A

Adaptive Resonance Theory

Gail A. Carpenter[1] and Stephen Grossberg[2]
[1]Department of Mathematics & Center for Adaptive Systems, Boston University, Boston, MA, USA
[2]Center for Adaptive Systems, Graduate Program in Cognitive and Neural Systems, Department of Mathematics, Boston University, Boston, MA, USA

Abstract

Computational models based on cognitive and neural systems are now deeply embedded in the standard repertoire of machine learning and data mining methods, with intelligent learning systems enhancing performance in nearly every existing application area. Beyond data mining, this article shows how models based on adaptive resonance theory (ART) may provide entirely new questions and practical solutions for technological applications. ART models carry out hypothesis testing, search, and incremental fast or slow, self-stabilizing learning, recognition, and prediction in response to large nonstationary databases (big data). Three computational examples, each based on the distributed ART neural network, frame questions and illustrate how a learning system (each with no free parameters) may enhance the analysis of large-scale data. Performance of each task is simulated on a common mapping platform, a remote sensing dataset called the Boston Testbed, available online along with open-source system code. Key design elements of ART models and links to software for each system are included. The article further points to future applications for integrative ART-based systems that have already been computationally specified and simulated. New application directions include autonomous robotics, general-purpose machine vision, audition, speech recognition, language acquisition, eye movement control, visual search, figure-ground separation, invariant object recognition, social cognition, object and spatial attention, scene understanding, space-time integration, episodic memory, navigation, object tracking, system-level analysis of mental disorders, and machine consciousness.

Adaptive Resonance Theory

Adaptive resonance theory (ART) neural networks model real-time hypothesis testing, search, learning, recognition, and prediction. Since the 1980s, these models of human cognitive information processing have served as computational engines for a variety of neuromorphic technologies (http://techlab.bu.edu/resources/articles/C5). This article points to a broader range of technology transfers that bring new methods to new problem domains. It describes applications of three specific systems, ART knowledge discov-

C. Sammut, G.I. Webb (eds.), *Encyclopedia of Machine Learning and Data Mining*,
DOI 10.1007/978-1-4899-7502-7_6-1

ery, self-supervised ART, and biased ART, and summarizes future application areas for large-scale, brain-based model systems.

ART Design Elements

In this article, *ART* refers generally to a theory of cognitive information processing and to an inclusive family of neural models. Design principles derived from scientific analyses and design constraints imposed by targeted applications have jointly guided the development of variants of the basic systems.

Stable Fast Learning with Distributed and Winner-Take-All Coding

ART systems permit fast online learning, whereby long-term memories reach their asymptotes on each input trial. With slow learning, memories change only slightly on each trial. One characteristic that distinguishes classes of ART systems from one another is the nature of their patterns of persistent activation at the coding field F_2 (Fig. 1). The coding field is functionally analogous to the hidden layer of multilayer perceptrons (*Encyclopedia* cross reference). At the perceptron hidden layer, activation is distributed across many nodes, learning needs to be slow, and activation does not persist once inputs are removed. The ART coding field is a competitive network where, typically, one or a few nodes in the normalized F_2 pattern $\mathbf{y}$ sustain persistent activation, even as their generating inputs shift, habituate, or vanish. The pattern $\mathbf{y}$ persists until an active reset signal (Fig. 1c) prepares the coding field to register a new F_0-to-F_2input. Early ART networks (Carpenter and Grossberg 1987; Carpenter et al. 1991a, 1992) employed *localist*, or *winner-take-all*, coding, whereby strongly competitive feedback results in only one F_2 node staying active until the next reset. With fast as well as slow learning, memory stability in these early networks relied on their winner-take-all architectures.

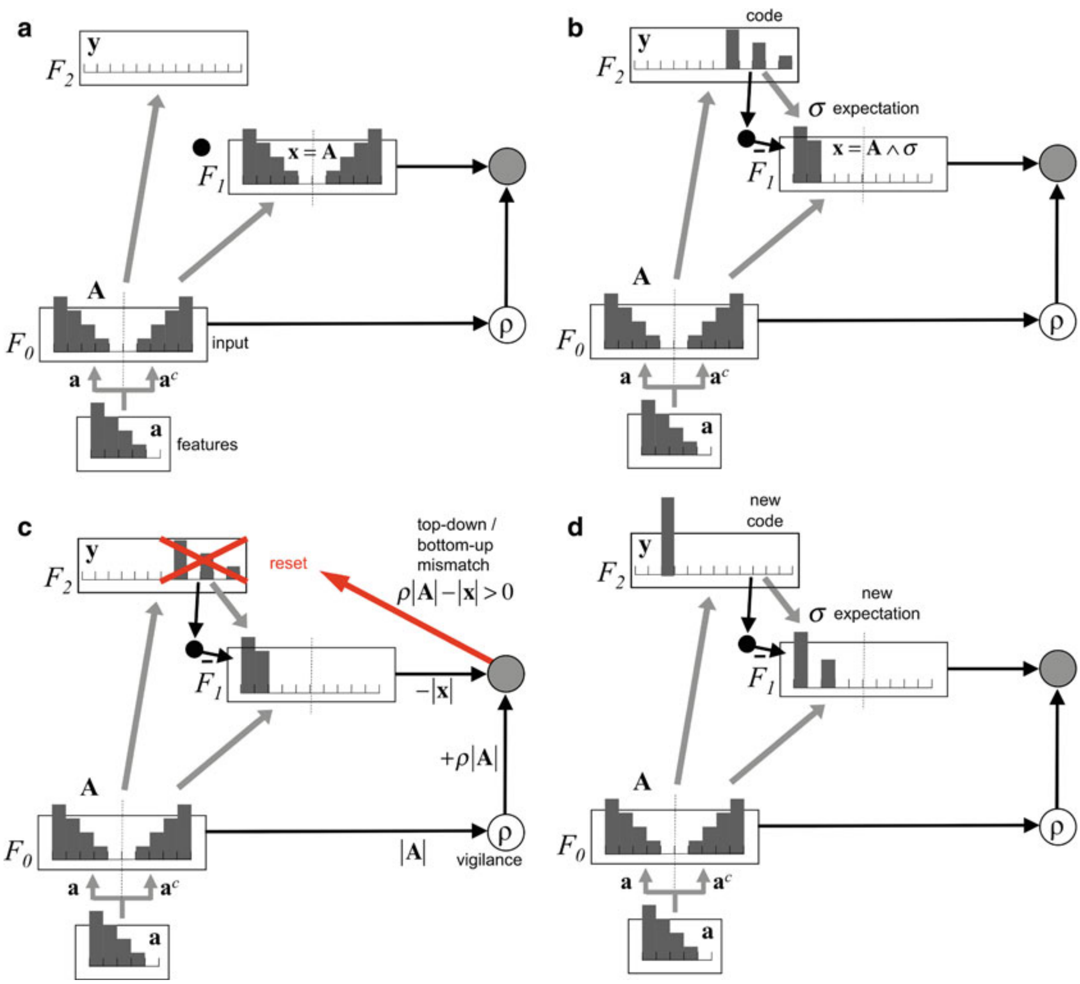
Achieving stable fast learning with distributed code representations presents a computational challenge to any learning network. In order to meet this challenge, distributed ART (Carpenter 1997) introduced a new network configuration (Fig. 1) in which system fields are identified with cortical layers (Carpenter 2001). New learning laws (*dInstar* and *dOutstar*) that realize stable fast learning with distributed coding predict adaptive dynamics between cortical layers.

Distributed ART (dART) systems employ a new unit of long-term memory, which replaces the traditional multiplicative weight (*Encyclopedia* cross reference) with a *dynamic weight* (Carpenter 1994). In a path from the F_2 coding node j to the F_1 matching node i, the dynamic weight equals the amount by which coding node activation y_j exceeds an *adaptive threshold* τ_{ji}. The total signal σ_i from F_2 to the i^{th} F_1 node is the sum of these dynamic weights, and F_1 node activation x_i equals the minimum of the top-down expectation σ_i and the bottom-up input A_i. During dOutstar learning, the top-down pattern σ converges toward the matched pattern $\mathbf{x}$.

When coding node activation $y_j$is below τ_{ji}, the dynamic weight is zero and no learning occurs in that path, even if y_j is positive. This property is critical for stable fast learning with distributed codes. Although the dInstar and dOutstar laws are compatible with F_2 patterns $\mathbf{y}$ that are arbitrarily distributed, in practice, following an initial learning phase, most changes in paths to and from a coding node j occur only when its activation y_j is large. This type of learning is therefore called *quasi-localist*. In the special case where coding is winner-take-all, the dynamic weight is equivalent to a multiplicative weight that formally equals the *complement* of the adaptive threshold.

Complement Coding: Learning Both Absent Features and Present Features

ART networks employ a preprocessing step called *complement coding* (Carpenter et al. 1991b), which models the nervous system's ubiquitous computational design known as *opponent processing* (Hurvich and Jameson 1957). Balancing an entity against its opponent, as in opponent colors such as red vs. green or agonist-antagonist muscle pairs, allows a system to act upon relative quantities, even as absolute magnitudes fluctuate unpredictably. In ART systems, complement coding is analogous to retinal on-cells and off-cells (Schiller 1982).



Adaptive Resonance Theory, Fig. 1 Distributed ART (dART) (Carpenter 1997). (**a**) At the field F_0, complement coding transforms the feature pattern **a** to the system input **A**, which represents both scaled feature values $a_i \in [0, 1]$ and their complements $(1 - a_i)$ $(i = 1 \dots M)$. (**b**) F_2 is a competitive field that transforms its input pattern into the working memory code **y**. The F_2 nodes that remain active following competition send the pattern σ of learned top-down expectations to the match field F_1. The pattern active at F_1 becomes $\mathbf{x} = \mathbf{A} \wedge \sigma$, where $\wedge$ denotes the component-wise minimum, or fuzzy intersection. (**c**) A parameter $\rho \in [0, 1]$, called vigilance, sets the matching criterion. The system registers a mismatch if the size of **x** is less than ρ times the size of **A**. A top-down/bottom-up mismatch triggers a signal that resets the active F_2 code. (**d**) Medium-term memories in the F_0-to-F_2 dynamic weights allow the system to activate a new code **y**. When only one F_2 node remains active following competition, the code is maximally compressed, or winner-take-all. When $|\mathbf{x}| \geq \rho\,|\mathbf{A}|$, the activation pattern **y** persists until the next reset, even if input **A** changes or F_0-to-F_2 signals habituate. During learning, thresholds τ_{ij} in paths from F_0 to F_2 increase according to the dInstar law; and thresholds τ_{ji} in paths from F_2 to F_1 increase according to the dOutstar law

When the learning system is presented with a set of input features $\mathbf{a} \equiv (a_1 \dots a_i \dots a_M)$, complement coding doubles the number of input components, presenting to the network an input **A** that concatenates the original feature vector and its complement (Fig. 1a).

Complement coding produces normalized inputs **A** that allow a model to encode features that are consistently *absent* on an equal basis with features that are consistently *present*. Features that are sometimes absent and sometimes present when a given F_2 node is highly active are regarded as uninformative with respect to that node, and the corresponding *present* and *absent* top-down feature expectations shrink to zero. When a new input activates this node, these features are suppressed at the match field F_1 (Fig. 1b). If the active code then produces an error signal, attentional biasing can enhance the salience of input features that it had previously ignored, as described below.

Matching, Attention, and Search

A neural computation central to both scientific and technological analyses is the *ART matching rule* (Carpenter and Grossberg 1987), which controls how attention is focused on critical feature patterns via dynamic matching of a bottom-up sensory input with a top-down learned expectation. Bottom-up/top-down pattern matching and attentional focusing are, perhaps, the primary features common to all ART models across their many variations. Active input features that are not confirmed by top-down expectations are inhibited (Fig. 1b). The remaining activation pattern defines a focus of attention, which, in turn, determines what feature patterns are learned. Basing memories on attended features rather than whole patterns supports the design goal of encoding stable memories with fast as well as slow learning. Encoding attended feature subsets also enables one-to-many learning, where the system may attach many context-dependent labels (*Spot*, *dog*, *animal*) to one input. This capability promotes knowledge discovery ($Spot \Rightarrow dog$ and $dog \Rightarrow animal$) in a learning system that experiences one input at a time, with no explicit connection between inputs.

When the match is good enough, F_2 activation persists and learning proceeds. Where they exceed the corresponding bottom-up input components, top-down signals decay as expectations converge toward the attended pattern at F_1. The coding field F_2 contains a reserve of *uncommitted* coding nodes, which compete with the previously active *committed* nodes. When a previously uncommitted node is first activated during supervised learning, it is associated with its designated output class. During testing, the selection of an uncommitted node means *I don't know*. ART networks for supervised learning are called *ARTMAP* (Carpenter et al. 1991a, 1992).

A mismatch between an active top-down expectation and the bottom-up input leads to a parallel memory search (Fig. 1c). The ART matching criterion is set by a *vigilance parameter* ρ. Low vigilance permits the learning of broad classes, across diverse exemplars, while high vigilance limits learning to narrow classes. When a new input arrives, vigilance equals a baseline level. Baseline vigilance is set equal to zero to maximize generalization. ARTMAP vigilance increases following a predictive error or negative reinforcement (*Encyclopedia* cross reference). The internal computation that determines how far ρ rises to correct the error is called *match tracking* (Carpenter et al. 1991a). As vigilance rises, the network pays more attention to how well top-down expectations match the bottom-up input. The match tracking modification MT– (Carpenter and Markuzon 1998) also allows the system to learn inconsistent cases. For example, three similar, even identical, map regions may have been correctly labeled by different observers as *ocean* or *water* or *natural*. The ability to learn one-to-many maps, which can label a single test input as *ocean* and *water* and *natural*, is a key feature of the ART knowledge discovery system described below.

Applications

Three computational examples illustrate how cognitive and neural systems can introduce new approaches to the analysis of large datasets. Application 1 (self-supervised ART) addresses the question: how can a neural system learning from one example at a time absorb information that is inconsistent but correct, as when a family pet is called *Spot* and *dog* and *animal*, while rejecting similar incorrect information, as when the same pet is called *wolf*? How does this system transform scattered information into knowledge that *dogs are animals*, but not conversely? Application 2 (ART knowledge discovery) asks: how can a real-time system, initially trained with a few labeled examples and a limited feature set, continue to learn from experience, without supervision, when confronted with oceans of additional information, without eroding reliable early memories? How can such individual systems adapt to their unique application contexts? Application 3 (biased ART) asks: how can a neural system that has made an error refocus attention on features that it initially ignored?

The Boston Testbed

The Boston Testbed was developed to compare performance of learning systems applied to challenging problems of spatial analysis. Each multispectral Boston image pixel produces 41 feature values: 6 Landsat 7 Thematic Mapper (TM) bands at 30 m resolution, 2 thermal bands at 60 m resolution, 1 panchromatic band at 15 m resolution, and 32 derived bands representing local contrast, color, and texture. In the Boston dataset, each of 28,735 ground truth pixels is labeled as belonging to one of seven classes (*beach, ocean, ice, river, park, residential, industrial*). For knowledge discovery system training, some *ocean, ice*, and *river* pixels are instead labeled as belonging to broader classes such as *water* or *natural*. No pixel has more than one label, and the learning system is given no information about relationships between target classes. The labeled dataset is available from the CNS Technology Lab Website [http://techlab.bu.edu/classer/data_sets/].

A cross-validation procedure divides an image into four vertical strips: two for training, one for validation (if needed for parameter selection), and one for testing. Class mixtures differ markedly across strips. For example, one strip contains many *ocean* pixels, while another strip contains neither *ocean* nor *beach* pixels. Geographically dissimilar training and testing areas robustly assess regional generalization. In this article, spatial analysis simulations on the Boston Testbed follow this protocol to illustrate ART systems for self-supervised learning, knowledge discovery, and attentional control. Since each system in Applications 1–3 requires no parameter selection, training uses randomly chosen pixels from three strips, with testing on the fourth strip.

Application 1: Learning from Experience with Self-Supervised ART

Computational models of supervised pattern recognition typically utilize two learning phases. During an initial training phase, input patterns, described as specified values of a set of features, are presented along with output class labels or patterns. During a subsequent testing phase, the model generates output predictions for unlabeled inputs, and no further learning takes place.

Although supervised learning has been successfully applied in diverse settings, it does not reflect many natural learning situations. Humans do learn from explicit training, as from a textbook or a teacher, and they do take tests. However, students do not stop learning when they leave the classroom. Rather, they continue to learn from experience, incorporating not only more information but new types of information, all the while building on the foundation of their earlier knowledge. Self-supervised ART models such life-long learning.

An unsupervised learning system clusters unlabeled input patterns. *Semi-supervised* learning incorporates both labeled and unlabeled inputs in its training set, but all inputs typically have the same number of specified feature values. Without any novel features from which to learn, semi-supervised learning systems use unlabeled data to refine the model parameters defined using labeled data. Reviews of semi-supervised learning (Chapelle et al. 2006) have found that many of the successful models are carefully selected and tuned, using a priori knowledge of the problem. Chapelle et al. (2006) conclude that none of the semi-supervised models they review is robust enough to be general purpose. The main difficulty seems to be that, whenever unlabeled instances are different enough from labeled instances to merit learning, these differences could contain misinformation that may damage system performance.

The *self-supervised* paradigm models two learning stages. During Stage 1 learning, the system receives all output labels, but only a subset of possible feature values for each input. During Stage 2 learning, the system may receive more feature values for each input, but no output labels. In Stage 1, when the system can confidently incorporate externally specified output labels, self-supervised ART (Amis and Carpenter 2010) employs winner-take-all coding and fast learning. In Stage 2, when the system internally generates its own output labels, codes are distributed so that incorrect hypotheses do not abruptly override reliable "classroom learning" of Stage 1. The distributed ART learning laws, dInstar (Carpenter 1997) and dOutstar (Carpenter 1994), scale memory changes to internally generated measures of prediction confidence and prevent memory changes altogether for most inputs. Memory stability derives from the dynamic weight representation of long-term memories, which permits learning only in paths to and from highly active coding nodes. Dynamic weights solve a problem inherent in learning laws based on multiplicative weights, which are prone to catastrophic forgetting when implemented with distributed codes and huge datasets, even when learning is very slow.

In addition to emulating the human learning experience, self-supervised learning maps to technological applications that need to cope with huge, ever-changing datasets. A supervised learning system that completes all training before making test predictions does not adapt to new information and individual contexts. A semi-supervised system risks degrading its supervised knowledge. Self-supervised ART continues to learn from new experiences, with built-in safeguards that conserve useful memories. Self-supervised ART code is available from the CNS Technology Lab Website (http://techlab.bu.edu/SSART/).

A simulation study based on the Boston Testbed (Amis and Carpenter 2010) illustrates ways in which high-dimensional problems may challenge *any system* learning without labels. As in most ground truth datasets, labeled pixels consist primarily of clear exemplars of single classes. Because sensors have a 15–60 m resolution, many unlabeled pixels cover multiple classes, such as *ice* and *industrial*. Stage 2 inputs thus mix and distort features from multiple classes, placing many of the unlabeled feature vectors far from the distinct class clusters of the Stage 1 training set. Although the distributed ART learning laws are open to unrestricted adaptation on any pixel, the distributed codes of Stage 2 minimize the influence of mixed pixels. Most memory changes occur on unambiguous cases, despite the fact that the unlabeled pixels provide no external indices of class ambiguity. Self-supervised Stage 2 learning dramatically improves performance compared to learning that ends after Stage 1. On every one of 500 individual simulations, Stage 2 learning improves test accuracy, as unlabeled fully featured inputs consistently expand knowledge from Stage 1 training.

Application 2: Transforming Information into Knowledge Using ART Knowledge Discovery

Classifying terrain or objects may require the resolution of conflicting information from sensors working at different times, locations, and scales and from users with different goals and situations. *Image fusion* has been defined as "the acquisition, processing and synergistic combination of information provided by various sensors or by the same sensor in many measuring contexts" (Simone et al. 2002, p. 3). When multiple sources provide inconsistent data, fusion methods are called upon to appraise information components to decide among various options and to resolve inconsistencies, as when evidence suggests that an object is a *car* or a *truck* or a *bus*. Fusion methods weigh the confidence and reliability of each source, merging complementary information or gathering more data. In any case, at most one of these answers is correct.

The method described here defines a complementary approach to the information fusion problem, considering the case where sensors and sources are both nominally inconsistent and reliable, as when evidence suggests that an object is a *car* and a *vehicle* and *man-made* or when a *car* is alternatively labeled *automobile*. Underlying relationships among classes are assumed to be unknown to the automated system or the human user, as if the labels were encrypted.

The ART knowledge discovery model acts as a self-organizing expert system to derive consistent knowledge structures from such nominally inconsistent data (Carpenter et al. 2005). Once derived, a rule set can be used to assign classes to levels. For each rule $x \Rightarrow y$, class x is located at a lower level than class y. Classes connected by arrows that codify a list of rules and confidence values form a graphical representation of a knowledge hierarchy. For spatial data, the resulting diagram of the relationships among classes can guide the construction of orderly layered maps. ART knowledge discovery code is available from the CNS Technology Lab Website (http://techlab.bu.edu/classer/classer_toolkit_overview). On the Boston Testbed, the ART knowledge discovery system places each class at its correct level and finds all the correct rules for this example.

Application 3: Correcting Errors by Biasing Attention Using Biased ART

Memories in ART networks are based on matched patterns that focus attention on *critical features*, where bottom-up inputs match active top-down expectations. While this learning strategy has proved successful for both brain models and applications, computational examples demonstrate that paying too much attention to critical features that have been selected to represent a given category early on may distort memory representations during subsequent learning. If training inputs are repeatedly presented, an ART system will correct these initial errors. However, real-time learning may not afford such repeat opportunities. Biased ART (bART) (Carpenter and Gaddam 2010) solves the problem of overemphasis on early critical features by directing attention away from initially attended features after the system makes a predictive error.

Activity $\mathbf{x}$ at the ART field F_1 computes the match between the field's bottom-up and top-down input patterns (Fig. 1). A reset signal shuts off the active F_2 code when $\mathbf{x}$ fails to meet the matching criterion determined by vigilance ρ. Reset alone does not, however, induce a different code: unless the prior code has left an enduring trace within the F_0–F_2 subsystem, the network will simply reactivate the same pattern at F_2.

Following reset, all ART systems shift attention away from previously active *coding* nodes at the field F_2. As modeled in ART 3 (Carpenter and Grossberg 1990), biasing the bottom-up input to the coding field to favor previously inactive F_2 nodes implements search by enabling the network to activate a new code in response to a reset signal. The ART 3 search mechanism defines a medium-term memory in the F_0-to-F_2 adaptive filter so that the system does not perseverate indefinitely on an output class that had just produced a reset. A presynaptic interpretation of this bias mechanism is transmitter depletion or habituation.

The biased ART network (Carpenter and Gaddam 2010) introduces a second, top-down, medium-term memory which, following reset, shifts attention away from previously active *feature* nodes at the match field F_1. In Fig. 1, the first feature is strongly represented in the input $\mathbf{A}$ and in the matched patterns $\mathbf{x}$ at F_1 both before reset (Fig. 1b) and after reset (Fig. 1d). Following the same sequence as in Fig. 1a–c, biased ART would diminish the size of the first feature in the matched pattern. The addition of featural biasing helps the system to pay more attention to input features that it had previously ignored.

The biasing mechanism is a small modular element that can be added to any ART network. While computational examples and Boston Testbed simulations demonstrate how featural biasing in response to predictive errors improves performance on supervised learning tasks, the error signal that gates biasing could have originated from other sources, as in reinforcement learning. Biased ART code is available from the CNS Technology Lab Website (http://techlab.bu.edu/bART).

Future Directions

Applications for tested software based on computational intelligence abound. This section outlines areas where ART systems may open qualitatively new frontiers for novel technologies. Future applications summarized here would adapt and specialize brain models that have already been mathematically specified and computationally simulated to explain and predict large psychological and neurobiological databases. By linking the brain to mind, these models characterize both mechanism (how the model works) and function (what the model is for). Both mechanism and function are needed to design new applications. These systems embody new designs for autonomous adaptive agents, including new computational paradigms that are called Complementary Computing and Laminar Computing. These paradigms enable the autonomous adaptation in real time of individual persons or machines to nonstationary situations filled with unexpected events. See Grossberg (2013) for a review.

New Paradigms for Autonomous Intelligent Systems: Complementary Computing and Laminar Computing

Functional integration is essential to the design of a complex autonomous system such as a robot moving and learning freely in an unpredictable environment. Linking independent modules for, say, vision and motor control will not necessarily produce a coordinated system that can adapt to unexpected events in changeable contexts. How, then, should such an autonomous adaptive system be designed?

A clue can be found in the nature of brain specialization. How have brains evolved while interacting with the physical world and embodying its invariants? Many scientists have proposed that our brains possess independent modules. The brain's organization into distinct anatomical areas and processing streams shows that brain regions are indeed specialized. Whereas independent modules compute their particular processes on their own, behavioral data argue against this possibility. *Complementary Computing* (Grossberg 2000a,b, 2013) concerns the discovery that pairs of parallel cortical processing streams compute computationally complementary properties. Each stream has complementary strengths and weaknesses, much as in physical principles like the Heisenberg uncertainty principle. Each cortical stream can also possess multiple processing stages. These stages realize a *hierarchical resolution of uncertainty*. "Uncertainty" here means that computing one set of properties at a given stage prevents computation of a complementary set of properties at that stage. Complementary Computing proposes that the computational unit of brain processing that has behavioral significance consists of parallel and hierarchical interactions between complementary cortical processing streams with multiple processing stages. These interactions overcome complementary weaknesses to compute necessary information about a particular type of biological intelligence.

Five decades of neural modeling have shown how Complementary Computing is embedded as a fundamental design principle in neural systems for vision, speech and language, cognition, emotion, and sensory-motor control. Complementary Computing hereby provides a blueprint for designing large-scale autonomous adaptive systems that are poised for technological implementation.

A unifying anatomical theme that enables communication among cortical systems is *Laminar Computing*. The cerebral cortex, the seat of higher intelligence in all modalities, is organized into layered circuits (often six main layers) that undergo characteristic bottom-up, top-down, and horizontal interactions. As information travels up and down connected regions, distributed decisions are made in real time based on a preponderance of evidence. Multiple levels suppress weaker groupings while communicating locally coherent choices. The distributed ART model (Fig. 1), for example, features three cortical layers, with its distributed code (e.g., at a cortical layer 6) producing a distributed output. Stacks of match fields (inflow) and coding fields (outflow) lay the substrate for cortical hierarchies.

How do specializations of this shared laminar design embody different types of biological intelligence, including vision, speech, language, and cognition? How does this shared design enable seamless intercortical interactions? Models of Laminar Computing clarify how these different types of intelligence all use variations of the same laminar circuitry (Grossberg 2013; Grossberg and Pearson 2008). This circuitry represents a revolutionary synthesis of desirable computational properties of feedforward and feedback processing, digital and analog processing, and bottom-up data-driven processing and top-down attentive hypothesis-driven processing. Realizing such designs in hardware that embodies biological intelligence promises to facilitate the development of increasingly general-purpose adaptive autonomous systems for multiple applications.

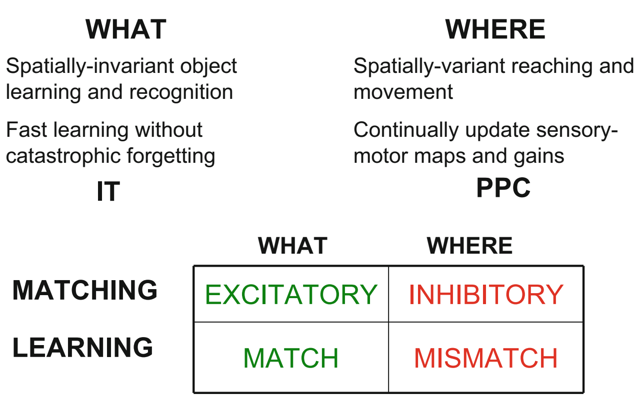
Complementary Computing in the Design of Perceptual/Cognitive and Spatial/Motor Systems

Many neural models that embody subsystems of an autonomous adaptive agent have been developed and computationally characterized. It remains to unify and adapt them to particular machine learning applications. Complementary Computing implies that not all of these subsystems could be based on variants of ART. In particular, accumulating experimental and theoretical evidence shows that perceptual/cognitive and spatial/motor processes use different learning, matching, and predictive laws for their complementary functions (Fig. 2). ART-like processing is ubiquitous in perceptual and cognitive processes, including excitatory matching and match-based learning that enables self-stabilizing memories to form. Vector Associative Map (VAM) processing is often found in spatial and motor processes, which rely on inhibitory matching and mismatch-based learning. In these modalities, spatial maps and motor plants are adaptively updated without needing to remember past maps and parameters. Complementary mechanisms create a self-stabilizing perceptual/cognitive front end for intelligently manipulating the more labile spatial/motor processes that enable our changeable bodies to act effectively upon a changing world.

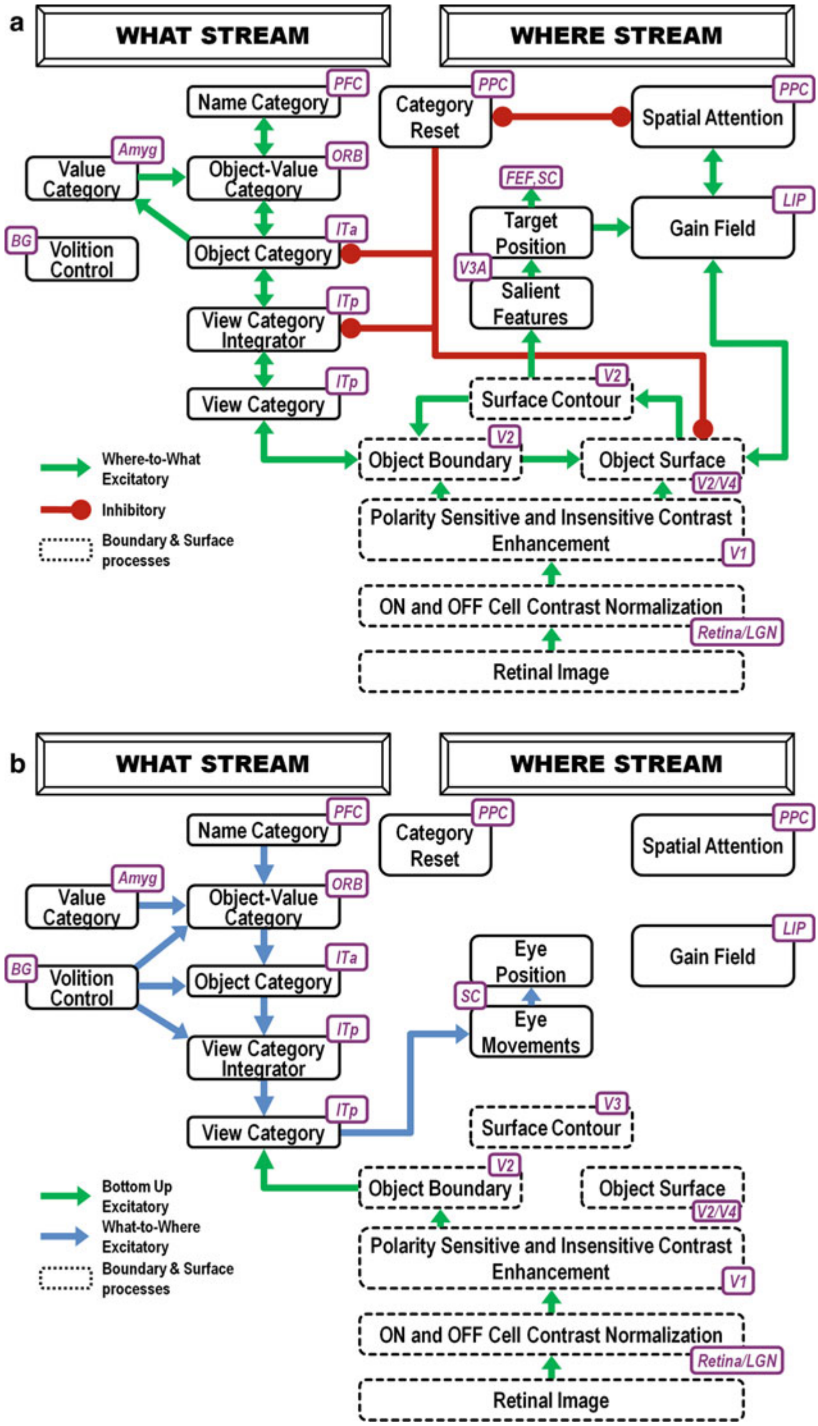
Some of the existing large-scale ART systems are briefly reviewed here, using visually based systems for definiteness. Citations refer to articles that specify system equations and simulations and that can be downloaded from http://cns.bu.edu/~steve.

Where's Waldo? Unifying Spatial and Object Attention, Learning, Recognition, and Search of Valued Objects and Scenes

ART models have been incorporated into larger system architectures that clarify how individuals autonomously carry out intelligent behaviors as they explore novel environments. One such development is the ARTSCAN family of architectures, which model how individuals rapidly learn to search a scene to detect, attend, invariantly recognize, and look at a valued target object (Fig. 3; Cao, Grossberg, and Markowitz 2011; Chang, Grossberg, and Cao 2014; Fazl, Gross-



Adaptive Resonance Theory, Fig. 2 Complementary What and Where cortical processing streams for spatially invariant object recognition and spatially variant spatial representation and action, respectively. Perception and recognition use top-down excitatory matching and match-based fast or slow learning without catastrophic forgetting. Spatial and motor tasks use inhibitory matching and mismatch-based learning to achieve adaptation to changing bodily parameters. *IT* inferotemporal cortex, *PPC* posterior parietal cortex



Adaptive Resonance Theory, Fig. 3 (continue) berg, and Mingolla 2009; Foley, Grossberg, and Mingolla 2012; Grossberg, Srinivasan, and Yazdanbakhsh 2014). Such a competence represents a proposed solution of the Where's Waldo problem.

The ventral What stream is associated with object learning, recognition, and prediction, whereas the dorsal Where stream carries out processes such as object localization, spatial attention, and eye movement control. To achieve efficient object recognition, the What stream learns object category representations that become increasingly invariant under view, size, and position changes at higher processing stages. Such invariance enables objects to be learned and recognized without causing a combinatorial explosion. However, by stripping away the positional coordinates of each object exemplar, the What stream loses the ability to command actions to the positions of valued objects. The Where stream computes positional representations of the world and controls actions to acquire objects in it, but does not represent detailed properties of the objects themselves.

ARTSCAN architectures model how an autonomous agent can determine when the views that are foveated by successive scanning movements belong to the same object and thus determine which view-selective categories should be associatively linked to an emerging view- , size-, and positionally-invariant object category. This competence, which avoids the problem of erroneously merging pieces of different objects, works even under the unsupervised learning conditions that are the norm during many object learning experiences in vivo. The model identifies a new role for spatial attention in the Where stream, namely, control of invariant object category learning by the What stream. Interactions across the What and Where streams overcome the deficiencies of computationally complementary properties of these streams.

In the ARTSCAN Search model, both Where-to-What and What-to-Where stream interactions are needed to overcome complementary weaknesses: Where stream processes of spatial attention and predictive eye movement control regulate What stream processes whereby multiple view- and positionally-specific object categories are learned and associatively linked to view- and positionally-invariant object categories through bottom-up and object-attentive top-down interactions. What stream cognitive-emotional learning processes enable the focusing of motivated attention upon the invariant object categories of desired objects (Brown, Bullock, and Grossberg 1999, 2004; Dranias, Grossberg, and Bullock 2008; Grossberg and Seidman 2006). What stream cognitive names or motivational drives can, together with volitional signals, drive a search for Waldo. Mediated by object attention, search proceeds from What stream

◄

Adaptive Resonance Theory, Fig. 3 ARTSCAN Search macrocircuit and corresponding brain regions. *Dashed boxes* indicate boundary and surface preprocessing. (**a**) Category learning system. *Arrows* represent excitatory cortical processes. Spatial attention in the Where stream regulates view-specific and view-invariant category learning and recognition, and attendant reinforcement learning, in the What stream. Connections ending in circular disks indicate inhibitory connections. (**b**) Where's Waldo search system. Search begins when a name category or value category is activated and subliminally primes an object-value category via the *ART matching rule*. A volition control signal enables the primed object-value category to fire output signals. Bolstered by volitional control signals, these output signals can, in turn, propagate through a positionally-invariant object category to all the positionally-variant view category integrators whose various views and positions are represented by the object category. The view category integrators can subliminally prime, but not fully activate, these view categories. All this occurs in the What stream. When the bottom-up input from an object's boundary/surface representation also activates one of these view categories, its activity becomes suprathreshold, wins the competition across view categories for persistent activation, and activates a spatial attentional representation of Waldo's position in the Where stream. *ITa* anterior part of inferotemporal cortex, *ITp* posterior part of inferotemporal cortex, *PPC* posterior parietal cortex, *LIP* lateral intraparietal cortex, *LGN* lateral geniculate nucleus, *ORB* orbitofrontal cortex, *Amyg* amygdala, *BG* basal ganglia, *PFC* prefrontal cortex, *SC* superior colliculus, *V1* striate visual cortex, *V2, V3*, and *V4* prestriate visual cortices positionally-invariant representations to Where stream positionally-specific representations that focus spatial attention on Waldo's position. ARTSCAN architectures hereby model how the dynamics of multiple brain regions are coordinated to achieve clear functional goals.

The focus of spatial attention on Waldo's position in the Where stream can be used to control eye and hand movements toward Waldo, after navigational circuits (see below) bring the observer close enough to contact him. VAM-type learning circuits have been developed for the control of goal-oriented eye and hand movements that can be used for this purpose (e.g., Bullock and Grossberg 1988, 1991; Bullock, Cisek, and Grossberg 1998; Contreras-Vidal, Grossberg, and Bullock 1997; Gancarz and Grossberg 1999; Grossberg, Srihasam, and Bullock 2012; Pack, Grossberg, and Mingolla 2001; Srihasam, Bullock, and Grossberg 2009).

The ARTSCENE system (Grossberg and Huang 2009) models how humans can incrementally learn and rapidly predict scene identity by gist and then accumulates learned evidence from scenic textures to refine its initial hypothesis, using the same kind of spatial attentional shrouds that help to learn invariant object categories in ARTSCAN. The ARTSCENE Search system (Huang and Grossberg 2010) models how humans use target-predictive contextual information to guide search for desired targets in familiar scenes. For example, humans can learn that a certain combination of objects may define a context for a kitchen and trigger a more efficient search for a typical object, such as a sink, in that context.

General-Purpose Vision and How It Supports Object Learning, Recognition, and Tracking

Visual preprocessing constrains the quality of visually based learning and recognition. On an assembly line, automated vision systems successfully scan for target objects in this carefully controlled environment. In contrast, a human or robot navigating a natural scene faces overlaid textures, edges, shading, and depth information, with multiple scales and shifting perspectives. In the human brain, evolution has produced a huge preprocessor, involving multiple brain regions, for object and scene representation and for target tracking and navigation. One reason for this is that visual boundaries and surfaces, visual form and motion, and target tracking and visually based navigation are computationally complementary, thus requiring several distinct but interacting cortical processing streams.

Prior to the development of systems such as ARTSCAN and ARTSCENE, the FACADE (Form-And-Color-And-DEpth) model provided a neural theory of form perception, including 3D vision and figure-ground separation (e.g., Cao and Grossberg 2005, 2012; Fang and Grossberg 2009; Grossberg, Kuhlmann, and Mingolla 2007; Grossberg and Swaminathan 2004; Kelly and Grossberg 2000). The 3D FORMOTION model provides a neural theory of motion processing and form-motion interactions (e.g., Baloch and Grossberg 1997; Baloch, Grossberg, Mingolla, and Nogueira 1999; Berzhanskaya, Grossberg, and Mingolla 2007; Grossberg, Leveille, and Versace 2011; Grossberg, Mingolla, and Viswanathan 2001; Grossberg and Rudd 1992). The FACADE model has just the properties that are needed for solving the Where's Waldo problem, and the 3D FORMOTION model has just the properties that are needed for tracking unpredictably moving targets. Their complementary properties enabled these extensions.

Visual and Spatial Navigation, Cognitive Working Memory, and Planning

In addition to being able to see, learn, recognize, and track valued goal objects, an animal or autonomous robot must also be able to navigate to or away from them and to interact with them through goal-oriented hand and arm movements. Navigation is controlled by two distinct and interacting systems: a visually guided system and a spatial path integration system.

Visually guided navigation through a cluttered natural scene is modeled using the 3D FORMOTION model as a front end. The STARS and ViSTARS neural systems (Browning, Grossberg, and Mingolla 2009a, b; Elder, Grossberg, and

Mingolla 2009) model how primates use object motion information to segment objects and optic flow information to determine heading (self-motion direction), for purposes of goal approach and obstacle avoidance in response to realistic environments. The models predict how computationally complementary processes in parallel streams within the visual cortex compute object motion for tracking and self-motion for navigation. The models' steering decisions compute goals as attractors and obstacles as repellers, as do humans.

Spatial navigation based upon path integration signals has been a topic of great interest recently. Indeed, the 2014 Nobel Prize in Physiology or Medicine was awarded to John O'Keefe for his discovery of place cells in the hippocampal cortex and to Edvard and May-Britt Moser for their discovery of grid cells in the entorhinal cortex. The GridPlaceMap neural system (Grossberg and Pilly 2012, 2014; Pilly and Grossberg 2012, 2014; Mhatre, Grossberg, and Gorchetchnikov 2012; Pilly and Grossberg 2014) proposes how entorhinal grid cells and hippocampal place cells may be learned as spatial categories in a hierarchy of self-organizing maps. The model responds to realistic rat navigational trajectories by learning both grid cells with hexagonal grid firing fields of multiple spatial scales, and place cells with one or more firing fields. Model dynamics match neurophysiological data about their development in juvenile rats. The GridPlaceMap model enjoys several parsimonious design features that will facilitate their embodiment in technological applications, including hardware: (1) similar ring attractor mechanisms process both linear and angular path integration inputs that drive map learning; (2) the same self-organizing map mechanisms can learn grid cell and place cell receptive fields in a hierarchy of maps, and both grid and place cells can develop by detecting, learning, and remembering the most frequent and energetic co-occurrences of their inputs; and (3) the learning of the dorsoventral organization of grid cell modules with multiple spatial scales that occur in the pathway from the medial entorhinal cortex to hippocampus seems to use mechanisms that are homologous to those for adaptively timed temporal learning that occur in the pathway from the lateral entorhinal cortex to hippocampus (Grossberg and Merrill 1989, 1992; Grossberg and Schmajuk 1989). The homologous mechanisms for representing space and time in this entorhinal-hippocampal system has led to the phrase "neural relativity" for this parsimonious design.

Finally, the GridPlaceMap model is an ART system. It proposes how top-down hippocampus-to-entorhinal attentional mechanisms may stabilize map learning and thereby simulates how hippocampal inactivation may disrupt grid cell properties and explains challenging data about theta, beta, and gamma oscillations.

Visual and path integration information cooperate during navigation. Cognitive planning also influences navigational decisions. More research is needed to show how learning fuses visual, path integration, and planning circuits into a unified navigational system. The design of a general planning system will be facilitated by the fact that similar circuits for short-term storage of event sequences (working memory) and for learning of sequential plans are used by the brain to control linguistic, spatial, and motor behaviors (Grossberg and Pearson 2008; Silver, Grossberg, Bullock, Histed, and Miller 2011).

Social Cognition

How can multiple autonomous systems interact intelligently? Individuals experience the world from self-centered perspectives. What we learn from each other is thus computed in different coordinates within our minds. How do we bridge these diverse coordinates? A model of social cognition that explains how a teacher can instruct a learner who experiences the world from a different perspective can be used to enable a single human or robotic teacher to instruct a large "class" of embodied robots that all experience the teacher from different perspectives.

Piaget's *circular reaction* notes the feedback loop between the eye and hand in the learning infant, laying the foundation for visually guided reaching. Similarly, feedback between babbled sounds and hearing forms the learned substrate of language production. These *intra*personal circular reactions were extended to *inter*personal circular reactions within the Circular Reactions for Imitative Behavior (CRIB) model (Grossberg and Vladusich 2010). This model shows how social cognition builds upon ARTSCAN mechanisms. These mechanisms clarify how an infant learns how to share joint attention with adult teachers and to follow their gaze toward valued goal objects. The infant also needs to be capable of view-invariant object learning and recognition whereby it can carry out goal-directed behaviors, such as the use of tools, using different object views than the ones that its teachers use. Such capabilities are often attributed to *mirror neurons*. This attribution does not, however, explain the brain processes whereby these competences arise. CRIB proposes how intrapersonal circular reactions create a foundation for interpersonal circular reactions when infants and other learners interact with external teachers in space. Both types of circular reactions involve learned coordinate transformations between body-centered arm movement commands and retinotopic visual feedback, and coordination of processes within and between the What and Where cortical processing streams. Specific breakdowns of model processes generate formal symptoms similar to clinical symptoms of autism.

Mental Disorders and Homeostatic Plasticity

Optimally functioning autonomous intelligent systems require properly balanced complementary systems. What happens when they become imbalanced? In humans, they can experience mental disorders.

Scientific literature on human mental disorders such as autism and schizophrenia is, of necessity, more anecdotal than parametric and is, therefore, an insufficient foundation for model construction. Real-time models of normal mental behavior that are based on the huge databases from decades of psychological and neurobiological experiments have, however, provided insights into the mechanisms of abnormal behaviors (e.g., Carpenter and Grossberg 1993; Grossberg 1984, 2000a, b; Grossberg and Seidman 2006).

Imbalanced processes across the complementary systems that control normal behaviors can produce constellations of model symptoms that strikingly resemble mental disorders. For example, fixing the ART vigilance parameter ρ at too high a level leads to symptoms familiar in autistic individuals, notably learning of hyperconcrete categories and difficulty paying attention to the meaning of a task. Underarousal of the model amygdala can lead to insensitivity to social meanings and also to intense emotional outbursts and coping strategies to reduce event complexity and unexpectedness. Damage to the model cerebellum can lead to defects of adaptively timed learning and thus a host of problems in socialization.

In both humans and robots, it remains an open problem to model how biologically based autonomous systems can discover and maintain their own optimal operating parameters in response to the challenges of an unpredictable world. An initial step toward solving this *homeostatic plasticity* problem was made in Chandler and Grossberg (2012).

Machine Consciousness?

An early ART prediction is that *all conscious states are resonant states,* though not all resonant states are conscious. Since that time, ART has predicted how specific resonances support different kinds of consciousness. These observations suggest the question: can machines that embody ART resonant dynamics experience a type of consciousness? For example, ART models predict that *surface-shroud resonances* subserve conscious percepts of visual qualia, *feature-category resonances* subserve recognition of familiar objects and scenes, *spectral-shroud resonances* subserve conscious percepts of auditory streams, spectral-pitch-and-timbe resonances subserve conscious recognition of auditory streams, *item-list resonances* subserve conscious percepts of speech and language, and *cognitive-emotional resonances* subserve conscious feelings and knowing the objects or events that cause them. ART models also identify the brain regions and interactions that would support these resonances.

These results about model correlates of consciousness emerge from ART analyses of the mechanistic relationships among processes of Consciousness, Learning, Expectation, Attention, Resonance, and Synchrony (the CLEARS processes). Recall, however, that not all resonant states are conscious states. For example, entorhinal-hippocampal resonances are predicted to dynamically stabilize the learning of entorhinal grid cells and hippocampal place cells, and parietal-prefrontal resonances are predicted to trigger the selective opening of basal ganglia gates to enable the read-out of context-appropriate actions. Grossberg (2013, 2016) reviews these and other aspects of ART as a cognitive and neural theory.

Recommended Reading

Amis GP, Carpenter GA (2010) Self-supervised ARTMAP. Neural Netw 23:265–282

Baloch AA, Grossberg S (1997) A neural model of high-level motion processing: line motion and formotion dynamics. Vis Res 37:3037–3059

Baloch AA, Grossberg S, Mingolla E, Nogueira CAM (1999) A neural model of first-order and second-order motion perception and magnocellular dynamics. J Opt Soc Am A 16:953–978

Berzhanskaya J, Grossberg S, Mingolla E (2007) Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. Spat Vis 20:337–395

Brown J, Bullock D, Grossberg S (1999) How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. J Neurosci 19:10502–10511

Brown JW, Bullock D, Grossberg S (2004) How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. Neural Netw 17:471–510

Browning A, Grossberg S, Mingolla M (2009a) A neural model of how the brain computes heading from optic flow in realistic scenes. Cogn Psychol 59:320–356

Browning A, Grossberg S, Mingolla M (2009b) Cortical dynamics of navigation and steering in natural scenes: motion-based object segmentation, heading, and obstacle avoidance. Neural Netw 22:1383–1398

Bullock D, Grossberg S (1988) Neural dynamics of planned arm movements: emergent invariants and speed-accuracy properties during trajectory formation. Psychol Rev 95:49–90

Bullock D, Grossberg S (1991) Adaptive neural networks for control of movement trajectories invariant under speed and force rescaling. Hum Mov Sci 10:3–53

Bullock D, Cisek P, Grossberg S (1998) Cortical networks for control of voluntary arm movements under variable force conditions. Cereb Cortex 8:48–62

Cao Y, Grossberg S (2005) A laminar cortical model of stereopsis and 3D surface perception: closure and da Vinci stereopsis. Spat Vis 18:515–578

Cao Y, Grossberg S (2012) Stereopsis and 3D surface perception by spiking neurons in laminar cortical circuits: a method of converting neural rate models into spiking models. Neural Netw 26:75–98

Cao Y, Grossberg S, Markowitz J (2011) How does the brain rapidly learn and reorganize view- and positionally-invariant object representations in inferior temporal cortex? Neural Netw 24:1050–1061

Carpenter GA (1994) A distributed outstar network for spatial pattern learning. Neural Netw 7:159–168

Carpenter GA (1997) Distributed learning, recognition, and prediction by ART and ARTMAP neural networks. Neural Netw 10:1473–1494

Carpenter GA (2001) Neural network models of learning and memory: leading questions and an emerging framework. Trends Cogn Sci 5:114–118

Carpenter GA, Gaddam SC (2010) Biased ART: a neural architecture that shifts attention toward previously disregarded features following an incorrect prediction. Neural Netw 23:435–451

Carpenter GA, Grossberg S (1987) A massively parallel architecture for a self-organizing neural pattern recognition machine. Comput Vis Graph Image Process 37:54–115

Carpenter GA, Grossberg S (1990) ART 3: hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. Neural Netw 4: 129–152

Carpenter G, Grossberg S (1993) Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. Trends Neurosci 16:131–137

Carpenter GA, Markuzon N (1998) ARTMAP-IC and medical diagnosis: instance counting and inconsistent cases. Neural Netw 11:323–336

Carpenter GA, Grossberg S, Reynolds JH (1991a) ARTMAP: supervised real-time learning and classification of nonstationary data by a self-organizing neural network. Neural Netw 4:565–588

Carpenter GA, Grossberg S, Rosen DB (1991b) Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system. Neural Netw 4:759–771

Carpenter GA, Grossberg S, Markuzon N, Reynolds JH, Rosen DB (1992) Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps. IEEE Trans Neural Netw 3:698–713

Carpenter GA, Martens S, Ogas OJ (2005) Self-organizing information fusion and hierarchical knowledge discovery: a new framework using

ARTMAP neural networks. Neural Netw 18:287–295
Chandler B, Grossberg S (2012) Joining distributed pattern processing and homeostatic plasticity in recurrent on-center off-surround shunting networks: noise, saturation, short-term memory, synaptic scaling, and BDNF. Neural Netw 25:21–29
Chang H-C, Grossberg S, Cao Y (2014) Where's Waldo? How perceptual cognitive, and emotional brain processes cooperate during learning to categorize and find desired objects in a cluttered scene. Front Integr Neurosci doi:10.3389/fnint.2014.0043
Chapelle O, Schölkopf B, Zien A (eds) (2006) Semi-supervised learning. MIT, Cambridge
Contreras-Vidal JL, Grossberg S, Bullock D (1997) A neural model of cerebellar learning for arm movement control: cortico-spino-cerebellar dynamics. Learn Mem 3:475–502
Dranias M, Grossberg S, Bullock D (2008) Dopaminergic and non-dopaminergic value systems in conditioning and outcome-specific revaluation. Brain Res 1238:239–287
Elder D, Grossberg S, Mingolla M (2009) A neural model of visually guided steering, obstacle avoidance, and route selection. J Exp Psychol Hum Percept Perform 35:1501–1531
Fang L, Grossberg S (2009) From stereogram to surface: how the brain sees the world in depth. Spat Vis 22:45–82
Fazl A, Grossberg S, Mingolla E (2009) View-invariant object category learning, recognition, and search: how spatial and object attention are coordinated using surface-based attentional shrouds. Cogn Psychol 58:1–48
Foley NC, Grossberg S, Mingolla E (2012) Neural dynamics of object-based multifocal visual spatial attention and priming: object cueing, useful-field-of-view, and crowding. Cognitive Psychology 65:77–117
Gancarz G, Grossberg G (1999) A neural model of the saccadic eye movement control explains task-specific adaptation. Vis Res 39:3123–3143
Grossberg S (1984) Some psychophysiological and pharmacological correlates of a developmental, cognitive, and motivational theory. In: Karrer R, Cohen J, Tueting P (eds) Brain and information: event related potential. New York Academy of Sciences, New York, pp 58–142.
Grossberg S (2000a) The complementary brain: unifying brain dynamics and modularity. Trends Cogn Sci 4:233–246
Grossberg S (2000b) The imbalanced brain: from normal behavior to schizophrenia. Biol Psychiatry 48:81–98
Grossberg S (2013) Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing World. Neural Netw 37:1–47
Grossberg, S. (2016). Towards solving the hard problem of consciousness: the varieties of brain resonances and the conscious experiences that they support. Submitted for publication
Grossberg S, Huang T-R (2009) ARTSCENE: a neural system for natural scene classification. J Vis 9(6):1–19
Grossberg S, Merrill JWL (1992) A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. Cogn Brain Res 1:3–38
Grossberg S, Merrill JWL (1996) The hippocampus and cerebellum in adaptively timed learning, recognition, and movement. J Cogn Neurosci 8:257–277
Grossberg S, Pearson L (2008) Laminar cortical dynamics of cognitive and motor working memory, sequence learning and performance: toward a unified theory of how the cerebral cortex works. Psychol Rev 115:677–732
Grossberg S, Pilly PK (2012) How entorhinal grid cells may learn multiple spatial scales from a dorsoventral gradient of cell response rates in a self-organizing map. PLoS Comput Biol 8(10):31002648. doi:10.1371/journal.pcbi.1002648
Grossberg S, Pilly PK (2014) Coordinated learning of grid cell and place cell spatial and temporal properties: multiple scales, attention, and oscillations. Philos Trans R Soc B 369:20120524
Grossberg S, Rudd ME (1992) Cortical dynamics of visual motion perception: short-range and long-range apparent motion (with Rudd ME). Psychol Rev 99:78–121
Grossberg S, Schmajuk NA (1989) Neural dynamics of adaptive timing and temporal discrimination during associative learning. Neural Netw 2:79–102
Grossberg S, Seidman D (2006) Neural dynamics of autistic behaviors: cognitive, emotional, and timing substrates. Psychol Rev 113:483–525
Grossberg S, Swaminathan G (2004) A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: development, attention and bistability. Vis Res 44:1147–1187
Grossberg S, Vladusich T (2010) How do children learn to follow gaze, share joint attention, imitate their teachers, and use tools during social interactions? Neural Netw 23:940–965
Grossberg S, Leveille J, Versace M (2011) How do object reference frames and motion vector decomposition emerge in laminar cortical circuits? Atten Percept Psychophys 73:1147–1170
Grossberg S, Kuhlmann L, Mingolla E (2007) A neural model of 3D shape-from-texture: multiple-scale filtering, boundary grouping, and surface filling-in. Vis Res 47:634–672
Grossberg S, Mingolla E, Viswanathan L (2001) Neural dynamics of motion integration and segmentation within and across apertures. Vis Res 41:2521–2553
Grossberg S, Srihasam K, Bullock D (2012) Neural dynamics of saccadic and smooth pursuit eye movement coordination during visual tracking of unpredictably moving targets. Neural Netw 27:1–20

Huang T-R, Grossberg S (2010) Cortical dynamics of contextually cued attentive visual learning and search: spatial and object evidence accumulation. Psychol Rev 117:1080–1112

Hurvich LM, Jameson D (1957) An opponent-process theory of color vision. Psychol Rev 64:384–390

Kelly FJ, Grossberg S (2000) Neural dynamics of 3-D surface perception: figure-ground separation and lightness perception. Percept Psychophys 62:1596–1619

Mhatre H, Gorchetchnikov A, Grossberg S (2012) Grid cell hexagonal patterns formed by fast self-organized learning within entorhinal cortex. Hippocampus 22:320–334

Pack C, Grossberg S, Mingolla E (2001) A neural model of smooth pursuit control and motion perception by cortical area MST. J Cogn Neurosci 13:102–120

Pilly PK, Grossberg S (2012) How do spatial learning and memory occur in the brain? Coordinated learning of entorhinal grid cells and hippocampal place cells. J Cogn Neurosci 24:1031–1054

Pilly PK, Grossberg S (2014) How does the modular organization of entorhinal grid cells develop? Front Hum Neurosci. doi:10.3389/fnhum.2014.0037

Schiller PH (1982) Central connections of the retinal ON and OFF pathways. Nature 297:580–583

Simone G, Farina A, Morabito FC, Serpico SB, Bruzzone L (2002) Image fusion techniques for remote sensing applications. Inf Fusion 3:3–15

Srihasam K, Bullock D, Grossberg S (2009) Target selection by frontal cortex during coordinated saccadic and smooth pursuit eye movements. J Cogn Neurosci 21:1611–1627