Statistical and clinical evaluation of JPEG2000 for digital medical images

Min-Mo Sung

Department of Medical Science

The Graduate School, Yonsei University

Statistical and clinical evaluation of JPEG2000 for digital medical images

Directed by Professor Hee-Joung Kim

The Master's Thesis
submitted to the Department of Medical Science,
the Graduate School of Yonsei University
in partial fulfillment of the requirements
for the degree of Master of Medical Science

Min-Mo Sung

December 2003

This certifies that the Master's Thesis of Min-Mo Sung is approved.

Thesis Supervisor : Hee-Joung Kim

Thesis Committee : Jin-Suck Suh

Thesis Committee : Nam-Hyun Kim

The Graduate School
Yonsei University

December 2003

ACKNOWLEDGEMENTS

본 논문이 완성되기까지 세심한 배려와 깊은 관심으로 이끌어 주신 김희중 지도 교수님께 진심으로 감사드립니다. 또한 연구를 진행하는 동안 많은 도움과 가르침을 주신 서진석 교수님과 김남현 교수님, 부족한 부분을 지적해 주시고 충고와 조언을 아끼지 않으셨던 전태주 교수님, 항상 맑은 웃음과 따스함으로 편안함을 주신 최병욱 교수님, 연구진행간 많은 아이디어와 도움으로 격려해주신 김은경 교수님께 감사드립니다. 그리고 대학원 생활동안 많은 도움과 배려로 학생들을 위해주신 유선국 교수님, 정해조 박사님, 지창룡 선생님, 인피니트에 최승욱 이사님, 선배로서 많은 힘이 되어주었던 진오형, 옆에서 지켜보는 것만으로도 보고 배울 것이 많았던 하규형, 항상 털털한 웃음을 지었던 동욱이 형, 열심히 책상 앞을 지키면서 무언가 할 일이 그렇게도 많았던 원석, 순일, 재훈, 상호, 동현, 새롬 우리 의학영상연구실 친구들 그리고 창룡이형, 기창이형을 비롯한 의공학 교실 선후배님들 모두 감사합니다. 동기사랑 나라사랑 민성, 응주, 덕헌, 진규, 한우, 종완, 기홍, 병철 동기들 그리고 순만이 형을 비롯한 선후배님들, 괴롭고, 힘들때 마다 곁에서 힘내라 응원해준 용준, 윤수, 진홍, 경환, 정환, 재연, 보람, 동관, 기원 우리 대치동 친구들 모두에게 고맙단 말을 꼭 하고 싶습니다. 형 노릇도 못하면서 항상 따라주고 함께해준 봉재, 지만, 영찬, 동호회 친구들, 저에게 소중한 친구 혜원, 연경, 은석, 히로끼, 다이스케, KJ 그리고 저를 아는 모든 분들께 감사드립니다. 끝으로 오늘의 제가 있기까지 저를 믿고 끝임 없는 사랑과 보살핌으로 아껴주신 부모님과 동생 병모에게 기쁨으로 감사의 맘을 전합니다.

2003. 11.

성민모 드림

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Statistical and clinical Evaluation of JPEG2000
for Digital Medical Images

Min-Mo Sung

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Hee-Joung Kim)

Digital medical Images such as Computed Radiography (CR), Digital Mammography (MG) and even Positron Emission Tomography (PET) images will require large storage facilities and long transmission times for Picture Archiving and Communications System (PACS) implementation. Recently, the DICOM Working-Group 4 adopted JPEG2000 compression algorithm in DICOM standard to better utilize medical images. The wavelet-based JPEG2000 can achieve higher compression ratio with less distortion than the DCT-based JPEG. However, the clinical and statistical evaluation of JPEG2000 compressed for medical images has not been comprehensively tested in terms of clinical perspective. The purpose of this study was to evaluate the compression ratios of JPEG2000 for digital medical images using Receiver Operating Characteristic (ROC) and t-test analysis. Sixty CR chest images, twenty mammography images and twenty-one PET brain images were selected at YUMC (Yonsei University Medical Center). All images were compressed using JJ2000 public domain software, as recommended by the Joint Photographic Expert Group. The compression ratios applied for the studies were : 5:1, 10:1, 15:1, 20:1, 25:1, 30:1, 35:1, 40:1, 45:1, 50:1, 55:1, 60:1, 65:1, 70:1, 75:1 and 80:1. The original and reconstructed images were displayed in random order, 10 seconds per image, on a viewing monitor and evaluated those images with ROC, t-test analysis by 3 radiologists. The traditional statistical quality measures such as Peak Signal-to-Noise Ratio (PSNR), which is a commonly used measure for the evaluation of reconstructed images, measures how the reconstructed image differs from the original image by making pixel-by-pixel comparisons. In case of digital medical images, however, the ability of how accurate to discriminate diseased cases from normal cases relative to the quality of the reconstructed images is evaluated using ROC analysis. ROC curves can be used to compare the diagnostic performance of two or more reconstructed images. The t-test can be also used to evaluate the subjective image quality of reconstructed images. The results of t-test suggested that the possible compression ratios using JPEG2000 for medical images such as a CR, MG and PET images may be as much as 20:1, 15:1 and 25:1 without visual loss or with preserving significant medical information at a confidence level of 99%, respectively. Otherwise, the results of ROC analysis for evaluating lesion detectability showed that there was almost no difference for compression ratios up to 10:1 and 20:1, at a 99% of confidence level, and lesion detectability was very close to that of the original images even for compression ratios up to 60:1, 80:1 in CR and MG images, respectively.

Key words : JPEG2000, JPEG, Computed Radiography (CR), Mammography (MG), Positron Emission Tomography (PET), ROC analysis, t-test

Statistical and Clinical Evaluation of JPEG2000 for Digital Medical Images

Min-Mo Sung

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Hee-Joung Kim)

I. INTRODUCTION

Picture Archiving and Communications System (PACS) requires image compression to speed up the image transmission rate and to save disk space to reduce the cost of PACS implementation, despite the rapid advances made in storage and transmission technology. Presently, Yonsei University Medical Center (YUMC) archives 20-30GB of digital images per day in Full-PACS of General Electric Company (GE) was installed last year[1]. Therefore, the envisaged YUMC PACS facility requires an efficient method for reducing the amount of data stored, which are likely to be mainly radiography image, either from scanned film or direct digital acquisition. These images require high image quality to be clinically useful.

Image compression is largely a process of discarding some information, while maintaining relevant diagnostic information. Lossless compression allows exact recovery of the original image, and is certainly the obvious choice for medical imaging applications, because it does not affect image quality. However, it achieves very modest compression ratios, typically from 1.5 to 3:1. On the other hand, lossy compression method doesn't allow exact recovery after compression, although it allows much higher compression ratios. Generally, higher compression ratios can be obtained at the expense of more image degradation. That is, image quality degrades as the compression ratio increases. Image degradation because of lossy compression may or may not be visually apparent. The digital images reconstructed from the lossy compressed images must be very close to the originals in image quality and must preserve all the significant information of the original images for clinical diagnostic applications. The related term used by the American College of Radiology and the National Electrical Manufacturing Association (ACR/NEMA) is "information preserving." The ACR/NEMA standard report defines a compression scheme to be information preserving if the resulting image retains all of the significant information of the original image[2]. Both "visually lossless" and "information preserving" are subjective definitions and great caution should be taken in their interpretation. Therefore, research efforts have been focused on lossy compression algorithms and the evaluation of low bit rates that discard image data of no diagnostic significance but retain medically relevant information.

The DICOM (digital imaging and communications in medicine) standard provides a compression mechanism for supporting the use of JPEG, Run-Length-Encoding, and JPEG-LS image compression[3]. In particular, the JPEG standard is the primary compression technology within DICOM. The JPEG standard includes lossy and lossless compression techniques, and both type of compression have been used in medical imaging applications. Despite its popularity, it has been long recognized that the current JPEG standard has certain limitations. One limitation is the relatively poor compression efficiency of its lossless method, and another is the presence of objectionable blocking artifacts that can occur at high compression ratios with its lossy method[4]. The lossy method of JPEG is also limited in terms of input images, as it has a maximum bit depth of 12bit/pixel. Therefore, the DICOM subcommittee, WG4, which is responsible for image compression has been reviewing JPEG2000. It provides a new compression algorithm based on the use of a wavelet technique for medical image compression with ISO/IEC JTC1/SC29/WG1[5].

JPEG2000 can provide significantly higher compression ratios than the JPEG technique with less degradation and distortion. JPEG2000 also provides various features, which may be advantageous for medical imaging applications. The functionalities of JPEG2000 include integrated lossy/lossless compression, region-of-interest (ROI) encoding, inherent multi-resolution capability, and progressive decoding[6,7]. Therefore, the JPEG2000 image compression algorithm will be adopted in DICOM standard by DICOM subcommittee, WG4, which is working Supplement 61(: JPEG2000 Transfer Syntaxes).

This study report upon the results of the compression ratios possible with JPEG2000 for clinical diagnostic applications on digital radiography image.

II. MATERIALS AND METHODS

How does one decide if a reconstructed image is good enough for a specific application, such as diagnosis? We describe two approaches to the measurement of medical image quality for clinical diagnostic applications in JPEG2000 algorithm : ROC analysis and the t-test.

1. JPEG2000 Alogrithm

JPEG2000 offers similar or improved image quality performance than the current JPEG standard for the compression of medical images and provides the richest available set of features efficiently within a unified algorithm. These features are multiple-resolution and quality scalability compression, region of interest encoding, client-side region of decoding, progressive display and various dynamic range (e.g. 1 bit to 16 bit) for each color component[6,7].

A. Resolution and Quality Scalability

JPEG2000 allows reduced resolution images to be quickly and efficiently extracted from a compressed image. With JPEG2000, sub resolution images of various size can be extracted from the compressed file and do not need to be stored separately. It also allows reduced quality images to be extractd from a compressed image, producing images with higher compressed ratios optimized for specific workflows. For example, a medium quality image can be extracted and transmitted for browing from an original image on the server.

B. Region of Interest Encoding

Region of Interest (ROI) encoding allows important image features to be compressed at a high quality than the background, trimming image size without sacrificing important details. When combined with progressive display, ROI encoding also allows the foreground to be transmitted and display first, followed by the image background.

C. Client-side Region of Interest Decoding

JPEG2000 enables a client like a radiologist to interactively select a Region of Interest to decode at a higher quality or resolution. The radiologist will retrieve only the portion of the JPEG2000 code stream needed to display the selected area at a higher quality on screen. The user can select any area or multiple area for enhanced viewing. Use of client-side Region of Interest decoding maximizes bandwidth efficiency by quickly providing just enough information.

D. Progressive Display

JPEG2000's progressive transmission format displays a low resolution version of an image after only a small portion of the file has been received. As more data arrives, the display is progressively refined until the full resolution image is shown. This feature allows a radiologist to quickly orient themselves to an image, reducing the time spent waiting for data to arrive.

Many of these features are valuable in a medical imaging PACS environment. To perform these tasks, JPEG2000 makes use of several recent advancements in compression technology as I mentioned above. In particular, a two-dimensional

DWT (discrete wavelet transform) is at the heart of JPEG2000. The application of the DWT to image compression has been the subject of extensive research over the past decade[8]. The other basic encoding process, which converts source image data files into a compressed form, is organized into scalar quantization and arithmetic encoding such as a EBCOT algorithm[6,8]. The decoding process, which converts compressed image data into a reconstructed image, is the inverse of the encoding process (Fig. 1).

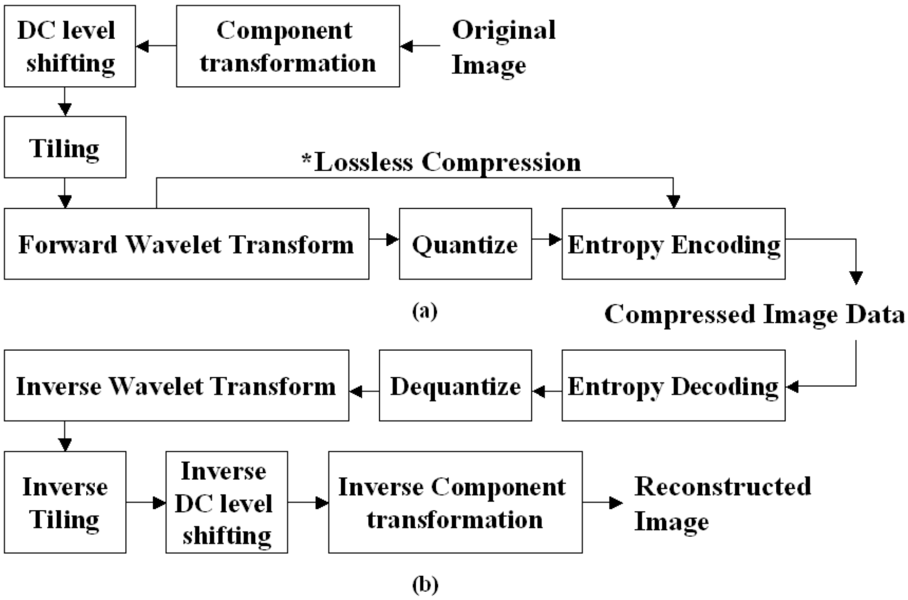


Fig. 1 JPEG2000 encoder/decoder. In encode procedure (a), the original image is decomposed into component, which are to be decomposed into rectangle tiles. If the components are unsigned, DC level shifting is performed. DWT is applied on each tile. After quantizing the sub bands of the transformed coefficients, the bit-planes of quantized coefficients are entropy coded. In decoding procedure (b), the decoding process is the inverse of the encoding process

2. ROC Analysis

The most common means of measuring image quality for reconstructed images is based on ROC analysis, because simulated measures resulting from PSNR (peak signal-to-noise ratio) do not always correlate well with human subjective assessments. Radiologists with expertise in diagnosing the subject diseases participate as observers to review the reconstructed images after compression[9]. For each reconstructed image, an individual radiologist is asked to give an ROC confidence rating on a scale of 1 to 5 representing his or her impression of the likelihood of the presence of the disease. Since the radiologists rated diagnostic usefulness rather than general appearance or edge patterns or simply line, the ROC study related detectability to compression level. The ratings of the original image are graded based on the "gold standard".

The area Az under the ROC curve (AUC) is an index of quantitative measure of the observer's performance on the reconstructed images. An area of 1 represents the perfect original images an area of 0.5 represents the worthless reconstructed images (Fig 2).

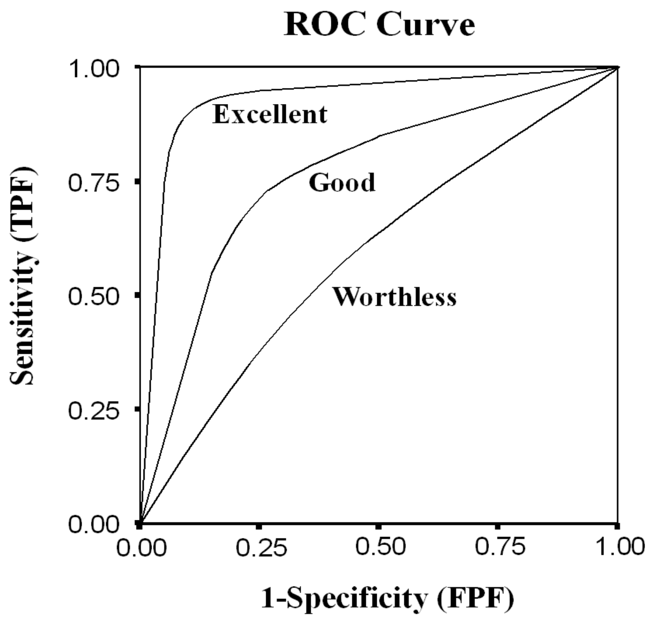


Fig. 2 ROC Curve Graph. This graph shows three ROC curves representing excellent, good, and poor tests plotted on the same graph. When the area of the ROC curve is very close to 1, we can say that the diagnostic performance will be as good as those made from the original image. The "Sensitivity" is the same as TPF (True Positive Fraction) the fraction of images with lesion that are correctly reported by the observer. The "1-Specificity"is the same as FPF (False Positive Fraction); that is, the fractions of the normal cases that is falsely reported as abnormal.

Thus, if the AUC is very close to 1, we can say that the diagnoses of this disease based on the reconstructed image with the predetermined compression ratio will be as good as those made from the original image. In other words, this compression ratio is acceptable for this image type[10,11]. A rough guide for the classification of the accuracy of diagnostic test is the traditional academic point system: That is, the AUC is (0.90 ~ 1.0 = excellent, 0.80 ~ 0.90 = good, 0.70 ~ 0.80 = fair, 0.60 ~ 0.70 = poor, 0.50 ~ 0.60 = fail) [12].

3. t-test Analysis

The subjective confidence rating of a reconstructed image can also be judged using the t-test. The t-test is the most commonly used method to evaluate differences in means of two groups. In this case, the hypotheses are:

H_o: $u_1 = u_2$ (u_1, u_2 : means of two groups)
Null hypothesis: There is no difference of image quality between the original and the reconstructed images.

H_a: $u_1 <> u_2$ (u_1, u_2 : means of two groups)
Alternative: There is difference of image quality between the original and the reconstructed images.

The reconstructed and original images can be presented to radiologists who rate them on a scale of 1 to 5 [13]. Therefore, two groups of images can be analyzed into means, ranges, variances, standard deviations, and standard errors and the presence of a significant difference determined by using the t-test of the rating scores.

The t-value is a ratio, which is used to statistically analyze the t-test. The top part of the ratio, the numerator, is simply the difference between the means of the reconstructed group (u_c) and the means of the original group (u_o). The bottom part of the ratio, the denominator, is a measure of the variability of the rated scores. In other words, the t-value is obtained by dividing the difference between the two means by the standard error of the difference, which is like the standard deviation in the rated score.

The t-value is give as

$$t\text{-}value = \frac{difference\ between\ group\ means}{variability\ of\ groups}$$

$$= \frac{u_c - u_o}{\sqrt{\frac{\mathrm{var}_c}{n_c} + \frac{\mathrm{var}_o}{n_o}}} \qquad \text{.............(1)}$$

var_c, var_o (variability of reconstructed and original group, respectively)

n_c, n_o (the number of reconstructed and original images, respectively)

If the t-value is not big enough, we will fail to reject the Null Hypothesis and conclude that there is no difference between the original and the reconstructed images. If the t-value is big enough, there will be a difference between the original and the reconstructed images and the Null Hypothesis will be rejected[14].

The p-value reported with a t-test represents the probability of the error involved when accepting this hypothesis about the existence of a difference. Technically speaking, this is the probability of an error associated with rejecting the hypothesis of no difference between the two evaluations of radiologists. Therefore, a low p-value can be interpreted as a rejection of the Null Hypothesis and to reflect real differences between the compressed and the original images. It is conventional in statistics to reject the null hypothesis at the 1% or 5% level.

4. Method

A. Acquisition of Computed Radiography images

Sixty CR chest images which were included 31 with abnormal lesions and 29 normal images were selected at YUMC. They were obtained with a pixel dimension of 1760 x 2140 and 10 bit-rate from a FUJI CR system. Window width and level were 783 and 1023, respectively.

B. Acquisition of Mammography images

Twenty screen-film mammographic images (8"x10") were selected at YUMC. They include 12 images with abnormal lesions with mass or micro-calcification clusters and 8 normal images. The twenty screen-film mammographic images were digitized as 200dpi (dots per inch) using a UMAX PowerLook 2100XL scanner and stored as grayscale images with 8 bit per pixel (bpp) and 1360 x 1760 pixel size. Automatic brightness/contrast settings were applied.

C. Acquisition of Positron Emission Tomography

PET brain image data of twenty-one adult patients were obtained with GE advance PET scanner after injecting approximately 341 MBq of FDG. Floating point PET brain images were converted into images with 12 bits/pixel and 16 bits/pixel integer images by Analyze software.

The computer platform used for this study was an IBM with Intel PIII-860, 512MB RAM and 32MB of video RAM. All images were compressed using

JJ2000 public domain software, as recommended by the Joint Photographic Expert Group. The compression ratios applied for the studies were : 5:1, 10:1, 15:1, 20:1, 25:1, 30:1, 35:1, 40:1, 45:1, 50:1, 55:1, 60:1, 65:1, 70:1, 75:1 and 80:1. All images were untiled to enhance image quality and were decomposed to the 5th level. The original and reconstructed images were displayed in random order, 10 seconds per image, on a viewing monitor and evaluated by 3 radiologists. All images were displayed using a TOTOKU MDL2102A Monochrome LCD 1.5K Monitor (TOTOKU Electric co. Ltd., Nagano, Japan) with an IrfanView v3.61 image viewing program (Irfan Skiljan, Graduate of Vienna University of Technology, Austria).

To evaluate the lesion detectability relative to image quality on CR and MG images, we performed ROC analysis. For all images, the 3 radiologists were independently asked to give a ROC confidence rating on a scale of 1 to 5, corresponding to the likelihood of lesion presence, with 5 indicating definite presence and 1 definite absence. Confidence values 4 and 2 indicated that the lesion was probably present and probably absent, respectively. A confidence value of 3 indicated that the lesion presence was equivocal or indeterminate. ROC analysis used the original images as a gold standard, because any difference could only favor the uncompressed images[9,14].

To evaluate the image quality relative to compression ratio on CR, MG and PET images, we used a paired sample t-test. Three radiologists were also independently asked to rate their subjective assessment from 1 to 5 of the quality of each image for lesion detectability. Confidence values 5 and 1 indicated that the image quality was definitely acceptable and definitely unacceptable for diagnosis, respectively, 4 and 2 that it was probably acceptable and probably unacceptable, respectively, and 3 indicated equivocal or indeterminate quality. The t-test also used the original images as a gold standard[9,14].

The reconstructed images were quantitatively evaluated by performing ROC analysis and t-test upon the results of the 3 radiologist's evaluations of lesion detectability and image quality, through a comparison of the reconstructed images with the original images.

The results of ROC and t-test were obtained using SPSS 9.0 statistics software (SPSS Inc, Chicago, USA) for evaluating reconstructed images.

III. RESULTS

1. ROC analysis results of CR images

Table 1. shows the results of ROC analysis, and includes estimates of the area, the standard error of the area, and the confidence limits of the area. The area and standard error have values near 1 and 0, respectively, which indicate that the degree of lesion detectability with reconstructed images is almost always close to that of the original images.

Compression Ratio	Area	Std. Error	Asymptotic Sig.	Asymptotic 99% Confidence Interval	
				Lower Bound	Upper Bound
5:1	0.986	0.010	0.00	0.959	1.012
10:1	0.986	0.010	0.00	0.959	1.012
15:1	0.954	0.017	0.00	0.909	0.999
20:1	0.946	0.017	0.00	0.903	0.988
30:1	0.952	0.018	0.00	0.905	0.999
40:1	0.951	0.016	0.00	0.909	0.992
60:1	0.907	0.024	0.00	0.845	0.969
80:1	0.899	0.025	0.00	0.834	0.965

Table 1. Results of ROC analysis on CR images for various compression ratios at a confidence level of 99 % (The standard errors were negligible).

The results of AUC for evaluating lesion detectability demonstrated that there was almost no difference for compression ratios up to 10:1 (AUC 0.986), at a 99% confidence level, and that lesion detectability was very close to that of the original images even for compression ratios up to 60:1 (AUC 0.954~0.907). (Fig. 3) Therefore, compressions of up to 60:1 may be acceptable for clinical diagnostic applications, since the AUC for lesion detectability differed only marginally between the reconstructed and original images.

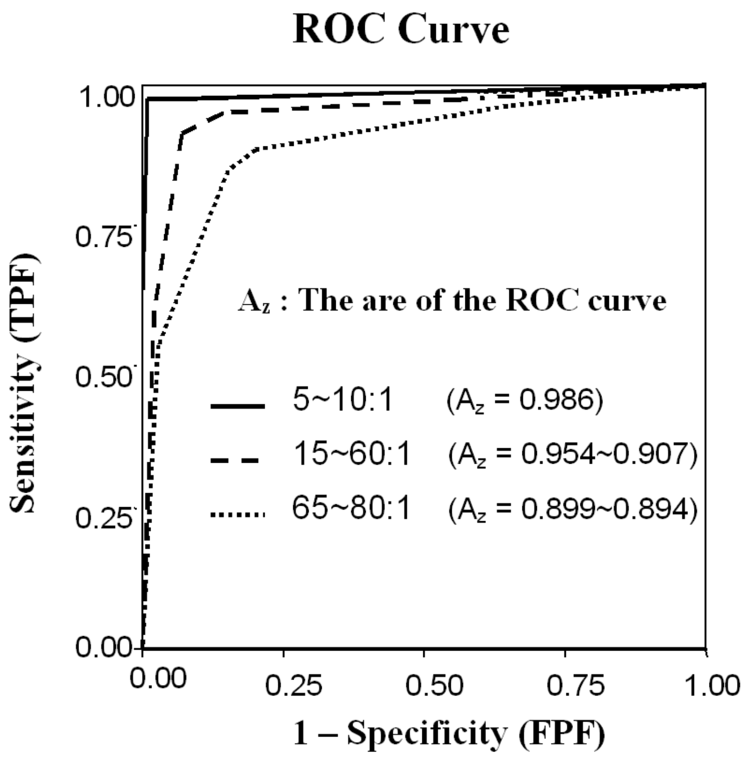


Fig. 3 Graph of ROC curve on CR images for various compression ratios at a confidence level of 99 %.

2. t-test results of CR images

Table 2. shows the results of the paired sample t-test obtained in the evaluation of image quality. These results indicated that the t-value (1.09) was not big enough at a compression ratio of ~15:1 and that no differences were found between the quality of the original and reconstructed images in the range 5~20:1, at the 99% confidence level (p-value>0.01). The t-test established that at a compression of 25:1~ there was a difference between the original and reconstructed images quality (p-value<0.01). Fig 4. shows an example of comparisons made between reconstructed images and the original image, using the JPEG2000 image compression technique.

	Paired Differences					t-value	p-value
	Mean	Std. Deviation	Std. Error Mean	99% Confidence Interval of the Difference			
				Lower	Upper		
Original 5:1	4.30 4.31	0.09	8.33e-03	-3.01e-02	1.35e-02	-1.0	**0.319**
Original 10:1	4.30 4.22	0.66	5.98e-02	-7.34e-02	0.240	1.39	**0.166**
Original 15:1	4.30 4.23	0.67	6.12e-02	-9.35e-02	0.227	1.09	**0.278**
Original 20:1	4.30 4.14	0.72	6.59e-02	-1.41e-02	0.331	2.40	**0.018**
Original 25:1	4.30 4.11	0.64	5.83e-02	3.89e-02	0.344	3.29	**0.001**
Original 30:1	4.30 4.07	0.65	5.88e-02	7.93e-02	0.387	3.96	**0.000**

Table 2. Results of paired sample t-test on CR images by comparing reconstructed images with original images at a confidence level of 99%.

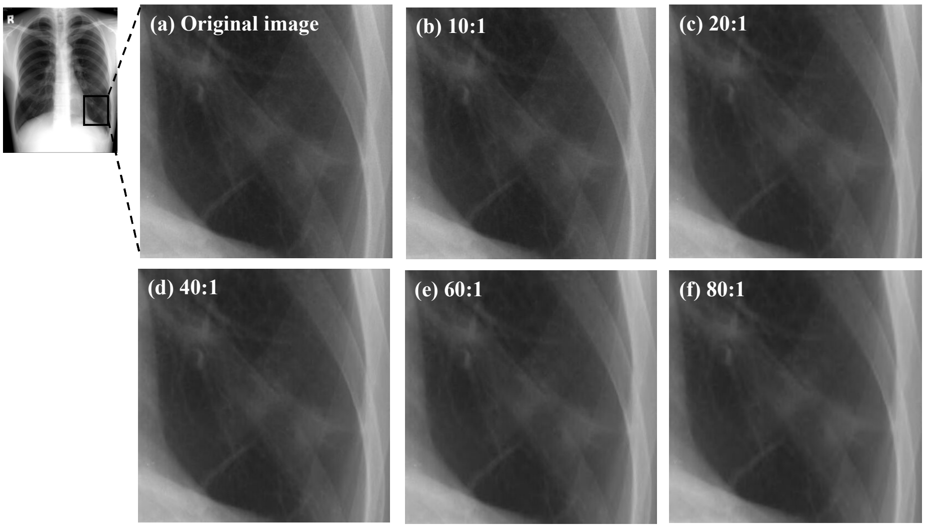


Fig. 4 The reconstructed CR images. It is apparent that distortion and degradation occurred as compression was increased.

3. ROC analysis results of MG images

Table 3. shows the results of ROC analysis, includes estimates of the area, the standard error of the area, and the confidence limits of the area. The results of ROC analysis for evaluating lesion detectability showed that there was almost no difference for compression ratios up to 20:1, at a 99% of confidence level, and lesion detectability was very close to that of the original images even for compression ratios up to 80:1 (Fig. 5).

Compression Ratio	Area	Std. Error	Asymptotic Sig.	Asymptotic 99% Confidence Interval	
				Lower Bound	Upper Bound
5:1	0.976	0.025	0.00	0.912	1.039
10:1	0.976	0.025	0.00	0.912	1.039
15:1	0.976	0.025	0.00	0.912	1.039
20:1	0.976	0.025	0.00	0.912	1.039
30:1	0.949	0.032	0.00	0.886	1.012
40:1	0.962	0.027	0.00	0.908	1.025
60:1	0.964	0.027	0.00	0.893	1.034
80:1	0.949	0.030	0.00	0.871	1.027

Table 3. Results of ROC analysis on MG images for various compression ratios at a confidence level of 99 % (The standard errors were negligible).

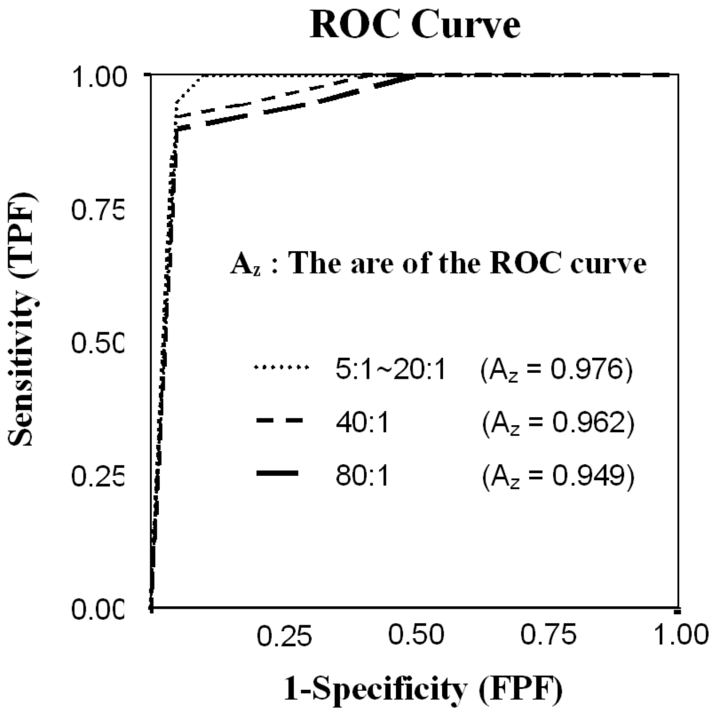


Fig. 5 Graph of ROC curve on MG images for various compression ratios at a confidence level of 99 %.

4. t-test results of MG images

Table 4. shows the results of the paired sample t-test obtain when evaluating subjective image quality. The mean value 4.03 of digitized original images quality showed that the original image quality was acceptable for diagnosis. The t-test showed that the t-value (1.22) is not big enough in ~ 15:1 and no difference was found between the quality of the original and reconstructed images in the range ~15:1 at the 99% confidence level (p>0.01). Fig 6. shows the results of comparisons made between reconstructed images using the JPEG2000 image compression technique and original image

	Paired Differences					**t-value**	**p-value**
	Mean	**Std. Deviation**	**Std. Error Mean**	**99% Confidence Interval of the Difference**			
				Lower	**Upper**		
Original 5:1	4.03 4.05	0.39	5.04e-02	-0.151	0.117	-0.33	**0.742**
Original 10:1	4.03 4.13	0.48	6.15e-02	-0.264	6.38e-02	-1.63	**0.109**
Original 15:1	4.03 3.95	0.53	6.84e-02	-9.88e-02	0.266	1.22	**0.228**
Original 20:1	4.03 3.83	0.55	7.05e-02	1.23e-02	0.388	2.84	**0.006**
Original 25:1	4.03 3.75	0.74	9.54e-02	2.95e-02	0.537	2.97	**0.004**
Original 30:1	4.03 3.45	0.74	9.59e-02	0.328	0.839	6.08	**0.000**

Table 4. The results of the paired sample t-test on MG images by comparing reconstructed images with original image.

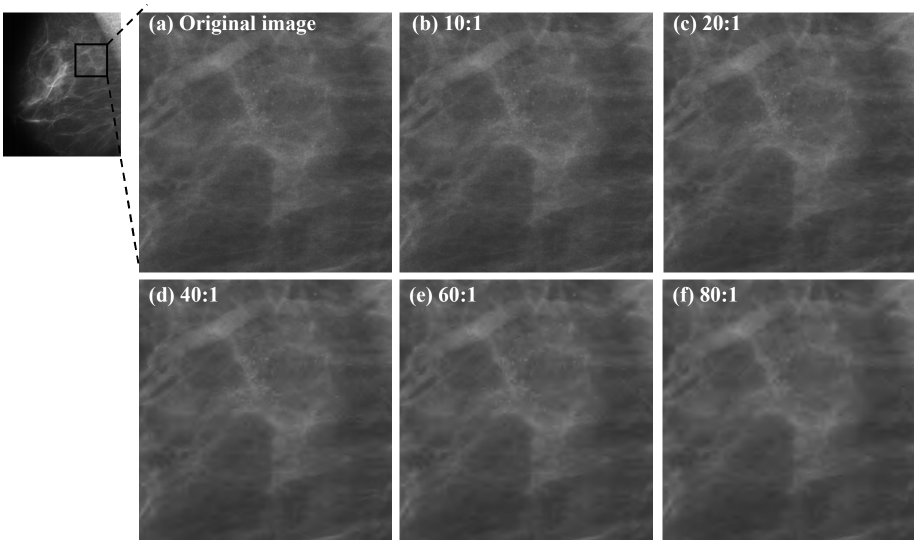


Fig. 6 The reconstructed MG images. The results of comparisons made between reconstructed images using the JPEG2000 image compression technique and original image, showing that distortion and degradation occurred as compression was increased.

5. t-test results of PET images

Table 5. shows the result of paired sample t-test on PET images by comparing compressed images quality with original image quality. In comparison with original image, There were no difference between 5:1~25:1 of compressed images and original image in 99% of confidence level (p>0.01, respectively). The results of t-test by comparing 30:1~ with original image (p<0.01) showed that there were some difference between original image and 30:1~.

	Paired Differences					t-value	p-value
	Mean	Std. Deviation	Std. Error Mean	99% Confidence Interval of the Difference			
				Lower	Upper		
Original 5:1	4.24 4.13	0.3333	0.1111	-0.4839	0.2617	0.000	**0.347**
Original 10:1	4.24 4.02	0.4410	0.1470	-0.7154	0.2710	1.512	**0.169**
Original 15:1	4.24 3.91	0.5000	0.1667	-0.8926	0.2259	2.000	**0.081**
Original 20:1	4.24 3.80	0.5270	0.1757	-1.0339	0.1450	2.530	**0.050**
Original 25:1	4.24 3.58	0.8660	0.2887	-1.6353	0.3020	2.309	**0.035**
Original 30:1	4.24 3.02	0.6667	0.2222	-1.9679	-0.4766	5.500	**0.001**

Table 5. The result of t-test on PET images by comparing compressed images with original image.

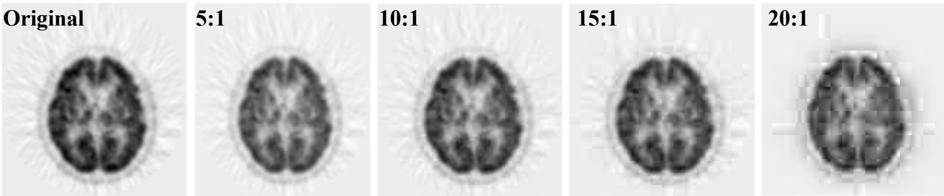
6. t-test results of PET images with JPEG

Table 6. shows the result of paired sample t-test on PET images which are compressed with JPEG by comparing compressed images quality with original image quality.

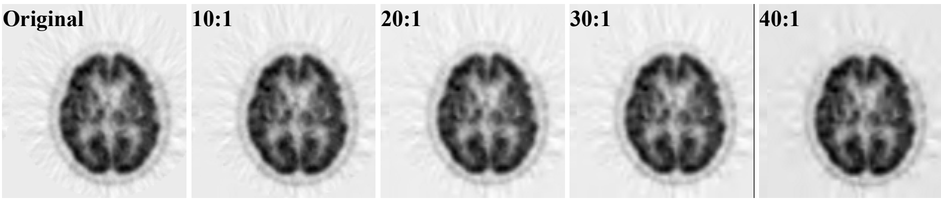
In comparison with original image, there were no difference between 5:1~10:1 of compressed images and original image in 99% of confidence level ($p>0.01$, respectively). The results of paired sample t-test by comparing 15:1 with original image quality ($p<0.01$) showed that there were some difference between original image and 15:1~. Fig. 7 shows the comparison of JPEG compressed images and JPEG2000 compressed images on PET brain images.

	Paired Differences						
				99% Confidence Interval of the Difference			
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t-value	p-value
Original 5:1	4.24 3.91	0.7071	0.2357	-1.1242	0.4574	1.414	**0.195**
Original 10:1	4.24 3.80	0.7265	0.2422	-1.2570	0.3681	1.835	**0.104**
Original 15:1	4.24 2.58	0.8660	0.2887	-2.635	-0.6980	5.774	**0.000**
Original 20:1	4.24 1.24	0.7071	0.2357	-3.7909	-2.2091	12.728	**0.000**

Table 6. The result of paired sample t-test on PET images which are compressed with JPEG by comparing compressed images with original image.



(a) JPEG Images



(b) JPEG2000 Images

Fig. 7 The comparison of JPEG compressed images and JPEG2000 compressed images on PET brain images. (a) These images show the result of comparing between compressed images by using JPEG image compression technique and original image. In this case, radiologists could find distortion and degradation in 10:1~. (b) These images show the result of comparing between compressed images by using JPEG2000 image compression technique and original image. In this case, radiologists could find distortion and degradation in 25:1~.

IV. DISCUSSION

Compression technology has advanced to a point that allows medical images to be compressed with very high compression ratios without affecting the clinical diagnostic performance. Highly correlated images such as CR and MG images could have very high compression efficiencies, because much of the image information can be concentrated into a small number of coefficients when transformed into the spatial frequency domain. The less significant coefficients of the transformed image can be quantized and dramatic file size reductions can be achieved with minimal image information loss. The ACR/NEMA standard report defines the term "information preserving" or "visually lossless", which means the reconstructed image retains all the significant information of the original image. The purpose of this study was to evaluate the "visually lossless" or "information preserving" compression ratios of JPEG2000 for CR, MG and PET images by performing ROC analysis and t-test. Many other compression studies undertaken in medical imaging applications indicate that it is conceivable to compress a radiography image to 10:1 or even higher without losing diagnostic quality. The maximum ratios achievable by the JPEG compression method has usually been found to be between 10-20:1 for projection radiography[15,16]. Recent studies using wavelet transforms have demonstrated the possibility of achieving compression ratios up to 20-30:1 in projection radiography without compromising image diagnostic quality[17,18]. For full-frame DCT method of JPEG, compression ratios of about 9.6:1 have been reported to cause no diagnostic degradation for subtle micro-calcifications of MG image[19]. In this study, the results of the statistically testing performed showed that there were no detectable differences at compression ratios up to 10:1 and 20:1 on CR and MG images at a confidence level of 99%, respectively . It also shows that lesion detectability is very good up to 60:1 and 80:1 on CR and MG images, respectively.

The results of the t-test suggested that the possible compression ratios using JPEG2000 for CR, MG and PET images may be as high as 20:1, 15:1 and 25:1, respectively. Each reconstructed image's quality is as same as the original image quality. These results indicate that at these compression levels visual loss did not occur and that significant medical information was conserved with confidence level of 99%. However, further studies will be needed to apply this conclusion for other imaging modalities and different medical applications.

V. CONCLUSION

The purpose of this study was to evaluate the compression ratios of JPEG2000 for digital medical images using Receiver Operating Characteristic (ROC) and t-test analysis. For this study, three kinds of digital medical images such as CR, MG and PET images were selected at YUMC and all images were compressed using JPEG2000 lossy compression algorithm. The compression ratios applied for the studies were : 5:1, 10:1, 15:1, 20:1, 25:1, 30:1, 35:1, 40:1, 45:1, 50:1, 55:1, 60:1, 65:1, 70:1, 75:1 and 80:1 and evaluated by 3 radiologists with ROC analysis and t-test. The results of t-test suggested that the possible compression ratios using JPEG2000 for medical images such as a CR, MG and PET images may be as much as 20:1, 15:1 and 25:1 without visual loss or with preserving significant medical information at a confidence level of 99%, respectively. Otherwise, the results of ROC analysis for evaluating lesion detectability showed that there was almost no difference for compression ratios up to 10:1 and 20:1, at a 99% of confidence level, and lesion detectability was very close to that of the original images even for compression ratios up to 60:1, 80:1 in CR and MG images, respectively.

Image compression facilitates PACS as an economically viable alternative to analog film-based systems by reducing the bit size required to store and represent images, while maintaining relevant diagnostic information. It also enables the fast transmission over a PACS network and workstation display of large medical images for diagnostic and review purposes. The JPEG2000 compression algorithm, which has been adopted as the DICOM standard for medical images, provides higher compression efficiency than JPEG, RLE and JPEG-Lossless compression algorithms.

REFERENCES

1. 성민모, 김희중, 홍진오, 정하규, 유형식. 연세의료원 Mini-Pacs system에서의 DICOM Validation 연구. 대한 PACS 학회지 2000;6:67-72.
2. American College of Radiology, National Electrical Manufacturers Association. Standards Publication for Data Compression Standard. Washington, (DC): NEMA publication; 1989.
3. National Electrical Manufacturers Association. Digital Imaging and Communications in Medicine (DICOM) PS 3.5-2000. Rosslyn (VA): NEMA Publication; 2000.
4. Persons K, Palisson P, Manduca A, Erickson BJ, Savcenko V. An analytical look at the effects of compression on medical images. J Digit Imaging 1997; 10-3: 60-66.
5. DICOM Standards Committee, Working Group 4 Compression. Supplement 61: JPEG2000 Transfer Syntaxes. Rosslyn (VA): NEMA Publication; 2001
6. Charilaos C, Skodras A, Ebrahimi T. The JPEG2000 still image coding system: An Overview. IEEE Trans. Consumer Electronics 2000; 46: 1103-1127.
7. Foos DH. JPEG2000 compression of medical imagery. Proc. SPIE 2000; 3980: 85-96.
8. Kim YM, Steven CH. Handbook of Medical Imaging vol. 3: Display and PACS. Bellingham (DC): SPIE Press; 2000
9. Sung MM, Kim HJ, Kim EK, Kwak JY, Yoo JK, Yoo HS. Clinical Evaluation of JPEG2000 compression Algorithm for Digital Mammography. IEEE Trans. on Nuclear Science. 2002; 49: 827-832.

10. Wong S, Zaremba L, Gooden D, Huang HK. Radiologic Image Compression-A Review. Proc. IEEE 1995; 83-2: 194-219.
11. Huang HK. PACS: Basic Principle and Applications. New York, (NY): Wiley-Liss; 1999
12. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29-36.
13. Perlmutter SM, Tseng C, Cosman PC, Li KCP, Olshen RA, Gray RM. Measurement accuracy as a measure of image quality in compressed MR chest scans. Proc. IEEE 1994; 1: 861-865.
14. Sung MM, Kim HJ, Yoo SK, Choi BW, Nam JE, Yoo HS. Clinical Evaluation of Compression Ratios using JPEG2000 on Computed Radiography Chest Images. J Digital Imaging 2002; 15: 78-83.
15. Good WF, Maits GS, Gur D. Joint photographic experts group (JPEG) compatible data compression of mammograms. J Digital Imaging 1994; 7: 123-132.
16. Yamamoto LG. Using JPEG image compression to facilitate telemedicine. Am J Emerg Med 1995; 13: 55-57.
17. Goldberg M. Application of wavelet compression to digitized radiographs. Am J Roentgenol 1994; 163: 463-468.
18. Saipetch P. Applying wavelet transforms with arithmetic coding to radiological image compression. IEEE Eng Med Biol 1995; 14: 587-593.
19. Chan HP. Image compression in digital mammography: effects on computerized detection of subtle microcalcifications. Med Phys 1996; 23: 1325-1336.

ABSTRACT(IN KOREAN)

JPEG2000 알고리즘을 이용한 Digital Medical 영상의 압축 비율별 임상적 평가 및 비교

<지도교수 김희중>

연세대학교 대학원 의과학과

성민모

의료영상의 표준인 DICOM (Digitial Imaging and Communications in Medicine)은 현재 RLE, JPEG, JPEG-Lossless 압축 알고리즘들을 의료영상의 압축표준으로 채택하고 있다. 가장 널리 쓰이고 있는 JPEG 압축알고리즘은 손실 압축방법으로 영상의 압축비율이 높은 반면, 영상의 quality가 떨어지고, 의료영상에 필요한 기능이 없다는 단점을 가지고 있다. 이에 DICOM 표준의 압축방식을 관리, 평가, 채택하는 소그룹, DICOM Working Group4는 새로운 압축방식인 JPEG2000압축 알고리즘을 DICOM 표준에 추가함으로써, 의료영상에서의 압축 효율성을 증대시키고자하였다 JPEG200은 의료영상에서 필요로 하는 기능을 대부분 지원하며, 우수한 압축 효율성을 지니고 있다. JPEG2000의 특징은 높은 압축 비율에서도 뛰어난 영상의 quality를 가지며, 하나의 부호화된 비트 스트림에서 무손실 및 손실 압축을 동시에 구현할 수 있고, Region of Interest (ROI) 기능을 제공한다. 또한 progressive display의 기능을 제공하며, 1bit-depth에서 16bit-depth까지 다양한 bit-depth를 영상의 압축이 가능하다. 본 연구의 목적은 디지털 의료영상들에 JPEG2000 lossy 알고리즘을 이용하여, 임상실험과 통계적인 평가를 통해 압축 비율별 영상의 화질을 비교하고자하는 것이다. 오리지날 영상과 비교하여 압축된 영상이 판독에 영향을 미치지 않는 범위의 압축비율을 알아보고자 ROC analysis 와 t-test analysis를 수행하였다. 본 실험을 위해 연세의료원에서 60개의 Computed

Radiography (CR) 영상, 20개의 Mammography (MG) 영상 그리고 21개의 Positron Emission Tomography (PET) 영상을 획득하였고, JPEG2000 lossy 알고리즘을 이용하여 5:1-80:1 까지 영상을 압축하였다. 영상은 10초 간격으로 random하게 display하였고, 각각 3명의 판독의에 의해 ROC analysis와 t-test 분석을 수행하였다. Lesion detectability를 평가 할 수 있는 ROC 분석과 오리지날 영상과 화질을 비교하기 위한 t-test은 1에서 5까지 lesion의 유무와 판독을 위한 영상의 적합성을 평가하여, 통계 분석을 수행하였다.
각각의 실험결과 t-test 분석에서 JPEG2000으로 압축한 CR영상은 20:1까지 오리지날 영상과 비교하여 화질의 차이가 없음을 확인하였으며, ROC 분석결과 60:1까지 lesion detectability가 매우 우수하였다. JPEG2000으로 압축한 MG 영상의 경우 t-test 분석 결과 15:1까지의 영상에서 오리지날 영상과의 화질 차이가 없었으며, 80:1까지 lesion detectablity가 매우 우수하였다. PET 영상은 25:1까지 압축된 영상에서의 화질의 차이가 없음을 알 수 있었다.

핵심되는 말 : JPEG2000, JPEG, Computed Radiography, Mammography, Positron Emission Tomography, ROC analysis, T-test.