MIPS SPECIAL FEATURE

Grey-scale inversion improves detection of lung nodules

¹J W ROBINSON, BAppSc, ¹J T RYAN, PhD, ¹M F McENTEE, PhD, ¹S J LEWIS, PhD, ²M G EVANOFF, MD, PhD, ³L A RAINFORD, PhD and ¹P C BRENNAN, PhD

¹*Discipline of Medical Radiation Sciences, Faculty of Health Science, University of Sydney, Sydney, NSW, Australia,*
²*American Board of Radiology, Tucson, AZ, USA, and* ³*School of Medicine and Medical Sciences, University College Dublin, Belfield, Dublin, Ireland*

Objective: The current study aims to establish whether detection of solitary pulmonary nodules can be improved by inverting the grey scale of posteroanterior (PA) chests.

Methods: 30 PA chest images were presented on 2 occasions to 16 senior radiologists on either primary or secondary class displays in the standard or inverted mode. 15 images within each group contained a single nodule positioned in a range of anatomical sites. A receiver operating characteristic (ROC) methodology was used to explore differences between the presentation modes.

Results: Improved ROC scores were evident with inverted (Az 0.77) compared with standard (Az 0.73) ($p=0.02$) images; however, this difference was seen only with the primary displays. The benefits seen are most likely owing to increased nodule luminance with the inverted images, particularly when using primary displays.

Conclusion: This study demonstrates that the inverted image can offer advantages in lung nodule detection over the standard presentation mode when images are viewed on high-specification viewing systems. The study has demonstrated that there is an improvement in the detectability of lung nodules on an inverted image with a primary display monitor that is not evident with secondary displays. This is likely to be the result of increased nodule luminance on primary displays when images are presented in the inverted mode.

Received 1 November 2011
Revised 10 December 2011
Accepted 14 December 2011

DOI: 10.1259/bjr/27961545

© 2013 The British Institute of Radiology

In radiology there is an increasing use of digital display technology, which has had an enormous impact on medical image presentation [1–3]. This is no longer limited to CT or MRI, but is equally relevant to projectional radiography in which digital images are acquired using computed radiography, direct and indirect digital radiography and scanning technologies [3].

In digital imaging systems raw image data are normally processed using look-up tables and other algorithms so that, when the image is first presented, it requires minimum reader time or effort to change contrast or brightness details. Nonetheless, further image processing techniques are often applied by the observer with the aim of increasing the diagnostic value of an image [4, 5], and one such feature, which is available with almost all image workstations, is inversion of the grey scale of an image. Even though inversion algorithms have been available for as long as digital images, the diagnostic advantages of the inverted mode over the standard mode remain unclear.

Previous work examining the efficacy of image inversion has focused on either dental or chest radiology. The dental research demonstrated higher levels of

accuracy when measuring endodontic wires inserted into extracted pre-molars for the inverted compared with the standard mode [6]; no differences in periodontal bone loss measurements between inverted and unprocessed images [7]; and low diagnostic agreement between standard and inverted panoramic images for the detection of mental foramen and mandibular canal [8]. The chest studies showed that inversion mode is commonly used by intensive care unit clinicians to view the end point of line and tube placements but was inconclusive regarding detection sensitivity [9]; no advantage in inversion mode for nodule detection of pulmonary nodules, but the effect of familiarity with standard image and observer experience needs to be accounted for [10]; and better detection in the inverted mode than in the standard mode using film–screen systems assessed by residents-in-training [11]. Although these previous studies suggest that inverting the images may offer certain advantages in specific situations, further studies involving expert thoracic radiologists and a better understanding of the mechanisms for potential improvement with inversion are required.

The current study focused on inversion of chest images and used a receiver operating characteristic (ROC) methodology to investigate the detectability of a solitary pulmonary nodule. Two classes of monitor displays were considered and images were assessed by expert chest radiologists.

Address correspondence to: Mr John Robinson, Discipline of Medical Radiation Sciences, Faculty of Health Science, University of Sydney, East Street, Sydney NSW 2141, Australia. E-mail: john.robinsonmrs@sydney.edu.au

Table 1. Comparison of the two monitor class displays used for the study

Specification type	Barco	ViewSonic
Maximum luminance ^a	450 cd m ⁻²	157 cd m ⁻²
Minimum luminance ^a	1.0 cd m ⁻²	0.6 cd m ⁻²
Contrast ratio ^a	450:1	261:1
Display resolution	1536×2048 pixels	1600×1200 pixels
Screen type	LCD	LCD

LCD, liquid crystal display.

^aAt settings used in this study.

Methods and materials

The study aimed to compare the efficacy of standard and inverted images when displayed on either a primary or a secondary display. The primary display was a Coronis 3MP (Barco, Kortrijk, Belgium) radiology reporting monitor driven by a Barco Med 5MP2FH video card. The secondary display was a View Panel VP 201m (ViewSonic, Walnut, CA) with an NVIDIA GeForce 7100/nForce 630i (Santa Clara, CA) video card. The specifications are shown in Table 1. On each day of reading, both monitors were calibrated to the digital imaging and communication in medicine grayscale standard display function (DICOM GSDF) using Verilum software and a luminance pod (Image Smiths Inc., Germantown, MD) and a calibrated photometer (model no. 07-621; Nuclear Associates, Hicksville, NY).

The bank of images consisted of 30 DICOM format images, which were presented to observers in both standard and inverted modes. Images were cropped to 1500×1500 pixels by removing the surrounding areas of the chest image, but all anatomical details remained. With the primary workstation monitor, the images were displayed at full resolution. When images were presented on the secondary display, images were fitted to the available screen size and each observer had the

option to present the image at full resolution using zoom and panning post-processing facilities. ViewDEX v. 1.0 (a viewer for digital evaluation of radiographic images; Southern Älvsborg Hospital, Sahlgrenska University Hospital, Sahlgrenska, and Gothenburg University, Gothenburg, Sweden) software, running on the Java Platform Standard Edition Runtime Environment (Sun Microsystems, Santa Clara, CA) and designed for observer performance studies, was used to display the images [12]. Observers were asked to identify lesions on the images and levels of confidence were scored on a scale of 1–7, with 7 representing complete confidence that a lesion was present. On both monitor types, to enable an interaction with the images under normal clinical conditions, observers were allowed to window (but not invert), pan and zoom the image and correct any error they felt they had previously made. Figure 1 demonstrates the graphical user interface.

In accordance with the Health Insurance Portability and Accountability Act of 1996, permission for the use of images was granted by an institutional review board and all patient identifiers were removed from the images. All images, according to the respective radiology report and subsequent verification by an expert thoracic radiologist (not involved in the study proper), did not demonstrate any nodular lesions or lung pathology. A single digitally imposed lesion with diffuse edge and a gaussian contrast profile [13], of varying subtlety and measuring between 6 and 12 mm in diameter, was subsequently inserted on each of 15 images using methods described elsewhere [14], and the authenticity of the lesions was confirmed by the same radiologist. The position of each lesion is shown in Table 2. The luminance of all lesions in both sets of images was recorded using a calibrated Hagner Universal Photometer (Hagner AB, Stockholm, Sweden) and compared between each image type using a paired Student's *t*-test and a *p*-value of <0.05 to describe significant differences.

16 examining radiologists with the American Board of Radiology (ABR) studied all images with both presentation modes on either the primary or secondary class

Figure 1. The graphical user interface used for image assessment. The image was displayed on the left of the monitor while the facilities—window, pan, zoom and score for the image—are shown on the right.

Table 2. Position of each lesion on the 15 “abnormal” chest posteroanterior images

Image number	Chest location	Anatomical background
1	Left upper zone	Intercostal space
2	Right lower zone	Vascular branching
3	Left upper zone	Rib
4	Left mid-zone	Intercostal space
5	Left lower zone	Heart apex
6	Right upper zone	Rib
7	Right lower zone	Diaphragm
8	Right lower zone	Diaphragm
9	Left upper zone	Intercostal space and scapula
10	Left upper zone	Mediastinum
11	Right upper zone	Mediastinum
12	Left upper zone	Clavicle
13	Left upper zone	Intercostal space
14	Right lower zone	Pulmonary artery
15	Right lower zone	Pulmonary artery

monitors (8 radiologists looked at the images on each monitor type) without any time restriction. Mean post-ABR certification experience across all radiologists for the primary and secondary display was 22 and 21 years, respectively. A demonstration of the software was given to all observers prior to any readings and support was available at all times throughout each observation. The illumination within the reading room remained within 10% of 30 lux [15] for the duration of the study.

ROC analysis was performed using the Dorfman, Berbaum and Metz multireader-multicase methodology (DBM, MRMC v. 2.2) (<http://perception.radiology.uiowa.edu> and Kurt Rossman Laboratories for Radiologic Image Research <http://xray.bsd.uchicago.edu/krl/>) [16]. ROC values were generated using the trapezoidal method, and statistical differences between Az scores for whitebone and blackbone modes were indicated at a p -value ≤ 0.05 using a one-way analysis of variance. For all analyses, only the readers were treated as random samples, whereas the cases were fixed, meaning that the results are only relevant

Table 3. Details for radiologists reading from each monitor class display

Radiologist number	Years certified	Age	Corrective lenses	Monitor
1	27	59	Yes	Primary
2	28	58	Yes	Primary
3	31	62	Yes	Primary
4	17	–	No	Primary
5	16	45	Yes	Primary
6	16	47	Yes	Primary
7	19	38	Yes	Primary
8	19	42	Yes	Primary
9	16	50	Yes	Secondary
10	19	49	Yes	Secondary
11	15	45	Yes	Secondary
12	33	–	Yes	Secondary
13	18	53	Yes	Secondary
14	19	52	Yes	Secondary
15	10	39	No	Secondary
16	38	68	Yes	Secondary

–, age was unavailable.

Table 4. Luminance values in cd m^{-2} for each lesion present for the standard and inverted abnormal images; the luminance difference between image presentations is shown in parentheses

Image	Primary class display		Secondary class display	
	Standard	Inverted	Standard	Inverted
1	77	260 (183)	71	165 (94)
2	94	145 (51)	77	118 (41)
3	83	218 (135)	71	136 (65)
4	118	130 (12)	83	94 (11)
5	106	118 (12)	83	94 (11)
6	83	201 (118)	71	130 (59)
7	83	165 (82)	77	118 (41)
8	83	189 (106)	71	118 (48)
9	83	177 (94)	71	130 (59)
10	89	153 (64)	77	112 (35)
11	83	201 (118)	71	130 (59)
12	112	118 (6)	77	118 (41)
13	71	307 (236)	59	171 (92)
14	94	165 (71)	77	106 (29)
15	83	260 (177)	71	153 (82)
Mean	89.5	187.1 (97.7)	73.8	126.2 (51.1)

to the specific images shown. This decision was made on the basis that any results were dependent on the lesion location, and therefore could only apply to the images used in this study. For each inverted *vs* standard comparison, the sample sizes used in the study enabled an effect size of 0.07, 0.10 and 0.10 to be detected for the combined, primary and secondary class displays, respectively, at 80% power.

In addition to the above analyses, to examine whether the combination of standard and inverted images would enhance the ROC score compared with either one of the two single formats, the best score for each image (using either format) for each reader was collated and compared with the individual presentation scores using the DBM approach described above. So, for example, if image 1 scored 0.50 for the inverted and 0.60 for the standard mode, 0.60 would be taken as the best score and collated with all the other best image scores. This group of best scores was then compared individually with the standard and inverted scores.

Results

Details for each radiologist participating in the study are shown in Table 3. Luminance values for each lesion are shown in Table 4, with the lesions having statistically higher values in the inverted images than in the standard images ($p=0.0001$) for both types of display. The increase in brightness for the inverted images compared with the standard images was greater for the primary display than for the secondary display ($p=0.02$).

The Az scores generated for all readers along with overall mean values are shown in Table 5.

The mean results for the combined, primary and secondary class display groups, along with p -values, are shown in Table 6. A statistical difference ($p=0.02$) was observed between image presentations when all observers were combined and for the observers who were allocated the primary class display ($p=0.04$).

Table 5. Receiver operating characteristic (ROC) scores for all readers, for both classes of monitor displays, for both image types along with difference scores and mean values; mean values and 95% confidence intervals (CIs) are shown study

Reader	Standard	ROC score	
		Inverted	Difference
1	0.73	0.68	0.05
2	0.70	0.75	-0.05
3	0.84	0.82	0.02
4	0.70	0.81	-0.11
5	0.62	0.66	-0.04
6	0.66	0.84	-0.16
7	0.80	0.75	0.05
8	0.60	0.73	-0.13
9	0.70	0.60	0.10
10	0.83	0.85	-0.02
11	0.82	0.81	0.01
12	0.73	0.78	-0.05
13	0.69	0.82	-0.13
14	0.70	0.80	-0.10
15	0.72	0.78	-0.06
16	0.89	0.79	0.10
Mean	0.73	0.77	-0.03
CI	0.68, 0.77	0.73, 0.80	-0.07, -0.001

When the best scores for each image were collected and compared with the individual presentation scores, a significant improvement was shown with combined best scores compared with the standard image presentation (Table 7).

Discussion

The current study presents a comparison of standard and inverted chest images to establish whether nodular lesions are more easily detected with a specific image presentation. It should be acknowledged that, since radiologists are much more likely to have greater experience using standard images as the default presentation, familiarity with this image type may have introduced some inherent bias; however, despite this, higher Az scores were noted for the inverted images. The ability of highly specialised individuals to compensate for differing viewing conditions has been shown elsewhere for varying ambient [15] and display [17] conditions and is possibly linked to well-established and fine-tuned search patterns that are relatively unaffected by technological or environmental variations,

unlike individuals asked to make judgements on images with which they are less familiar.

The advantage shown with the inverted images was only apparent with the primary workstations, which is interesting since both display types were calibrated to the grayscale display function (GSDF) issued by the American Association of Physicists in Medicine (AAPM 2005) [18]. One of the key aims of the GSDF, in line with the Barten model, is to perceptually linearise the display luminance [19] so that equal changes in pixel values result in equal changes in perceived brightness, regardless of whether these pixel values are high or low. To achieve this, a grey-scale map is used to amplify incremental changes with lower value pixels. However, to fully understand whether image inversion has a deleterious effect on GSDF calibration, one must examine the sequence of events from pixel value to image presentation, and, since calibration occurs after inversion, appropriate grey-scale mapping should still be applied. This being the case and assuming that the GSDF calibration procedure is correct, an alternative to a calibration-based explanation should be explored to explain the inversion advantage, particularly with the primary monitor.

Two key interdependent visual characteristics are responsible for lesion discriminability: brightness resolution, which is the ability to detect differences in brightness levels, and spatial resolution, which is the ability to detect fine detail through space. With regard to the former it has been shown that, at a given spatial frequency, the contrast threshold, which is the relative luminance increment required to detect a signal, is lower at higher luminances. In other words, the brighter the image the easier it is to detect grey-level change. With regard to spatial changes, maximum spatial acuity (in terms of cycles per millimetre) increases with higher luminances at given viewing distances and contrast thresholds. All this points to increased detection of signals at higher brightness levels. As demonstrated by the luminance values in Table 4, it is evident that all lesions on inverted images are brighter than on the standard images. This indicates that, with the normally viewed standard image, even though the lesion looks bright (compared with the surrounding lungs), it is actually darker than the median value across all pixels in the image. In particular, the level of brightness increase is largest for the primary monitor compared with the secondary class display, and it is with the primary class device that statistical changes are noted. It should be observed, however, that the figures in Table 4 represent

Table 6. Mean receiver operating characteristic (ROC) and difference scores for monitor class display combinations

Monitor display class	ROC image score			p-value
	Standard	Inverted	Difference	
Primary and secondary	0.73 (0.68, 0.77)	0.77 (0.73, 0.80)	-0.03 (-0.07, -0.001)	0.02 ^a
Primary	0.71 (0.63, 0.76)	0.76 (0.71, 0.80)	-0.05 (-0.10, -0.01)	0.04 ^a
Secondary	0.76 (0.70, 0.82)	0.78 (0.72, 0.84)	-0.02 (-0.07, 0.03)	0.54

Confidence intervals are shown in parentheses.

^aSignificant difference.

Table 7. Mean values when the best scores from the whitebone and blackbone are combined along with any differences from the standard and inverted image-specific scores

Monitor display class	Mean	Difference	
		Standard	Inverted
Primary and secondary	0.78	+0.05 (0.006 ^a)	+0.03 (0.4)
Primary	0.76	+0.05 (0.02 ^a)	+0.01 (0.06)
Secondary	0.80	+0.04 (0.17)	+0.02 (0.54)

+ or – indicates that the combined best score is higher or lower, respectively, than the image-specific scores. *p*-values are shown in parentheses.

^aSignificant difference.

the unmanipulated state, and, with windowing, the lesion brightness levels at which decisions were made are quite likely to be different from those shown here.

Although locations of all lesions were recorded in this study (Table 2), the relatively low numbers of nodules in specific chest regions did not support a region-specific analysis. It is, however, important to establish whether the density and the complexity of the background on which a lesion is found has an impact upon detection [20], and therefore whether inversion only works effectively for specific thoracic locations. Additionally, although visually the simulated lesions appeared to behave in the same way as actual lesions, it is important to validate this and future research is ongoing in this field.

In summary, the paper provides some evidence that inverting images may offer some advantages for lesion detection on a primary display monitor. This is possibly a function of the increased brightness of the lesion in the inverted mode and contrast ratio of the primary class monitor, and therefore may only offer benefit to those images in which the lesion is at a relatively low luminance value in the original image. The findings of this study demonstrate the need to explore the utility of this cost-effective post-processing tool for other image and modality types.

References

- Korner M, Weber CH, Wirth S, Pfeifer K, Reiser MF, Treitl M. Advances in digital radiography: physical principles and system overview. *Radiographics* 2007;27:675–86.
- Al Khalifah K, Brindhaban A. Comparison between conventional radiography and digital radiography for various kVp and mAs settings using a pelvic phantom. *Radiography* 2004;10:119–25.
- Cao XH, Huang HK. Current status and future advances of digital radiography and PACS. *IEEE Eng Med Biol* 2000;19:80–8.
- Lee KR, Siegel EL, Templeton AW, Dwyer SJ III, Murphey MD, Wetzel LH. State-of-the-art digital radiography. *Radiographics* 1991;11:1013–25.
- Prokop M, Schaefer-Prokop CM. Digital image processing. *Eur Radiol* 1997;7(Suppl. 3):S73–82.
- Kal BI, Baksi BG, Dundar N, Sen BH. Effect of various digital processing algorithms on the measurement accuracy of endodontic file length. *Oral Surg Oral Med Oral Pathol Oral Radiol Endosc* 2007;103:280–4.
- Scaf G, Morihisa O, Loffredo Lde CM. Comparison between inverted and unprocessed digitized radiographic imaging in periodontal bone loss measurements. *J Appl Oral Sci* 2007;15:492–4.
- Sakakura CE, Loffredo Lde CM, Scaf G. Diagnostic agreement of conventional and inverted scanned panoramic radiographs in the detection of the mandibular canal and the mental foramen. *J Oral Implantol* 2004;30:2–6.
- Seshadri SB, Arenson RL, Kundel HL, Feingold E, Kishore S, Horii SC, et al. High-resolution image workstation for intensive care units. *Radiographics* 1994;14:863–73.
- Kheddache S, Månsson LG, Angelhed JE, Denbratt L, Gottfridson B, Schlossman D. Digital chest radiography: should images be presented in negative or positive mode? *Eur J Radiol* 1991;13:151–5.
- Sheline ME, Brikman I, Epstein DM, Mezrich JL, Kundel HL, Arenson RL. The diagnosis of pulmonary nodules: comparison between standard and inverse digitized images and conventional chest radiographs. *AJR Am J Roentgenol* 1989;152:261–3.
- Börjesson S, Hakånsson M, Båth M, Kheddache S, Svensson S, Tingberg A, et al. A software tool for increased efficiency in observer performance studies in radiology. *Radiat Prot Dosimetry* 2005;114:45–52.
- Samei E, Flynn MJ, Eyler WR. Simulation of subtle lung nodules in projection chest radiography. *Radiology* 1997;202:117–24.
- Ryan JT, McEntee MF, Barrett S, Evanoff MG, Manning DJ, Brennan PC. A software system for the simulation of chest lesions. *Proc SPIE* 2007;6515:J1–8.
- Brennan PC, McEntee M, Evanoff M, Phillips P, O'Connor WT, Manning DJ. Ambient lighting: effect of illumination on the soft-copy reporting of radiographs of the wrist. *AJR Am J Roentgenol* 2007;188:W177–8.
- Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis: generalization to the population of readers and patients with the jackknife method. *Invest Radiol* 1992;27:723–31.
- Toomey RJ, Ryan JT, McEntee MF, Evanoff MG, Chakraborty DP, McNulty JP, et al. Diagnostic efficacy of handheld devices for emergency radiologic consultation. *AJR Am J Roentgenol* 2010;194:469–74.
- Samei E, Badano A, Chakraborty D, Compton K, Cornelius C, Corrigan K, et al. Assessment of display performance for medical imaging systems: Executive summary of AAPM TG18 report. *Med Phys* 2005;32:1205–25.
- Park S, Badano A, Gallas BD, Myers KJ. Incorporating human contrast sensitivity in model observers for detection tasks. *IEEE Trans Med Imaging* 2009;28:339–47.
- Brett DS, Berry E, Smith MA. The effect of detectability in local area anatomical noise. *Br J Radiol* 2007;80:186–93.