

Four-Dimensional Compression of fMRI Using JPEG2000

Hariharan G. Lalgudi¹, Ali Bilgin¹, Michael W. Marcellin¹, Ali Tabesh¹, Mariappan S. Nadar²
and Theodore P. Trouard^{3,4}

¹Signal Processing and Coding Lab, Department of Electrical and Computing Engineering,
The University of Arizona, Tucson, AZ, USA;

²Siemens Corporate Research, Imaging and Visualization Department, Princeton, NJ, USA;

³Department of Radiology, The University of Arizona, Tucson, AZ, USA;

⁴Biomedical Engineering Program, The University of Arizona, Tucson, AZ, USA

ABSTRACT

Many medical imaging techniques available today generate 4D data sets. One such technique is functional magnetic resonance imaging (fMRI) which aims to determine regions of the brain that are activated due to various cognitive and/or motor functions or sensory stimuli. These data sets often require substantial resources for storage and transmission and hence call for efficient compression algorithms.

fMRI data can be seen as a time-series of 3D images of the brain. Many different strategies can be employed for compressing such data. One possibility is to treat each 2D slice independently. Alternatively, it is also possible to compress each 3D image independently. Such methods do not fully exploit the redundancy present in 4D data. In this work, methods using 4D wavelet transforms are proposed. They are compared to different 2D and 3D methods. The proposed schemes are based on JPEG2000, which is included in the DICOM standard as a transfer syntax. Methodologies to test the effects of lossy compression on the end result of fMRI analysis are introduced and used to compare different compression algorithms.

Keywords: JPEG2000, fMRI, 4D-Compression.

1. INTRODUCTION

As the medical imaging industry advances toward a filmless environment, the amount of digital data that need to be managed presents a significant challenge. There is increased use of medical imaging techniques that produce four dimensional (4D) datasets such as fMRI. fMRI data can be seen as a time-series of 3D images of the brain and are used to detect activations in the brain pertaining to a particular task. These 4D datasets require tremendous storage space and transmission bandwidth. Hence there is a need for efficient compression techniques.

Picture Archiving and Communication Systems (PACS) have achieved tremendous development in today's healthcare environment. PACS utilize the Digital Imaging and COmmunications in Medicine (DICOM) standard. Implementation of these 4D medical datasets in the existing PACS has been discussed in Ref. 1. Here we investigate efficient compression techniques to compress 4D datasets using JPEG2000,² which has been included in DICOM as a transfer syntax. Both raw and pre-processed fMRI data are tested. Due to the low resolution and low SNR nature of fMRI data, activation is found by statistical analysis rather than by visual means. Thus it is of interest to explore the effects of lossy compression on fMRI analysis.

This work was supported in part by Siemens Corporate Research.

1.1. Prior Work

There has been much research done on efficient compression schemes for 3D medical volumes such as CT, MRI, Ultrasound and SPECT. Volumetric compression using zero trees is developed in Ref. 3 and Ref. 4. Extension of the 2D-EBCOT (Embedded block coding with optimal truncation) paradigm to 3D ESCOT (Embedded subband coding with optimal truncation) is done in Ref. 5. A Quad-tree based method is discussed in Ref. 6. Algorithms that give high decoder throughput are developed in Ref. 7 and Ref. 8.

In contrast, very little work has yet been done in the field of 4D medical image compression. Compression of fMRI using 3-D wavelets is introduced in Ref. 9. No entropy coding is done on the wavelet coefficients and hence the lossy results presented do not give the true picture. In Ref. 10, fMRI compression is done using a scheme where in the intensity change of each voxel from that of the first volume is stored using one byte. This method is based on the fact that images do not change much across time. The coordinates where the difference exceeds 1 byte are indicated in header information. The two-dimensional embedded zerotree wavelet (EZW)¹¹ algorithm was extended to four dimensions and applied to echocardiographic data in Ref. 12.

2. FUNCTIONAL MAGNETIC RESONANCE IMAGING

2.1. Principle

In fMRI, a time series of 3D images of the brain is acquired while the subject is performing some task. Changes in blood oxygen level increase the ratio of oxygenated to depleted hemoglobin. This change can be detected using Blood Oxygen Level Dependent (BOLD) contrast.¹³ Thus, the increase in the ratio of oxy-Hb to deoxy-Hb results in a brighter region in the functional MR image in areas of higher neural activity. The change in the MR signal is extremely small which demands powerful MRI machines. Even then the signal to noise ratio is too low for direct imaging, and statistical methods and averaging are required to show activity.

A simple and commonly used task design is the block paradigm where alternating rest and activation states prolong over time (approximately 30-60 s per epoch) averaging multiple responses.^{13,14,15} A block paradigm series usually consists of 50-150 volumes in time. These scans are repeated over different sessions to get 3-4 runs of the same experiment which are used to get consistent results. Another type of task design is the so called event related fMRI which requires increased scan time and thus an increased number of images. This is because the activated brain region has to relax and return to physiological baseline before the event can occur. Event related paradigms need 200-600 volumes per experiment.

As an fMRI experiment needs a fast sampling of the brain's physiological response, the usual slice matrix size of a single fMRI image is kept small to 64×64 which sacrifices in-plane resolution to $3-7 \text{ mm}^2$.^{13,14} The SNR of fMRI signals are generally very low. This makes activation detection difficult.

2.2. fMRI analysis

There has been extensive research during the past years in the field of fMRI analysis. The basic steps carried out in typical fMRI processing are illustrated in Fig. 1. The procedure basically consists of some pre-processing steps followed by analysis. The analysis of fMRI can be divided broadly into two types, one that uses a model for the response while the other that is based on a non-parametric/data-driven approach. In the first approach, usually referred to as "classical data analysis"/General linear model (GLM),¹⁶ each measured time series is modeled as a linear combination of signal and noise. The second approach is the model free analysis known

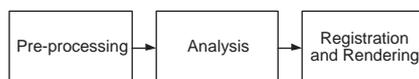


Figure 1. Steps in fMRI analysis.

as Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC).¹⁵ It uses independent component analysis to separate the noise and the signal and does not use any model of the signal. After analysis, the functional activation map is registered to a high resolution volumetric image and color coded for visualization.

3. fMRI COMPRESSION USING JPEG2000

We first employ existing methods to compress 4D datasets. Each ‘xy’ slice is treated as a 2D image and compressed using JPEG2000. This gives a 2D baseline to compare with our results in 3D and 4D.

Next, we compress the dataset as 3D cubes. JPEG2000 Part 2 describes the use of the wavelet transform along the third dimension to decorrelate the data along that dimension. A 1D wavelet is first applied across slices and the resulting transformed slices are compressed using JPEG2000. Fig. 2 shows two different pictorial representations of the 3D wavelet data that are ultimately compressed using this procedure. Rate allocation can be employed jointly for all slices so that we can have constant quality across slices. This procedure has been used to compress 3D medical images and has been proven to be very effective.¹⁷

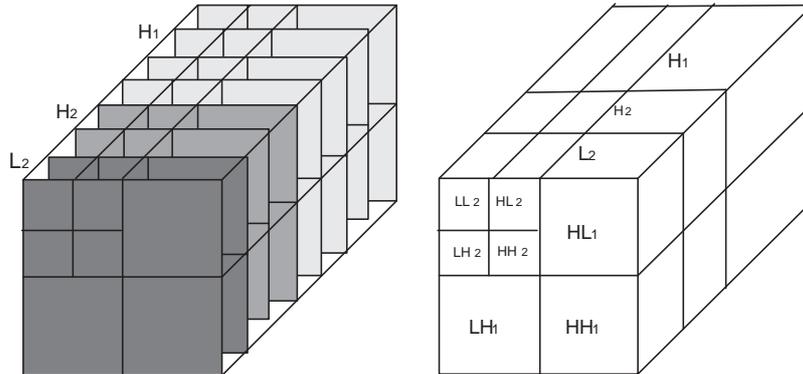


Figure 2. 3D wavelet decomposition specified by JPEG2000 Part 2.

For 4D datasets, we can employ 3D compression on either ‘xyz’ cubes or ‘xyt’ cubes. We note here that results obtained for ‘xyt’ are much superior to those of ‘xyz’. Extensive results are presented later in the paper.

3.1 Compression as Hypercube(4D)

The 2D/3D methods do not exploit the redundancy across all four dimensions. Hence we also explore use of wavelet transform to decorrelate the data along the fourth dimension. Specifically we apply a 1D wavelet transform along t followed by a 1D wavelet transform along z and compress the resulting wavelet coefficients as 2D slices using JPEG2000. Little or no gain was obtained in this case over the ‘xyt’ results. We believe, this is because the thickness of the slices (in the z dimension) is high (around 5 mm). There are also other artifacts such as head movement which makes the correlation in z smaller compared to ‘xy’ or ‘t’. These factors motivates exploration of different wavelet decompositions as discussed next.

3.2 Study of Different Decompositions

A modified 3D wavelet decomposition (M3DW) introduced in Ref. 8 first performs a standard 2D dyadic wavelet transform on each slice. This would generate four 3D subbands L_xL_y , L_xH_y , H_xL_y and H_xH_y each of which has a reduced spatial resolution in x and y by a factor of two, but preserves the same spatial resolution in z. In the next step only L_xL_y is decomposed into $L_xL_yL_z$ and $L_xL_yH_z$ as opposed to decomposing all the subbands. Due to spatial smoothing by the dyadic wavelet that is applied first, the curved boundaries of human organs register well along z.⁸ This decomposition is illustrated in Fig. 3. Extensive experiments in 4D employing similar ideas yielded the decomposition structure of Fig. 4 as having the best results.

4. COMPRESSION AFTER PRE-PROCESSING

fMRI data go through significant pre-processing before they are analyzed to determine the activation map of the brain for a particular task. Fig. 5 shows the pre-processing steps that are done before the analysis. Some steps are optional, based on the statistics of the data. These pre-processing steps can improve compression efficiency. Registration (motion correction and slice time correction) and spatial smoothing increase correlation and make

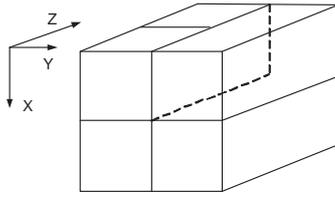


Figure 3. Modified 3D wavelet decomposition (M3DW).

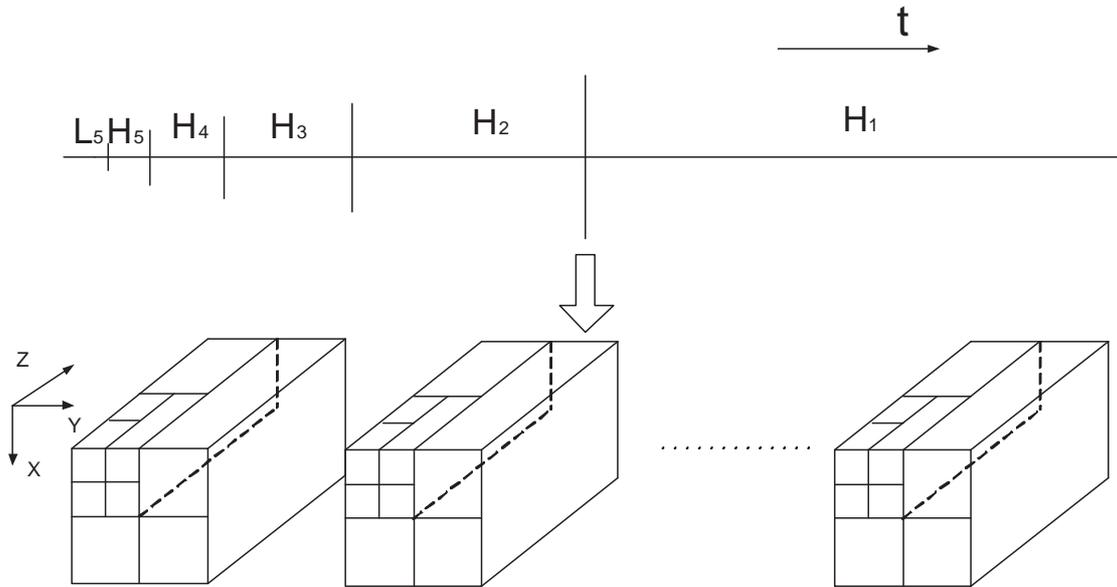


Figure 4. Wavelet decomposition in 4D.

the data more compressible. Since the first three steps are carried out before the analysis is done, it may be profitable to compress the pre-processed data. Results of the lossless and lossy compression of both raw and pre-processed fMRI data are presented in Section 6.

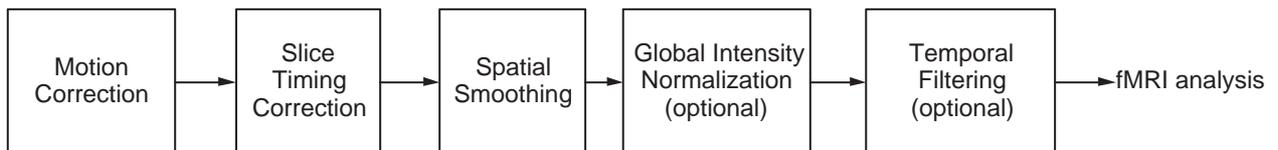


Figure 5. Pre-processing steps in fMRI.

5. EVALUATION OF LOSSY COMPRESSION ON fMRI ANALYSIS

The pre-processed data is fed to the analyzer which determines potential activation pertaining to the task the subject performs. The output of the analysis first gives a functional activation map. This would basically contain the statistical z-score for each voxel in the 'xyz' cube.

To explore the effects of lossy compression on the analysis performance, we would like to compare the activation map of the original data with the map obtained after lossy compression. Unfortunately, the result obtained with the original data is not the ground truth. fMRI analysis is still an ongoing research and statisticians

working in this field have not agreed on the choice of best strategy. Many of the statistical results which conform to the expectations of the neuroscience community are more likely to be believed and accepted for publication. However there are also strategies that have been used¹⁸ to evaluate different statistical methods of analyses quantitatively. Hence we follow two approaches to the problem of evaluating the effects of lossy compression on fMRI analysis.

The first method uses one of the best statistical methods for fMRI analysis and assumes the output of that to be the ground truth. We can then compare the lossy results with this “ground truth” to know the effects of compression. fMRI Software Library (FSL) maintained by Oxford Center for functional magnetic resonance imaging of the brain¹⁵ has become a well known tool and is being used by many fMRI researchers. We analyze the fMRI dataset with the FSL utility included in the medical image processing software ‘medx’. A pictorial representation of the set up for this method is given in Fig. 6. Path 1 follows the usual analysis procedure of the original data, the output (activation map) of which is assumed to be the ground truth. In path 2 the raw data is compressed lossily prior to any pre-processing and the activation map is obtained. In path 3 we compress the pre-processed data obtained at the intermediate point. Thus the activation maps obtained by paths 2 and 3 reflect the effects of compression and are compared with original activation map.

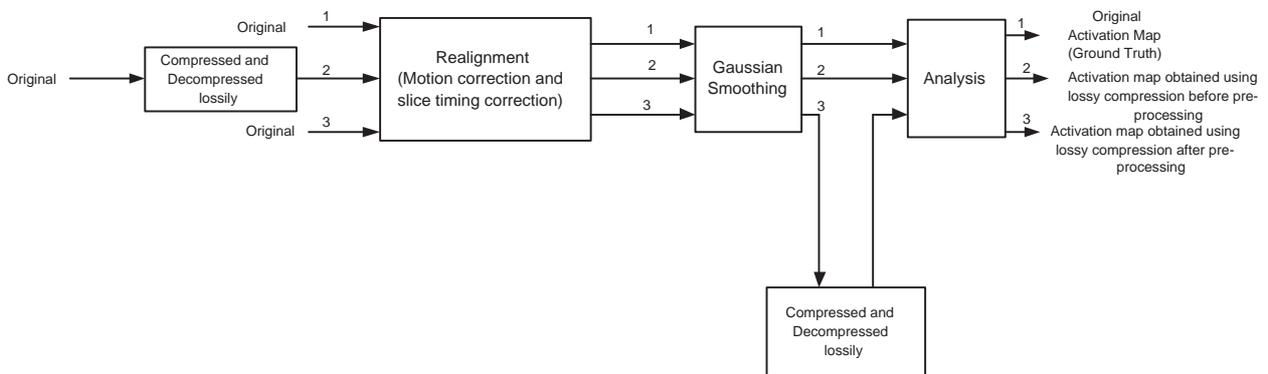


Figure 6. Lossy compression effects on fMRI.

In the second approach we add artificial activations (that realistically simulate fMRI activations)¹⁸ to the datasets obtained from a real fMRI experiment. By adding the artificial activation, we know the ground truth and hence can compare to the detected activations to measure the accuracy obtained. We follow the same signal specifications as in Ref. 18. The procedure is depicted in Fig. 7. The ROC plots obtained for the original and the compressed data are then compared to evaluate the effects of lossy compression.

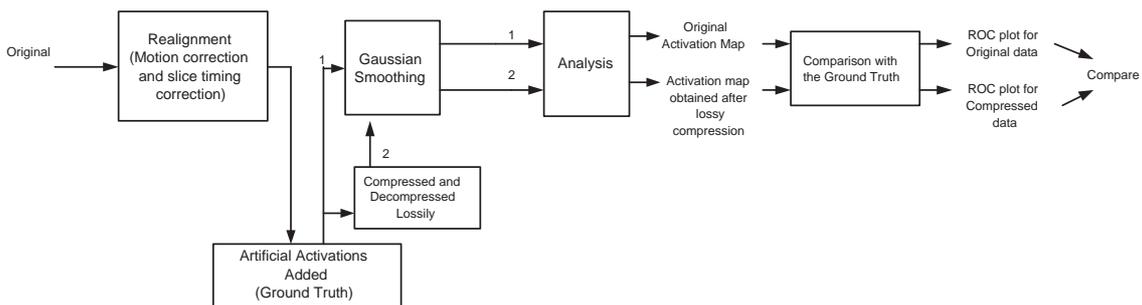


Figure 7. fMRI analysis with artificial activation.

6. EXPERIMENTAL RESULTS

We test the algorithms on three different fMRI datasets. A brief description of the datasets is tabulated below (Table 1). All results in this section were obtained using the (5,3) reversible transform in all dimensions.

Table 1. fMRI datasets.

File Name	Task	Source	Dimension (x,y,z,t)	Original Bits/pixel	File Size (in MB)	Resolution in 'xy' and z respectively (mm)
mb01	Finger Tapping (Block Design)	Radiology Dept, U of A	64x64x21x100	13	17.1	3.75 and 5
siem	Unknown	Siemens Corporate Research	64x64x16x120	8	7.5	unknown
3T	Event Related	fMRI Data center	64x64x29x246	15	30.7	3.75 and 5

Table 2 gives the lossless results for fMRI using 2D, 3D and 4D JPEG2000. Different decompositions were explored and the best results are presented.

Table 2. Lossless compression performance using JPEG2000 (in bits/pixel).

Dataset \Rightarrow Original bit depth \Rightarrow Method \Downarrow	mb01 13	Pre-processed mb01 13	3T 15	siem 8
2D JPEG2000	7.2003	4.7266	9.9033	4.7772
3D JPEG2000 'xyz' cubes	7.1734	4.5317	9.8253	4.7155
3D JPEG2000 'xyt' cubes	6.0149	3.4228	8.3502	4.0860
4D JPEG2000	5.9623	3.3846	8.3570	4.0560

To evaluate lossy performance, we first use mean squared error (MSE) as the distortion measure to evaluate the effects of lossy compression on fMRI. Fig. 8 shows the plot of the Peak Signal to Noise Ratio (PSNR) versus bitrate. Rate allocation is done across the 'xyt' cubes so that we get constant distortion across cubes. The variation of PSNR across slices (for different 'xyt' cubes) with and without rate allocation is shown in Fig. 9.

As mentioned previously, we take two approaches to evaluate the effects of lossy compression on the analysis result. In the first case, the activation map of the original dataset is considered to be the ground truth. The activation map consists of the z-score of each voxel, which is related to the probability of the voxel being active. The z-score is thresholded to classify a voxel as active or not. Root Mean Squared Error (RMSE) between the original z-score and the z-score obtained after lossy compression is the first measure that we use here to quantify the effect of lossy compression, and is plotted as a function of compressed bit rate in Fig. 10. Typically a threshold of 3 is used for classifying the activation. Fig. 11 shows a plot of False Postive Rate (FPR) vs True Positive Rate (TPR) at different rates (threshold at 3). There is considerable drop in the TPR as opposed to increase in FPR due to compression. This might be attributed to the fact that the compression noise reduces the variance of the output, which weakens the activation signal in fMRI and hence reduces the TPR.

Secondly, we add artificial activations to the fMRI so that we know the ground truth. The Receiver Operating Characteristic (ROC) curve is then used to validate the activation points obtained by the analysis procedure. The closeness of the ROC curve of the original data with that of the compressed data will give a good picture of the compression effects on analysis. The ROC plot for the original and the compressed data (at different rates) are shown in Fig. 12. The numbers in the plot indicate the threshold that is used on the z-map.

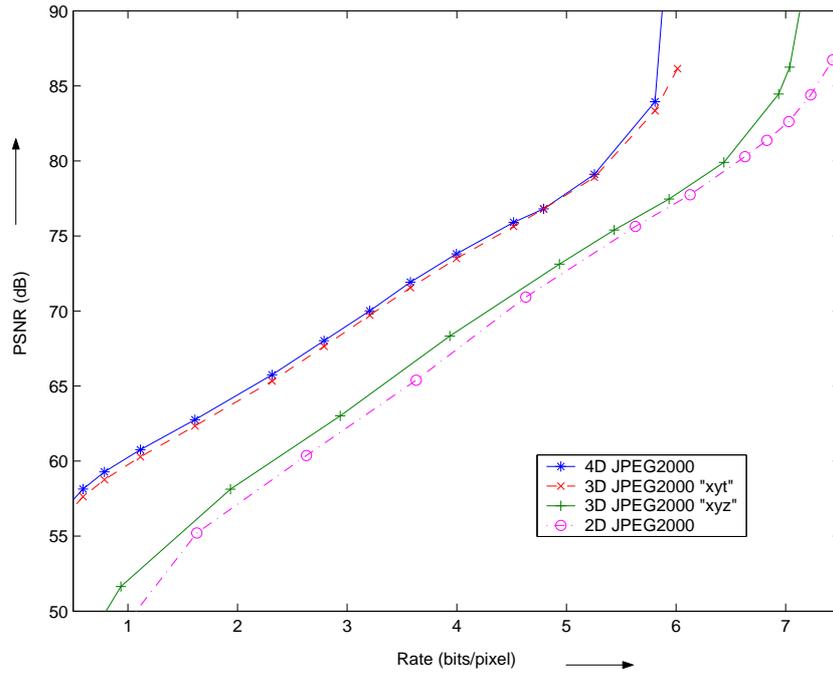


Figure 8. Rate vs distortion for mb01.

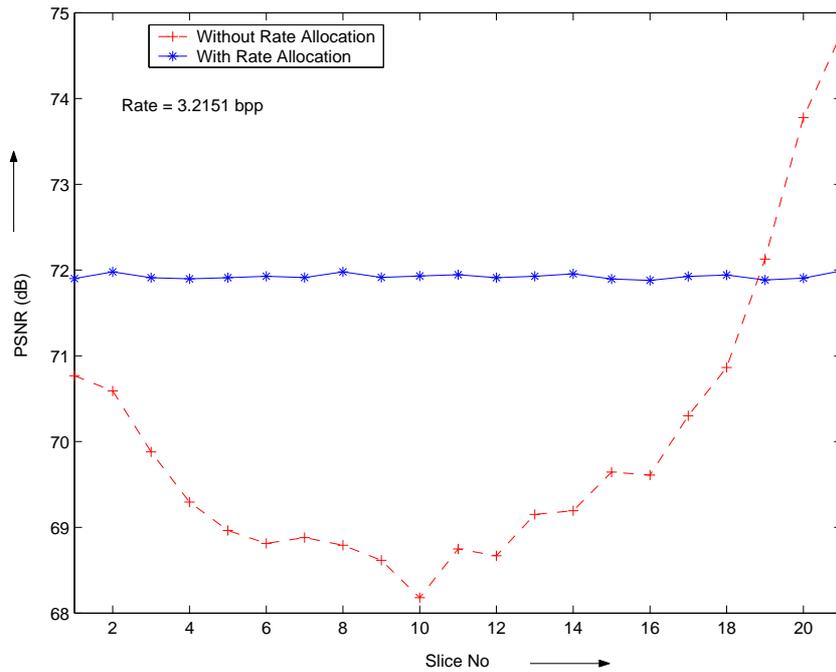


Figure 9. Distortion variation for "xyt" cubes across slices for mb01 (using 3D JPEG2000).

REFERENCES

1. S. G. Erberich, "PACS-based functional magnetic resonance imaging," *Computerized Medical Imaging and Graphics* **27**, pp. 229–240, Mar-Jun 2003.

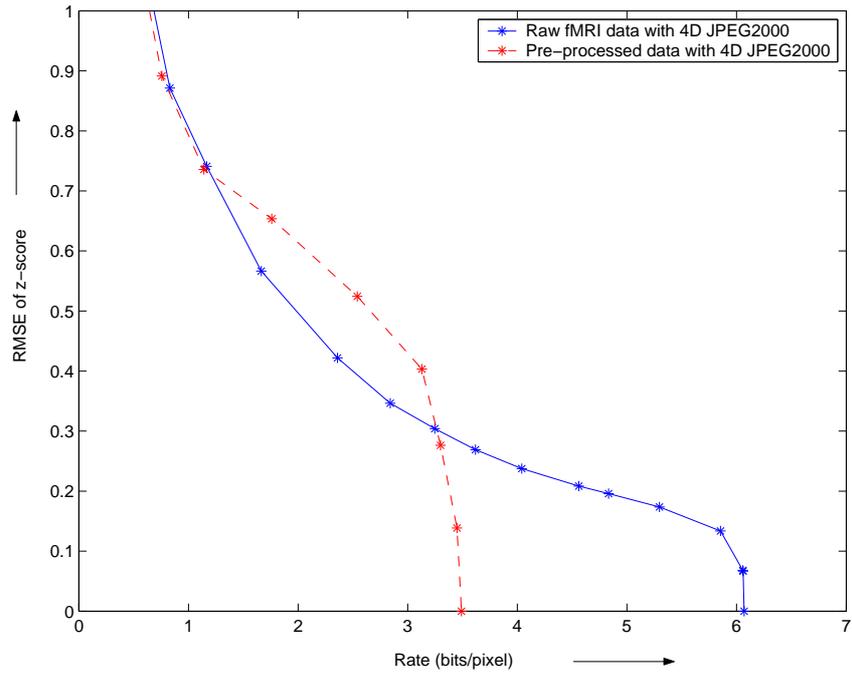


Figure 10. RMSE of z-scores.

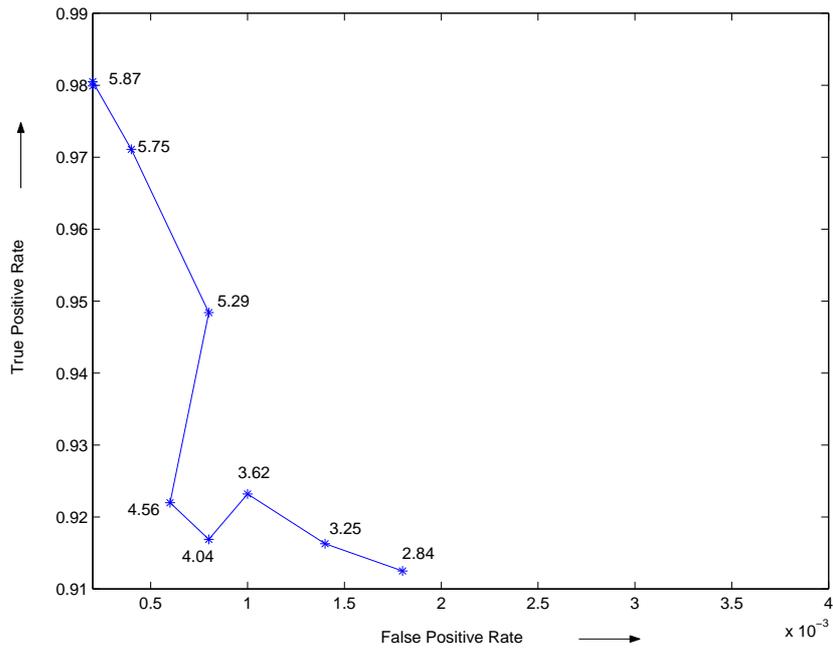


Figure 11. Plot of FPR vs. TPR at different rates.

2. D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Massachusetts: Kluwer Academic, 2002.
3. A. Bilgin, G. Zweig, and M. W. Marcellin, "Three-dimensional compression using integer wavelet transforms," *Applied Optics: Information Processing, Special Issue on Information Theory in Optoelectronic*

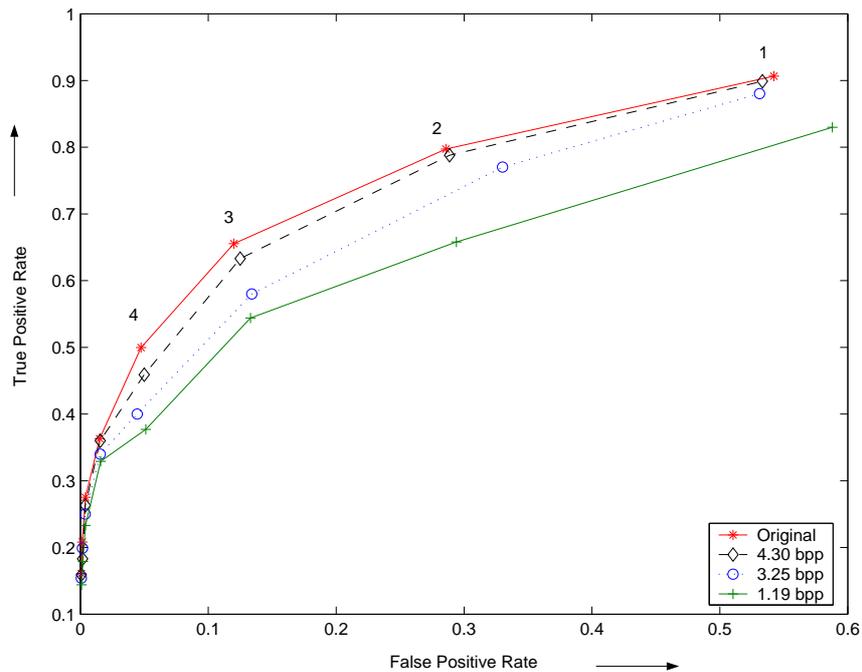


Figure 12. ROC plot of original and compressed data with 4D JPEG 2000.

systems **39**, pp. 1799–1814, Apr 2000.

4. Y. Kim and W. A. Pearlman, “Lossless volumetric medical image compression,” in *Applications of Digital Image Processing XXII, Proc. SPIE* **3808**, pp. 305–312, Jul 1999.
5. Z. Xiong, X. Wu, S. Cheng, and J. Hua, “Lossy-to-lossless compression of medical volumetric data using three-dimensional integer wavelet transforms,” *IEEE Transactions on Medical Imaging* **22**, pp. 459–470, Mar 2003.
6. P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. G. Nieto, and J. Cornelis, “Wavelet coding of volumetric medical datasets,” *IEEE Transactions on Medical Imaging* **22**, pp. 441–458, Mar 2003.
7. G. Menegaz and J. P. Thiran, “Three-dimensional encoding/two-dimensional decoding of medical data,” *IEEE Transactions on Medical Imaging* **22**, pp. 424–440, Mar 2003.
8. X. Wu and T. Qiu, “Wavelet coding of volumetric medical images for high throughput and operability,” to appear in *IEEE Transactions on Medical Imaging*.
9. C. Taswell, “Wavelet transform compression of functional magnetic resonance image sequences,” in *Proceedings of the IASTED International Conference on Signal and Image Processing*, pp. 28–31, Oct 1998.
10. M. S. Cohen, “A rapid and efficient data compression method for image time series,” *Human Brain Mapping* **12**, pp. 20–24, Jan 2001.
11. J. M. Shapiro, “Embedded image coding using zerotrees of wavelet coefficients,” *IEEE Trans. Signal Processing* **41**, pp. 3445–3462, Dec 1993.
12. L. Zeng, C. P. Jansen, S. Marsch, M. Unser, and P. R. Hunziker, “Four-dimensional wavelet compression of arbitrarily sized echocardiographic data,” *IEEE Transactions on Medical Imaging* **21**, pp. 1179–1187, Sept 2002.
13. R. B. Buxton, *Introduction to Functional Magnetic Resonance Imaging*, Cambridge University, 1992.
14. P. M. M. P. Jezzard and S. M. Smith, *Functional MRI: An Introduction to Methods*, Oxford University, 2001.
15. <http://www.fmrib.ox.ac.uk/fsl/>.
16. K. J. Worsley, C. H. Liao, J. Aston, V. Petre, G. H. Duncan, F. Morales, and A. Evans, “A general statistical analysis for fMRI data,” *NeuroImage* **15(1)**, pp. 1–15, Jan 2002.

17. K. M. Siddiqui, E. L. Siegel, B. I. Reiner, O. Crave, J. P. Johnson, Z. Wu, J. C. Dagher, A. Bilgin, M. W. Marcellin, and M. Nadar, "Improved compressibility of multi-slice CT datasets using 3D JPEG2000 compression," in *Computer Assisted Radiology and Surgery*, pp. 28–31, Jun 2004.
18. P. Skudlarski, R. T. Constable, and J. C. Core, "ROC analysis of statistical methods used in functional MRI: individual subjects," *NeuroImage* **9**, pp. 311–329.