, at node D, each compressed image is decompressed independently.

In the proposed method, the outcome of the learning phase, called $\mathcal{M}$, of memory-assisted compression is available at both S and D. Then, using $\mathcal{M}$, just the residuals of other test images are encoded at node S and decoded at node D. The proposed memory-assisted lossless compression method consists of two main phases : 1) Learning(memorization), 2) Memory-assisted Compression (testing).

The main question now is how can we consider and model the memorization concept from a set of gray-scale medical images? The simple answer comes from the Karhunen Loeve transform (KLT) [9]. KLT is shown to be the optimal orthogonal transform through which the energy (information) contained in the signal is compacted. Indeed, with KLT, most energy is redistributed over a small number of components or simply called eigenimages. These eigenimages are obtained from the decomposition of the estimated covariance matrix. For our experiments, we simply use the PCA transformation matrix to decorrelate the images and remove inter-image redundancy. PCA is a statistical approach used to find an orthogonal transformation to decorrelate random variables. The PCA technique has extensively been used in diverse signal and image processing applications. It has originally been introduced as a dimension reduction technique. The technique starts with a set of observation vectors of dimension N. For images, the columns are concatenated into a large vector of size N (number of pixels). Let M be the number of observations in the training set:

$$\mathrm{x_i} = [p_1, ..., p_N]^T, i = 1, ...M \tag{1}$$

From these observations, the mean vector and covariance matrix are estimated:

$$\mathrm{m} = \frac{1}{M}\sum_{i=1}^{M} x_i \tag{2}$$

$$\mathbf{C} = \frac{1}{M}\sum_{i=1}^{M}(x_i - m)(x_i - m)^T \tag{3}$$

Let's represent the mean centered observations by: $w_i = x_i - m, i = 1, 2, ..., M$.

The goal is to find a subspace whose basis vectors correspond the maximum-variance directions in the orthogonal space. Let **W** represent this linear transformation that maps the original N-dimensional space onto a K-dimensional feature subspace where normally K≪N. The columns of **W** are the eigenvectors, e_i, of the covariance matrix C. The eigenvectors are obtained using eigen-decomposition of **C**, $\lambda_i e_i = \mathbf{C}e_i$ and λ_i is the eigenvalue associated with e_i.

For a given observation vector x_i, the transformation results in a new vector y_i given by: $y_i = \mathbf{W}^T(x_i - m)$. As the first few eigenvalues represent most energy in the data, we usually select K to be much smaller than N. The original observation vectors can then be reconstructed using the inverse transform. $\hat{x}_i = \mathbf{W}^T + m$. Note that since **C** is an NxN matrix where N is the total number of pixels in the image, finding the eigenvectors may not be easy to do. Luckily, a number of techniques have been proposed to get around this difficulty.

For our experimental setup, the PCA procedure discussed above was first applied to the training set of images. Once the training stage is completed, we move to the coding stage. In this stage, test images are first projected over the PCA space. Using the reduced PCA space, the test images are reconstructed. These reconstructed images are close approximation of the original images. An error image is obtained by subtracting the reconstruction image from the original test image. It is simple to show that the pixel values of such an error image are uniformly distributed.

As such, we can compress error images in a more optimized way using the traditional CTW, JP-LS, CALIC and bzip2 algorithms. This means that we only need to send the feature vectors as well as the new compressed error images to the receiver. At the receiver, we first reconstruct the image projection then add to it the decompressed error image.

To evaluate the performance of the proposed approach, we considered two scenarios and 4 generic compression techniques. The compression techniques considered include: the CTW, JPEG-LS, CALIC, and bzip2 algorithms. The two scenarios are explained below:

- **Scenario 1(Comp):** It denote the case of using the CTW, bzip2, JPEG-LS, and CALIC algorithms directly on the test set.
- **Scenario 2(PCAComp):** Here, PCA is applied on a train set of images. Then, for testing, the images are first projected and reconstructed using the PCA. Second, the residual images are encoded using the CTW, bzip2, JPEG-LS, and CALIC algorithms. At the receiver, the images are reconstructed using the decoded residuals added to the PCA-reconstructed images.

5. SIMULATION RESULTS

In this section, we will discuss our experimental results. For lossy compression, we usually use the RMSE and a measure of performance, but for lossless coding, compression efficiency is usually measured using compression ratio. Computational complexity is another factor that determines the efficiency of the method. This can be the number of CPU cycles, number of hardware gates, or memory bandwidth, etc. These are usually application dependent. In our work, we focused mainly on the compression ratio defined as the compressed image size S_{comp} over the original image size S_o, i.e. $CR = \frac{S_{comp}}{S_o}$

In other words, the *CR* represents the number of bits the scheme uses to represent each bit in the uncompressed image. We calculated the *CR* for various compression schemes discussed in Section 4. The image database used is the JRST database which can be downloaded from [10]. It contains 154 nodule and non-nodule 8-bit Chest X-ray images with matrix size of 2048x2048 pixels. We selected a subset containing 20 Chest X-ray images. 10 images are selected as our training set while the other 10 images were used for testing the algorithm. The experiments were repeated 10 times by randomly changing the 20 images.

The CTW, CALIC, bzip2 and JPEG-LS algorithms were applied on the test set as our benchmark lossless image compression algorithms. We display in figures 2 and 3 the histograms of 2 typical images before and after applying PCA. The entropy for the original image (Fig. 2) was 4.83 and decreased to 2.35 for the image after PCA decorrelation. We also show the histogram of the first eigen image (Fig.4), which clearly proves the maximum variance that such an image can model. Fig. 5 displays compression ratio for different lossless image compression algorithms using both the Comp and the PCAComp scenarios.

The compression ratios (*CR*) for CALIC, CTW, bzip2, JPEG-LS were 0.18, 0.19, 0.24, 0.27, respectively. After applying our memory-assisted algorithm, the compression ratio was improved by 17.79%, 27.67%, 14.90% and 14.84%, respectively. As can be seen, similar trends are observed on the compression ratios of the CTW, CALIC and JPEG-LS algorithms. The CALIC, in particular, achieved a slightly better *CR* than the others.

We also considered a number of other experiments on different sets of training images with various sizes. Fig. 6

presents the average compression ratio on 3 different sets of training images including 10, 20 and 30 chest X-ray images. As can be seen, there is only a minor improvement in compression ratio when we increase the size of the training set (number of images) from 10 to 30 images. The highest compression efficiency is achieved for the CALIC and the JPLS algorithms when used with proposed PCAComp algorithm.

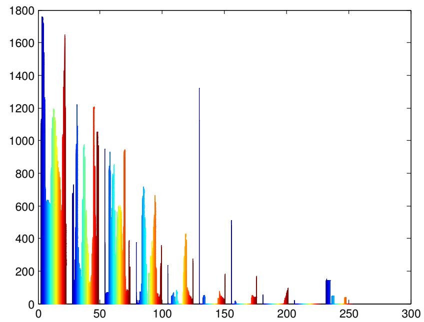

Fig. 2. Histogram of a single image.

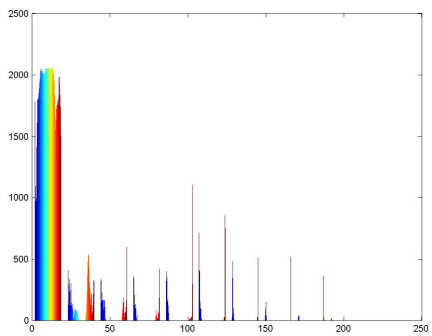

Fig. 3. Histogram of a single image after applying PCA.

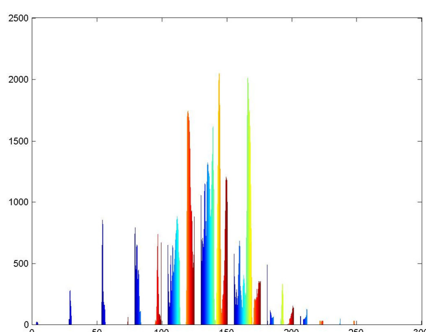

Fig. 4. Histogram of first eigenimage.

By applying PCA on top of CALIC and JPEG-LS, we obtained gains of 16.97% and 14.55%, respectively. These reported gains are for training sets of 10 images. The same trend can be observed for the image set of 20 and 30. To further analyze the energy compaction property of PCA, we display in Fig. 7 the compression ratios (*CR*) of memory-assisted techniques for CTW and CALIC using different number of eigenimages. As expected, few PCA components are indeed important in the reconstruction. Actually, our experiments showed that more than 97% of the total energy is contained in the first 5 eigenvalues (i.e. $\frac{\sum_{i=1}^{5}}{\sum_{i=1}^{M}\lambda_i} > 0.95$)

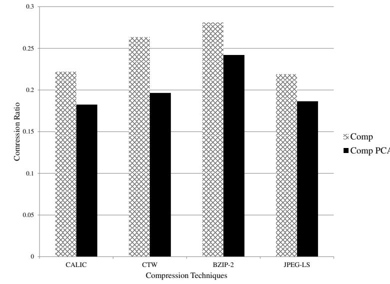

Fig. 5. Compression ratio (*CR*) for traditional and memory-assisted algorithms with different compression techniques.

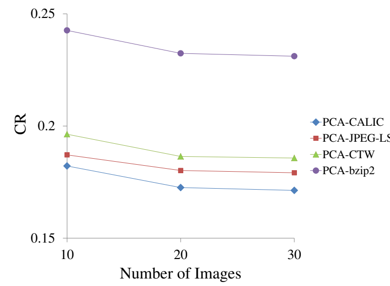

Fig. 6. Average compression ratio (*CR*) of the proposed algorithm for different training set sizes of 10, 20 and 30.

6. CONCLUSION

We discussed a new method for lossless compression using the concept of memory-assisted universal coding. The proposed approach is well suited to compress large datasets of medical images especially for recurrent usage. The algorithm consists of a learning phase followed by a testing phase. In the learning phase, PCA is performed on training images to extract a set of eigenimages which are used to reconstruct the different test images. The reconstructed images are simply represented (coded) by low dimensional feature vectors. The error (or residual) images are then compressed using traditional lossless compression algorithms such as the CALIC, JPEG-LS, bzip2 and CTW algorithms. Our experimental results using the JRST database showed that the performance of traditional lossless algorithms can be improved by an average of 20% using the proposed algorithm. The proposed concept of using memory to enhance the performance of universal coders is expected to have a major impact in areas where images exhibit high correlation.

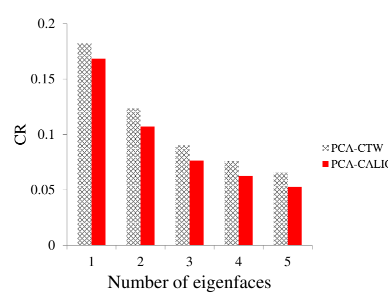

Fig. 7. Average compression ratio (*CR*) of memory assisted compression algorithms for different number of eigenfaces.

Acknowledgments

The authors would like to acknowledge the support of KFUPM (under project IN121012), Georgia Institute of Technology, Universidade de Aveiro, and KACST (project NSTIP 11-MED1612-04).

7. REFERENCES

[1] A. Beirami, M. Sardari, and F. Fekri, “Results on the fundamental gain of memory-assisted universal source coding,” in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, july 2012, pp. 1087 –1091.

[2] H. MacMahon, K. Doi, S. Sanada, S. Montner, M. Giger, C. Metz, N. Nakamori, F. Yin, X. Xu, H. Yonekawa, and et. al., “Data compression: Effect on diagnostic accuracy in digital chest radiography,” vol. 178, no. 1, pp. 175–9, Jan 1991.

[3] N. Ekstrand, “Lossless compression of grayscale images via context tree weighting,” in *Data Compression Conference, 1996. DCC ’96. Proceedings*, mar/apr 1996, pp. 132 –139.

[4] R. Asraf, M. Akbar, and N. Jafri, “Statistical analysis of difference image for absolutely lossless compression of medical images,” in *Engineering in Medicine and Biology Society, 2006. EMBS ’06. 28th Annual International Conference of the IEEE*, 30 2006-sept. 3 2006, pp. 4767 –4770.

[5] S. Kil, J. Lee, D. Shen, J. Ryu, E. Lee, H. Min, and S. Hong, “Lossless medical image compression using redundancy analysis,” *IJCSNS International Journal of Computer Science and Network Security*, vol. 6, no. 1, pp. 50–56, Jan 2006.

[6] S. Ghrare, M. M. Ali, K. Jumari, and M. Ismail, “An efficient low complexity lossless coding algorithm for medical images,” vol. 6, no. 8, pp. 1502–1508, 2009.

[7] S.-G. Miaou, F.-S. Ke, and S.-C. Chen, “A lossless compression method for medical image sequences using jpeg-ls and interframe coding,” *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 5, pp. 818 –821, sept. 2009.

[8] I. Jolliffe, “Principal component analysis,” *Springer*, 1986.

[9] J. Burl, “Estimating the basis functions of the karhunen-loeve transform,” *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 37, no. 1, pp. 99 –105, jan 1989.

[10] [Online]. Available: www.jsrt.or.jp