Original Investigation

Fixated and Not Fixated Regions of Mammograms: A Higher-Order Statistical Analysis of Visual Search Behavior

Suneeta Mall, BE, Patrick Brennan, PhD, Claudia Mello-Thoms, PhD

Rationale and Objectives: Visual search is an inhomogeneous yet efficient sampling process accomplished by the saccades and the central (foveal) vision. Areas that attract the central vision have been studied for errors in interpretation of medical imaging. In this study, we extend existing visual search studies to understand what characterizes areas that receive direct visual attention and elicit a mark by the radiologist (True and False Positive decisions) from those that elicit a mark but were captured by the peripheral vision. We also investigate if there are any differences between these areas and those that are never fixated by radiologists.

Materials and Methods: Eight radiologists participated in this fully crossed multi-reader multi-case visual search study of digital mammography (DM) involving 120 two-view cases (59 cancers). From these DM images, 3 types of areas, namely Fixated Clusters (FC), Marked Peripherally Fixated Clusters (MPFC) and Never Fixated Clusters (NFC), were extracted and analysed using statistical information theory (in the form of third and fourth-order cumulants and polyspectrum [specifically bispectrum and trispectrum]) in addition to traditional second-order statistics (in the form of power spectrum) and other nonspectral features to characterize these types of areas.

Results: Our results suggest that energy profiles of FC, MPFC, and NFC areas are distinct. We found evidence that energy profiles and dwell time of these areas influence radiologists' decisions (and confidence in such decisions). We also noted that foveated areas are selected on the basis of being most informative.

Conclusion: We show that properties of these areas influence radiologists' decisions and their confidence in the decisions made.

Key Words: Visual search; perception; mammography; high-order analysis; spatial frequency analysis.

INTRODUCTION

Because of the large complexity of the world that surrounds us, and the impossibility of sampling it all with a high-resolution lens, the human visual system has evolved to apply inhomogeneous sampling to any given scene. In this system, high detail is supported in the central vision, the *fovea centralis*, and resolution decays rapidly as one moves toward the periphery (with visual acuity dropping by 50% at about 10° of visual angle (1)). To be able to avoid missing important information in the world, the eyes move in a ballistic fashion from one point to the next, in a process called saccades (2), reaching speeds of 700°/sec (3). This allows the deployment of the foveal vision to all parts of a given scene, and efficient sampling of data to be carried out. Although many aspects of this system are well understood (such as its timing parameters, motor control, etc (4)), the selection of the regions to which the foveal vision is deployed is still an active area of research. This is a critical problem not only for understanding human behavior but also for developing efficient and effective computer-based search strategies.

Moreover, in the human visual system, the visual stimuli of various spatial locations and orientations (5) are processed by three different types of cortical cells—classified as simple, complex, and hypercomplex (6,7). Hypercomplex cells are known to process signals of different orientations and length and are considered of high order as opposed to simple cells that process signals linearly (6,7). To simulate the processing of hypercomplex cells, we aimed to perform high-order statistical analysis. Higher-order statistics can also reveal complex differences such as structural properties (eg, corners, junctions, curved lines, curved edges) of a region (4,8) that cannot be described by properties such as luminance, contrast, or spatial variance (obtained through autocorrelation or power spectrum).

In mammography, previous studies have shown that radiologists' eyes are attracted to the locations of cancers that are not reported in 70% of the cases (9–11). This suggests that something "interesting" caught the visual system's attention,

Acad Radiol 2017; ■:■■–■■

From the Medical Image Perception and Optimization Group (MIOPeG), Faculty of Health Sciences, University of Sydney, Lidcombe, NSW 2141, Australia. Received August 30, 2016; revised November 19, 2016; accepted November 21, 2016. **Address correspondence to:** S.M. e-mail: smal5514@uni.sydney.edu.au

http://dx.doi.org/10.1016/j.acra.2016.11.020

but processing of the acquired information led to the area being dismissed (ie, being ruled out as containing a cancer). Using spatial frequency analysis, it has been shown that cancers that are correctly reported (True Positive decisions) differ in their wavelet packets energy profile from cancers that attract visual attention but are not reported (False Negative decisions) (12–14). Both of these differ from areas of the background that are sampled but neither contain a lesion nor receive a report of containing one (True Negative decisions) (12–14). However, what characterizes areas that receive direct visual attention and elicit a mark by the radiologist (True and False Positive decisions) from those that were never fixated? And what differences (if any) exist between these areas and those that are only captured by peripheral vision?

In this paper we will use statistical information theory (in the form of third and fourth-order cumulants and polyspectrum [specifically bispectrum and trispectrum]) in addition to traditional second-order statistics (in the form of power spectrum) and nonspectral features to characterize these types of areas. In other words, we will characterize areas that attracted attention (as measured by either direct or peripheral fixation involvement) from those that never interested the radiologists. We hypothesize that foveated areas are selected on the basis of being most informative (ie, they have the least amount of redundant features) (8).

MATERIALS AND METHODS

Ethical approval required to conduct this study was obtained from the Institutional Review Board of the University of Pittsburgh (IRB #PRO09040434) where data were collected.

Study Design

Eight Mammography Quality Standards Act-certified breast radiologists participated in this study. The cases were obtained from a routine screening program using a Selenia full-field digital mammography system (Hologic Inc, Bedford, MA). A total of 120 two-view (craniocaudal [CC] and mediolateral oblique [MLO]) cases, including 59 biopsy-proven cancers and 61 lesion-free (stable for 2 years) noncancer cases were selected. Forty-three of the 59 cancer cases had a malignancy appearing in both MLO and CC view whereas the remaining malignant cases had cancer being depicted in one view only (6 cases in the CC view and 10 cases in the MLO view). In this test set, each case depicted at most one malignancy, resulting in a total of 59 cancer lesions that appeared in 102 views whereas the remaining 138 views were normal. Participants of the study were unaware of the number of lesions present in the case set and were allowed to mark as many cancers as they deemed appropriate. In this fully crossed study design, each radiologist interpreted the 120 two-view cases in a different randomized order in two separate sessions that lasted about an hour.

Other Mammography Quality Standards Act-certified breast radiologists, who did not participate as an observer in this study, using pathology reports and additional imaging, established ground truth. Along with the truth, the center of the malignant lesion in their respective views was recorded in a "truth table", which was used to evaluate the accuracy of the participating radiologists' markings.

Study Protocol

The radiologists were seated 60 cm from a workstation that contained two calibrated medical-grade 5 megapixel flat-panel portrait-mode displays (model C5i, Planar Systems Inc., Beaverton, OR), with a resolution of 2048×2560 pixels, typical brightness of 146 ftL, and 3061 unique shades of gray. The radiologists wore a head-mounted eye-position tracking (ET) system (ASL Model H6, Applied Sciences Laboratory, Bedford, MA) that used an infrared beam (at temporal resolution of 60 Hz) to calculate line of gaze by monitoring the pupil and the first corneal reflection. A magnetic head tracker was used to monitor head position, and this allowed the radiologists to freely move their heads from side to side as well as toward the displays, up to 20 cm, at which point they were outside the range of the head tracker. The ET integrates eye position and head position to calculate the intersection of the line of gaze and the display plane. The system has an accuracy (measured as the difference between true eye position and computed eye position) of less than 1° of visual angle, and it covers a visual range of 50° horizontally and 40° vertically.

Radiologists were instructed to identify malignant lesions only and score such lesions on a 5-point confidence scale, with 1 indicating a 1%–20% and 5 indicating an 81%–100% confidence that a cancer was present at the location.

Prior to the beginning of each reading session, a calibration of ET was performed wherein a 3×3 grid was shown on both the displays. After every five cases, the ET system was rechecked and if necessary, it was recalibrated, but this was only required twice at most during each reading session.

After the calibration, the first (or next) case appeared on the displays wherein the left- and right-hand side monitors would respectively display CC and MLO view of the case. The eye tracker captured the X and Y co-ordinates of fixation location on ASL plane, dwell time, view, radiologists' distance to the monitor, and other details. Radiologists were advised to mark the location of malignant lesions on the screen if and when they found it, along with providing a confidence score. The software had the capability to capture both these pieces of information on screen with the help of pop-up dialog boxes. Upon termination of search for a given case, the radiologists used a mouse-controlled cursor to click on a button in the display to select the next case of their reading sequence and were not allowed to come back to previously assessed cases.

Data Processing

The raw data obtained from the ET system required further processing not only to clean it from blinks and convert it into the same coordinate system as the display but also to cluster