The British Journal of Radiology, 80 (2007), 389–391

COMMENTARY

The radiology task: Bayesian theory and perception

T DONOVAN, MSc and D J MANNING, PhD

School of Medical Imaging Sciences, St Martin's College, Bowerham Road, Lancaster LA1 3JD, UK

ABSTRACT. The use of a Bayesian framework to understand how radiologists search images for pathology is important as it formalizes, mathematically, how visual and cognitive processes control eye movements by modelling the ideal searcher against which human performance can be compared. It is important that the interpretation of medical images is understood so that new developments in the ways images are presented and the use of image processing software are matched to human abilities and limitations.

Received 20 September 2006
Revised 13 December 2006
Accepted 14 December 2006

DOI: 10.1259/bjr/98148548

Visual search and eye movements

Radiology is accustomed to change. Technology has imposed this requirement on practitioners in over a century of development. However, recent innovations in display and image presentation – for example, the change to soft-copy viewing of digital images and to sectional stacks, as well as the use of computer aided diagnosis (CAD) – have transformed the task of the observer in ways that few previous developments have. Radiologists are now working in a fundamentally new visual environment when they make their diagnostic decisions, and it is reasonable at this point of change to take stock of their unique abilities and the nature of the task.

In this short commentary, we reflect on the visual task and decision-making process in medical image interpretation and comment on how radiology acts as a source of endless fascination for vision scientists.

The interpretation of medical images and the development of expertise require both perceptual and cognitive skills. Visual search is a fundamental perceptual skill, whereas diagnostic reasoning and decision making are cognitive skills acquired over time through education and experience. At its most basic, the information processing flow to achieve expert interpretation of medical images is: visual search – object recognition – decision making [1]. During the search process, such as looking for a face in a crowd, the visual system has to solve some difficult problems. This is no different for the radiologist when presented with a medical image. How does the radiologist make sense of the different ambiguous shapes, shades and contours when searching for pathology? The use of eye-tracking as a research methodology has produced some interesting data about this and there are, for example, different patterns of eye movements between novices and experts, *i.e.* experts will exclude large areas of the image during a search for lung nodules and they concentrate on regions where, in their experience, nodules are more likely to occur [2]. Eye movements are not generally something we are aware of, but the types of eye movement in visual search are not involuntary. Rather, they can be described in terms of target selection, which in turn is related to the motivational state of the radiologist and to higher cognitive processes [3]. Early vision research in medical imaging found that although most radiologists advocate some form of systematic search, eye-tracking studies do not reflect that radiologists actually do this. It was suggested that the radiologist who has to use a deliberate search strategy is paying attention to search and not to the task of image interpretation [4].

The reason why visual search is necessary relates to the variable spatial resolution of the retina; when presented with an image, the visual system will very quickly perform a global analysis before using high-speed eye movements called saccades to direct the highest resolution region of the retina, in order to fixate potential target locations in the visual scenes. It is during these fixations, which average 200–300 ms, that detection and identification processes are applied across the visual field, and when eye movements to subsequent fixation applications are planned and programmed [5]. Understanding search performance and eye movements is important because they relate directly to determining how errors are made in the detection and location of pathology, and to the understanding of how developing expertise changes the way neural resources are used. There is currently no formal theory of optimal eye movement strategies in conducting visual search [6], but recently there have been a number of important research articles advocating a Bayesian approach to perception, which is directly applicable to medical image perception. The Bayesian approach is attractive because it aims to formalize these ideas mathematically by modelling the ideal searcher and comparing human and ideal search quantitatively [7].

Address correspondence to: Mr Tim Donovan, Senior Lecturer, Medical Imaging Sciences, St Martin's College, Bowerham Road, Lancaster LA1 3JD, UK. E-mail: t.donovan@ucsm.ac.uk

The Bayesian framework

The formulae of the Bayesian framework can represent any biological/psychological system that can be characterized by an input and an output. Scenes and stimuli are not all presented with an equal probability of their detection or of their understanding, so perception can be considered a process of unconscious probabilistic inference [8]. Visual perception is an active process resulting from an interaction between information entering the eyes and mechanisms that relate the incoming information to previous visual experiences and current expectations. This involves a knowledge-rich inferential process [9]. In a Bayesian model of visual decision-making, the brain is seen to represent information probabilistically by coding and computing with probability density functions. It is statistically optimal in this model for the brain to compute the probability of a scene or feature given the prior knowledge of the statistical structure of the visual world and the current input information from the eye (input) [10]. Bayesian theory implies that the posterior density function (output), which is the probability of each possible true state of the scene given the retinal image stimulus, is proportional to the product of three functions. These are the likelihood functions associated with each cue, the prior density function representing the relative probability of each possible state of the scene prior to receiving the stimulus and the costs associated with making different types of errors. This last function can be thought of as a utility function that represents the costs and benefits.

Regarding perception, Bayesian theory takes account of the trade-off between feature reliability and priors [8]. The visual system is forced to guess if it has to make sense of an ambiguous image/scene where several different objects could have produced the same image. The visual system can, however, make intelligent guesses by biasing its guesses towards typical objects or interpretations. Bayes' formula implies that these guesses, and hence perception, are a trade off between image feature reliability as embodied by the likelihood and the prior probability. Some perception may be more prior driven and others more data driven. The less reliable the image features, the more the perception is influenced by the prior, such as geometry, shape and lighting.

How might the Bayesian framework apply to the radiological task? A radiologist is presented with a chest radiograph from a test bank which may or may not have a lung nodule. Depending on experience and expertise, the perceptual system will be primed towards a particular strategy as not all locations on a chest radiograph are equally probable in possible nodule locations. These are the prior probabilities. Initially, radiologists gain a global impression during a first fixation that uses peripheral vision. Experts have an advantage because they will have a strong prototypical idea of normal anatomy, so perturbed regions of anatomy are identified as possible nodule locations. The posterior probabilities of these locations are calculated and an eye movement (saccade) is made to a fixation location that will maximally increase the likelihood of identifying a nodule. This is repeated until search is completed. The Bayesian framework takes account of the criterion used to indicate whether an area of interest is actually a nodule or not.

The Bayesian framework is complementary to other theories such as signal detection theory. This theory can be explained as an explicit quantitative application of statistical decision theory in perception and cognition, as the theory specifies Bayesian ideal observers and optimal decision rules for simple detection and discrimination tasks [11].

The ideal observer

Image quality can be defined in terms of the performance of an observer on a clinically relevant task. Such tasks can be divided into classification and estimation tasks. The simplest classification task is detection of a non-random signal on a non-random background (usually called SKE/BKE (signal known exactly/background known exactly)), although this rarely represents the normal radiological problem. Estimation tasks are concerned with extraction of numerical information from the images, but for both kinds of tasks the idea is to optimize the performance of the observer to the task of interest. We concentrate on classification for the present argument because it best represents the radiologist's main activity. The performance of a human observer is measured by psychophysical studies and analyzed by ROC (receiver operating characteristic) curves; mathematical-model observers are either ideal observers or human models. Ideal observers set an upper limit to the performance of any observer on the task of interest and include the true ideal or Bayesian observer, which sets an absolute upper limit to task performance.

The Bayesian ideal observer is useful as a benchmark against which to compare human performance [5]. This approach also divides the perceptual task up into pieces and then combines them to understand the whole. They are also a useful point for developing realistic models [10]. In Bayesian networks, a generative model specifies the causal relationship between random variables (*e.g.* objects, lighting and viewpoint); one can also specify the task, *i.e.* the costs and benefits associated with different possible errors in the perceptual decision. The models can then allow the inference solution, *i.e.* the Bayesian ideal observer selects the interpretation that has the maximum expected utility.

In the paper by Najemnik and Geisler [6], they developed their ''ideal searcher'' by first characterizing visibility maps of the visual system under consideration for targets and backgrounds of interest. The ideal searcher works by simply collecting responses from possible target locations, updating the posterior probabilities, and then moving the eyes to maximize the new information. To optimally integrate responses across fixations, the ideal searcher accumulates the weighted responses from each potential target location.

Given the prior probabilities of possible target locations and the visibility maps which specify all relevant values of performance, one can simulate behaviour of the ideal searcher. The ideal searcher will make centre-of-gravity fixations. Saccade lengths of the ideal searcher are moderate in size because the posterior probabilities at nearby locations are suppressed, so that recently fixated locations are not returned to and posterior probabilities at distant locations tend not to be increased. Najemnik and Geisler [6] found that this was reflected in humans. Occasional exclusion saccades are made. The ideal searcher does well with parallel detection, integration of information across fixations and selection of the next fixation. Humans, however, are not very efficient at integration of information across fixations as apparently little can be gained by posterior probability information more than one or two fixations in the past. However, a coarse memory is needed to reduce the likelihood of returning to the same display region, and this explains inhibition of return.

Ideal observers can be quite useful in identifying the constraints in perceptual tasks, such as image noise or neural noise, and much of the measured variation in human perception may be due to pre-neural factors [11]. An example of where human performance does not match the performance of the ideal observer, and where the human may be limited by pre-neural factors, is the decline in the identification of familiar features with eccentricity from the fovea, which drops much more quickly than that predicted by the ideal observer [12].

Perceptual learning has also been studied in humans by comparison with an optimal Bayesian algorithm [13]. This showed that humans learn rapidly, and that this rapid initial learning reflects a property inherent to the visual task and not a property particular to the human perceptual learning mechanism.

Relevance for medical image perception

What vision science has demonstrated is that when humans interact with a visual display, the information is not passively absorbed; it is, in fact, cognitive factors that play the dominant role in gaze control. Medical image perception is a complicated task. A radiologist looking for pathology does not simply cover the image in the minimum time using the smallest possible visual field. Research looking at oculomotor strategies in other real world tasks [14, 15] has shown that eye movement patterns must be learned, as observers have learnt what objects in a scene are relevant because almost no fixations fall on irrelevant objects.

These findings are relevant in the design of viewing and reporting stations, and the use of image processing software/CAD so that it is matched to human abilities and limitations. It is pertinent that medical image perception research should concentrate on the observer as well as the technology as it seems interobserver differences are often greater than differences between imaging techniques, of which an excellent example is the study by Weatherburn et al [16]. This work found that the detection of chest lesions did not vary between conventional film, CR (computed radiography) hard-copy and PACS (picture archiving and communication systems) soft-copy images, but there were statistically significant differences between observers.

Bayesian theory is a useful starting point for trying to explain some of the eye-tracking patterns we see, and to provide insight into understanding image interpretation. This is because it provides a mathematical framework for representing the properties of the image, describing the image interpretation task and taking account of the costs and benefits associated with different perceptual decisions.

References

1. Nodine CF, Kundel HL. Using eye movements to study visual search and to improve tumor detection. Radiographics 1987;7:1241–50.
2. Manning DJ, Ethell SC, Donovan T. Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. Br J Radiol 2004;77:231–5.
3. Findlay JM, Gilchrist ID. Active vision: the psychology of looking and seeing. Oxford, UK: Oxford University Press, 2003.
4. Kundel HL, Nodine CF, Thickman D, Toto L. Searching for lung nodules. A comparison of human performance with random and systematic scanning models. Invest Radiol 1987;22:417–22.
5. Geisler WS, Kersten D. Illusions, perception and Bayes. Nat Neurosci 2002;5:508–10.
6. Najemnik J, Geisler WS. Optimal eye movement strategies in visual search. Nature 2005;434:387–91.
7. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 2004;27:712–9.
8. Kersten D, Mamassian P, Yuille A. Object perception as Bayesian inference. Annu Rev Psychol 2004;55:271–304.
9. Guo K, Nevado A, Robertson RG, Pulgarin M, Thiele A, Young MP. Effects on orientation perception of manipulating the spatio-temporal prior probability of stimuli. Vision Res 2004;44:2349–58.
10. Friston K. Functional integration and inference in the brain. Prog Neurobiol 2002;68:113–43.
11. Geisler WS, Diehl RL. Bayesian natural selection and the evolution of perceptual systems. Philos Trans R Soc Lond B Biol Sci 2002;357:419–48.
12. Beckmann PJ, Legge GE. Preneural limitations on letter identification in central and peripheral vision. J Opt Soc Am A Opt Image Sci Vis 2002;19:2349–62.
13. Eckstein MP, Abbey CK, Pham BT, Shimozaki SS. Perceptual learning through optimization of attentional weighting: human versus optimal Bayesian learner. J Vis 2004;4:1006–19.
14. Turano KA, Geruschat DR, Baker FH. Oculomotor strategies for the direction of gaze tested with a real-world activity. Vision Res 2003;43:333–46.
15. Hayhoe M, Ballard D. Eye movements in natural behavior. Trends Cogn Sci 2005;9:188–94.
16. Weatherburn GC, Ridout D, Strickland NH, Robins P, Glastonbury CM, Curati W, et al. A comparison of conventional film, CR hard copy and PACS soft copy images of the chest: analyses of ROC curves and interobserver agreement. Eur J Radiol 2003;47:206–14.