Journal of Digital Imaging

Evaluation of Stylus for Radiographic Image Annotation

Gautam S. Muralidhar,[1] Gary J. Whitman,[2] Tamara Miner Haygood,[2] Tanya W. Stephens,[2] Alan C. Bovik,[3] and Mia K. Markey[1]

We evaluated the use of a stylus as a computer interface for radiographic image annotation. Our case study concerned the annotation of spiculated lesions on mammograms. Three experienced radiologists annotated 20 mammograms depicting spiculated lesions. We evaluated the interobserver agreement in annotations marked with a stylus versus those marked with a mouse using the intraclass correlation coefficient. Better agreement in annotating spicule width was observed with the stylus, suggesting that it is easier to accurately annotate subtle regions on an image using a stylus.

KEY WORDS: Radiography, mammography, imaging informatics, image display

BACKGROUND

Computers and digital technology have become ubiquitous in hospitals and healthcare facilities in North America. One of the key developments in this digital revolution over the last decade has been the emergence of digital radiology. Images produced by modalities including magnetic resonance imaging (MRI), computed tomography (CT), X-ray mammography, and ultrasound are now digital and can be stored, processed, and displayed on computers without the use of printed film.

Another recent development has been the increasing use of a stylus. A stylus is an electronic pen that enables the user to write on the computer screen in much the same way as on a sheet of paper. The data written on the computer screen can then be digitally stored onto the computer hard drive. A stylus makes data reporting easy and is one of the reasons why mobile computing devices based on a stylus interface, such as tablet personal computers, are being regarded with interest in healthcare, especially in nursing and data reporting.

Another area in healthcare where we envisage the use of a stylus is annotation in radiology. Radiologists often annotate images with china markers on printed films. With digital displays, annotations can be made on the computer screen using a mouse and image annotation software installed on clinical workstations. However, as discussed by Reiner and Siegel,[1] reporting in radiology is yet to make full use of advances in digital displays and devices. As part of the changes to adopt to enhance quality of reporting, Reiner and Siegel advocate the use of devices such as stylus for annotation and reporting in radiology, especially for cross-sectional imaging studies that involve a large number of images. Their reasoning behind this is that devices like mouse and keyboard distract the radiologist and do not allow maintaining 100% eye contact with the imaging data. In addition to this, we believe that the mouse and keyboard interfaces provide limited flexibility when annotating images and that a stylus could be

---

[1]From the Department of Biomedical Engineering, The University of Texas at Austin, 1 University Station C0800, Austin, TX, 78712-0238, USA.

[2]From the Department of Diagnostic Radiology, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX, 77030, USA.

[3]From the Department of Electrical and Computer Engineering, The University of Texas at Austin, 1 University Station C0803, Austin, TX, 78712-1084, USA.

Correspondence to: Gautam S. Muralidhar, Department of Biomedical Engineering, The University of Texas at Austin, 1 University Station C0800, Austin, TX, 78712-0238, USA; tel: +1-512-7994336; fax: +1-512-4710616; e-mail: gautam@mail.utexas.edu

Online publication 26 August 2009

doi: 10.1007/s10278-009-9232-6

a very useful tool for accurate freehand annotations on images.

In this work, we present a case study concerning the annotation of spiculated lesions on digitized mammograms using a stylus. The main motivation behind our work is to investigate the effectiveness of the stylus for the task of image annotation. Reliable digital annotation in mammography has received attention from the viewpoint of accumulating annotated case libraries that can be used by medical residents and fellows for research and education.[2] Reliable annotation of spiculated lesions has also shown to be important in the development of computer-aided detection systems (e.g.,[3],[4]). The study described in [4] has shown that radiologists can reliably annotate spiculated lesions on digitized mammograms using a mouse. The purpose of this study was to investigate the use of a stylus interface for the same annotation task.

METHODS

Dataset

Twenty cases of spiculated mammographic lesions were used for the annotation study. All 20 cases were selected at random from the Digital Database for Screening Mammography, a publicly available collection of deidentified proven cases.[5] A region of interest (ROI) around each lesion on the mediolateral oblique view was used in this analysis. The ROI was defined such that the central mass of the lesion and all the spicules were included.

Readers

Three experienced radiologists took part in this study as readers (GW, TH, and TS). Two of the three radiologists (GW and TS) are breast imaging specialists, while the third (TH) has a more varied practice and spends approximately 15% of clinical time in mammography. The readers have a collective experience of more than 40 years in reading mammograms. All of the radiologists read more than 100 mammograms per month. Institutional review board approval was obtained prior to this study.

Annotation Experiment

The image annotations were performed using the ROI manager plugin of the NIH ImageJ software[6,7] installed on a tablet PC. The ROI manager plugin has options to select the appropriate line width for the annotations and display the annotations in color. All three radiologists opted for the color display and set the line widths based on their comfort level. They were allowed to adjust the zoom on the ROI as they wished.

It was agreed between the three radiologists that the first annotation on each ROI would always correspond to the major axis of the central mass region of the spiculated lesion. The next two annotations corresponded to the width and length of a spicule, respectively, and these annotations were repeated for every spicule. Since the resolution of the images was known beforehand, the pixel annotations were converted to physically meaningful quantities (e.g., millimeters). The radiologists were given unlimited time to complete the annotation task.

Annotations were performed twice by each radiologist on the same set of mammograms, once with the mouse and once with the stylus. During the annotation phase, each radiologist was free to choose which interface he/she wanted to use first. Each radiologist performed the annotations using the stylus and the mouse in two sessions for a total of four sessions for each radiologist. The stylus and mouse annotations were performed separately under similar experimental conditions. On completion of the annotation study, there were two sets of annotations on 20 cases, one set annotated using the stylus and the other set annotated using the mouse by each of the three radiologists. Figures 1 and 2 illustrate examples of a spiculated lesion annotated by one of the radiologists using the mouse and the stylus interfaces.

We assessed the interobserver agreement across three radiologists using the intraclass correlation coefficient (ICC) method.[8] The ICC analysis was performed on both the mouse and the stylus annotations to assess agreement in measurements of major axis, average spicule length, and average spicule width. We did not assess the interobserver agreement in the number of spicules annotated on each mammogram. Our reasoning behind this is that the variability in number of spicules is independent of the interface used to annotate a spiculated lesion and is not an appropriate param-

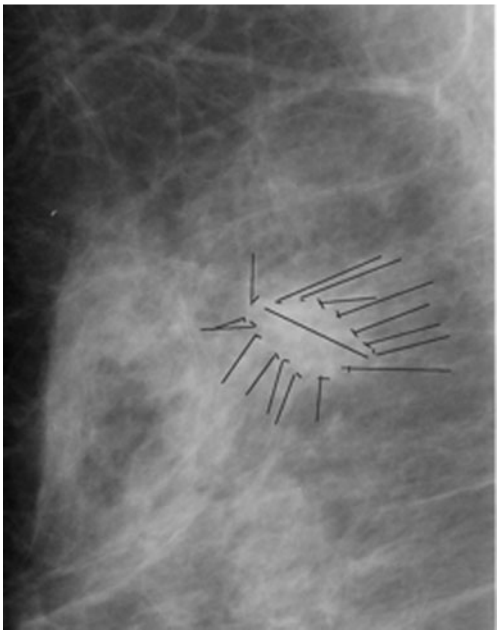

Fig. 1. Annotations made using the mouse interface by one of the radiologists on a mammogram depicting a spiculated lesion.

eter to assess the effectiveness of an interface for image annotation. The ICC results for interobserver agreement were interpreted using the ICC interpretation guideline.[9] According to this guideline, an ICC value lower than 0.4 suggests that the observers are in poor agreement. An ICC value between 0.4 and 0.75 suggests that the level of agreement is fair to good, while an ICC value greater than 0.75 suggests excellent agreement.

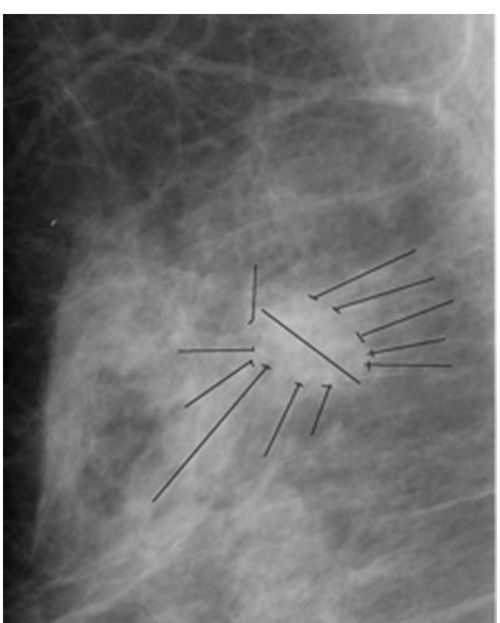

Fig. 2. Annotations made using the stylus interface by the same radiologist on a mammogram depicting a spiculated lesion.

We also conducted a study to assess the productivity of a radiologist when using a stylus. The time taken to annotate each of 12 randomly selected cases with both the interfaces by the same radiologist was compared using the Wilcoxon signed-rank test. This study was completed in two sessions, one each for mouse and stylus. We chose the radiologist who had the most experience in annotating images using both the interfaces. Since this was a timed study that we wanted to be completed in one session for each interface, we chose only 12 out of 20 cases to ensure that the radiologist does not feel pressurized. The timing data is summarized in Table 1.

RESULTS

The ICC result comparing the measurements of major axis obtained from annotations made by the three radiologists using the mouse and stylus suggests excellent agreement for both the interfaces. The degree of agreement in measurement of average spicule length was fair to good regardless of the interface used. However, the degree of interobserver agreement in measurement of average spicule width changed from poor to good when the annotations were made using the stylus. Each of the three ICC values summarized in Table 2 has to be viewed independently and there is no relationship between the ICC values comparing one set of measurements (e.g., major axis) with

Table 1. Annotation Time for Each Case Annotated Using Mouse and Stylus by the Same Radiologist

Case number	Annotation time for mouse (s)	Annotation time for stylus (s)
1	103	86
2	121	72
3	108	70
4	78	74
5	109	91
6	55	57
7	78	77
8	90	54
9	113	60
10	89	85
11	101	58
12	81	41

Table 2. ICC Coefficients for Annotations Made by Three Radiologists Using Mouse and Stylus

Interface used	Number of cases	ICC score—major axis	ICC score—width	ICC score—spicule length
Mouse	20	0.871–excellent	0.280–poor	0.634–fair to good
Stylus	20	0.953–excellent	0.474–fair/good	0.643–fair/good

another set of measurements (e.g., average spicule length). Our analysis suggests that it is easier to annotate spicule width more consistently using a stylus than using a mouse. It is important to note that the spicule width is a much smaller annotation than the major axis and spicule length, which probably makes it hard to achieve good consistency in annotating it with a mouse due to limited flexibility offered by a mouse.

Analysis of the timing data using the Wilcoxon signed-rank test revealed that the time taken to annotate each case using a stylus is significantly lower than the time taken to annotate each case using a mouse (p value=0.0015). The time taken to annotate each case with a stylus was lower than the time taken to annotate the same case with a mouse for 11 out of the 12 cases. We attribute this to the greater ease of use offered by the stylus for the image annotation task. Since the analysis of timing data obtained from one radiologist was conclusive, we did not repeat the timed analysis with other radiologists.

DISCUSSION

With increasing use of digital displays in radiology, a computer accessory that can provide good flexibility to perform image annotations on the computer screen is desirable. The stylus is one such tool that has been advocated for use in radiology reporting as it enables the radiologist to focus 100% on the imaging data than be distracted by computer interfaces such as a keyboard and a mouse.[1] However, there is a need to quantitatively assess the effectiveness of new computer interfaces like the stylus before they are adopted in routine radiological practice. We have conducted one such study that quantitatively assesses the effectiveness of the stylus for radiographic image annotation. We believe that our study shows the stylus to be a promising interface to use for image annotations, especially when subtle regions like spicule width and small nodules have to be annotated. Such annotations can be carried out with greater precision and reliability using the stylus over a traditional interface like the mouse.

An important point to note is that we used a tablet PC in our study to assess the effectiveness of the stylus for radiographic image annotation. We acknowledge that the resolution of a tablet PC does not compare to the resolution of actual mammography workstation monitors. However, as discussed by Krupinski and Kallergi,[10] not all radiology modalities require high-resolution workstation monitors. For example, while full-field digital mammography and computed radiography require high-resolution monochrome displays, the same does not apply to modalities like MRI and CT, which are often viewed with lower-resolution displays that support color.[10] The radiologists that took part in our study were unanimous in their opinion that a stylus could be a useful tool for accurately annotating images on actual workstations in much the same way as annotations have traditionally been made with a china marker on printed films.

Annotation accuracy to a large extent also depends on the human marking the annotations. However, a computer interface that is more ergonomically favorable for marking annotations might help achieve better annotation accuracy. We believe that our study has shown the stylus to be a promising interface that is ergonomically friendly and can be used with image annotation software[2,6] for research and educational purposes. However, for the stylus to be routinely used in radiological practice, our study will have to be extended to high-resolution workstations by considering a larger pool of human users (radiologists, residents, and fellows), different annotation tasks in radiology, and other subspecialties like image-guided radiation therapy (IGRT) planning. At higher resolutions, annotation of subtle regions might not be a problem. However, a human user could use the stylus as an all-in-one annotation tool to mark free text and image annotations, which might be ergonomically more favorable and certainly less distracting than using multiple interfaces such as keyboard and mouse.

CONCLUSION

We conducted a study in which we evaluated the use of a stylus as a computer interface for annotating spiculated lesions on mammograms. We noticed a significant improvement in the degree of interobserver agreement in the annotation of spicule width. The results of our study provide strong evidence that, in order to make subtle annotations on an image that is digitally displayed, the stylus might be a better interface than the mouse. The width of a spicule is a small annotation in terms of number of pixels, and using a stylus gives more flexibility to the observer to annotate subtle regions. Analysis of the timing data suggests that it is far more productive to use a stylus for image annotation tasks than a mouse. We strongly believe that our case study demonstrates the potential use of a stylus for image annotation in radiology especially for research and educational purposes. However, for the stylus to be routinely used in radiological practice, its effectiveness for different annotation tasks in radiology and subspecialties like IGRT planning has to be evaluated on high-resolution workstation monitors.

ACKNOWLEDGEMENTS

This work was supported in part by an Early Career Award from the Wallace H. Coulter Foundation. We thank the reviewers for their insightful comments and suggestions that have helped improve the quality of this manuscript.

REFERENCES

1. Reiner B, Siegel E: Radiology reporting: returning to our image-centric roots. AJR Am J Roentgenol 187:1151–1155, 2006
2. Zheng Y, Wu M, Cole E, Pisano ED: Online annotation tool for digital mammography. Acad Radiol 11:566–572, 2004
3. Sampat MP, Bovik AC, Whitman GJ, Markey MK: A model-based framework for the detection of spiculated masses on mammography. Med Phys 35:2110–2123, 2008
4. Sampat MP, et al: The reliability of measuring physical characteristics of spiculated masses on mammography. Br J Radiol 79:S134–S140, 2006
5. Heath M, Bowyer KW, Kopans D, Moore R, Kegelmeyer P, Jr.: The digital database for screening mammography. Proceedings of the 5th International Workshop on Digital Mammography
6. ImageJ. Available at http://rsb.info.nih.gov/ij/
7. Abramoff MD, Magelhaes PJ, Ram SJ: Image processing with ImageJ. Biophoton Int 11:36–42, 2004
8. McGraw KO, Wong SP: Forming inferences about some intraclass correlation coefficients. Psychol Methods 1:30–46, 1996
9. Rosner B: Fundamentals of Biostatistics, Belmont: Duxbury, 2005
10. Krupinski EA, Kallergi M: Choosing a radiology workstation: technical and clinical considerations. Radiology 242:671–682, 2007