SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

March 2014, Vol. 57 032114:1–032114:15
doi: 10.1007/s11432-013-4911-9

A computational cognition model of perception, memory, and judgment

FU XiaoLan[1]*, CAI LianHong[2], LIU Ye[1], JIA Jia[2], CHEN WenFeng[1],
YI Zhang[3], ZHAO GuoZhen[4], LIU YongJin[2] & WU ChangXu[4]

[1]*State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing* 100101*, China;*
[2]*TNLIST, Department of Computer Science and Technology, Tsinghua University, Beijing* 100084*, China;*
[3]*College of Computer Science, Sichuan University, Chengdu* 610064*, China;*
[4]*Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing* 100101*, China*

Received January 9, 2013; accepted June 5, 2013; published online January 25, 2014

Abstract The mechanism of human cognition and its computability provide an important theoretical foundation to intelligent computation of visual media. This paper focuses on the intelligent processing of massive data of visual media and its corresponding processes of perception, memory, and judgment in cognition. In particular, both the human cognitive mechanism and cognitive computability of visual media are investigated in this paper at the following three levels: neurophysiology, cognitive psychology, and computational modeling. A computational cognition model of Perception, Memory, and Judgment (PMJ model for short) is proposed, which consists of three stages and three pathways by integrating the cognitive mechanism and computability aspects in a unified framework. Finally, this paper illustrates the applications of the proposed PMJ model in five visual media research areas. As demonstrated by these applications, the PMJ model sheds some light on the intelligent processing of visual media, and it would be innovative for researchers to apply human cognitive mechanism to computer science.

Keywords perception, memory, judgment, computational cognition model

Citation Fu X L, Cai L H, Liu Y, et al. A computational cognition model of perception, memory, and judgment. Sci China Inf Sci, 2014, 57: 032114(15), doi: 10.1007/s11432-013-4911-9

1 Introduction

The mysteries of the human mind have attracted considerable attentions in natural sciences. For nearly half a century, cognitive science has emerged as a new discipline that focuses on the various scientific issues of the human mind. Cognitive science is the interdisciplinary scientific study of human perception and thinking process, which includes all of cognitive processes from sensory input to complex problem solving, individual human being to intelligent activity of human society, as well as the nature of human

*Corresponding author (email: fuxl@psych.ac.cn)

intelligence and machine intelligence. Cognitive science research not only promotes the understanding of the nature of the human mind, but also promotes the development of modern science and technology. Recently, computer science has become increasingly prominent in cognitive science, and the knowledge of theoretic computer science provides a solid basis for considering what the functional architecture of a computational brain is [1].

The visual media, including digital images, video and three-dimensional models, contains superfluous visual information. Since the intelligent processing of visual media utilizes a combination of cognitive mechanism and computation, visual media is a good example to integrate cognitive mechanism into an intelligent computation [2–4]. Marr's visual computing theory is the most representative cognitive computing model, which plays an important role in guiding intelligent computer image processing [5]. The algorithm Marr proposed is not only in line with the results of neurophysiology experiments conducted in primate animals, but also explains the characteristics of the human visual system [5]. Marr's model was the most successful model that combined human cognitive mechanisms and computer algorithms. However, with the rapid development of science and technology, the pending visual media information from the Internet is massive, unordered, uncertain, and interactive in social groups. Thus, it is imperative to propose new theories and methods to process massive amounts of visual media.

Over the past decade, a large number of neurophysiological and cognitive neuroscience researches have provided in-depth and detailed experimental data and theoretical models to reveal the brain's information processing mechanisms. Because of the complexity of information processing in the human brain, cognitive scientists recognize that computational models can enhance our understanding of the cognitive system functions and provide a theoretical foundation and technical support [6]. For example, *Science*, *Nature*, and *Neuron* recently published a series of studies [7–12] that showed the role of the bottom-up and top-down visual attention selection in the process of human visual perception. Different neural pathways as well as corresponding computational models that successfully simulated their neural mechanisms were discussed. Further neurophysiology research and computational modeling research indicated that the perceptual significance of stimuli depends on the background information in the environment [13]; background information is also shown to be very important in object recognition process [14]. Poggio et al. proposed a systematical computational model based on the perception principles of biological visual system [15–17]. However, the researches on the neural mechanisms of visual information processing still lack a quantifiable cognitive model and a corresponding mathematical theory to explain the internal mechanism clearly. All of these problems obstruct the practical application in engineering. How could cognitive mechanisms and computational models simulating human cognitive functions be applied to the intelligent sensing of machine perception in the natural environment? How could they be applied to solve practical problems of intelligent computing? All of these questions would be answered by the fundamental scientific exploration of intelligent information process computing.

This paper addresses the intelligent processing of massive amounts of visual media and makes the processing of perception, memory, and judgment in cognition correspond to the steps of analysis, modeling, and decision in computing, respectively. We review both human cognitive mechanism and cognitive computability of visual media at three levels: neurophysiology, cognitive psychology, and computational modeling. Then, the Computational Cognition Model of Perception, Memory, and Judgment (PMJ model) is proposed, consisting of three stages and three pathways integrating cognitive mechanism and computing in its framework based on the basic mechanisms of human cognition. In the framework of PMJ model, we study the important cognitive mechanisms of human information processing on mass visual media, build a neural network model based on PMJ model, achieve a quantitative description of the visual cognition load, and further explore the mathematical formulation of the model. Finally, the model is applied in the field of affective forecasting based on Internet images and image retargeting. PMJ model would provide realizable cognitive basis for improving the efficiency of mass visual media processing from the Internet and realizing visual media interaction, integration and presentation in accordance with human perception and cognition. Furthermore, the model would effectively promote cognitive computing from qualitative research to quantitative research, and enhance the research level of intelligent processing of Internet visual media.

2 PMJ model

Psychological research has reached a consensus that most cognitive processes are composed of a series of successive processing stages [5]. These processing stages mainly include the following steps: When a stimulus is presented, the cognitive system processes it by sensory and perceptual processing first, after perceptual processing the information is then transferred to short-term memory; by rehearsal, some of the information in short-term memory is transferred to long-term memory; finally, with the interaction between cognitive system and the outside world, knowledge and experience in long-term memory, and the perceptual processing information from the outside world influence the response of cognitive system to the outside world in conjunction. However, could the processes described above be computed? To answer this question, we need to discuss the theoretical basis of cognition computation first.

2.1 The computability of cognition

Computationalists from cognitive psychology proposed that cognition is a kind of computational form [18], and the primary function of the brain is to process information [14]. The received information can be represented in the brain. If such a representation of the brain is absent, it is impossible for the brain to communicate with the world [1]. Representations of the brain are functioning homomorphisms; that is, there are structure-preserving mappings (homomorphisms) from states of the outside world (the represented system) to symbols in the brain (the representing system) [1]. Symbols are the physical manifestation of computation and representation in cognitive processes. They carry information and embody the results of those computations [1]. Therefore, it is essential to understand that the symbols of the brain are physical entities and cognition is the computation of symbols and the information processing of the brain [1]. Good symbols in a computational system must be distinguishable, constructible, compact, and efficacious [1].

The computability of cognition could bind the mechanism of human cognition and computational models realized by computers. It is the theoretical foundation for the research in which human behaviors could be explained by computational processes, and the basic principles of cognitive modeling for instructing engineered computing (categorization, identification, and encoding) based on cognitive hypotheses. The computability of cognition not only makes the quantification of cognitive properties possible, but also serves as the basis for quantified data to be computed in the computing processes of modeling and judgment.

2.2 The definition of PMJ model

Based on the fundamental principle of human cognition and the computability of cognition, we propose the PMJ model in which perception, memory, and judgment of cognitive process correspond to analysis, modeling, and decision of computing, respectively. The theoretical framework consists of three stages, multiple pathways, and a series of cognitive strategies which are the combination of cognition and computation. PMJ model is shown in Figure 1.

As shown in Figure 1, the framework of the cognitive model is highlighted with the dotted lines that include the main stages of cognition, such as perception, memory, and judgment [5]. In each stage, the cognitive system would complete certain information processing tasks and provide information as input for other stages, or receive the output of the other stages. All stages would interact with each other to complete the cognitive processing tasks [5]. In the model, the arrowed lines with numbers indicate the pathways of a certain cognitive function. The processings of perception, memory, and judgment of cognition in PMJ model correspond to the stages of analysis, modeling, and decision in computing, respectively.

In the stage of perception, through the processes of pre-attention selection [19,20] and selective attention [21–26], the cognitive load of cognition system is reduced [27,28], and then salient visual features are extracted (as indicated by 1 in Figure 1) [13]. In the stage of memory, dynamic memory system is achieved (as indicated by 2 in Figure 1) through the mechanisms of encoding and storing processes [29–32] and the mechanisms of updating and consolidation [33–38]. In the stage of judgment, judgments and decisions are made efficiently (as indicated by 3 in Figure 1) through categorization learning [39–43] and encoding based on action coding or abstract coding [44–53].

Cognition consists of a series of complex processes, and there are multiple processing pathways between the various stages of cognition. The cognition system chooses pathways dynamically, depending on the difficulty [54] and the goal [55,56] of the information processing tasks. These processing pathways complete the transfer of information between the three stages of processing, ultimately achieving judgment efficiently, and output the decision results. There are three kinds of pathways, summarized as the fast processing pathway, the fine processing pathway, and the feedback processing pathway.

(1) The fast processing pathway is the process from perception to judgment (the arrowed line of 8 as shown in Figure 1), which achieves judgment based on the output of perception. The process of this pathway does not require too much knowledge nor experience involved [57,58] in which global features and contour information of the input stimulus as well as the low spatial frequency information are processed [59–61]. Based on coarse and primary processing of the input information, the cognition system makes a fast categorization judgment [59–61].

(2) The fine processing pathways are the processes from perception to memory, and then from memory to perception and judgment (the arrowed lines of 4+5 and 7 as shown in Figure 1), which achieves both perception and judgment based on the knowledge stored in the memory system. Knowledge and experience play an important role in the pathways [62–65], in which local features and detailed information of the input stimulus as well as the high spatial frequency information are processed [59–61]. Based on fine processing of the input information and matching them with the knowledge in long-term memory, the cognition system makes a judgment.

(3) The feedback pathways are the processes from judgment to memory, and from judgment to perception (the arrowed lines of 6 and 9 as shown in Figure 1), which updates perception and memory based on the results of judgment [66]. Based on the results of judgment the cognition system updates knowledge stored in long-term memory. The output of judgment would serve as a clue for future perception processes, and further improve their efficiency and accuracy [66].

This paper focuses on the intelligent process of massive amounts of visual media and investigates the cognitive mechanisms, neural network, and application in computing of PMJ model. The mappings from perception, memory, and judgment of cognitive processing to analysis, modeling, and decision of computing are discussed. For example, the processing of judgment based on the output of perception in the fast processing pathway (the arrowed line of 8 in Figure 1) is mapped to feature-based modeling and decision, while the processing of perception and judgment based on the knowledge in memory system in the fine processing pathways (the arrowed lines of 4+5 and 7 in Figure 1) is mapped to knowledge-based learning and decision. Perception and memory updating based on the results of judgment in the feedback processing pathways (the arrowed lines of 6 and 9 in Figure 1) is mapped to decision-based modeling and optimization. In the remainder of this paper, the applications of PMJ model in five different visual media research areas are presented.

3 Understanding cognitive psychology research based on PMJ model—visual search

Visual search is a visual behavior to detect or locate a target item presented within a specified field. Like other cognitive processes, visual search is a process consisting of several PMJ sub-processes.

In a rapid search, the observer first attends to the global layout, assessing the rough contents of the search field (as shown in Figure 2). These processes are rapid and parallel processes carried out simultaneously across the search field and generally can be completed within a single glance [67]. Treisman's feature integration theory (FIT) [68–70] proposed that feature maps in mental representations encode the simple attributes in pre-attentive stage (such as color, motion, orientation, and coarse aspects of shape, where a single map is selective for a particular value such as "red" along a given stimulus dimension "color"). If a target is highly conspicuous, like the case where the target is a single feature object, the features of the target are quickly extracted leading to the immediate detection of the target. As such, it is

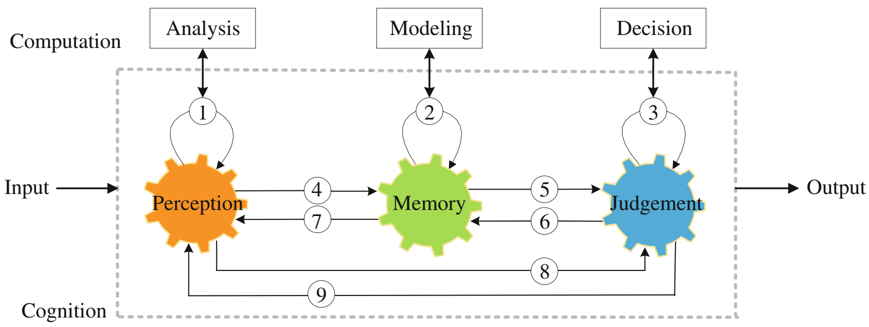

Figure 1 The schematic representation of PMJ model.

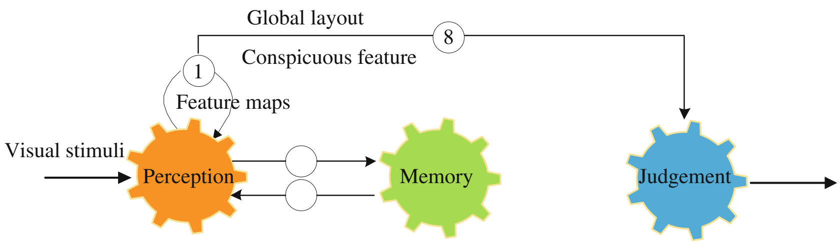

Figure 2 Rapid search.

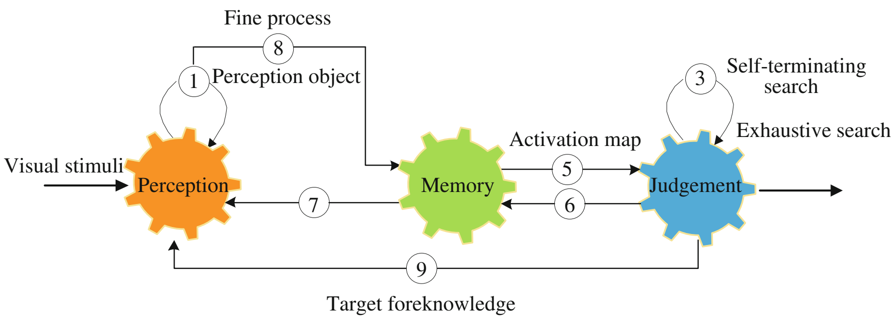

Figure 3 Optimized search.

much simpler to complete the search in the relevant feature map, if the detected target is a single feature object. The underlying mechanism of rapid parallel search is the fast processing pathway of cognition.

If the target is not conspicuous (e.g., the target is a multidimensional object), the observer will have to focally scan the image [68]. To detect such a specific multidimensional target, it is necessary to scan the displayed items of the target one by one, where focal attention is necessary to bind stimulus properties together [68]. The goal of scanning is to bring the target within the searcher's visual lobe that is the region around the point of regard within which information is gathered during each fixation [71]. A successful target-present judgment requires accurate detection, by which the observer matches a pattern extracted from the search field to a stored mental representation of the target stimulus in the memory and reaches a yes–no decision [71]. These serial processes reflect the pathways as indicated by the arrowed lines of 4+5 and 7 in Figure 1.

In guided search, the information acquired in the parallel process can be used to guide the serial process. Guided search is restricted to items with at least one dimension similar to the target [72]. Via the feedback pathway, observers perform such a search by restricting attention to stimuli that possess at least one of the target properties. The guided attention occurs within a "master map" of locations. The master map of locations contains all of the locations in which features have been detected, with each location in the master map having access to the multiple feature maps [68].

In optimized search (as shown in Figure 3), the fine process during the serial stage may form the memory representation of the stimuli such as properties and locations. For example, the recognition performance of the target and distracter items in the search task is above chance [73]. In addition, the target foreknowledge may enable top-down or knowledge-driven search processes. When the target is well specified, knowledge-driven processes can amplify or attenuate the activation within feature maps, allowing the searcher to bias attentional scanning towards those objects within a scene that contain known target features [71,74]. The preview search, where selection of new items is impaired when these items share features with the old items, provides another piece of evidence for top-down search processes. Such top-down processes are modulated by the target properties (e.g., the angry expression of the target face) [75,76]. The processes in optimized search reflect the role of memory representation, depicting the top-down process as indicated by the arrowed line 7 in Figure 1.

As mentioned above, visual search is accompanied with the processes of decision and judgment. The visual system needs to make a decision on target-present or target-absent judgment to terminate or continue the search. A search can be classified as either self-terminating or exhaustive [77]. In self-terminating searches, a search ends once a target has been discovered. Exhaustive searches, in contrast, continue through the full search field even if a target is already discovered. Search on target-present trials can be either self-terminating or exhaustive, regardless of whether the processing is parallel or serial. But in both the parallel and serial searches, exhaustive processing is required to determine that a target is absent—that is, all items must be inspected.

4 Neural computing based on PMJ model—Neocortex

It is well known that behavior could be explained by the activity of neurons. Intelligent activity could be taken as a kind of behavior; thus, intelligence could be also explained by the activity of neurons. A brain contains around 100 billion neurons, among these neurons, the neurons located in the neocortex play important roles in intelligence. Research on neural science shows that almost everything we think of as intelligence such as perception, language, imagination, mathematics, art, music, planning, etc. all occurs in the neocortex [78]. The neocortex is the set of intelligence. Generally speaking, the typical human neocortex is with size around 1,000 cm^2 and 2 mm thick; it contains about thirty billion neurons. About 100,000 neurons are contained in a tiny square millimeter consisting of 100 trillion of synapses [79]. Memories, knowledge, skills, and one's life experience are all stored in the neurons of the neocortex. The neocortex is divided into many neocortex functional regions. Physically, these regions are arranged in an irregular patchwork quilt with nearly identical architecture. Functionally, the regions are connected in a branching hierarchy [78,80]. Mountcastle's proposal suggests that there may have a single powerful algorithm implemented by every region of cortex. The way the cortex processes signals from the ear is the same as the way it processes signals from the eyes [81]. Each region could be looked upon as an information processing unit.

The information processing algorithm implemented by each region should complete three cognitive tasks: perception, memory, and judgment. To understand how a neocortex region completes these three tasks, it is necessary to know that the information flows in the neocortex are actually the temporal and spatial patterns. Each region uses the input temporal and spatial patterns continuously to conduct perception, memory and judgment. Figure 4 shows how the three cognitive tasks are completed in a region.

Firstly, a region will make a prediction based on its existing knowledge. However, is such a prediction correct? It is only the event itself that can answer this question. Then, after the event, information of what really happened and the prediction will intersect and generate a feedback, which answers the question whether the prediction is correct. Next, neocortex region updates its knowledge based on this response. Updated knowledge will be used for the next prediction, and repeat the whole process. Take the weather forecast as an example. Brian will first generate a prediction based on the weather knowledge of the past. Correctness of this prediction will be judged in the next day when it really comes. The real weather condition in the next day together with the prediction will generate a feedback, which is exactly the state of the region. The brain can use this feedback for updating its ability of prediction, or

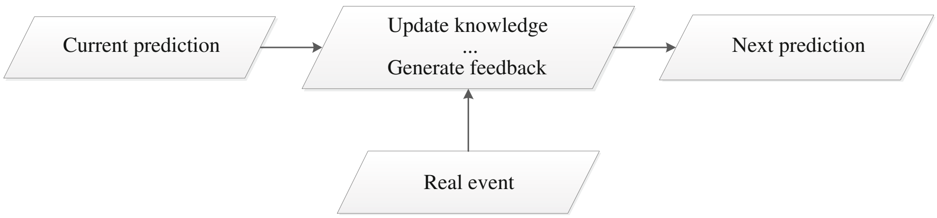

Figure 4 Procedures of completing the perception, memory and judgment in neocortex.

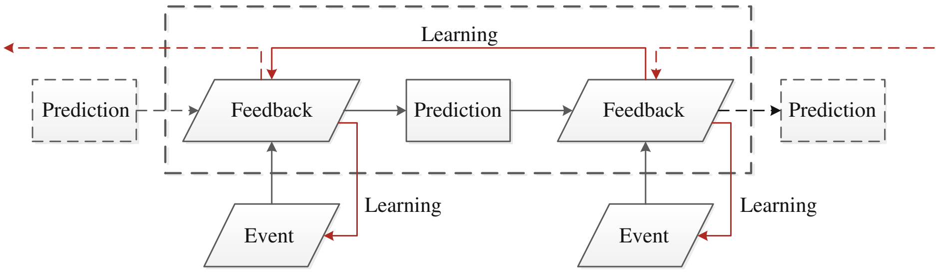

Figure 5 The PMJ model in neocortex.

knowledge, and then make further prediction.

Figure 5 shows a PMJ model for cortex region algorithm. Perception takes place when input information gets sparsely represented by cortex neurons as spatio-temporal pattern. Cortex neuron network memorizes input patterns constantly by learning. Knowledge is stored on the connection weight of the network. Judgment part contains three functions: prediction, learning, and feedback. New input pattern and previous prediction intersect and generate feedback information. A region learns from the feedback information to update its prediction ability, i.e. update connection weights. The process will be repeated until the region has satisfactory prediction ability.

In this PMJ model, perception part generates the input to the region. Such a process is also encoding data. Brain uses sparse coding for representing data. A mathematical model for sparse representation can be formulated by

$$\min \|z\|_0, \text{ s.t. } y = f(z),$$

where y is the input vector and z is the sparse code from perception part, $\|z\|_0$ denotes the number of non-zero entries in z, and f is the perception mapping between input and output. In [82,83], many research results can be found in the case where f is linear mapping.

Three variables can be defined in each region: neuron state variable, prediction variable and memory variable. The three variables interact with each other over time. In this way, the evolution of neurons in a region naturally forms a dynamic system. In a region at time t, denote neuron state by $x(t)$, prediction by $p(t)$, memory by $m(t)$. Dynamic modeling of the region can be described as

$$\begin{cases} \dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = X\left(z(t), p(t-\tau(t))\right), \\ p(t) = P\left(x(t), \displaystyle\int_0^t u(t,s)m(s)\mathrm{d}s\right), \\ \dfrac{\mathrm{d}m(t)}{\mathrm{d}t} = M\left(m(t), p(t-\tau(t)), \displaystyle\int_0^t v(t,s)x(s)\mathrm{d}s\right), \end{cases}$$

where $\tau(t) \geqslant 0$ is time delay, $u(t,s)$, $v(t,s)$ are kernel functions of some kind, and X, P, M are linear or non-linear functionals defined on corresponding function space. In the PMJ model described above, we can get a computable model by choosing some suitable non-linear mappings. It is certainly not easy to choose such mapping. In real applications, they can be chosen according to actual data.

The virtual significance of response in regions is that neuron state of the region is determined by both previous prediction and current input of the region. New prediction closely depends on current feedback.

Learning is the change on memory. It happens when updating acquired knowledge for more accurate prediction in a region. Past predictions and states will influence learning. Neuron science believes that memory is stored as attractors in the brain [84,85]. Attractors can be categorized by discrete attractors, continuous attractors, strange attractors, etc. Discrete attractors are suitable for memory of isolated events. There is a wealth of research with sophisticated mathematic tools in this field. Continuous attractors are attractors distributed in a continuous manner. Some recent research results on neuron science show that continuous attractor depicts many important nature of information process in brain to a great degree. Continuous attractor model has been successfully used for storing and representing the process of continuous variables in the brain, e.g. orientation of moving objects, space position information etc. [86–89]. Not much has been done on the relationship between memory and strange attractors so far.

5 Formalized model based on PMJ model—speed control

Speed control consists of speed perception, memory, speed selection, and control. We have studied human visual information processing speed and characteristics involved in the process of speed control. Most of the time, in realistic driving situations, drivers are aware of their traveling speed relying on perceptual cues, combined with occasional speedometer inspection. These perceptual cues may be visual, auditory, or kinesthetic cues. While each category plays an important role in assessing traveling speed, visual cues (e.g., optical flow), serve as the predominant reference that drivers use to estimate their traveling speed. Optical flow is one of the key research areas in computer vision and its related fields. It plays an important role in the study of the target object segmentation, identification, tracking, and robot navigation. In a real driving situation, optical flows are transformed into physiological electrical signals via the visual cells of the retina, transmitted to various regions of the brain through the optic nerve and processed in corresponding regions. There are two major neural networks involved in the visual information transmission. One mainly processes the changes in spatial visual information with high-density spatial distribution but slow response time; the other primarily deals with the changes in temporal visual information with low-density spatial distribution but fast response time. Image processing algorithms in computer vision such as filtering, enhancement, and restoration can simulate the process of how humans perceive optical flow.

Speed selection and memory involve multiple information processing pathways of the PMJ model. First, when new drivers begin to drive for the first time, they must deliberate over each speed choice presented to them and then make a quick decision using only basic attribute information (see the pathway as indicated by 8 in Figure 1). For example, new drivers tend to adjust their traveling speeds according to the posted speed limits only, ignoring the changes of traffic flow, road and weather conditions. With repeated experience, complex cognitive activities involving multiple decision choices, multiple-attribute weighting, attributes and rules competition begin to dominate the process of speed choice. Most drivers begin to internalize a set of rules (e.g., decrease speed when snowing) that become applicable when deliberating over a target speed (see the pathway as indicated by 4 and 5 in Figure 1). Finally, speed control is a dynamic memory-based reinforcement learning process. Each deliberation process of speed choice and its associated consequence will reinforce or update memory (see the pathway as indicated by 6 in Figure 1). For instance, after a driver speeds and receives a speeding ticket, the driver may change his/her speed choices under similar driving conditions.

We integrated the Queuing Network - Modal Human Processor (QN-MHP) [90] with the Rule-based Decision Field Theory (RDFT) [91] to quantify the processes of speed perception, memory, speed selection and control. As illustrated in Figure 6, QN-MHP consists of perceptual, cognitive (involving memory and decision making), and motor subnetworks. In QN-MHP, brain regions with similar functions are represented as servers (e.g., Servers 1–4, Server A). Specific entities represent pieces of information that pass through and are processed by the servers. An entity travels on routes which represent neural pathways connecting the different brain regions. According to the QN-MHP, Server F in the cognitive subnetwork performed complex cognitive functions such as multiple-choice decision, visuomotor choices, anticipation of stimuli in simple reaction tasks, etc. Thus, we incorporated the deliberation process in

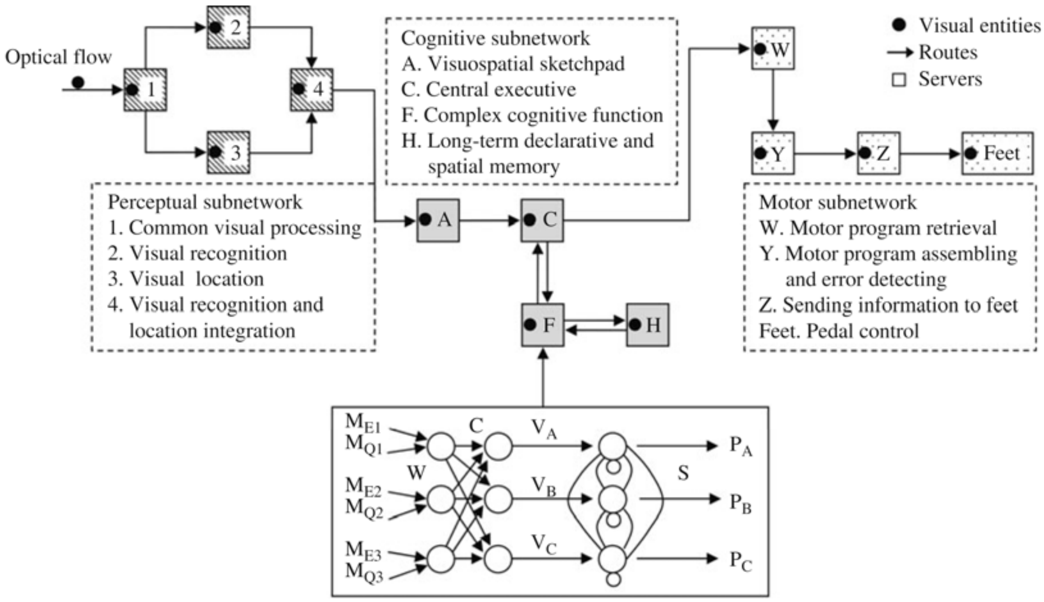

Figure 6 Human speed perception and decision making model (effective servers and routes in QN-MHP that were used in the model were highlighted).

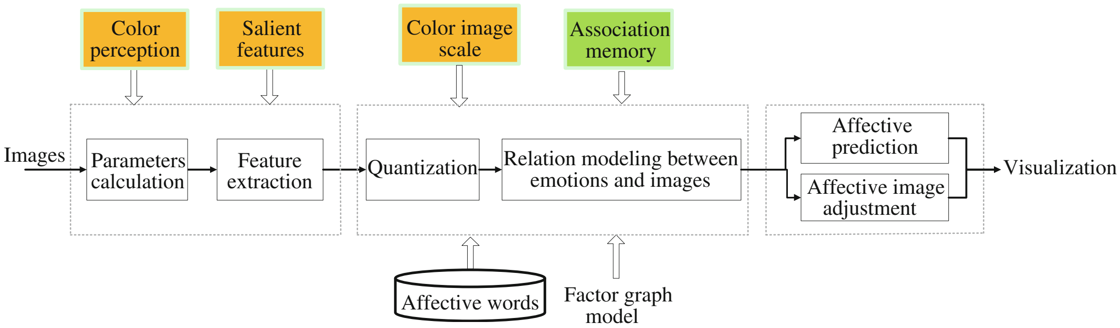

Figure 7 The framework of image affective prediction based on PMJ model.

RDFT in this Server F. We first formulated a metric to demonstrate the cognitive process of speed choice and assumed that at each point in time during deliberation, the attention weights specifically selected a single attribute which focused on the values of a single attribute from the metric (equal probability). Based on the formulation methods of the momentary valence and preference, we calculated a driver's accumulated preference for each speed choice at each point. If any preference exceeded the pre-defined decision threshold, such speed choice was selected without further deliberative process.

Speed perception and decision making model takes the task difficulty and individual driver differences in the information processing speed and capacity into account. It can provide quantitative predictions of a driver's perceived speed and desired target speed (e.g., if a driver follows the speed limit or drives over the speed limit at 10 mph or 20 mph). Additionally, this work can be further extended to model human behaviors of speed perception, selection and control when they are moving (e.g., walking, running). For a detailed description of the mathematical deduction and parameter settings, see Zhao et al. [92].

6 PMJ application—image affective prediction

As an efficient communication medium, images convey a wealth of information, especially emotions. Based on the PMJ model, we first propose the emotion related color features and a relationship model between the color features and image emotions. Then this model is applied to understand the emotional impact of images [93,94]. Finally, we implement an affective image adjustment system that automatically adjusts image color to meet a desired emotion [95,96], completing the precision processing of the PMJ model. The framework is illustrated in Figure 7.

According to the theories of color perception, there is an inherent connection between colors and emotions. Artists and amateurs have the same emotional experience to art works. In art design, a color theme which is a template of colors and also called a color combination is commonly used to describe the color composition of a painting. Color themes play a more important role than color itself in human visual perception. For instance, no matter what a color theme actually is, one with a strong contrast tends to express the intense emotion; on the other hand, that with a weak contrast tends to express the mild emotion. Based on the long-term psychophysical investigations, Kobayashi summarized the relationship between color themes and affective words [97]. So based on the above foundation, we adopt color themes as salient features. Color themes are extracted by considering two factors: the area and the contrast. In other words, colors which have large areas and strong contrast with neighboring colors tend to be selected. We propose an optimization algorithm to extract color themes.

We propose a partially labeled factor graph model (PFG) for modeling the relationship between emotions and images. This model better utilizes the Internet images and their connections, and can overcome the noise.

The relationship model can be used for the common affective cognition on images. For example, we use our proposed model to infer the mood from van Gogh's paintings "Starry Night over the Rhone" and "Wheatfield with Crows". The top three prediction categories are casual (probability: 19.3%), modern (15.03%), chic (10.94%) and casual (19.09%), dapper (13.17%), chic (12.33%), respectively. These results are almost consistent with the original users' comments to the photos. The relationship model can also be used for inferring emotions around special events. For example, we downloaded images from Flickr around Thanksgiving 2011, and use our model to predict the affective category of each image. Figure 8 shows affective distributions before and during Thanksgiving.

Furthermore, benefited from the model, we implement a system called affective image adjustment. The system supports automated changes on the emotional impact of an image driven by a single word. Figure 8 shows some examples.

7 PMJ application—image retargeting quality assessment

Image retargeting (also known as image resizing) technique can adjust an input image into one of arbitrary image size without serious content distortion to fit different displaying devices. Image retargeting has a wide range of applications in image transformation over the Internet, adaptive mobile displaying, webpage design, etc. A number of image retargeting methods have been proposed [98–100] and now the evaluation of these methods becomes important. The most accurate method of evalution is to invite human participants with normal color vision to assess the quality using the mean opinion scores (MOSs) metric. However, subject assessment using MOS is time-consuming and expensive. Thus an objective image retargeting assessment (OIRA) that uses computer program to simulate the human vision system (HVS) is much desired.

Based on the proposed PMJ model, we develop an HVS-simulated objective assessment algorithm in [101]. As illustrated in Figure 9, in this algorithm, at the perception stage, we extract the SIFT features in a scale-space on both original and retargeting images. At the top level of the scale space, the SIFT features are used to establish a coarse match between original and retargeting images. This coarse match provides a structure matching which can be constructed very fast due to the few SIFT features at the top level, and this behavior well simulates the fast processing pathway (the arrowed line of 8 in Figure 1) in the proposed PMJ model (cf. Figure 1). SIFT structure matching is important but not sufficient alone for a high quality objective assessment. Then at the memory stage, we adapt the SSIM metric to establish a memory unit that takes illumination, contrastness and structural information for each pixel into account. The fine granularity correspondence using SSIM metric can provide accurate correspondence for each image pixel and well simulates the fine processing pathways (the arrowed lines 4+5 in Figure 1) in the proposed PMJ model. Finally at the judgment stage, by combining the bottom-up

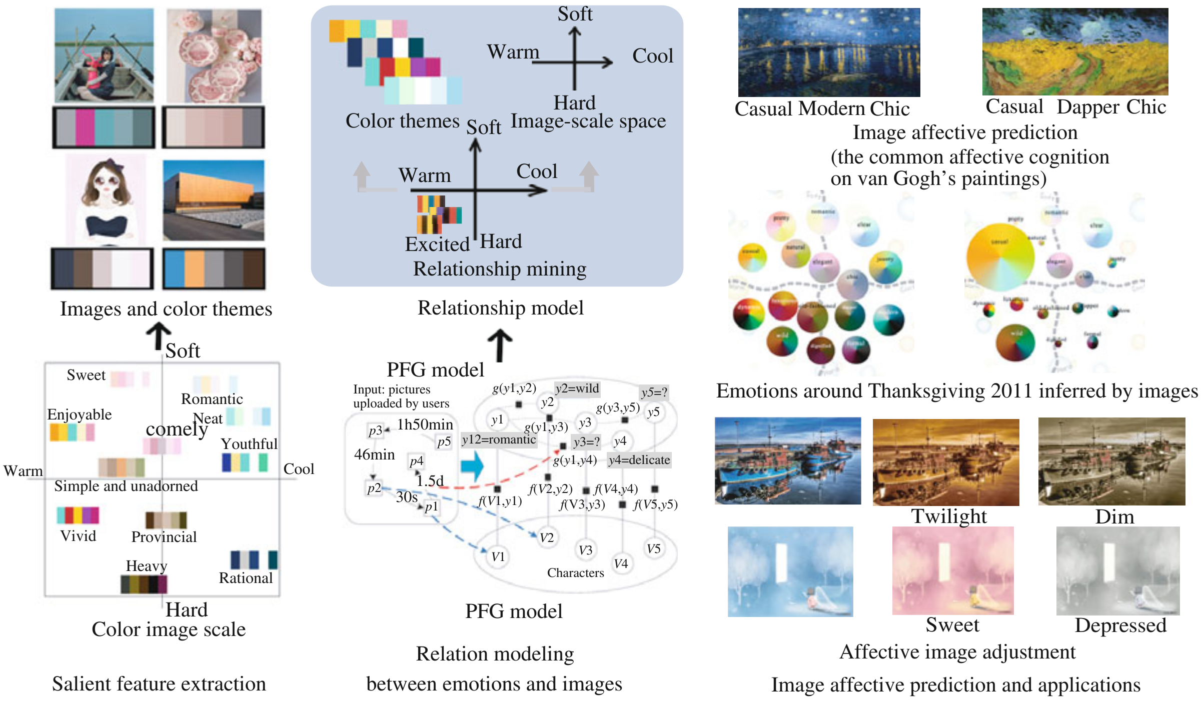

Figure 8 Image affective prediction and affective image adjustment.

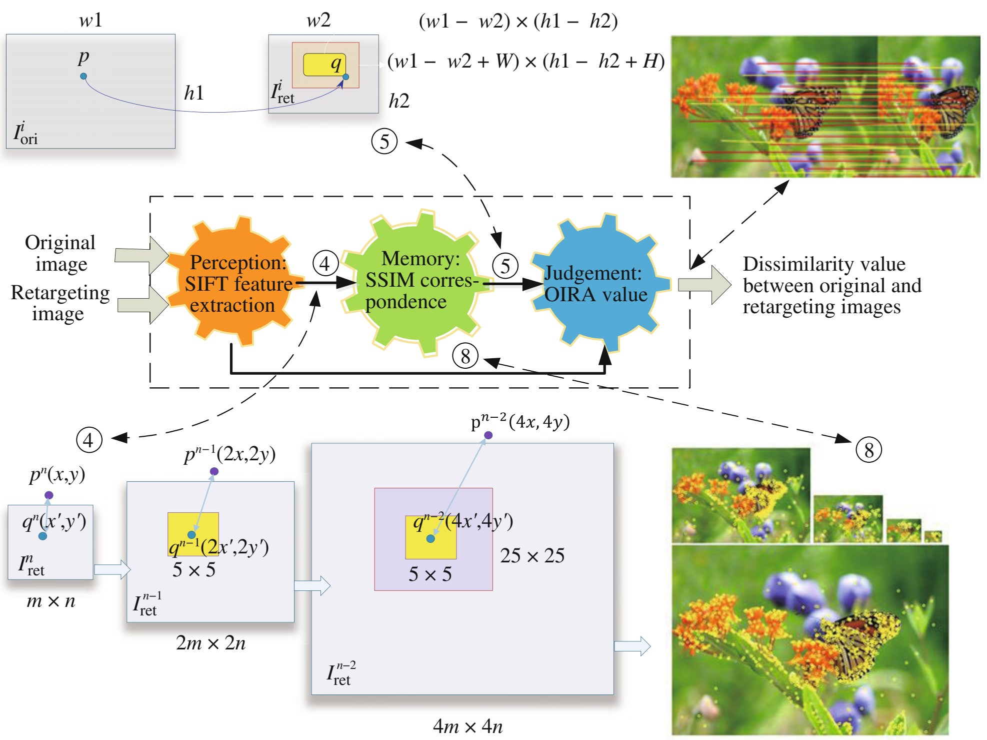

Figure 9 PMJ-model-oriented image retargeting quality assessment.

SIFT correspondence for structure matching and the top-down SSIM correspondence for each pixel, we obtain a high quality image retargeting assessment value using the algorithm proposed in [101].

The above-mentioned objective image assessment method has two distinct features in applying the proposed PMJ computational cognition model. First, we take the hypothesis in [102]: *the human visual system is sensitive to global topological properties and extraction of global topological properties is a basic* *factor in perceptual organization.* In a spatio-temporal continuous scene, the topological properties include spatial relationships in a geometric structure and the structural stability under temporal changes, in a manner similar to Klein's hierarchy of geometries. For static images, we limit the scope with spatial geometric structures which is reflected in the bottom-up SIFT correspondence for structure matching. Secondly, we take the hypothesis on human vision system that *its intermediate or high level process seems to selectively focus on salient regions.* Accordingly, every pixel in an image needs not to have the same importance for assessment: We reflect this hypothesis in the top-down SSIM correspondence for each pixel. Experiments on the measurement of the performance of objective assessments with the proposed OIRA metrics were developed in [101] and the results show good consistency between the proposed objective metric OIRA and subjective assessments by human observers.

8 Conclusion

With the recent rapid development of computer technology, it is a major breakthrough for cognitive science to apply cognitive psychology research in computer science to enhance the level of intelligence processing of modern day massive visual media information. Based on a review of cognitive science research over the past two decades, we proposed a computable model for the intelligent processing of visual media.

A PMJ model is proposed in this paper, with the applications of the PMJ model in five different visual media research areas. Future work on the PMJ model includes referring to and learning from other cognitive models. For example, adaptive control of thought–rational model (ACT-R Model) proposed by Anderson et al. [103] simulated nearly all of cognitive tasks of human brains, even predicted the activation patterns of the brain.

In the proposed PMJ model, cognitive processes are divided into perception, memory, and judgment. Three kinds of pathways among these stages were introduced, namely the fast process pathway, the fine process pathway, and the feedback process pathway. Perception, memory, and judgment of cognitive process in PMJ model correspond to analysis, modeling, and decision of computing, respectively. The proposed model provides a new direction for the intelligent process of visual media information, and it would be innovative for researchers to apply human cognitive mechanism to computer science. The future work on PMJ model will focus on the hypotheses underlying the stages and pathways in the model. Furthermore, it is crucial to transform the psychological hypotheses to principles for computation. All of these principles for computation will shed light on analysis, modeling, and decision of visual media computing.

Acknowledgements

This work was supported by National Basic Research Program of China (973 Program) (Grant No. 2011CB3022-01). We thank ZHANG JingYu, FU QiuFang, XUAN YuMing, WANG XiaoHui, YAN WenJing, CHEN Yu-Hsin, and ZHANG Lei for useful discussions and helpful comments on earlier drafts of this manuscript.

References

1 Gallistel C R, King A. Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience. New York: Blackwell/Wiley, 2009. iiv–xvi

2 Hu S M, Chen T, Xu K, et al. Internet visual media processing: a survey with graphics and vision applications. Vis Comput, 2013, 29: 393–405

3 Hulusic V, Debattista K, Aggarwal V, et al. Maintaining frame rate perception in interactive environments by exploiting audio-visual cross-modal interaction. Vis Comput, 2011, 27: 57–66

4 Vazquez P-P, Marco J. Using normalized compression distance for image similarity measurement: an experimental study. Vis Comput, 2012, 28: 1063–1084

5 Eysenck M W, Keane M T. Cognitive Psychology: a Student's Handbook. 6th ed. New York: Psychology Press, 2010. 1–50

6 National Institute on Drug Abuse. Computational neuroscience at the NIH. Nat Neurosci, 2000, 3: 1161–1164
7 Buschman T J, Miller E K. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science, 2007, 315: 1860–1862
8 Navalpakkam V, Itti L. Search goal tunes visual features optimally. Neuron, 2007, 53: 605–617
9 Katsuki F, Constantinidis C. Early involvement of prefrontal cortex in visual bottom-up attention. Nat Neurosci, 2012, 15: 1160–1166
10 Corbetta M, Shulman G L. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci, 2002, 3: 201–215
11 Zanto T P, Rubens M T, Thangavel A, et al. Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. Nat Neurosci, 2011, 14: 656–661
12 Tomita H M, Ohbayashi K, Nakahara I, et al. Top-down signal from prefrontal cortex in executive control of memory retrieval. Nature, 1999, 401: 699–703
13 Itti L, Koch C. Computational modelling of visual attention. Nat Rev Neurosci, 2001, 2: 194–203
14 Cox D, Meyers E, Sinha P. Contextually evoked object-specific responses in human visual cortex. Science, 2004, 303: 115–117
15 Kouh M, Poggio T. A canonical neural circuit for cortical nonlinear operations. Neural Comput, 2008, 20: 1427–1451
16 Poggio T, Bizzi E. Generalization in vision and motor control. Nature, 2004, 431: 768–774
17 Hung C P, Kreiman G, Poggio T, et al. Fast readout of object identity from macaque inferior temporal cortex. Science, 2005, 310: 863–866
18 Pylyshyn Z W. Computation and Cognition: toward a Foundation for Cognitive Science. Cambridge: The MIT Press, 1984. 1–16
19 Hasselmo M E, Sarter M. Modes and models of forebrain cholinergic neuromodulation of cognition. Neuropsychopharmacology, 2010, 36: 52–73
20 Tamietto M, de Gelder B. Neural bases of the non-conscious perception of emotional signals. Nat Rev Neurosci, 2010, 11: 697–709
21 Fries P, Reynolds J H, Rorie A E, et al. Modulation of oscillatory neuronal synchronization by selective visual attention. Science, 2001, 291: 1560–1563
22 Roberts M, Delicato L S, Herrero J, et al. Attention alters spatial integration in macaque V1 in an eccentricity-dependent manner. Nat Neurosci, 2007, 10: 1483–1491
23 Qiu F T, Sugihara T, von der Heydt R. Figure-ground mechanisms provide structure for selective attention. Nat Neurosci, 2007, 10: 1492–1499
24 Hübner R, Steinhauser M, Lehle C. A dual-stage two-phase model of selective attention. Psychol Rev, 2010, 117: 759–784
25 Gondan M, Blurton S P, Hughes F, et al. Effects of spatial and selective attention on basic multisensory integration. J Exp Phychol-Hum Percep Perf, 2011, 37: 1887–1897
26 Schafer R J, Moore T. Selective attention from voluntary control of neurons in prefrontal cortex. Science, 2011, 332: 1568–1571
27 Couperus J W. Perceptual load influences selective attention across development. Develop Psychol, 2011, 47: 1431–1439
28 Cosman J D, Vecera S P. Object-based attention overrides perceptual load to modulate visual distraction. J Exp Phychol-Hum Percep Perf, 2012, 38: 576–579
29 Chen C C, Wu J K, Lin H W, et al. Visualizing long-term memory formation in two neurons of the drosophila brain. Science, 2012, 335: 678–685
30 Fell J, Axmacher N. The role of phase synchronization in memory processes. Nat Rev Neurosci, 2011, 12: 105–118
31 Fusi S, Abbott L F. Limits on the memory storage capacity of bounded synapses. Nat Neurosci, 2007, 10: 485–493
32 Eichenbaum H. A cortical-hippocampal system for declarative memory. Nat Rev Neurosci, 2000, 1: 41–50
33 McGaugh J L. Memory—a century of consolidation. Science, 2000, 287: 248–251
34 Tronson N C, Taylor J R. Molecular mechanisms of memory reconsolidation. Nat Rev Neurosci, 2007, 8: 262–275
35 Edelson M, Sharot T, Dolan R J, et al. Following the crowd: brain substrates of long-term memory conformity. Science, 2011, 333: 108–111
36 Frankland P W, Bontempi B. The organization of recent and remote memories. Nat Rev Neurosci, 2005, 6: 119–130
37 Nadel L, Hardt O. Update on memory systems and processes. Neuropsychopharmacology, 2011, 36: 251–273
38 Nader K, Hardt O. A single standard for memory: the case for reconsolidation. Nat Rev Neurosci, 2009, 10: 224–234
39 Gonzalez C, Dutt V. Instance-based learning: Integrating sampling and repeated decisions from experience. Psychol Rev, 2011, 118: 523–551
40 Homa D, Hout M C, Milliken L, et al. Bogus concerns about the false prototype enhancement effect. J Exp Psychol-Learn Mem Cogn, 2011, 37: 368–377
41 Smith J D, Minda J P. Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning.

J Exp Psychol-Learn Mem Cogn, 2002, 28: 800–811
42 Smith J D, Redford J S, Haas S M. Prototype abstraction by monkeys (Macaca mulatta). J Exp Psychol-Gen, 2008, 137: 390–401
43 Lewandowsky S, Palmeri T J, Waldmann M R. Introduction to the special section on theory and data in categorization: integrating computational, behavioral, and cognitive neuroscience approaches. J Exp Psychol-Learn Mem Cogn, 2012, 38: 803–806
44 Freedman D J, Assad J A. A proposed common neural mechanism for categorization and perceptual decisions. Nat Neurosci, 2011, 14: 143–146
45 Gold J I, Shadlen M N. The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. J Neurosci, 2003, 23: 632–651
46 Kable J W, Glimcher P W. The neurobiology of decision: consensus and controversy. Neuron, 2009, 63: 733–745
47 Freedman D J, Assad J A. Experience-dependent representation of visual categories in parietal cortex. Nature, 2006, 443: 85–88
48 Freedman D J, Riesenhuber M, Poggio T, et al. Categorical representation of visual stimuli in the primate prefrontal cortex. Science, 2001, 291: 312–316
49 Freedman D J, Riesenhuber M, Poggio T, et al. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci, 2003, 23: 5235–5246
50 Williams Z M, Elfar J C, Eskandar E N, et al. Parietal activity and the perceived direction of ambiguous apparent motion. Nat Neurosci, 2003, 6: 616–623
51 Toth L J, Assad J A. Dynamic coding of behaviourally relevant stimuli in parietal cortex. Nature, 2002, 415: 165–168
52 Stoet G, Snyder L H. Single neurons in posterior parietal cortex of monkeys encode cognitive set. Neuron, 2004, 42: 1003–1012
53 Gottlieb J. From thought to action: the parietal cortex as a bridge between perception, action, and cognition. Neuron, 2007, 53: 9–16
54 Chen Y, Martinez-Conde S, Macknik S L, et al. Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. Nat Neurosci, 2008, 11: 974–982
55 Asplund C L, Todd J J, Snyder A P, et al. A central role for the lateral prefrontal cortex in goal-directed and stimulus-driven attention. Nat Neurosci, 2010, 13: 507–512
56 Solway A, Botvinick M M. Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. Psychol Rev, 2012, 119: 120–154
57 Purcell B A, Heitz R P, Cohen J Y, et al. Neurally constrained modeling of perceptual decision making. Psychol Rev, 2010, 117: 1113–1143
58 Palmeri T J, Gauthier I. Visual object understanding. Nat Rev Neurosci, 2004, 5: 291–304
59 Peyrin C, Michel C M, Schwartz S, et al. The neural substrates and timing of top-down processes during coarse-to-fine categorization of visual scenes: a combined fMRI and ERP study. J Cognitive Neurosci, 2010, 22: 2768–2780
60 Gao Z F, Bentin S. Coarse-to-fine encoding of spatial frequency information into visual short-term memory for faces but impartial decay. J Exp Phychol-Hum Percep Perf, 2011, 37: 1051–1064
61 Goffaux V, Peters J, Haubrechts J, et al. From coarse to fine? Spatial and temporal dynamics of cortical face processing. Cereb Cortex, 2011, 21: 467–476
62 Griffiths O, Mitchell C J. Selective attention in human associative learning and recognition memory. J Exp Psychol-Gen, 2008, 137: 626–648
63 Deng W, Aimone J B, Gage F H. New neurons and new memories: how does adult hippocampal neurogenesis affect learning and memory? Nat Rev Neurosci, 2010, 11: 339–350
64 De Fockert J W, Rees G, Frith C D, et al. The role of working memory in visual selective attention. Science, 2001, 291: 1803–1806
65 Saalmann Y B, Pigarev I N, Vidyasagar T R. Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. Science, 2007, 316: 1612–1615
66 Sigala N, Logothetis N K. Visual categorization shapes feature selectivity in the primate temporal cortex. Nature, 2002, 415: 318–320
67 Kundel H L, Nodine C F. Interpreting chest radiographs without visual search. Radiology, 1975, 116: 527–532
68 Treisman A M, Gelade G. A feature-integration theory of attention. Cog Psychol, 1980, 12: 97–136
69 Liu Y J, Fu Q F, Liu Y, et al. 2D-line-drawing-based 3D object recognition. In: Computational Visual Media, Beijing, 2012. 146–153
70 Liu Y J, Luo X, Joneja A, et al. User-adaptive sketch-based 3D CAD model retrieval. IEE Trans Autom Sci Eng, 2013, 99: 1–13
71 Wolfe J M. Guided Search 4.0: current progress with a model of visual search. In: Integrated Models of Cognitive Systems. New York: Oxford, 2007. 99–119
72 Wolfe J M, Cave K R, Franzel S L. Guided search: an alternative to feature integration model for visual search. J Exp

Phychol-Hum Percep Perf, 1989, 15: 419–433

73 Williams C C, Henderson J M, Zacks R T. Incidental visual memory for targets and distractors in visual search. Percept Psychophys, 2005, 67: 816–827

74 Wolfe J M. Guided search 2.0: a revised model of visual search. Psychonomic Bull Rev, 1994, 1: 202–238

75 Hao F, Zhang H, Fu X L. Modulation of attention by faces expressing emotion: evidence from visual marking. In: Tao J H, Tan T N, Picard R W, eds. Affective Computing and Intelligent Interaction. Berlin/Heidelberg: Springer-Verlag, 2005. 127–134

76 Hao F, Fu X L. Visual marking: a mechanism of prioritizing selection. Adv Psychol Sci, 2006, 14: 7–11

77 Sternberg S. High-speed scanning in human memory. Science, 1966, 153: 652–654

78 Hawkins J, Blakeslee S. On Intelligence. New York: Times Books, 2004

79 Bear M F, Connors B W, Paradiso M A. Neuroscience: exploring the brain. 3rd ed. Philadelphia: Lippincott Williams & Wilkins, 2006

80 Rinkus G J. A cortical sparse distributed coding model linking mini- and macrocolumn-scale functionality. Front Neuroanat, 2010, 4: 17

81 Mountcastle V. An Organizing Principle for Cerebral Function: the Unit Model and the Distributed System. Cambridge: MIT Press, 1978

82 Elad M. Sparse and Redundant Representations: from Theory to Applications in Signal and Image Processing. New York: Springer, 2010

83 Bruckstein A M, Donoho D L, Elad M. From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM Rev, 2009, 51: 34–81

84 Yi Z, Tan K K. Convergence Analysis of Recurrent Neural Networks. Dordrecht: Kluwer Academic Publishers, 2004

85 Tang H J, Tan K C, Yi Z. Neural Networks: Computational Models and Applications. Heidelberg: Springer-Verlag, 2007

86 Seung H S. How the brain keeps the eyes still. Proc Nat Acad Sci USA, 1996, 93: 13339–13344

87 Wu S, Amari S, Nakahara H. Population coding and decoding in a neural field: a computational study. Neural Comput, 2002, 14: 999–1026

88 Zhang K. Representation of spatial orientation by the intrinsic dynamics of head-direction cell ensembles: a theory. J Neurosci, 1996, 16: 2112–2126

89 Yu J, Yi Z, Zhang L. Representations of continuous attractors of recurrent neural networks. IEEE Trans Neural Netw, 2009, 20: 368–372

90 Wu C, Liu Y. Queuing network modeling of the psychological refractory period (PRP). Psychol Rev, 2008, 115: 913–954

91 Johnson J G, Busemeyer J R. Rule-based decision field theory: a dynamic computational model of transitions among decision-making strategies. In: Betsch T, Haberstroh S, eds. The Routines of Decision Making. Mahwah: Lawrence Erlbaum, 2005. 3–19

92 Zhao G, Wu C, Qiao C. A mathematical model for the prediction of speeding with its validation. IEEE Trans Intell Transp Syst, 2013, 14: 828–836

93 Wang X H, Jia J, Hu P Y, et al. Understanding the emotional impact of image. In: ACM Multimedia, Nara, 2012. 1369–1370

94 Jia J, Wu S, Wang X H, et al. Can we understand van Gogh's mood? Learning to infer affects from images in social networks. In: ACM Multimedia, Nara, 2012. 857–860

95 Wang X H, Jia J, Cai L H. Affective image adjustment with a single word. Vis Comput, 2013, 29: 1121–1133

96 Wang X H, Jia J, Liao H Y, et al. Affective image colorization. J Comput Sci Technol, 2012, 27: 1119–1128

97 Kobayashi S. Art of Color Combinations. Tokyo: Kodansha International, 1995

98 Zhang Y F, Hu S M, Martin R R. Shrinkability maps for content-aware video resizing. Comput Graph Forum, 2008, 27: 1797–1804

99 Zhang G X, Cheng M M, Hu S M, et al. A shape-preserving approach to image resizing. Comput Graph Forum, 2009, 28: 1897–1906

100 Dahan M J, Chen N, Shamir A, et al. Combining color and depth for enhanced image segmentation and retargeting. Vis Comput, 2012, 28: 1181–1193

101 Liu Y J, Luo X, Xuan Y M, et al. Image retargeting quality assessment. Comput Graph Forum, 2011, 30: 583–592

102 Chen L. Topological structure in visual perception. Science, 1982, 218: 699–700

103 Anderson J R, Bothell D, Byrne M D, et al. An integrated theory of the mind. Psychol Rev, 2004, 111: 1036–1060