Perceptually Based FROC Analysis[1]

Rachna Arora, MBBS, MPH, Harold L. Kundel, MD, Craig A. Beam, PhD

Rationale and Objectives. Analysis of reading data when cases have multiple targets and/or the reader is required to localize targets is difficult. One approach to this free-response operating characteristic (FROC) problem is for images to be segmented (eg, with quadrants) by the investigator and a segment-level analysis be conducted with the case as a nesting factor. In this report, we introduce an alternative method that uses the visual scan path of the reader to segment the image. We evaluate the new method by applying it to data from a mammography reading experiment.

Materials and Methods. The gaze scan path of one radiologist was recorded as she scanned 40 mammograms for masses and microcalcifications. The observer is an experienced mammographer and was not one of the authors. In addition, the reader provided a rating indicating the degree of suspicion for any suspected targets she identified and localized. We then established "perceptual regions" by using a clustering algorithm on the visual fixations. We combined ratings given to specific locations indicated by the reader with the segmentation from the visual scan to generate a series of ratings classified for whether the perceptually based region associated with the rating contained or did not contain a known target. We analyzed data generated by our method from all 40 cases by using the conventional maximum-likelihood method based on the binormal model. Finally, we tested goodness-of-fit of the binormal model to the data by using chi-square.

Results. Maximum-likelihood estimation led to a model that did not fit the data ($P < .001$). However, examination of the observed and expected counts suggests that the binormal assumption does not hold for segments that contain targets and a bimodal distribution model might be preferred.

Conclusion. Our new method provides an alternative approach to analysis of the FROC experiment. It needs to be developed further. Specifically, we propose that a mixture model extension of the binormal model be developed for ratings data arising from perceptually based FROC experiments. A disadvantage to our method is the requirement to record the scan path of the reader. However, we believe that adding such information to receiver operating characteristic (ROC) curve analysis will pay off when appropriate statistical models have been identified because we believe our data support our hypothesis that the perceptual scanning of images by humans deconvolves interpretation correlation. If true, this hypothesis implies that conventional statistical methods for ROC analysis based on independent data can be applied to the analysis of FROC data after conditioning on the scan path of the observer.

Key Words. Receiver operating characteristic (ROC) curves; eye-position recording; binormal model; maximum-likelihood estimation.

In recognition of the ability of humans to alter decision thresholds in experimental and clinical settings, receiver operating characteristic (ROC) analysis has become the preferred method for evaluating sources of interpretive error and comparing performance. The ROC approach to characterizing diagnostic accuracy offers a fundamental advantage because the ability to discriminate abnormal from normal cases is unconfounded with decision threshold (1). However, the current ROC method is very limited for the analysis of images with multiple responses and/or multiple targets (referred to as "free-response" reading

Acad Radiol 2005; 12:1567–1574

[1] From Biostatistics Core, H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Drive, Tampa, Florida 33612 (R.A.); Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania (H.L.K.); and Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Illinois (C.A.B.). Received April 8, 2005; revision received June 28; accepted June 29. **Address correspondence to:** R.A. e-mail: arorar@moffitt.usf.edu.

doi:10.1016/j.acra.2005.06.015

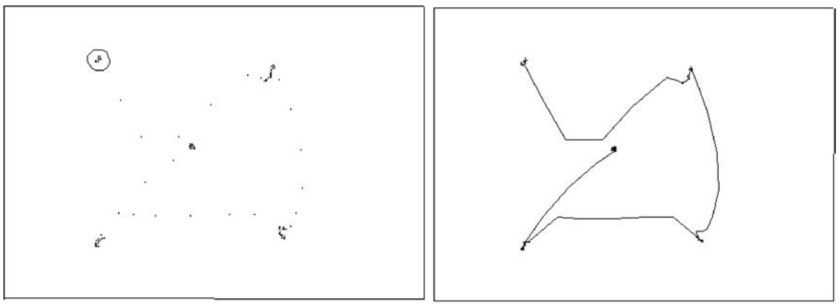

Figure 1. Raw data as dots and scan path.

experiments), a situation often reflective of clinical reality (2–6).

In this report, we introduce a model for the analysis of free-response operating characteristic (FROC) curves that uses the scan path (7,8) of a reader to create perceptually defined regions to partition the image before conventional ROC analysis. This "perceptual-segmentation" model then provides a probability model with a structure that can be exploited to derive a joint likelihood function conditional on the observed segmentation in a set of images. Our method can be contrasted with other segmentation approaches (9) that use segmentation independent of the reader's perceptual scanning of the image (eg, the use of quadrants).

MATERIALS AND METHODS

One reader, a full time mammographer who is not a study author, read 40 mammograms while the gaze was tracked by a head-mounted eye-head tracker (ASL model 4000SU; Applied Science Laboratories, Bedford, MA). The recordings were part of a larger study of the effect of expertise on eye-scanning behavior. Readers were instructed to search the mammograms for cancer as they would in a screening situation. Eye-scanning data and interpretation were recorded by one of the authors (H.L.K.).

The purpose in recording eye position is to determine where the reader is directing visual attention in the displayed image (8). It is assumed that the center of attention on the image is indicated by the axis of the gaze. Properly calibrated eye-position recording can relate the location of the axis of the gaze to locations in the displayed image. The resolving power of the retina is greatest in the fovea, which is a small central region of the retina on the axis of the gaze. Resolving power decreases exponentially toward the periphery. Consequently, detail is seen best on the axis of the gaze. Accurate location of the position of the axis of the gaze on the displayed image requires careful calibration and either monitoring eye position with the head immobilized or monitoring both head and eye position.

The scan path traced over the scene by the axis of the gaze consists of a series of rapid jumps (traditionally called saccades or macrosaccades) with intervening fixations when the gaze is relatively stationary. During fixation, the axis of the gaze drifts and there are small corrective microsaccades.

Converting Raw Data to Fixations

Raw eye-position data consist of a stream of (x, y) coordinates acquired at 50 to 60 samples/s. Figure 1 shows a person deliberately looking successively at five disks that subtend a visual angle of 0.5° when viewed at 38 cm. This is a typical calibration pattern. Raw data are shown on the left as dots, and the scan path is shown on the right as a line connecting the dots. A circle, 1° of visual angle in diameter, is drawn around the upper left fixation in the left illustration.

Raw data are reduced to "fixations" (x, y, t) by using an algorithm that finds the running mean of (x, y) and sums the sample time (t). If it is less than 60 milliseconds (three raw data points), data are considered part of a saccade and not part of a fixation. The fixation scan path on the calibration display is shown in Figure 2. Blinks produce short bursts (three to three data points) of x, y data with coordinates well beyond the range of the display that are easily recognized and removed.