Reverse Engineering the Visual System Via Genetic Programs

Diglio A. Simoni

Bioinformatics Program, Research Computing Division
RTI International, Research Triangle Park, NC 27709 USA
dsimoni@rti.org

Abstract. We propose a datamining based method for automated reverse engineering of search strategies during active visual search tasks. The method uses a genetic program (GP) that evolves populations of fuzzy decision trees and selects an optimal one. Previous psychophysical observations of subjects engaged in a simple search task result in a database of stimulus conditions and concomitant measures of eye gaze information and associated psychophysical metrics that globally describe the subjects search strategies. Fuzzy rules about the likely design properties of the components of the visual system involved in selecting fixation location during search are defined based on these metrics. A fitness function that incorporates both the fuzzy rules and the information in the database is used to conduct GP based datamining. The information extracted through the GP process is the internal design specification of the visual system vis-à-vis active visual search.

Keywords: active visual search, eye tracking, psychophysics, fuzzy logic, datamining, knowledge discovery, genetic programming, reverse engineering.

1 Introduction

A fuzzy logic control process is proposed that automatically selects the next fixation location during the target discovery periods of active visual search. This paper describes an automated procedure for extracting from a database the rules that drive this control process in the form of an optimal fuzzy decision tree.

Section 2 provides the motivation of the work within the context of active visual search. Section 3 briefly introduces a *perception-based* model that incorporates the decision tree structure. Section 4 briefly discussed the ideas of fuzzy logic and the decision tree approach. Section 5 describes how genetic programs can be used to obtain an optimal decision tree for active visual search. Finally Section 6 provides a summary.

2 Active Visual Search

Visual search is accomplished through a cycle of fixations and visual scene analysis interrupted by saccades. A saccade produces a rapid shift of gaze, redirecting the

D.D. Schmorrow, L.M. Reeves (Eds.): Augmented Cognition, HCII 2007, LNAI 4565, pp. 195–200, 2007.

fovea onto a new point in the visual scene. As the visual system reacquires the image data, the visual scene is remapped onto primary visual cortex governed by the physical limits imposed by the retinal photoreceptor layout and the cortical magnification factor. These limits constrain the representational power of cortex and, therefore, also constrain the computational capabilities of the visual system. Given that the results of these visual computations lead to the behaviors that we observe, it is important to understand how these and other physical constraints affect performance. Previous search studies have generated many insights into this question using various visual search tasks [1][2][3][4][5][6][7][8][9]. Although some of these studies permitted eye movements, the focus of the studies has been to characterize processes that occur during a fixation. More recent studies have specifically addressed what processes characterize visual search when the eye is allowed to move freely about the image [10][11][12][13][14][15][16][17][18][19].

With each new fixation, the amount of cortical machinery associated with the various stimuli, as well as the linear distances between their cortical representations, varies as the image information is mapped in a nonlinear fashion onto primary visual cortex to reflect the new foveal direction in the visual scene. Previous studies of active visual search in the monkey have shown that target detection probability is invariant with respect to set size after applying a proper scaling for stimulus density. This has been shown by appropriately normalizing the probability functions for different set sizes using a metric constructed from an average measure of local stimulus density obtained from each set size [16]. This observation, together with those obtained from lateral masking or crowding studies that describe the severe degradation in identification performance that results from the introduction of flanking distractors around a given target [20][21][22], suggests the importance of local stimulus density upon target detection. In addition, evidence exists from monkey studies that active, feature-selective, attentive mechanisms can adjust the effective stimulus density. Under conditions where the display can be segmented by color differences, the nearest neighbor distances and/or cortical separations that account for performance are determined by stimuli containing the target color—essentially discounting the remaining stimuli from consideration [16][23].

3 *Perception-Based* Model of Active Visual Search

A characterization of these types of constraints has been used to derive an extensible *perception-based* model of visual search that incorporates a novel selection mechanism for new fixation location based on hybrid neural network and fuzzy logic representations. The system's design principles are derived from psychophysical observations of human search performance during active visual search tasks via the use of real-time infrared eye trackers [24]. Psychophysical experiments [25] were used to obtain probabilistic measures of both stimulus and neuroanatomical features that constrain the human visual system's real-time selection of image regions during the target discovery periods of active visual search. Mathematical precisiation tools were used to recast the psychophysical metrics as fuzzy predicates in order to develop a rule set which drives a robust model of human search performance that takes into account the intrinsic uncertainty of sensory processing. The final result is a search mechanism composed of a bottom-up neural network-based sensory processing model (which computes a saliency across the image) coupled to a top-down fuzzy expert system (FES) model of search decision processes which can help both test as well as predict human search performance given precisely controlled sets of visual stimuli [26].

A question that remains is whether the search model optimally describes the eye movement strategies that the human foveated visual system uses when faced with the problem of finding a target in a visual scene. One method for assessing performance is to compare the computational model against an ideal Bayesian observer model [27][28][29][30] for search tasks. An ideal observer model uses precise knowledge about the statistics of the visual stimuli that contain the same spectral characteristics as natural scenes, and about its own sensory constraints, to make eye movements that gain the most information about target location. It has been found that humans achieve nearly optimal search performance when compared with a Bayesian ideal observer [17]. Ideally, a computational model of visual search should approach the same level of performance. In this paper we describe a genetic programming approach to evolve populations of fuzzy decision trees and select an optimal model that reverse engineers the search strategies employed by humans engaged in visual search tasks.

4 Datamining a Fuzzy Decision Tree

The particular approach to fuzzy logic used by the *perception-based* model is the fuzzy decision tree [31]. Fuzzy decision trees are extension of the classical AI concept of decision trees. The leaf nodes (those with degree one) are labeled with co-called root concepts. Nodes of degree greater than one are labeled with composite concepts, which are constructed from the root concepts using logical operations like “AND”, “OR” and “NOT”. Each root concept has a fuzzy membership function associated with it. These membership functions are derived from examination of the psychophysical metrics obtained during visual search task experiments. The membership functions for composite concepts are constructed from those assigned to the root concepts using fuzzy logic connectives and modifiers. The structure of the decision tree determines its function. This structure is equivalent to determining the rules that characterize the control process that drives the selection of fixation location during active visual search. Discovering this structure can be recast as a datamining problem because it is equivalent to efficiently extracting valuable non-obvious information embedded in a large quantity of data [32]. Data mining consists of three steps: a) the construction of a database that represents truth which in this case consists of the list of stimuli presented to subjects during active visual search tasks and their behaviors, including eye gaze information and associated psychophysical measures; b) the use of a data mining function to extract the valuable information which in this paper is a genetic program; and c) the determination of the value of the information obtained in the second step, which in our case is the comparison of the resulting model against both the ideal observer model and human performance measures.

5 Discovering the Fuzzy Decision Tree's Structure Using a Genetic Program

A genetic program is a problem independent method for automatically creating graphs that represent computer programs, mathematical expressions, digital circuits, and the like. The procedure evolves a solution using Darwin's principle of survival of the fittest. The method is based on an optimization procedure that manipulates a string of numbers in a manner loosely similar to how chromosomes are modified in biological evolution [33]. An initial population made up of strings of numbers is selected arbitrarily, perhaps at random. A string of numbers is a "chromosome" and each number in the string is a "gene." A set of chromosomes forms a population. Each chromosome represents parameters that optimize a "fitness function". In our case the chromosomes are the fuzzy predicates of a decision tree that describes search strategies during active visual search. The fitness function is a performance index that is to be maximized. In our case the fitness function is a comparison of the performance of the decision tree, given a particular stimulus display and a current fixation location in that display, with the behavioral data recorded during experiments using the same experimental conditions.

The operation of the genetic algorithm proceeds in steps. Beginning with the initial population, "selection" is used to choose which chromosomes should survive to form a "mating pool." Chromosomes are chosen based on how fit they are relative to the other members of the population. More fit individuals end up with more copies of themselves in the mating pool so that they will more significantly effect the formation of the next generation. Next, two operations are taken on the mating pool. First, "crossover" (which represents mating, the exchange of genetic material) occurs between parents. In crossover, a random spot is picked in the chromosome, and the genes after this spot are switched with the corresponding genes of the other parent. In our case, this is equivalent to sub-tree exchange between two decision trees. Following this, "mutation" occurs. Mutation is a change of the value of a randomly selected gene. After the crossover and mutation operations occur, the resulting strings form the next generation and the process is repeated. Another process known as "elitism" is also employed. Elitism consists of copying a certain number of fittest individuals into the next generation to make sure they are not lost from the population. Finally, a termination criterion is used to specify when the algorithm should stop, e.g., when a preset maximum number of generations has been reached or the fitness has not changed significantly in a certain number of generations.

6 Summary

A fuzzy logic based algorithm for optimal construction of decision trees that describe search strategies during active visual search is under is under development. A method for automatically determining fuzzy decision tree structure, and hence the related fuzzy if-then rules from a large behavioral database is discussed. This method uses a genetic program, an algorithm that automatically evolves other computer programs or mathematical expressions.

References

1. Cameron, E.L., Tai, J.C., Eckstein, M.P., Carrasco, M.: Signal detection theory applied to three visual search tasks—Identification, yes/no detection and localization. Spatial Vision 17, 295–325 (2004)
2. Duncan, J., Humphreys, G.W.: Visual search and stimulus similarity. Psychological Review 93, 433–458 (1989)
3. Eckstein, M.P., Thomas, J.P., Palmer, J., Shimozaki, S.S.: A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. Perception & Psychophysics 62, 425–451 (2000)
4. Palmer, J., Verghese, P., Pavel, M.: The psychophysics of visual search. Vision Research 40, 1227–1268 (2000)
5. Strasburger, H., Harvey, L.O., Harvey Jr., L.O., Rentschler, I.: Contrast thresholds for identification of numeric characters in direct and eccentric view. Perception & Psychophysics 49, 495–508 (1991)
6. Treisman, A.: Features and objects: The fourteenth Bartlett memorial lecture. Quarterly Journal of Experimental Psychology: A Human Experimental Psychology 40, 201–237 (1988)
7. Treisman, A.M., Gelade, G.: A feature-integration theory of attention. Cognitive Psychology 12, 97–136 (1980)
8. Wolfe, J.M., Cave, K.R., Franzel, S.L.: Guided search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance 15, 419–433 (1989)
9. Wolfe, J.M., O'Neill, P., Bennet, S.C.: Why are there eccentricity effects in visual search? Visual and attentional hypotheses. Perception & Psychophysics 60, 140–156 (1998)
10. Findlay, J.M., Brown, V., Gilchrist, I.D.: Saccade target selection in visual search: The effect of information from the previous fixation. Vision Research 41, 87–95 (2001)
11. Findlay, J.M., Gilchrist, I.D.: Eye guidance and visual search. In: Underwood, G. (ed.) Eye guidance in reading, driving and scene perception, pp. 295–312. Elsevier, Oxford (1998)
12. Geisler, W.S., Chou, K.L.: Separation of low-level and high-level factors in complex tasks: Visual search. Psychological Review 102, 356–378 (1995)
13. Hooge, I.T., Erkelens, C.J.: Adjustment of fixation duration in visual search. Vision Research 38, 1295–1302 (1998)
14. Maioli, C., Benaglio, I., Siri, S., Sosta, K., Cappa, S.: The integration of parallel and serial processing mechanisms in visual search: Evidence from eye movement recording. European Journal of Neuroscience 13, 364–372 (2001)
15. Motter, B.C., Belky, E.J.: The guidance of eye movements during active visual search. Vision Research 38, 1805–1815 (1998)
16. Motter, B.C., Belky, E.J.: The zone of focal attention during active visual search. Vision Research 38, 1007–1022 (1998)
17. Najemnik, J., Geisler, W.S.: Optimal eye movement strategies in visual search. Nature 434, 387–391 (2005)
18. Shen, J., Reingold, E.M., Pomplun, M., Williams, D.E.: Saccadic selectivity during visual search: The influence of central processing difficulty. In: Hyönä, J., Radach, R., Deubel, H. (eds.) The mind's eyes: Cognitive and applied aspects of eye movement research, pp. 65–88. Elsevier, Amsterdam (2003)
19. Zelinsky, G.J., Rao, R.P.N., Hayhoe, M.M., Ballard, D.H.: Eye movements reveal the spatiotemporal dynamics of visual search. Psychological Science 8, 448–453 (1997)

20. Bouma, H.: Interaction effects in parafoveal letter recognition. Nature 226, 177–178 (1970)
21. Pelli, D.G., Palomares, M., Majaj, N.J.: Crowding is unlike ordinary masking: Distinguishing feature integration from detection. Journal of Vision, 4 12(12), 1136–1169 (2004)
22. Toet, A., Levi, A.: The two-dimensional shape of spatial interaction zones in the parafovea. Vision Research 32, 1349–1357 (1992)
23. Motter, B.C., Holsapple, J.W.: Cortical image density determines the probability of target discovery during active search. Vision Research 40, 1311–1322 (2000)
24. Simoni, D.A., Motter, B.C.: Human search performance is a threshold function of cortical image separation. Journal of Vision 3(9), 228 (2003)
25. Motter, B.C., Simoni, D.A.: The roles of cortical image separation and size in active visual search performance. Journal of Vision, 7 2(6), 1–15 (2007)
26. Simoni, D.A.: Augmented Search for Clinical and Research Images. In: Dylan, D., Morrow, K., Stanney, M., Reeves, L. (eds.) Foundations of Augmented Cognition, 2nd edn., pp. 329–332. Strategic Analysis, Inc. (2006)
27. Green, D.M., Swets, J.A.: Signal Detection Theory and Psychophysics. Wiley, New York (1966)
28. Burgess, A.E., Ghandeharian, H.: Visual signal detection. II. Effect of signal-location identification. Journal of the Optical Society of America A 1, 906–910 (1984)
29. Geisler, W.S., Diehl, R.L.A: A Bayesian approach to the evolution of perceptual and cognitive systems. Cognitive Science 27, 379–402 (2003)
30. Kersten, D., Mamassian, P., Yuille, A.L.: Object perception as Bayesian inference. Annual Review of Psychology 55, 271–304 (2004)
31. Tsoukalas, L.H., Uhrig, R.E.: Fuzzy and Neural Approaches in Engineering (Chapter 5). John Wiley and Sons, New York (1997)
32. Bigus, J.P.: Data Mining with Neural Nets (Chapter 1). McGraw-Hill, New York (1996)
33. Goldberg, D.E.: Genetic Algorithms. In: Search, Optimization and Machine Learning, Addison-Wesley, London (1989)