

Low-Bitrate Medical Image Compression

Antonius Darma Setiawan
Institute Technology of Bandung
Jl. Ganesha 10 Bandung, Indonesia
adsetiawan@students.itb.ac.id

Andriyan Bayu Suksmono
Institute Technology of Bandung
Jl. Ganesha 10 Bandung, Indonesia
suksmono@stei.itb.ac.id

Tati LR Mengko
Institute Technology of Bandung
Jl. Ganesha 10 Bandung, Indonesia
tmengko@itb.ac.id

Hendra Gunawan
Institute Technology of Bandung
Jl. Ganesha 10 Bandung, Indonesia
hgunawan@math.itb.ac.id

Abstract

Medical imaging is an important element in the medical diagnostic process. However, storing medical images in a certain format is quite challenging. The size of a digital medical image is in the range of 3–32 MB per examination, depending on the imaging modality. Image compression will be a suitable solution when it comes to minimizing the size of stored medical images. The quest to find a compression method that has very good quality image reconstruction and a high compression ratio is limited by the Shannon/Nyquist sampling theorem and image entropy. However, in recent years, compressive sampling (CS) has arisen as an alternative sampling method. It guarantees the reconstruction of a high-quality signal from a small number of samples, and can be carried out if the signal is sparse enough. A set of trained bases grouped as a dictionary will guarantee the sparsest representation of the signal. This paper propose a low bit-rate medical image compression scheme based on CS and an overcomplete bases set to represent the medical images as sparse as possible.

1. Introduction

Medical records are an important dataset of patients' medical history, and are needed by physicians to engage in an accurate diagnostic process. Medical images are part of the medical record. Therefore, they also need to be ready or stored for diagnostic purposes for present or future use. Nowadays, we find that most medical images are not store in medical institutions such as hospitals. The reason for this is that it is hard to deal with the size of medical images. Most digital medical images are between 3 and 32 MB per examination, depending on the imaging modality [9]. Thus, the repository for medical images in the medical institution, also called a Picture Archiving and Communication System (PACS), will require huge storage capacities. This kind of system is very expensive. It also takes time to transfer the medical images through narrowband telecommunication lines, such as exist in the developing countries and remote areas.

Medical image compression will eliminate the problems described above. There are two image compression methods. The first method is called *lossless compression*,

and results in a reconstructed image with the same quality as the source image, but a smaller compression ratio (3–7 times smaller). The second method is *lossy compression*, which will give a much better compression ratio (more than 12 times smaller), but results in error between the reconstructed image and the source one. The design of lossy compression must maintain the medical information that is needed to conduct diagnostic activity. Otherwise, the result of diagnosis will be inaccurate.

These considerations raise the following question: Is it possible to have a low-bitrate lossy compression method in which the reconstructed image has almost the same quality as the image source? The answer is no, because the compression ratio and the reconstructed image quality go to opposite direction according to Shannon/Nyquist information theory. Thus, a new approach to compression is proposed in this paper. This compression scheme is built based on the compressive sampling (CS) method. The CS method is a new approach that has been developed over the past few years by Candes, Tao, Romberg, and Donoho [2,3,4]; CS samples only the important information instead of everything by means of sparsity and measurement matrix. It ensures high-quality signal reconstruction from highly incomplete measurements (as a replacement for sampling) with overwhelming probability. This property allows CS to provide a possible solution to the problem stated above.

The proposed compression scheme is founded on vector quantization (VQ) image compression; it is simply a generalization of the VQ image compression scheme. A comparison of the proposed scheme and other state-of-the-art compression algorithms such as JPEG, JPEG2000, and Scalable Fuzzy Vector Quantization (SFVQ) [15] will be presented in this paper.

2. Proposed Medical Image Compression Scheme Based on Compressive Sampling

2.1. Compressive Sampling

In the past four years, a prospective sampling theorem has been developed as an alternative to the Shannon/Nyquist sampling theorem; this is known as compressive sensing or compressive sampling (CS), and was developed by David Donoho, Emmanuel Candes,

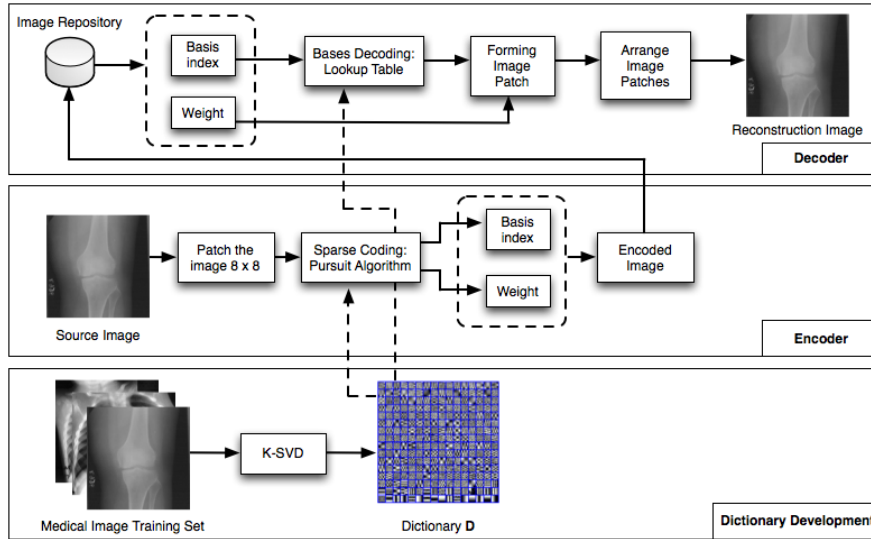


Figure 1. Medical image compression scheme based on CS and K-SVD

Justin Romberg, and Terence Tao [3,4]. These researchers state that if a signal is sampled (or measured, to be exact) above a minimum value, it can with overwhelming probability be reconstructed to resemble the source signal exactly. A reconstructed signal is represented through the linear combination of several bases. This is different from a traditional reconstruction method, which uses only one base to represent a reconstructed signal.

The forward transformation in current signal or image acquisition is inefficient. It takes a lot of samples, but loses a lot of the signal in the encoding process. In contrast, CS samples only the important information. The assumption in CS is that most signals are sparse in a particular transformation basis. Thus, the basic principles of CS are sparsity and incoherency: If we have a sparse signal and expand it to a certain transformation basis $\mathbf{Y}=\Psi\mathbf{x}$, then we will find that the coefficient \mathbf{x} has a small number of non-zero entries. Incoherency introduces measurement as a replacement for sampling. The measurement matrix Φ is chosen to be highly incoherent with the transformation matrix Ψ . The measurement (dimensionality reduction) result will be expressed as $\mathbf{y}=\Phi\Psi\mathbf{x}$.

How many measurements have to be taken in order to reconstruct the measurement exactly as the source signal? Candes has defined the number of measurements by the following:

$$m \geq \mu^2(\Phi, \Psi) \cdot k \cdot \log N, \quad (1)$$

where $\mu(\Phi, \Psi)$ is the coherency between measurements and the transformation matrix, while k is the number of coefficients that are not zero and N is signal length. Let the incoherency to be a constant; then, the number of measurements required in (1) depends on k and N . If we can represent the signal as sparsely as possible in the transformation domain, then the value of k will be small. This will affect the number of measurements taken. If we measure not less than m , then we can reconstruct a signal that, with overwhelming probability, will be the same as the source signal by using ℓ_1 minimization.

A dictionary $\mathbf{D}=\Phi\Psi$ is an overcomplete dictionary. As the reconstruction signal is the linear combination of

several bases in the dictionary, it will make the overcomplete dictionary guaranties less involved bases to represent a signal compared with the complete one. The encoded signal will end up being smaller [6,7,8,10]. Therefore, the development of an overcomplete dictionary that has the capacity to represent the signal as sparsely as possible is very important.

As learned from CS, performance will be optimal if the signal can be represented as sparsely as possible. The quest for the sparsest dictionary will help this to occur. This kind of dictionary can be achieved by using a trained overcomplete dictionary. An overcomplete dictionary will perform better than a complete one, because a larger vocabulary will represent the signal more simply than a smaller one. Furthermore, a trained dictionary will work better if we use a specific kind of signal, such as a medical image.

2.2. K-SVD Learned Dictionary

In Aharon, Elad, and Bruckstein's work on sparse representation [12], they have succeeded in developing a trained based dictionary that outperforms predefined dictionaries such as DCT and Haar. The algorithm to develop a trained dictionary is derived from K-Means. The algorithm is a generalization of K-Means, and is called K-SVD. If a set of sample images \mathbf{Y} is known, then the overcomplete dictionary, which represents \mathbf{Y} as sparsely as possible, can be developed by solving the objective function expressed in (2).

$$\min_{\mathbf{D}, \mathbf{x}} \{\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2\} \quad \text{s.t.} \quad \forall i, \|x_i\|_0 \leq T_0. \quad (2)$$

The number of non-zero entries coefficient in matrix \mathbf{X} is determined to be equal or less than T_0 . The K-SVD algorithm is described in [12].

This algorithm contains two major steps: the first is sparse coding, while the second is codebook updating. Sparse coding tries to find the coefficient corresponding to the previous dictionary with constraint T_0 . For the first iteration, the dictionary, $\mathbf{D}^{(0)}$, is initiated. The codebook or dictionary update stage involves updating the previous

Table 1. The Peak Signal to Noise Ratio (PSNR) value of image reconstruction from various medical image compression algorithms compared to the proposed method based on compressive sampling (CS) in decibels (dB)

X-ray	JPEG 0.2 bpp	JPEG2000 0.2 bpp	SFVQ 0.125 bpp	CS 0.125 bpp	
				DCT	K-SVD
XrayT1	30.6	32.9	30.3	34.2	36.6
XrayT2	31.1	33.2	30.6	37.7	37.8
XrayF1	38.7	41.0	38.6	41.3	41.4
XrayF2	40.4	43.0	40.0	42.5	42.5
XrayH1	36.7	41.1	35.0	41.5	41.6
XrayH2	35.8	40.6	34.1	41.3	41.4
Average	35.6	38.6	34.8	39.8	40.2

dictionary using the updated coefficient from the sparse coding stage. This process is conducted by means of the Single Value Decomposition (SVD) algorithm. The iteration is stopped when the convergence condition is fulfilled. The stopping condition is set to:

$$\|\mathbf{D}^j - \mathbf{D}^{(j-1)}\|_F^2 \leq \epsilon, \quad (3)$$

where ϵ is the error between the present iteration and the previous one. The complete algorithm of K-SVD is shown in Figure 1.

2.3. Proposed Scheme

The proposed medical image compression method is based on the VQ scheme, in a generalization of VQ. The dictionary in VQ is developed using the K-Means algorithm to create K clusters from image sample set \mathbf{Y} . Every sample \mathbf{y}_i is a $p \times p$ pixel image patch. K-Means is used to develop dictionary or codebook \mathbf{C} , consisting of K codewords. The signal \mathbf{Y} can be represented as:

$$\mathbf{Y} = \mathbf{C}\mathbf{X}, \quad (4)$$

where \mathbf{X} is the set of vector coefficients \mathbf{x}_i used to select the proper codeword to represent \mathbf{y}_i . Vector \mathbf{x}_i contains zero entries, except that in the j th position, and can be expressed as $\mathbf{x}_i = \mathbf{e}_j$ [12]. The position of j will select which codeword is going to be used to represent \mathbf{y}_i , and the value e in j is 1. The codeword with the minimum error as a distance function with \mathbf{y}_i is selected. In ℓ_2 -norm; this can be expressed as in [6]:

$$\forall_{k \neq j} \|\mathbf{y}_i - \mathbf{C}\mathbf{e}_j\|_2^2 \leq \|\mathbf{y}_i - \mathbf{C}\mathbf{e}_k\|_2^2. \quad (5)$$

The index j is saved as the encoded file. The size of the encoded file depends on the patch and image size. The compression ratio of this method is related to the patch size, which is $p \times p$. VQ-based medical image compression is developed and presented in [11].

VQ-based compression only uses one basis (codeword) to represent signal \mathbf{y}_i . The basic idea of the proposed method is to generalize this through the representation of \mathbf{y}_i with several k -weighted bases as a linear combination of them. The overcomplete dictionary \mathbf{D} contains either pre-specified bases or trained result bases. The number of involved bases k is intended to be small to gain a high compression ratio. The linear combination of several

weighted bases from the overcomplete dictionary will ensure a minimum error of $\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2$. An overcomplete dictionary \mathbf{D} that can represent the signal as sparsely as possible is needed to accomplish this. The model of the proposed method is expressed in (6):

$$\mathbf{Y} = \mathbf{D}\mathbf{X}. \quad (6)$$

The complete proposed compression scheme is shown in Figure 1. This scheme is divided into three major parts. The first part is dictionary development. K-SVD is used to develop the dictionary from a set of medical image samples. The other two parts are encoding and decoding. The encoding process is begun by patching the medical image into an 8×8 pixel image patch; the next process involves sparse coding using a pursuit algorithm to get the weight (w) and bases index (i). According to [6], any pursuit algorithm can be used in this process. The Orthogonal Matching Pursuit (OMP) algorithm was used in this experiment. All weight and basis indexes from all the image patches were saved as the encoded file (\mathbf{J}). The encoder can be expressed as

$$E: \mathcal{R}^n \rightarrow \mathbf{J}(\mathbf{i}, \mathbf{w}) \quad (7)$$

Finally, the decoding part is very simple. The first step is finding the appropriate bases (\mathbf{b}) according to bases index (i). This process uses the same dictionary as in the encoding process. The next step is forming the image patch (\mathbf{J}) using a linear combination of weight and the involved bases as in (8). The last step is arranging all the patches into a reconstructed image.

$$\mathbf{J}_q = \sum_{p=1}^k \mathbf{b}_p \mathbf{w}_p, \quad q = 1, 2, \dots, M \quad (8)$$

where M is the number of image path, and can be expressed as:

$$M = (W \times H)/d \quad (9)$$

Where W and H are the height and the width of the image respectively, whereas d is the dimension of the image patch, in this case is 8×8 .

3. Experiment and Result

Two experiments will be reported in this paper. The first is a comparison of the performance of the proposed compression scheme with other state-of-the-art compression

sion algorithms, specifically JPEG, JPEG2000, and SFVQ. The software used to test JPEG and JPEG2000 was taken from other sources and modified to meet the experiment condition. The JPEG encoder decoder was taken from MATLAB central. The encoder was developed by Yu Hen Hu. Meanwhile, the decoder was developed by Ravi Rakkund. JJ2000 software was downloaded from <http://jj2000.epfl.ch/> and used to encode and decode images in JPEG2000 format. The SFVQ and proposed compression scheme were self-developed and partially used the K-SVD Toolbox for MATLAB developed by Aharon et al.

The medical images used in the experiment comprise X-ray images of the thorax, knee joint, and hand. The images were resized to a dimension of 512×512 . Grayscale images with 8-bit color depth were used. There were six medical images employed, two for each X-ray image type.

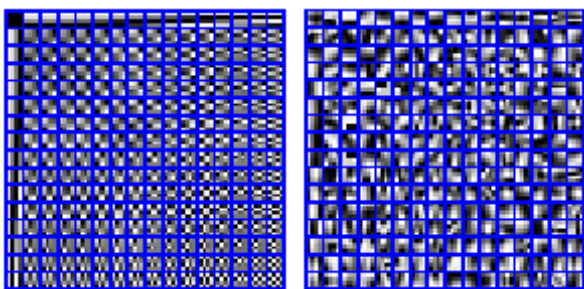


Figure 2. DCT (left) and trained (right) Dictionary

The second experiment compared the performance of the predefined bases dictionary and the trained bases dictionary using K-SVD in the proposed scheme. The DCT basis was chosen for the predefined bases dictionary. Here, the result of the reconstructed image will determine the performance of each dictionary. The result of the image reconstruction is denoted by Peak Signal to Noise Ratio (PSNR), which is expressed by (10).

$$\text{PSNR} = 20 \log_{10} \left(\frac{255}{\sqrt{\text{MSE}}} \right). \quad (10)$$

The PSNR shows how close the image reconstruction is to the source image in logarithmic scale; MSE is the mean square error. The MSE is the square error between the image reconstruction and the source image, normalized by image dimension (width \times height).

The results of the first and the second experiment are combined in Table 1. The table shows the PSNR value of each image compression algorithm for each image at almost the same bitrate. The first three columns are the state-of-the-art image compression algorithms used in comparison with the proposed method. The last two columns show the pre-specified and learned dictionary developed by K-SVD. The bitrate of the encoded image is 0.125 bpp for CS and SFVQ, and 0.2 bpp for JPEG and JPEG2000. This difference occurred because of software limitations.

As shown in Table I, the proposed method performs better image reconstruction than others at almost the same bitrate. The second experiment shows that the performance of the learned dictionary developed by K-SVD was better than that of the predefined DCT dictionary. This

approach is suitable when we use a particular type of image, such as a medical image. The result might be different if applied to natural images. The DCT and trained dictionary is shown in Fig. 2.

4. Conclusion

The proposed scheme of medical image compression is presented. The result shows that the proposed scheme is a promising compression method, and that the trained-based dictionary performs better than a predefined bases dictionary. It means that the trained-based dictionary represents the medical image more sparse and lead to better image reconstruction quality.

References

- [1] A. Zandi, J.D. Allen, E.L. Schwartz, and M. Boliek: "CREW: Compression with reversible embedded wavelets", *Data Compression Conference*, p. 212-221, 1995.
- [2] E. Candès and J. Romberg: "Quantitative Robust Uncertainty Principles and Optimally Sparse Decompositions", *Foundations of Computational Mathematics*, vol. 6, no. 2, pp. 227-254, 2006.
- [3] E. Candès, J. Romberg, and T. Tao: "Robust Uncertainty Principles: Exact Signal Reconstruction From Highly Incomplete Frequency Information", *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489-509, 2006.
- [4] E. Candès, J. Romberg, and T. Tao: "Stable Signal Recovery From Incomplete and Inaccurate Measurements", *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207-1223, 2006
- [5] M. Aharon, M. Elad, and A.M. Bruckstein: "On The Uniqueness of Overcomplete Dictionaries, and A Practical Way To Retrieve Them", *Journal of Linear Algebra and Applications*, vol. 416, pp. 48-67, 2006.
- [6] M. Aharon, M. Elad, and A.M. Bruckstein: "The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation", *IEEE Trans. On Signal Processing*, vol.54, no.11, pp. 4311-4322, November 2006.
- [7] Michael Elad, Roman Goldenberg, Ron Kimmel: "Low Bit-Rate Compression of Facial Images", *Image Processing, IEEE Transactions*, vol.16, no. 9, pp. 2379 - 2383, 2007
- [8] Michal Aharon, Michael Elad, Alfred Bruckstein: "K-SVD: Design of Dictionaries for Sparse Representation", *IEEE Trans. On Signal Processing*, vol.54, no.11, pp.4311-4322, 2006
- [9] S. Wong et. al.: "Radiologic image compression-A review", *Proc. of the IEEE*, Vol. 83, No. 2, 1995.
- [10] Richard G. Baraniuk: "Compressive Sensing", *IEEE Signal Processing Magazine*, vol.124, pp. 118 - 120, 2007
- [11] Setiawan, A. D. Suksmono, A. B. & Dabarsyah, B.: "Scalable Radiology Image Transfer and Compression Using Fuzzy Vector Quantization", *Journal of eHealth Tech. & Applications*, vol. 5, no. 2, 2007