〔美〕吕忠林（Zhong-lin Lu）
罗跃嘉（Yue-jia Luo） 主编

Progress in Cognitive Science: From Cellular Mechanisms to Computational Theories

认知科学进展：从分子机制到计算理论

北京大学出版社
PEKING UNIVERSITY PRESS

〔美〕吕忠林（Zhong-lin Lu）
罗跃嘉（Yue-jia Luo） 主编

Progress in Cognitive Science: From Cellular Mechanisms to Computational Theories

认知科学进展：从分子机制到计算理论

北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

认知科学进展:从分子机制到计算理论 = Progress in cognitive science:from cellular mechanisms to computational theories:英文/(美)吕忠林,罗跃嘉主编.—北京:北京大学出版社,2013.8

ISBN 978-7-301-22154-9

Ⅰ.①认… Ⅱ.①吕… ②罗… Ⅲ.①认知科学-英文 Ⅳ.①B842.1

中国版本图书馆 CIP 数据核字(2013)第 028368 号

书　　名:**Progress in Cognitive Science: From Cellular Mechanisms to Computational Theories**(认知科学进展:从分子机制到计算理论)

著作责任者:〔美〕吕忠林　罗跃嘉　主编

责 任 编 辑:谢　娜　陈小红

标 准 书 号:ISBN 978-7-301-22154-9/B·1110

出 版 发 行:北京大学出版社

地　　址:北京市海淀区成府路 205 号　100871

网　　址:http://www.pup.cn

新 浪 微 博:@北京大学出版社

电 子 信 箱:zpup@pup.cn

电　　话:邮购部 62752015　发行部 62750672　编辑部 62752021
出版部 62754962

印 刷 者:北京大学印刷厂

经 销 者:新华书店

730 毫米×980 毫米　16 开本　27.75 印张　506 千字

2013 年 8 月第 1 版　2013 年 8 月第 1 次印刷

定　　价:160.00 元

Preface

Cognitive science is the interdisciplinary study of mind and intelligence, in which psychology, neuroscience, artificial intelligence, anthropology, optometry, speech and hearing sciences, linguistics, biology, information science, and philosophy play major roles, computer science, statistics, informatics, and physics contribute important techniques and technical methods, and education, business, and government agencies are major areas of application. It is concerned with all our mental abilities — perceiving, learning, remembering, thinking, reasoning, and understanding. The research methods in cognitive science are similar to and compatible with the natural sciences, and especially physics, with behavioral experiments and brain measurements providing data that are fit with mathematical and computer simulation models. As brain imaging techniques have become increasingly available in recent years, cognitive science has seen especially pronounced growth of cognitive neuroscience. In addition to addressing the profound scientific problems of the nature of mind and brain, cognitive science also develops practical technology for constructing intelligent systems.

There are currently three main approaches in cognitive science, although the three overlap extensively: experimental psychology, computational cognition, and cognitive neuroscience. Experimental psychology applies experimental methods to collect behavioral data to investigate human cognition. Human choices, psychophysical responses, response times, and eye tracking are often measured and collected in experimental cognitive psychology. Computational cognition develops formal mathematical and computational models of human cognition based on either symbolic and subsymbolic representations, dynamical systems, or some combination of the two. Cognitive neuroscience investigates the biological mechanisms underlying cognition, with a specific focus on the neural substrates of mental processes and their behavioral manifestations that are typically based on measurements of brain activity. It addresses the questions of how psychological/cognitive functions are produced by the neuralcircuitry and even chemical activity. The three approaches are often inter-linked and provide both independent and complementary insights in every sub-domain of cognitive science.

The intellectual origins of cognitive science arose in the mid-1950s with the advent of mathematical and computational modeling, theoretical formalisms for language and thought, and theories of mind based on complex representations. Since 1970, more than sixty universities in North America and Europe have established cognitive science programs and many others have instituted courses in cognitive science. Cognitive science research has produced an extensive body of principles, representations, and algorithms. Successful applications range from treatment of mental illness through economics, business, and custom-built expert systems to mass-produced software and consumer electronics.

Realizing that cognitive science is not only one of the frontiers in scientific research, but also critical for economic development, new-generation technologies, and human health and well-being, the Chinese government declared cognitive science one of the five top national priorities of scientific research in 2005. This declaration takes notice of the fact that cognitive science is a relatively new discipline, making it possible for Chinese scientists to develop first-rate programs in a relatively short period of time. Chinese developments have thus far focused on neural approaches: Much like what happened in the United States, every major Chinese University now views MRI as an instrument establishing scientific credibility in the field, and one required to maintain and/or improve their academic status. Thus many universities have, or are about to purchase, state of the art MRI instruments. A number of brand new cognitive science centers, institutes, and key laboratories have emerged in the recent years with a focus on neural measurements and associated techniques.

Supported by the US National Science Foundation, US Asian Office of Aerospace Research and Development, University of Southern California, and Beijing Normal

University, a workshop, "Cognitive Science: From Cellular Mechanisms to Computational Theories", was held in Beijing, China from May 25th to 27th, 2009. The workshop was organized by Drs. Zhong-Lin Lu, YuejiaLuo, Xiaoping Hu, Guoqiang Bi, RichardShiffrin. The workshop has several related aims, each designed to promote Cognitive Science in China: (1) promote interchanges across multiple sub-disciplines of cognitive science that investigate cognition at different levels and using different tools, (2) survey the current state of cognitive science research in China and showcase to scientists and students in China the quantitative methods, successes, and interdisciplinary nature of cognitive science, and (3) establish channels for collaborative research, and initiate exchange programs at multiple levels.

The workshop brought researchers from a broad spectrum of distinct and connected sub-disciplines together to seek cross-disciplinary interactions and collaborations. The workshop was a tremendous success. We had 17 distinguished speakers and more than 180 participants (postdocs and graduate and undergraduate students) from many major universities, research institutes and organizations. The poster session had 27 presentations. In addition, there was a two-hour discussion session in the end of each day during the workshop.

While this book originated as the proceedings of the conference, it has been organized and expanded to cover Cognitive Science in a coherent manner. The editors solicited new chapters from outstanding scholars. Contributors include distinguished scientists from many sub-areas of cognitive science. Some are widely known outside the field of cognitive science. The list of authors includes many who are members of their nations academies of science, recipients of prestigious scientific prizes, and authors of some of the most widely cited papers in cognitive science. Most of the authors have made significant contributions to the field of cognitive science. We are extremely pleased that this distinguished group agreed to participate in the workshop, and to contribute chapters to this book.

This book covers recent progress in cognitive science from cellular mechanisms to computational theories. We think it would be suitable for upper-level undergraduate and graduate courses on cognitive science. But primarily, the book is intended for scientists, graduate and postdoctoral students working in areas of cognitive science.

All the contributors and many other people have given us a great deal of help in organizing the workshop and producing this book. In particular, we thank the staff and students, especially Wanjun Lin and Jinchuan Chen, in the Key Laboratory of Cognitive Neuroscience and Learning for their help in organizing the workshop. We wish to thank Xiaohong Chenfrom Peking University Press. We also thank NSF, AOARD, USC, and Beijing Normal University for their generous support.

Zhong-Lin Lu, Columbus, Ohio, USA

YuejiaLuo, Beijing, China

February, 2013

Contents

Contributions of Ideal Observer Theory to Vision Research

Wilson S. Geisler

(Center for Perceptual Systems and Department of Psychology, University of Texas at Austin)

Abstract: An ideal observer is a hypothetical device that performs optimally in a perceptual task given the available information. The theory of ideal observers has proven to be a powerful and useful tool in vision research, which has been applied to a wide range of problems. Here I first summarize the basic concepts and logic of ideal observer analysis and then briefly describe applications in a number of different areas, including pattern detection, discrimination and estimation, perceptual grouping, shape, depth and motion perception and visual attention, with an emphasis on recent applications. Given recent advances in mathematical statistics, in computational power, and in techniques for measuring behavioral performance, neural activity and natural scene statistics, it seems certain that ideal observer theory will play an ever increasing role in basic and applied areas of vision science.

CA email: geisler@psy. utexas. edu

Introduction

The major goal of basic vision research is to understand and predict visual performance. Empirical progress toward this goal has come from measurements of natural stimuli, physiological optics, anatomy and neurophysiology of visual pathways, and behavioral performance in adult and developing organisms. Empirical findings in vision research have been interpreted and driven by a wide array of qualitative and quantitative theories and models. Of the quantitative theories, the theory of ideal observers has played a unique and fundamental role, especially during the last 25 years.

There are many different visual tasks human and non-human primates perform under natural conditions and can perform under laboratory conditions. What is ultimately desired is a general theory that parsimoniously explains and quantitatively predicts visual performance in arbitrary natural and laboratory visual tasks. The field is a very long way from such a theory. Instead, vision researchers have been forced to identify specific well-defined tasks, or families of tasks, and then attempt to develop informal or formal models that can explain and predict performance in those specific tasks. For each task or family of tasks the field typically attempts to address a number of fundamental questions, which include: What are the properties of the stimuli in a given task that contribute to the measured performance? How and where are those properties encoded into neural activity along the visual pathway? How are the different sources of task-relevant sensory information combined by the visual system? What are the relative contributions of peripheral and central mechanisms in the task? What are the contributions of "bottom-up" and "top-down" mechanisms in the task? How is the task-relevant information in the neural activity along the visual pathways decoded into behavior?

Twenty-five years ago ideal-observer theory had only been worked out and applied to a very narrow range of simple tasks. In the intervening years it has been applied to much wider range of tasks. This article attempts to summarize some of the different kinds of tasks where ideal observer theory has played a major role in developing models of visual performance and in answering one or more of the questions listed above. Due to space limitations, the primary focus is on behavioral performance, even though ideal observer theory has also played an important role in studies of the underlying neurophysiology. Before getting down to specific tasks, there are some general points to make about the theory of ideal observers.

Ideal Observers

An ideal observer is a hypothetical device that performs a given task at the optimal level possible, given the available information and any specified constraints. If the ideal observer can be derived for a given task, then it can serve vision research in several important ways:

1. Identifying task-relevant stimulus properties

The ideal observer performs its task optimally; thus, in deriving the ideal observer one is forced to identify, at least implicitly, all the task-relevant properties of the stimuli. This makes it possible to rigorously evaluate and test which relevant stimulus properties real observers exploit when they perform the task.

2. Describing how to use those properties to perform the task

The ideal observer explicitly specifies one set of computations that is sufficient to achieve optimal performance in the task. Although there may be other sets of computations that are sufficient to achieve optimal or near-optimal performance, an ideal observer often provides deep insight into the computational requirements of the task.

3. Providing a benchmark against which to compare the performance of real or model vision systems

The performance of the ideal observer is a precise "information measure" that describes how the task-relevant information varies across stimulus conditions. In general, real and model (heuristic) observers do not efficiently use all the task-relevant information and hence do not reach the performance levels of the ideal observer. However, if a real or model observer is exploiting the same stimulus properties as the ideal observer, then its performance should parallel that of the ideal observer (e.g., stimulus conditions that are harder for the ideal observer should be harder for the real or model observer). When human performance approaches ideal performance, then the implications for neural processing can become particularly powerful; specifically, all hypotheses (model observers) that cannot approach ideal performance can be rejected. When human performance is far below ideal, there are generally a greater number of models than could explain human performance.

4. Suggesting principled hypotheses and models for real performance

Natural selection and learning during the lifespan necessarily drive perceptual systems in the direction of optimal performance in the tasks the organism normally performs in its natural environment. Although perceptual systems may not reach optimum, it is a good bet that they are closer to ideal than to the simple models one might generate from intuition or to explain some experimental result. Thus, a powerful research strategy is to use the ideal observer to guide the generation of hypotheses and models of real performance. This is often done by degrading the ideal observer with hypothesized neural noise or with hypothesized heuristic computations that approximate ideal computations. Models generated this way are principled and often have very few free parameters.

In all visual tasks, performance is limited at least in part by various sources of random variability. These include variability in the stimuli (e. g. , photon noise, heterogeneity of the objects defining a category, variability in scene illumination, variability due to the projection from a 3D environment to the 2D retinal images), variability in the sensory neural representation (e. g. , sensory neural noise), and variability in the decoding circuits (e. g. , decision and motor neural noise). Thus, ideal observers are properly defined in probabilistic terms, using statistical decision theory and information theory. Most of the ideal observers described here fall within the framework of Bayesian statistical decision theory.

The logic and structure of a Bayesian ideal observer is relatively straight forward. In most visual tasks, there is some actual unknown state of the world ω (e. g. , a particular class of physical object) that gives rise to a particular (random) received stimulus S reaching the eyes. The observer's goal is to make the response r_{opt} that maximizes the utility (or equivalently minimizes loss) averaged over all possible states of the world (in that task), given the stimulus S. If some biological constraints are included, then the goal becomes maximizing utility given a neural representation of the stimulus $Z = g(S;\theta)$, where $g(S;\theta)$ is the constraint function that specifies the mapping of the stimulus into a neural representation. For example, Z might represent the number of photons absorbed in each photoreceptor, and $g(S;\theta)$ the mapping from the stimulus at the eyes to photons absorbed in each photoreceptor. (The *symbol* θ is included because in some applications of ideal observer theory it is useful to allow unknown parameters in the mapping from stimulus to neural representation; see later.) Formally, the ideal observer's response is given by

$$r_{\mathrm{opt}}(Z) = \underset{r}{\operatorname{argmax}}\Big(\sum_{\omega} \gamma(r,\omega)p(\omega \mid Z)\Big) \tag{1}$$

where $p(\omega \mid Z)$ is the posterior probability of each state of the world given the received signal Z, and $\gamma(r,\omega)$ is the utility (gain or loss) of making response r when the true state of the world is ω. If there is no constraint function, then Z in equation (1) is replaced by S. The performance of the ideal observer (e.g., accuracy and/or reaction time) can sometimes be determined by direct calculation, but often can be determined only by Monte Carlo simulation (i.e., applying equation (1) to random samples of the signal Z).

Equation (1) is fairly general; in fact, all of the examples of ideal observers described here are special cases. However, as a concrete example, consider a task where there are just two categories of object and the observer's task is to be as accurate as possible in identifying which object was presented. In this case, the state of the world can take on only two values ($\omega = 1$ and $\omega = 2$) and observer's responses can take on only two values ($r = 1$ and $r = 2$). Because the goal is to be as accurate as possible, the proper utility function rewards correct responses ($\gamma(r,\omega) = 1$ if $r = \omega$ and does not reward (or punishes) incorrect responses ($\gamma(r,\omega) = 0$ if $r \neq \omega$). Substituting into equation (1) shows that the ideal decision rule is simply to make response $r = 1$ if $p(\omega = 1 \mid Z) < p(\omega = 2 \mid Z)$ and otherwise make response $r = 2$. In other words, the rule is simply to pick the object with the highest posterior probability.

Although the ideal observer framework as described above is sufficient for present purposes, there are a number of useful elaborations of the framework that should be mentioned here. One conceptual elaboration is the *influence graph* (or *Bayesian network*), which describes the qualitative mapping between states or properties of the world ω and properties of the stimulus S (e.g., see Kersten, Mamassian & Yuille, 2004; Jacobs & Kruschke, 2010). Influence graphs specify the task relevant properties of the world (local environment) and stimulus, and their causal relationships, and they imply how those properties should be treated in computing posterior probabilities for the task. A second elaboration of the framework is to incorporate mechanisms (including ideal Bayesian mechanisms) for learning posterior probability distributions, utility functions, or simple decision rules equivalent to equation (1), either on short (Jacobs & Kruschke, 2010) or evolutionary (Geisler & Diehl, 2003) time scales. A third elaboration is to take into account biophysical costs (e. g. , energy) of neural computations (Laughlin & Sejnowski, 2003; Koch et al. , 2004; Manning & Brainard, 2009) and motor responses (Körding & Wolpert, 2006), or more generally fitness (Geisler & Diehl, 2003).

Pattern Detection, Discrimination and Identification

The earliest applications of ideal observer theory in vision were concerned with understanding how detection is limited by photon noise and how the performance of real observers compares that of an ideal observer that is limited only by photon noise (e. g. , Rose, 1948; DeVries, 1943; Barlow, 1957; Cohn & Lashley, 1974). For this ideal observer, the threshold for detecting an increment (or decrement) in intensity increases in proportion to the square root of the background (baseline) intensity. Early studies showed that there are some conditions in which human increment detection performance parallels that of the photon noise limited ideal observer, but, on an absolute scale, humans are substantially less efficient than the ideal observer.

Shortly after the 25th anniversary of Vision Research, photon-noise-limited ideal observers were derived and applied to a wider range of tasks, including various acuity tasks (Geisler, 1984; 1989), contrast sensitivity and contrast discrimination tasks in adults (Banks, et al. , 1987; Geisler 1989; Banks et al. 1991; Sekiguchi et al. 1993; Arnow & Geisler, 1996) and in infants (Banks & Bennett, 1988), color discrimination (Geisler, 1989), and letter identification (Beckman & Legge, 2002). These studies also evaluated the additional effects on ideal observer performance of biological constraints such as the optics of the eye, the spatial and chromatic sampling by the photoreceptors, photoreceptor noise, and ganglion cell spatial summation. This work provides insight into how optics, photoreceptors, photon noise, and retinal spatial summation contribute to human performance. The general finding is that human performance is suboptimal, but often parallels ideal observer performance qualitatively (and sometimes quantitatively) for a surprising number of detection and discrimination tasks. In other words, for these tasks the variation in human performance across conditions is often predicted by the information available in the retinal responses (see Geisler, 2003 for a review). Nonetheless, the suboptimal performance of human observers implies substantial contributions of central factors.

Barlow (1978) reasoned that it may be possible in psychophysical experiments to largely bypass the effects of photon noise and retinal factors, and hence isolate the effects of some of the central factors, by adding high levels of external noise. This proved to be a powerful insight that spawned a number of studies measuring target detection and identification in Gaussian or Poisson pixel noise. Importantly, using statistically independent Gaussian or Poisson pixel noise makes it is relatively easy to derive and determine ideal observer performance. For example in simple detection (where the goal is to maximize accuracy) the ideal observer applies a template matching the shape of the target and then compares the template response to a criterion. Adding external noise raises detection and identification thresholds; however, as expected from bypassing low-level factors, performance generally moves closer to that of the ideal observer (i. e. , efficiency increases).

For an ideal observer limited by external noise, the square of contrast detection (or identification) threshold increases linearly with the square of the root-mean-squared (rms) contrast of the external noise. Human thresholds match this prediction approximately both in the fovea (Burgess et al. , 1981; Legge et al. , 1987; Pelli, 1990) and in the near periphery (Najemnik & Geisler, 2005), once the external noise contrast exceeds a certain level (see Figures 1a and 1b). Measuring contrast thresholds as a function of external noise contrast allows one to estimate an equivalent internal noise, which can be interpreted as the combined effect of those low-level factors that are swamped (dominated) by the external noise as external noise contrast increases (for review see Pelli & Farell, 1999).

Although efficiency is higher with moderate to high levels of external Gaussian or Poisson noise, performance is still generally well below ideal. Several factors probably contribute to this suboptimal performance. One factor is internal uncertainty (Tanner, 1961; Nachmias & Kocher, 1970; Cohn & Lashley, 1974; Pelli, 1985), which may include uncertainty about the spatial location of the target (spatial uncertainty) or uncertainty about certain target feature properties such as orientation or shape (channel uncertainty). These are forms of internal noise that necessarily limit performance. Another factor is contrast nonlinearities (e. g. , contrast gain control), which may produce masking effects above and beyond those due to the similarity of the target and external noise (Foley & Legge, 1981; Foley, 1994; Geisler & Albrecht, 1997). A third factor is inefficient pooling of target feature information. If the features that the real observer uses to detect the target do not correspond to the template that matches the shape of the target, then performance will be suboptimal. The image features that an observer uses in performing a detection or identification task can be estimated using the *classification image* technique, which is based on ideal observer theory and measures the trial-by-trial correlation between the image noise pixels and the observer's behavioral responses (Ahumada, 1996). Measurements of classification images for various kinds of target reveal non-optimal pooling of feature information (Ahumada, 1996; 2002; Eckstein et al., 2002; Gold et al., 2000; Murray et al., 2005). Some of this non-optimal pooling is due to uncertainty and contrast nonlinearities. However, these factors can only blur (or sharpen) the classification image; whereas measured classification images frequently reveal missing target features and sometimes added illusory features (Figure 1c; Murray et al., 2005).

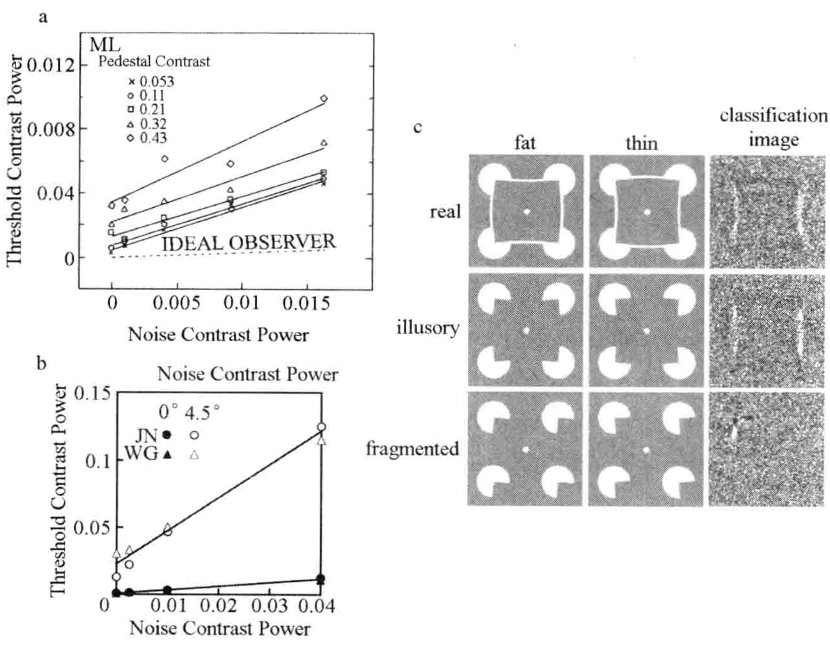

Figure 1 Detection and discrimination in Gaussian noise. a. Detection threshold as a function of white noise power for one observer for five pedestal contrasts. (adapted from Legge et al., 1987). b. Detection threshold for a 6 cpd target as a function of 1/f noise contrast for two observers at two retinal eccentricities (adapted from Najemnik & Geisler, 2005). c. Classification images for shape discrimination in white noise (adapted from Gold et al., 2000).