University of Cumbria

Dong, Leng, Chen, Yan, Gale, Alastair and Phillips, Peter (2016) Eye tracking method compatible with dual-screen mammography workstation. Procedia Computer Science, 90 . 206 - 211.

Downloaded from: http://insight.cumbria.ac.uk/2438/

International Conference On Medical Imaging Understanding and Analysis 2016, MIUA 2016, 6-8 July 2016, Loughborough, UK

Eye Tracking Method Compatible with Dual-Screen Mammography Workstation

Leng Dong[a]*, Yan Chen[a], Alastair Gale[a], Peter Phillips[b]

[a]Loughborough University, Loughborough, LE11 3TU, UK
[b]University of Cumbria, Lancaster, LA1 3JD, UK

Abstract

In this paper a new approach is proposed to track the perceptual behaviour of radiologists when they examine mammographic images displayed on large dual clinical monitors. Zooming and panning are inevitably performed by the radiologist to examine such large images by using the DICOM viewing software. Such image manipulating movements on the target displays makes eye tracking techniques difficult to perform and also the size of the dual clinical monitors makes existing eye tracking techniques generally inadequate. Hence a method using the Smart Eye Pro eye tracker and optical character recognition techniques was designed to relate the recorded radiologists' eye gaze behaviour on the monitors to the actual zoomed and panned medical image areas. This then allows clinical studies involving radiologists interacting with these mammographic images to be successfully carried out.

Keywords: Medical Imaging, Image Perception, Eye Tracking, DICOM, Mammography, Optical Character Recgonition;

1. Introduction

Medical imaging research often examines the performance of radiologists when they examine different types of images. A fairly common behavioural approach is to use eye tracking, where the radiologist's visual search

---

* Leng Dong. Tel.: +44-01509-635737;
   E-mail address: l.dong@lboro.ac.uk

behavior is recorded as they examine the displayed images and information concerning where they fixated or did not fixate in the image is then linked to their performance. Typically, three types of user error are elaborated by such an approach: errors which are made where the radiologist clearly did not look at or near an abnormality (a visual search error), abnormalities looked at but not detected (a detection error) or abnormalities looked at and information reported as being detected (e.g. micro-calcifications identified) but then this information is not interpreted appropriately (interpretation error).

In general, eye tracking can be undertaken using several different commercially available systems. These are typically either head-mounted glasses or systems which are affixed beneath the medical monitor. For most medical images either approach can work well, however particular difficulties arise with mammography images due to the large clinical monitor sizes. For instance, by monitoring a radiologist's eye movements on a digital mammography image, it is possible to tease out any performance differences between naïve breast screeners and experienced breast radiologists as well as how errors occurred and why experienced radiologists perform better. Previous research has examined the visual behaviour when participants read mammographic images on a clinical dual-screen workstation[1]. This was achieved by using a head mounted eye tracker to monitor the participants' eye movements. As this study reported, difficulties can occur when using a head mounted eye tracker to examine visual search behaviour with such large displays, such as the participant's head movements may cause the head mounted scene camera to lose track of the display. Also, due to the complex set up of the digital mammographic workstation, popular remote eye trackers which can be mounted below the monitors would not be able to be configured compatibly with the dual-screen display system. That is, the large size of the dual monitors exceeds the size of the visual display that can be recorded accurately by these eye trackers.

An additional problem arises with the medical images themselves which are viewed using a DICOM viewer. Radiologists typically use zooming and panning of images to examine fine details; this is especially the case with mammographic images where often the interest is in perceiving whether very small calcifications are present and where these manipulations are a key factor. Currently popular eye trackers on the market will only allow an observers' eye gaze position to be recorded according to a fixed co-ordinate system (i.e. the clinical display) which is defined and calibrated before the actual recording takes place. Any area of interest in the displayed image is then normally defined in relation to the full image being examined. Therefore, when zooming and panning is performed, it is difficult to relate the recorded observer's eye gaze position accurately (which is based on the coordinate model defined by the clinical display) to the actual eye gaze location on the image.

Based on the above issues, a new approach is presented which optimises a current eye tracking method when a clinical mammographic workstation is used and allows for accurate eye tracking enabling appropriate panning and zooming of the medical images.

1. Experimental Set Up

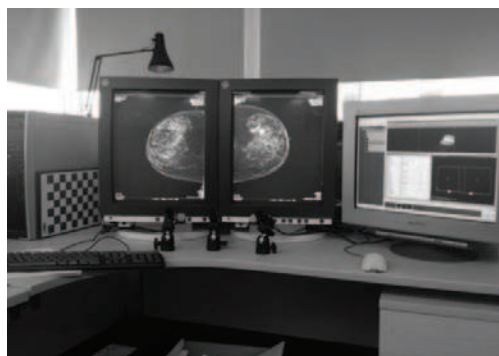Fig. 1. GE Digital Mammography Workstation with Synedra View DICOM viewer

The workstation used was a GE Digital Mammography Workstation (Fig. 1.) consisting of a desktop tower PC with one monitor and two 5MP diagnostic displays. The hardware configuration of this mammography workstation is no different to a standard compatible PC except that a special graphic card which supports the dual-screen output

of 12-bit grayscale images is installed onto the motherboard. Synedra View had been installed to enable PACS archiving and display DICOM images. The software has been approved to be used for clinical use. For our purposes the software is useful as it permits the user to manipulate each mammography image (on different monitors) without affecting the image on the other screen. Also if the user operates zooming and panning controls then the related zooming and panning information is displayed at the bottom right corner of each display. Utilising that information linked to the user's eye gaze data can be leveraged to calculate accurate gaze location information in relation to the displayed images. The following section describes a computer graphic method designed to collect that screen information.

2. Manipulation Behaviour Analysis Tool

In order to gauge where a user is looking when they are examining a mammographic image it is important to know whether the image being viewed is the full size image or some zoomed or panned part of that image. Unfortunately, there is no direct electronic way to acquire that information. However, because zooming and panning information is displayed at the bottom left corner of the diagnostic display it is possible to utilise a screen capture technique to record this information. Then with the help of an optical character recognition technique, it is hoped that images containing this zooming and panning information can be read and saved in a plain text format for subsequent detailed analysis.

2.1. Sub Screen Capture Tool

An application called Sub Screen Capture Tool[2] had been previously developed by one of the authors to capture a specified area of a computer screen repeatedly at a user-defined frequency. Here, the zooming and panning information as displayed by the Synedra View software was the screen target to capture. The default screen shot interval was set to 10 ms. Fig. 2 (a) shows the graphic user interface of Sub Screen Capture where the precise area of the Synedra display to be captured was defined. Fig. 2 (b) shows an example of the captured image. In this captured image example, 'Z: 2.84' indicates that the zoom factor was 2.84. This means the image displayed on the workstation was zoomed in to a new size of 2.84 times the original image size. The additional information 'P: 724/-3261' indicates the precise position information (726, -3261) of the x, y coordinates (measured in screen pixels) of the displayed image centre and the origin is the centre of the workstation screen.

(a)   (b)
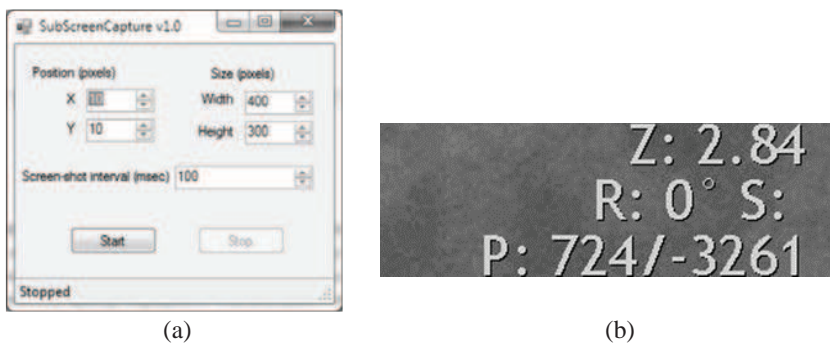
Fig. 2. (a) Sub Screen Capture GUI (b) example of captured picture

As a user viewed images on the workstation and panned and zoomed these images then this tool captured images every 10ms and saved them as jpg files.
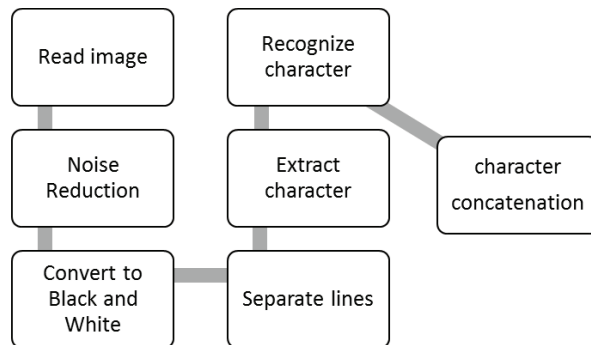
2.2. Optical Character Recognition (OCR)

Fig. 3. flowchart of OCR process

An optical character recognition (OCR) application was then used to extract the required information out of the image and saved this into a text file. The process of OCR involves several steps including segmentation, feature extraction, and classification[3]. A Matlab script using Image Processing Toolbox was developed to achieve this. The flowchart in Fig. 3 shows the brief procedure of the OCR. The captured JPG format image is first read into the computer memory stack and saved as a matrix. Each element in the matrix is equivalent to a pixel in the image. The value of the element is essentially the 8-bit greyscale value. As can be seen in Fig. 4 (a) that captured image is likely to have a noisy background, therefore a high pass threshold filter was applied to get rid of those pixels with lower greyscale value and clean up the image. Therefore, only text pixels with a high degree of greyscale value were kept (Fig. 4.b). In order to optimize image processing time, the filtered image was then compressed into a binary black and white image. After the above initializing process, a horizontal scan of each row in the matrix is performed with the aim of looking for any blank row in the image. If a blank row was spotted, a sub image with the height between the previous identified blank row and the currently identified blank row was extracted to present a single text line (Fig. 4.c). A similar process was then performed in the vertical direction with the aim of extracting a single character (Fig. 4.d).
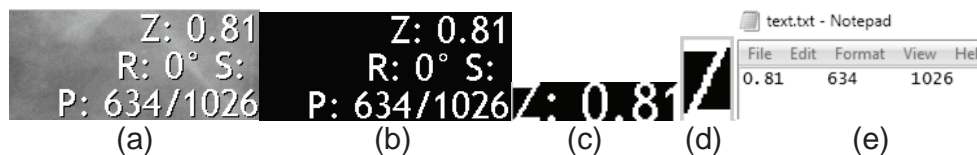

Fig. 4. An example demonstrating the OCR process

After a character is extracted, by computing the correlation coefficient between the extracted character and existing templates, the character with the largest coefficient value is the recognized one. Finally, the recognized numbers and decimals were written into a tab delimited text file (Fig. 4.e). This results in three numbers which represent the zoom factor with the panning x and y values respectively.

2.3. Coordinate Mapping

This information is then used to determine what information is displayed on screen at any point in time. Assume that the panning information is denoted by (x, y) and the zoom factor is denoted by z. The width and length of the mammography image is denoted by w and l. The resolution of the mammography diagnostic screen is p x q. If we want to display a zoomed area on the original image, the centre of the zoomed area (X, Y) is:

$$X = \frac{x}{z} + \frac{w}{2}$$

$$Y = \frac{y}{z} - \frac{l}{2}$$

The zoomed area then is represented by the rectangle which has the same length-width-ratio as the diagnostic screen. And the length (L) and width (W) is

$$W = \frac{p}{z}$$

$$L = \frac{q}{z}$$

2.4. Visualization

To test the accuracy of this function a calibration image was drawn. Five grey circles were drawn at four corners and the centre of the image. This image was then displayed on the mammography diagnostic screen. Each circle was then zoomed to fit the screen and the zoom and position information were manually recorded. The information was collected and plotted onto the original image using coordinate mapping equations. Fig. 5 shows the calibration image with the five zoomed areas. Each zoomed area is represented by red rectangles. These display the content shown on the screen after zooming and panning are applied. It can be seen clearly that all the zoomed areas exactly found their target circles. This means that this approach satisfies the accuracy requirement.
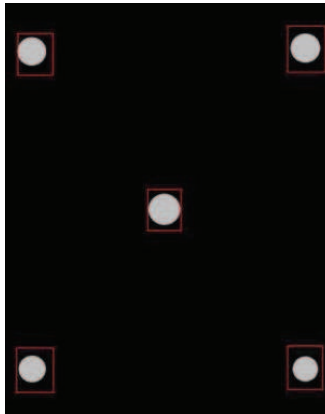

Fig. 5. Calibration image with the five zoom areas

3. Smart Eye Pro System

The Smart Eye Pro system is a head and gaze tracking system well suited for demanding environments, such as a vehicle cockpit where a lot of head movement occurs, and has the flexibility to cope with most research projects. The system measures the subject's head pose and produces eye gaze direction in full 3D. Additionally eye lid opening values and pupil dilation measurements can also be obtained. The system can be used with up to six cameras with different lenses, allowing for a very large field of view.

For this research, a Smart Eye Pro system consisting of 3 cameras (Fig. 6(a)) was used to track a user's eye movements on the two digital mammography screens. As the mammography workstation has two diagnostic screens, this set up is able to provide a large enough tracking angle. The disadvantage of using the Smart Eye Pro system is that it does not provide enough support for on screen tracking. This means that the user has to create a real world model of the computer screen with very accurate spatial data. This is even more challenging for this research as the surface of the diagnostic screen is not fully flat. A real world model of the two diagnostic screens was created by measuring the length between the cameras and the surface of the screens. A pixel is used as the unit of length to obtain the coordinate data. Besides this, measuring the angle of the screen surface is also a complex process. As the surface of the diagnostic screen is actually a curved surface, this will inevitably cause some inaccuracy happening when tracking someone's eye gaze. Fig. 6.b shows the real world model created using pixel information. The two rectangles represent the two diagnostic screens. The origin is the central camera. The two yellow circles are the two LED flashes installed on the two wing cameras. Fig. 6.c shows the monitoring panel when the eye tracking mode is

on. The Smart Eye Pro software analyses the participant's gaze direction from three cameras and decides the intersection point on the real world model of the diagnostic screen. The log information shows in the dialog window at the bottom left corner. The coordinate data of the intersection point can be exported from the log window.
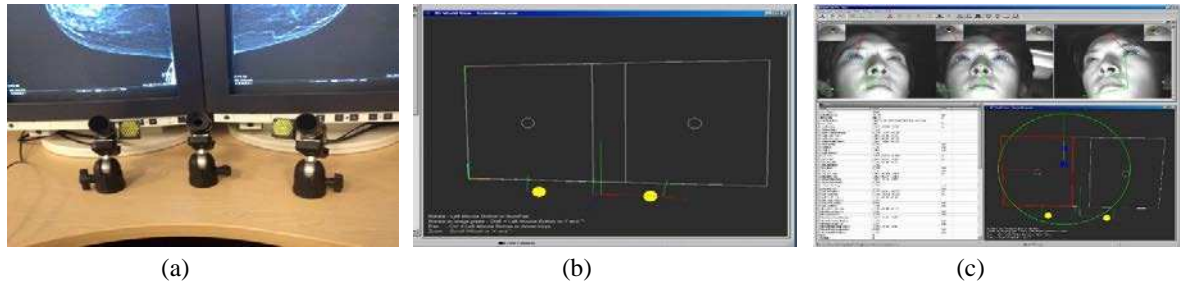

| (a) | (b) | (c) |

Fig. 6.(a) Smart Eye Pro system consists of 3 cameras (b) real world model (c) Smart Eye Pro software monitoring panel

5. Timestamp matching

The Smart Eye Pro system provides a way to obtain participants' gaze position on the screen coordinate plane. The zoom analysis tool described earlier has the ability to obtain position information of the image displayed on the screen after zooming and panning are applied. By combining the two systems, it is then possible to locate a participant's gaze position on the mammography image even when this image is zoomed and panned.

The Sub Screen Capture tool operates at a frequency of 10Hz and the Smart Eye Pro log file uses a 60Hz frequency. Therefore it is possible to match the timestamp by reducing the Smart Eye Pro log file's frequency to 10Hz. Both sets of data use the UNIX timestamp, which provides a more convenient way to match the timestamp.

6. Discussion and conclusion

In this paper how to accurately track the visual search behaviour of radiologists as they examine mammographic images displayed on large dual clinical monitors is addressed. The radiologist will inevitably zoom and pan these images. Such movement of the displayed target images is a problem for eye tracking. Whilst various eye tracking techniques exist, none can easily handle such large displays and also allow for the radiologist to move their heads about a lot – hence the Smart Eye Pro system was used here.

Relating the recorded eye gaze behaviour on the monitors to the actual zoomed and panned displayed image was handled by a newly designed method. Overall the new techniques described here allows clinical studies to be successfully carried out of how radiologists interact with these large mammographic images. On the other hand, limitations were also spotted when conducting this approach. The mammography workstation used in this study consists of two curved CRT (Cathode ray tube) monitors and as demonstrated in Fig. 6 (b) only flat surfacescan be defined as the tracking target in the Smart Eye Pro configuration tool. This issue inevitably will cause inaccuracy when relating the eye movement data onto the simulated coordination map. It is hoped that by replacing CRT monitors with flat panel displays, a better eye movement data accuracy can be achieved. Another limitation about this approach is that two processes (screen capture and OCR) are used to monitor participants' manipulating behaviour which provide an indirect way of information extraction and consume high computing power. Future improvements also includes seeking deeper cooperation with manufactures or DICOM viewer developers to discover a more straightforward way to acquire such information.

References

1. Chen, Y., 2010. An investigation of workstation image manipulation usage when examining FFDM images. Breast Cancer Research, 3(12).
2. Phillips, P., 2012. Sub Screen Capture Tool
3. Schantz, H. F., 1982. The history of OCR, optical character recognition. In: Recognition Technologies Users Association. Manchester
4. Smart Eye, 2016. Smart Eye Pro. [Online] Available at: http://smarteye.se/products/smart-eye-pro/ [Accessed May 2016].