Journal of Digital Imaging

Evaluation of the Image Quality of Temporal Subtraction Images Produced Automatically in a PACS Environment

Shuji Sakai, M.D., Ph.D.,[1] Hiroyasu Soeda, M.D.,[2] Akio Furuya, M.D.,[2] Hidetake Yabuuchi, M.D., Ph.D.,[2] Takashi Okafuji, M.D.,[2] Keiji Yamamoto, B.S.,[3] Hiroshi Honda, M.D., Ph.D.,[2] and Kunio Doi, Ph.D.[4]

The aim of the study is to evaluate the reliable production of temporal subtraction images in a picture archiving and communication system environment and to establish objective criteria for the evaluation of image quality. A total of 117 temporal subtraction chest images (55 in the upright position, 62 in the supine position) were obtained in five consecutive days. In all of these, we confirmed that there were no interval changes on the original images, and cases with diffuse lung disease were excluded. The temporal subtraction images were classified by three chest radiologists into five levels: 5, excellent; 4, good; 3, acceptable; 2, poor; and 1, very poor. The following were examined: (1) the yield of adequate quality of the temporal subtraction images; (2) whether the temporal subtraction images were obtained in the warping or nonwarping mode; and (3) the correlation of the overall subjective image quality with the relative shift angles, relative shift distances, and the standard deviation of gray levels in the temporal subtraction images. The percentages of acceptable temporal subtraction images were 100% and 66% in the upright and supine positions, respectively. Sixteen (26%) of the 62 supine-position images were made in nonwarping mode, whereas all upright images were made in warping mode. Significant correlations were obtained in the relative shift angle ($P < 0.05$), relative horizontal shift distance ($P < 0.05$), and standard deviation of gray levels ($P < 0.0001$). Temporal subtraction images with acceptable image quality were obtained in the upright position. The objective criteria may be useful for the evaluation of image quality.

KEY WORDS: Chest radiography, temporal subtraction, computer-assisted diagnosis, computer-assisted image interpretation, PACS

INTRODUCTION

Some computer-aided diagnosis (CAD) systems have already been manufactured, and, in many cases, evaluation has commenced at the clinical level.[1-3] The purpose of CAD is to achieve efficiency in diagnostic work by reducing reading time and improving accuracy. CAD has a great potential for future development by linking imaging equipment such as computed radiography (CR) equipment and flat panel detectors (FPDs) to a picture archiving and communications system (PACS).[4] The temporal subtraction technique is recognized as one of the most useful in CAD. Temporal subtraction is a technique in which a previous chest radiograph is automatically registered with and subtracted from a current radiograph to enhance interval changes.[3,5] It is difficult for radiologists to identify subtle interval changes, particularly in lesions that overlap normal anatomic structures such as ribs, great vessels, the heart,

and the diaphragm. In such cases, temporal subtraction images can be useful for enhancing changes in local opacities.[6] Some reports have indicated the usefulness of temporal subtraction images in the early detection of lung cancer or infection.[3,7,8]

Since January 2001, we have developed a system for the comparative interpretation of chest radiographs while referring to temporal subtraction images. A commercially available system for temporal subtraction was integrated into our hospital's PACS. We applied our technique not only to upright positioning images, but also to portable chest radiographs, primarily in the intensive care unit. In clinical practice, it is necessary that the system reliably produce temporal subtraction images of good quality; if the quality is not sufficient, excessive misregistration artifacts will make it difficult for radiologists to differentiate between true interval changes and simple artifacts. Furthermore, for temporal subtraction images, there is no objective method of evaluating image quality, and we must therefore depend exclusively on subjective methods for such evaluation.[9]

One purpose in this study was to evaluate the reliable production of temporal subtraction images in a PACS environment and to establish objective criteria for the evaluation of image quality by using both original images and temporal subtraction images.

MATERIALS AND METHODS

Data Acquisition and Storage of Digital Chest Radiographs

In January 2001, we integrated a commercially available system for temporal subtraction into the PACS of our hospital. The image data were obtained from CR (FCR 5501D and FCR 5000, Fuji Photo Film, Tokyo, Japan) or from digital radiography with an FPD (CXDI-11, Canon, Tokyo, Japan, and Thorax FD, Siemens, Erlangen, Germany) and recorded on a PACS (TOSPACS, Toshiba, Tokyo, Japan) with digital imaging and communications in medicine (DICOM) protocol. The image data from the CR and FPD were stored with 1,760 × 2,140 matrix and 10-bit grayscale, and a 2,200 × 2,675 matrix and 12-bit grayscale, respectively. All radiographs obtained in our hospital were transferred to the PACS server mentioned above as well as to an exclusive server (Truedia X/
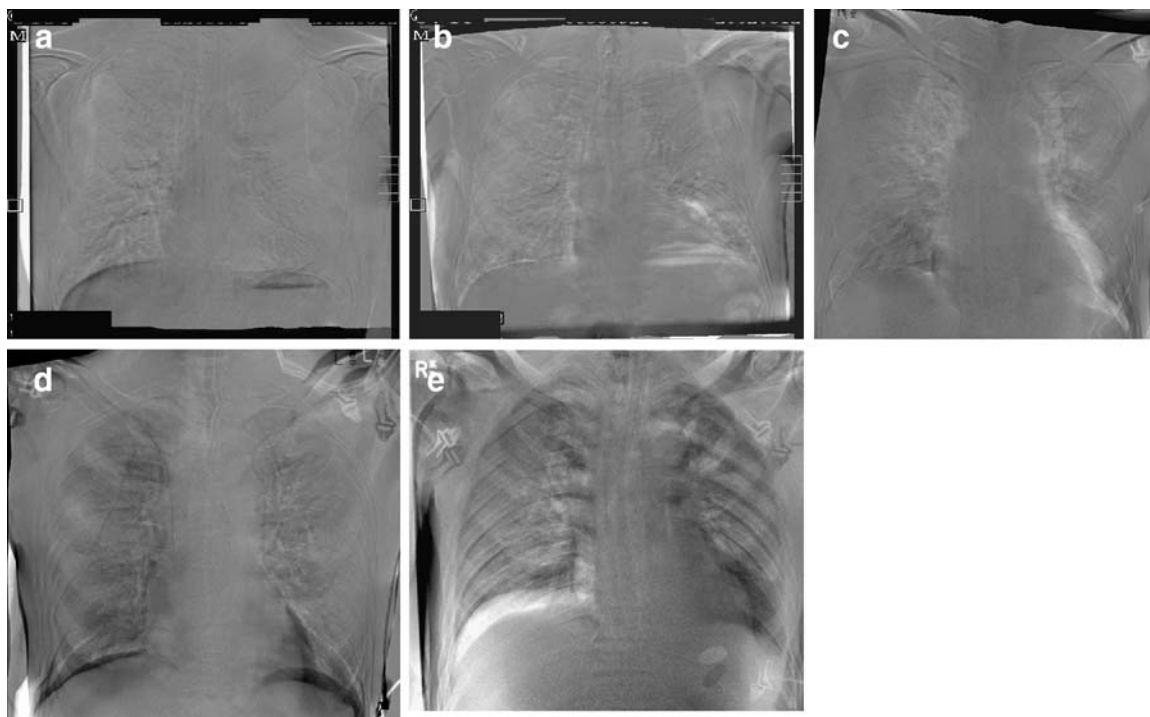

Fig 1. Examples of temporal subtraction images with each of the ve different scores. (a) Excellent (5 points); (b) good (4 points); (c) acceptable (3 points; (d) poor (2 points), and (e) very poor (1 point). Misregistration artifacts come into prominence in images with the lesser scores.
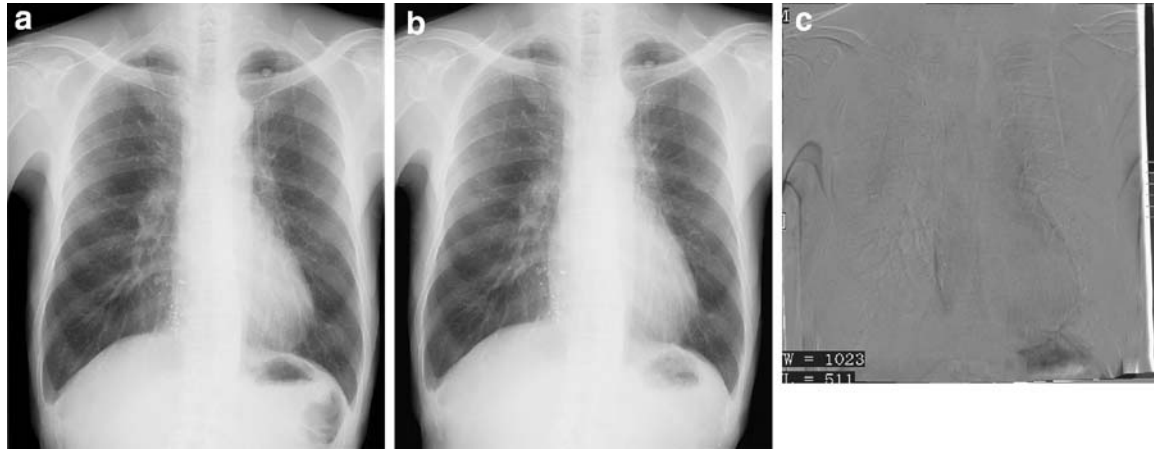
Fig 2. (a) Previous chest radiograph, (b) current radiograph, and (c) temporal subtraction image produced by warping mode and scored at 5 points. The temporal subtraction image shows only slight misregistration artifacts on the margin of the heart. The standard deviation of the grayscale of the temporal subtraction image was 41. The relative shift angle, vertical shift, and horizontal shift between the previous and current chest radiographs were 0.3°, 16 pixels, and 15 pixels, respectively.

R, Mitsubishi Space Software, Amagasaki, Japan) to produce a temporal subtraction image simultaneously. A previous image of the same patient with the same type of examination was selected by header information from DICOM and automatically loaded with a current image to a display terminal (TWS-2500, Toshiba) with two monochromatic cathode ray tubes (2,560 × 2,048 pixels, Barco, Kortrijk, Belgium) before interpretation. The temporal subtraction images were transferred to the same display terminal shortly thereafter. The link between PACS, the hospital information system (HIS), and the radiology information system (RIS) allows information about images and patients to be transferred automatically to a display terminal. We applied this technique not only to upright chest images, but also to portable chest radiographs, which are obtained primarily for intensive care unit or coronary care unit patients.

Production and Display of Temporal Subtraction Images

The server for producing temporal subtraction images can distinguish posteroanterior and anteroposterior chest radiographs from other radiographs such as lateral chest, abdomen, and bone radiographs, according to DICOM header information. In addition, the server can compress and store temporal subtraction images with a 586 × 586 matrix size. This server was designed to produce five temporal subtraction images automatically from the current and five previous radiographs. The CPU time for producing a single temporal subtraction image was 6 s; thus, the server requires 30 s to produce five temporal subtraction images upon the arrival of a new chest radiograph if it already has five or more previous radiographs of the same patient. When extraction of the thorax was possible
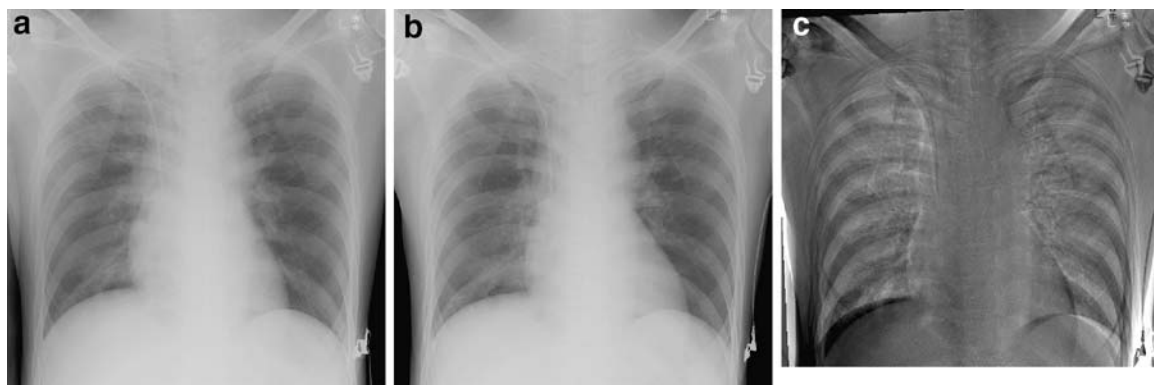
Fig 3. (a) Previous chest radiograph, (b) current radiograph, and (c) temporal subtraction image produced by warping mode and scored at 2 points. The temporal subtraction image shows remarkable misregistration artifact image of the thoracic cage. The standard deviation of the grayscale of the temporal subtraction image was 96. The relative shift angle, vertical shift, and horizontal shift between the previous and current chest radiographs were 1.8°, 0 pixels, and 12 pixels, respectively.

Table 1. Ratio of warping mode to nonwarping mode and sub-jective image quality in two positions

Position	Mode	Number	Subjective image quality
Upright			4.2
	Warping	55 (100%)	4.2
	Nonwarping	0 (0%)	
Supine			2.9
	Warping	46 (74%)	3.1 $P < 0.01$
	Nonwarping	16 (26%)	2.7

Mann–Whitney U test.

on both previous and current images, the present system produced temporal subtraction images using the warping mode after global matching. Namely, the warping mode is a registration method by using nonlinear distortion in addition to linear image shift and rotation. However, if this was not possible, the warping mode was not applied and only global matching was performed (nonwarping mode). In addition, when extraction of the thorax was possible on both the previous and current images, the relative shift angles and relative shift distances (vertical and horizontal) between current and previous images were calculated. Temporal subtraction images were displayed with a 586 × 586 matrix and 8-bit grayscale. The appearance of a new opacity was shown in black, and the disappearance of an existing opacity was shown in white.

Evaluation of Subjective Image Quality

A total of 218 temporal subtraction images were obtained on five consecutive days. Our institutional review board approved our study, and informed consent from all study subjects was obtained. Of these images, 76 temporal subtraction images obtained with FPDs were excluded. Because portable chest radiographs were obtained with CR only, the yield of subtraction images with reliable image quality can be compared between the two positions. We investigated the temporal subtraction images with particular attention to certain interval changes by using original images. Cases with diffuse lung disease were excluded by two chest radiologists. Consequently, 19 images were omitted because of the existence of certain interval changes, and six images were omitted because of diffuse lung disease. Finally, a total of 117 temporal subtraction images (upright position: 55 images, supine position: 62 images) were evaluated. The FCR 5501D was used for cases in the upright position and the FCR 5000 for those in the supine position. Three chest radiologists evaluated every temporal subtraction image at the same time based on the subjective image quality, and thus arrived at a consensus in scoring. These three observers, each of whom had at least 1 year of experience in interpreting chest radiographs using temporal subtraction images, rated the quality of each image at one of five levels: 5, excellent; 4, good; 3, acceptable; 2, poor; or 1, very poor (Fig. 1).[9] All temporal subtraction images were evaluated on the display terminal of our hospital by soft copy. We examined the yield of temporal subtraction images for each position, noting whether the images had been obtained in the warping or nonwarping mode. Differences in subjective image quality between these two modes were analyzed statistically by Mann–Whitney U test.

Establishment of Criteria for Objective Image Quality

All 117 temporal subtraction images and the corresponding original images were examined with an internal algorithm by use of an exclusive work station. First, the relative shift angles and relative shift distances (vertical and horizontal) between the previous and current images were calculated in cases in which the thorax could be extracted by the temporal subtraction system on all images. Specifically, the axis of thorax was extracted in previous and current image, and the relative shift angle between these axes was calculated. The locations of the thorax in the whole chest radiograph were identified on previous and current images, and the relative shift distances were calculated between them. Second, the standard deviation of the grayscale of each temporal subtraction image was calculated. Finally, the correlation of these parameters and the overall subjective image quality were examined. Statistical analysis was carried out by Spearman's correlation coefficient by rank (S.S.).

RESULTS

The average scores for overall subjective image quality in the upright and supine positions were 4.2 and 2.9, respectively. Subtraction images with scores of 3 or more (acceptable, good, and excellent) were considered to be acceptable for clinical use. The percentages of acceptable temporal subtraction images were 100% (55/55) in the upright and 66% (41/62) in the supine position in terms of the subjective overall image quality (Fig. 2). Sixteen of the 62 (26%) supine position images were made using the nonwarping mode, whereas all upright images were made using the warping mode (Fig. 3). There was a significant difference in subjective image quality between the 46 images produced with the warping mode and the 16 images obtained with the nonwarping mode (Table 1).

Table 2 shows the values of each parameter for objective image quality in the upright and supine positions. The means and standard deviations of the relative shift angle and the distance between the

Table 2. Value of each parameter for objective image qualities in two positions

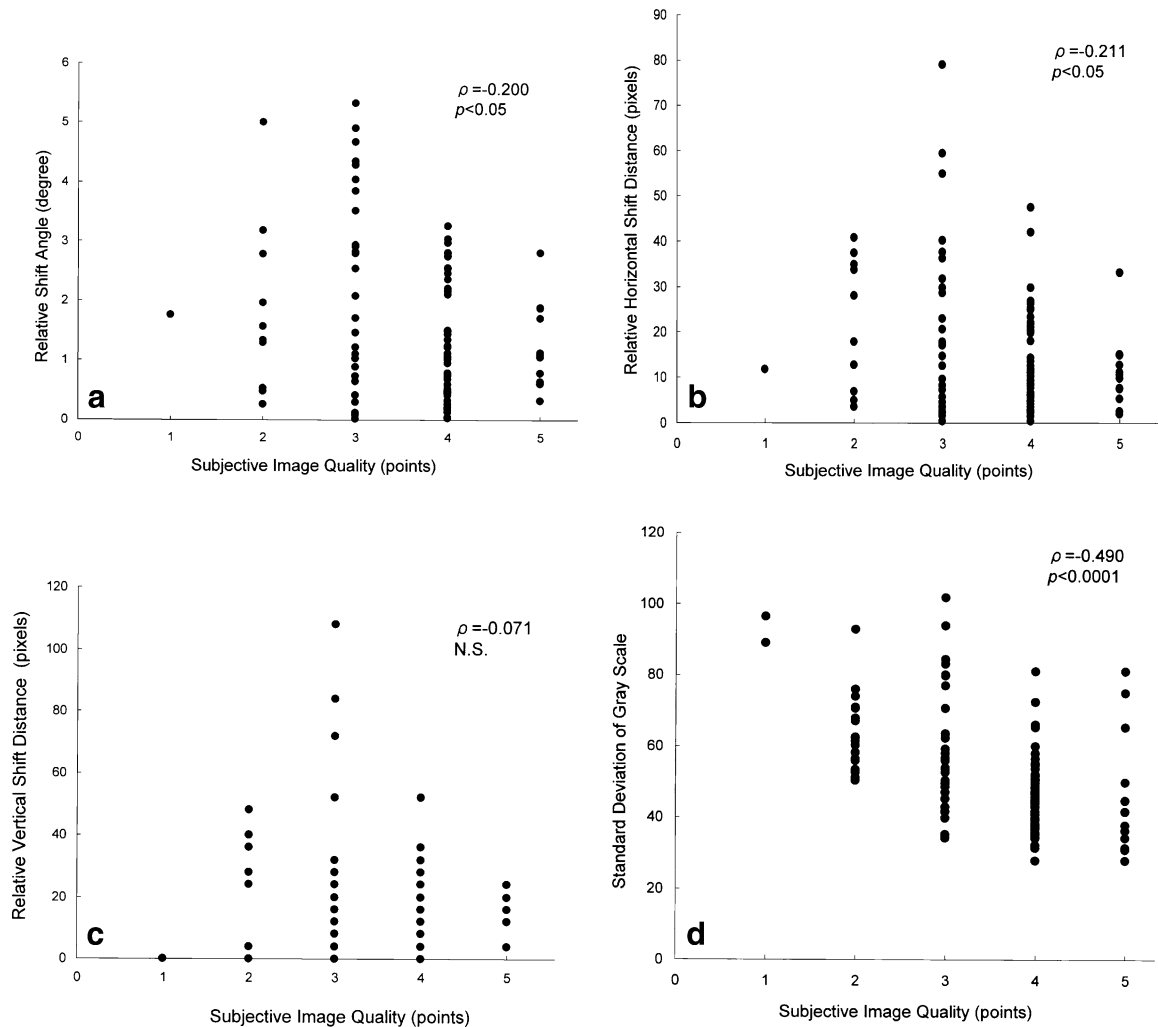|	Relative shift angle (°)	Relative shift, vertical distance (pixels)	Relative shift, horizontal distance (pixels)	Standard deviation of grayscale
Upright	1.2 ± 0.9	13.2 ± 8.5	10.1 ± 8.8	44.3 ± 14.4
Supine	2.2 ± 1.7	22.7 ± 22.9	22.8 ± 16.1	57.2 ± 16.6

Fig 4. Correlation between overall subjective image quality and the following parameters: (a) relative shift angle, (b) relative horizontal shift distance, and (c) relative vertical shift distance between previous and current chest radiographs. (d) Correlation between subjective image quality and standard deviation in temporal subtraction images.

current and previous images in the upright position are larger than those in the supine position. Additionally, the standard deviation of the gray-scale of the temporal subtraction images in the upright position is larger than that of images in the supine position. The Spearman's correlation coefficients by rank between subjective overall image quality and each parameter were as follows (Fig. 4): -0.200 for relative shift angle, -0.211 for relative horizontal shift, -0.071 for relative vertical shift, and -0.490 for standard deviation of gray levels. No significant correlation was found in the relative vertical shift distance, although significant correlations were found in the relative shift angle ($P < 0.05$), relative horizontal shift distance ($P < 0.05$), and standard deviation of gray levels ($P < 0.0001$).

DISCUSSION

Interpretation of digital chest radiographs in soft-copy images has been introduced at many hospitals with the rapid development of PACS in radiology. Several previous studies have shown no definite differences between soft-copy and hard-copy image quality with respect to the accuracy of diagnosis for various conditions.[10,11] The present

study shows that interpretation based on soft-copy chest radiographs is sufficiently practicable. Additionally, the interpretation of soft-copy images in a PACS environment has the potential for great future development because it links digital chest radiography technologies such as CR or FPD to CAD. We installed a server for producing temporal subtraction images in PACS, and the images made by this server are automatically loaded into the display terminal. This image-loading technique can easily be applied to other CAD systems such as CAD for mammography, gastrointestinal examinations, computed tomography (CT), etc.[1,2,12,13] The linkage of CR or FPD to PACS and CAD systems provides opportunities for significant technological advancement.[4,14]

Conventional radiographs must be converted to digital data by a film digitizer to yield temporal subtraction images; however, this process requires personnel costs. The present system, on the other hand, requires only a part of the maintenance costs for PACS and does not require personnel costs for producing temporal subtraction images because the system is automatic. Furthermore, because the images are obtained together with the original chest radiographs identified by patient code, the examination time, and the date, human error will be reduced in the production and display of temporal subtraction images using this system. In our system, temporal subtraction images were displayed with a 586×586 matrix and 8-bit grayscale because these values of matrix and grayscale of temporal subtraction were enough to enhance interval changes between previous and current images.[3,5,7]

The usefulness of temporal subtraction images in detecting a number of abnormal lesions such as lung nodules and interstitial infiltrates has previously been reported.[3,5,7,8] This usefulness is based on reliable production because excessive misregistration artifacts of temporal subtraction images often cause great problems for radiologists. Katsuragawa et al[9], reported from comparison of the qualities of subtraction images obtained with warping and non-warping methods that clinically acceptable temporal subtraction images were produced in 81.2% (147/181) and 97.2% (176/181) of all cases with non-warping and warping subtraction method, respectively. The original images in their study were obtained in the upright position from patients in lung cancer

screening programs. The yield of acceptable upright images in the present study was 100%. In the past, no study evaluated the image quality of subtraction images of patients in the supine position. The yield of temporal subtraction images with acceptable image quality was 66% in portable chest radiographs in our study. This result is unsatisfactory for routine clinical use, but may be a consequence of radiographic filming conditions in the supine position. Because the x-ray target and detector are synchronized in chest radiography in the upright position, the x-ray target may be located at the median of the patient's thorax. On the other hand, radiologic technologists can hardly center the x-ray target on a film because they must locate the x-ray target and the film independently. Thus, the software of our current system has a limitation to produce consistent temporal subtraction images of good quality by using supine chest radiographs.

Usually, an objective parameter used to define the image quality of chest radiographs is the signal-to-noise ratio (SNR).[15,16] However, the SNR cannot be used on temporal subtraction images because there is no signal that is registered completely except for the true interval change. When there is no interval change between current and previous images, a temporal subtraction image registered completely shows a monotonous medium gray; in this case, darker or lighter areas on an image are caused by misregistration artifacts. In other words, because true interval changes are signals and misregistration artifacts are noise in temporal subtraction images, it is difficult to distinguish between true interval changes and misregistration artifacts without referring to original images. Thus, it is not desirable in terms of quality assurance that temporal subtraction images do not have an objective parameter of image quality. In the present study, we used the standard deviation of the grayscale of temporal subtraction images as an objective parameter of image quality. However, when there were massive or diffuse interval changes between current and previous images, the standard deviation of the grayscale of temporal subtraction images could not be used. On the other hand, the relative shift angle and the vertical distance between current and previous images show a weak correlation with overall subjective image quality, and this correlation was found to

be statistically significant. Although the relative shift angle and the vertical distance between current and previous images cannot directly indicate objective image quality, they can nevertheless be used as objective parameters of image quality. Furthermore, the points of subjective image quality were not along a line with negative slope according to expectation on the standard deviation of grayscale. Chest radiologists evaluate mainly image quality of the intrathoracic areas, whereas the standard deviation of grayscale was calculated in the whole area including extrathoracic area. This disparity may be the reason why the points were not along a negative line in Figure 4d. A score of 3 was given to more images than the other scores because among the chest radiographs in the supine position, many images in which the thorax could be extracted by the temporal subtraction system were given a score of 3.

The present study had several limitations. Although the observers in our study had enough experience to interpret chest radiographs using temporal subtraction images, we had a low number of total observers. The case selection in our study was based on chest radiography, the existence of true interval changes was not confirmed on CT, and analysis of the internal algorithms of the present CAD system required an exclusive work station. For all these reasons, it would be difficult for other investigators to conduct a similar study. The primary purpose of temporal subtraction techniques is to enhance interval changes. However, our study examined only the cases without interval change between previous and current images because study design would become complex. Further study would be needed on cases with interval changes. Finally, we did not examine the influence of individual interpretation on the quality rating of temporal subtraction images. Thus, further investigation seems warranted. Nevertheless, the present findings clearly show the benefits of integrating the temporal subtraction technique into PACS.

In conclusion, we integrated PACS with a commercial system capable of automatically producing temporal subtraction images. All temporal subtraction images obtained with patients in the upright position were of acceptable quality for clinical use, as were 66% of the images taken in the supine position, which produces worse filming

conditions. Objective parameters such as the standard deviation of the grayscale of temporal subtraction images and the relative shift angle and distance between original images may also be useful in evaluating image quality.

ACKNOWLEDGMENT

REFERENCES

1. Warren Burhenne LJ, Wood SA, D'Orsi CJ, et al: Potential contribution of computer-aided detection to the sensitivity of screening mammography. Radiology 215:554-562, 2000

2. Nawano S, Murakami K, Moriyama N, et al: Computer-aided diagnosis in full digital mammography. Invest Radiol 34:310-316, 1999

3. Tsubamoto M, Johkoh T, Kozuka T, et al: Temporal subtraction for the detection of hazy pulmonary opacities on chest radiography. AJR 179:467-471, 2002

4. Morishita J, Katsuragawa S, Kondo K, et al: An automated patient recognition method based on an image-matching technique using previous chest radiographs in the picture archiving and communication system environment. Med Phys 28:1093-1097, 2001

5. Kano A, Doi K, MacMahon H, et al: Digital image subtraction of temporally sequential chest images for detection of interval change. Med Phys 21:453-461, 1994

6. Shiraishi J, Katsuragawa S, Ikezoe J, et al: Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. AJR 174:71-74, 2000

7. Johkoh T, Kozuka T, Tomiyama N, et al: Temporal subtraction for detection of solitary pulmonary nodules on chest radiographs: evaluation of a commercially available computer-aided diagnosis system. Radiology 223:806-811, 2002

8. Kakeda S, Nakamura K, Kamada K, et al: Improved detection of lung nodules by using a temporal subtraction technique. Radiology 224:145-151, 2002

9. Katsuragawa S, Tagashira H, Li Q, et al: Comparison of the quality of temporal subtraction images obtained with manual and automated methods of digital chest radiography. J Digit Imaging 12:166-172, 1999

10. Weatherburn GC, Ridout D, Strickland NH, et al: A comparison of conventional film, CR hard copy and PACS soft copy images of the chest: analyses of ROC curves and inter-observer agreement. Eur J Radiol 47:206-214, 2003

11. Balassy C, Prokop M, Weber M, Sailer J, Herold CJ, Schaefer-Prokop C: Flat-panel display (LCD) versus high-resolution gray-scale display (CRT) for chest radiography: an observer performance study. AJR 184:752-756, 2005

12. Kanazawa K, Kawata Y, Niki N, et al: Computer-aided diagnosis for pulmonary nodules based on helical CT images. Comput Med Imaging Graph 22:157-167, 1998

13. Yoshida H, Masutani Y, MacEneaney P, et al.:

Computerized detection of colonic polyps at CT colonography on the basis of volumetric features: pilot study. Radiology 222:327-336, 2002

14. MacMahon H: Digital chest radiography: Practical issues. J Thorac Imaging 18:138-147, 2003

15. Rill LN, Brateman L, Arreola M: Evaluating radio-graphic parameters for mobile chest computed radiography: phantoms, image quality and effective dose. Med Phys 30:2727-2735, 2003

16. Roehrig H, Krupinski AE, Chawla SA, Fan J, Gandhi K: Spatial noise and threshold contrasts in LCD displays. Proc SPIE 5034:174-186, 2003