The Computational Manifold Approach to Consciousness and Symbolic Processing in the Cerebral Cortex

Douglas S. Greer
General Manifolds

***Abstract*—A new abstract model of computation, the computational manifold, provides a framework for approaching the problems of consciousness, awareness, cognition and symbolic processing in the central nervous system. Physical properties involving space, time and frequency, such as the surface of the skin, sound spectrograms, visual images, and the muscle cross-sections are included in the state of manifold automata. The Brodmann areas are modeled as recurrent image associations implemented as neurotransmitter fields. Symbols are defined by state transition behavior near reciprocal-image attractors in dynamical systems. Control masks overlay the images and regulate awareness of the environment and cognitive reflection.**

***Index Terms*—natural intelligence, consciousness, perception, cognition, models of computation.**

I. Introduction

Time, space, and frequency all have the topology of a manifold. Consequently, these mathematical structures, the curves, surfaces, and solids that describe the characteristics and actions of plant and animal life, form the proper foundation for a formal definition of natural intelligence. We discuss a new approach to cognition where the inputs, outputs and internal state are expressed solely in terms of relationships between functions defined on continuous manifolds. This approach contrasts with discrete mathematics and many traditional techniques used in computer science, where real-world objects are forced to fit into a framework of finite sets. Rather than axiomatically assuming the existence of symbols, this new approach permits their construction and expression in a way that is consistent with everyday experience.

Physical quantities such as the intensity of light impinging on the surface of the retina, the air pressure on the eardrum over time, the position and temperature on the skin surface and the forces over entire muscle cross-sections, are all functions defined on manifolds. We refer to these functions as *images* or *fields*.

Manuscript received March 3, 2008. This document contains material that is patent pending by General Manifolds LLC, http://www.gmanif.com/ip.

D. S. Greer is with General Manifolds, Carlsbad, CA 92009 USA. (phone: +1-760-804-1615; fax: +1-760-431-0245).

A common assumption is that for cognition and symbolic processing to take place, "low-level" images must be converted into some form of "high-level" discrete representation. We propose that this need not be the case; a complete cognitive system can be based solely on the image data, without the need to reduce the dimension of the data or extract features. Following the common practice in computer science, this proposition can be divided into two parts: a practical mechanism for the construction of the machine, and a theoretical definition of its operation. We address both of these issues.

Manifolds are the basis of differential geometry where they are defined as spaces which are locally homeomorphic to $\mathbb{R}^n$ [1]. For example, curved lines are one-dimensional manifolds and smooth surfaces are two-dimensional manifolds. We present a formal definition of *Computational Manifold Automata* (CMA) as an abstract model of computation. The definition is similar to the specification of finite state machines [2], formal language theory [3] or Turing Machines [4] but with the additional mathematical structure of continuous topologies. Within this framework it is possible to unify the immediate awareness of perception and motor control with the reflective processing of language and cognition.

The memory capacity of an abstract Turing machine is unlimited, and consequently, these are often referred to as infinite automata. However, Turing machines have a discrete topology – like that of the integers, which are infinite but countable. In terms of measure theory, the set of all computable numbers is recursively enumerable (countable) and therefore has measure zero. Manifold automata on the other hand are based on sets with non-zero Borel measure, such as intervals of real numbers $\{x: a < x < b; x \in \mathbb{R}\}$.

The state of a CMA is a collection of images. If we consider a display screen as a manifold, the current state is the image currently displayed on the screen. However, in contrast to many definitions of automata, the image or state *is not* a symbol. Symbols may be defined as sets or as an attractor in a dynamical system.

Analysis of a digital logic circuit may help to illustrate the distinction between state and symbol. If the supply voltage V_{dd} for an integrated circuit is 1.0 volt, we may designate voltages between 0.0 and 0.2 volts as "zero" and those between 0.8 and

1.0 volts as "one". Voltage levels between 0.2 and 0.8 are allowed only temporarily when "switching" between the valid ranges. In the case of computational manifold automata, the "symbol" is a set of real numbers (that is a *range* of voltages) whereas the "state" is a single real number (that is a specific voltage level on the output). This comparison is similar to viewing the resistors and transistors of a digital circuit as an analog circuit where the voltage level is the state.

In a finite state automata, the voltage levels are abstracted away and the state is defined only by the symbols "0" or "1". However, the extension of this concept to images may be intractable. Consequently, the CMA state is expressed in terms of the definitive image values and the symbols are expressed in terms of sets or attractors.

Two NAND or NOR gates can be recursively connected to construct a Set-Reset (SR) flip-flop. This bi-stable memory element can be viewed as a dynamical system with two f*ixed-point attractors* that correspond to "0" and "1". The set of all values that converge toward a particular attractor is referred to as an *attractor basin* [5].

We construct a dynamical system analogous to an SR flip-flop where the bits are replaced by images and the logic gates are replaced by feed-forward image association processors. The convergence of the system toward a *reciprocal-image attractor* is then demonstrated.

Physics & Biology

In the title of his book "Chance and Necessity", Monad [6] refers to the random genetic mutations that happen by chance, and the necessity of conforming to the laws of physics and chemistry that remain constant over time. This relationship has several implications in neuroscience.

The conservation of linear and angular momentum implies that space is homogeneous and anisotropic [7], that is, physical space has the structure of Euclidian $\mathbb{R}^3$ for translations and rotations. Errors that result from its discretization will tend to be minimized by natural selection, and consequently, manifold automata is often the appropriate level of computational abstraction.

Because of metabolic advantages such as reduced size and lower energy consumption, storing information at the molecular level rather than the cellular level may be advantageous. This hypothesis leads to the theory of neurotransmitter fields [8] which describes how information is encoded in the density of neurotransmitters in the synaptic cleft rather than neuron action potentials. We briefly review the main concepts of this theory and show how this can serve as an implementation mechanism for computational manifold automata in natural intelligence.

Before formally defining CMA and demonstrating the use of image attractors in symbolic processing, we summarize the environmental, physiological and developmental topologies that are part of a unification of sensory processing and consciousness.

II. Perception and Motor Control

The phenomena sometimes referred to as the perceptual constancy of a stable visual world [9] is consistent with the notion that information about the physical matter currently surrounding an organism is so important to its survival, that the nervous system will construct a matching internal isomorphic model. The relationship between this model and the visual image projected on the retina is illustrated in Fig. 1. The right hand side of the diagram depicts an *internal* representation of the external physical world. The required transformation, which is between functions defined on manifolds, maps from the image on the surface of the retina to the model of three-dimensional physical space. This transformation is the shape extraction problem of computer vision. It has been described as the inverse of computer graphics, which transforms three-dimensional spatial representations to a two-dimensional visual image.

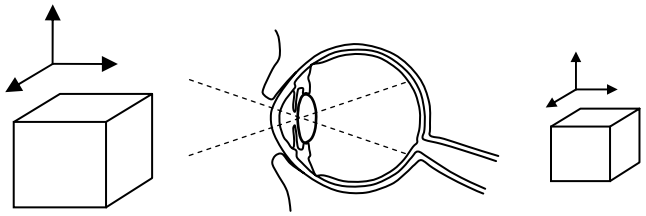

Fig. 1. The extraction of shape information from the visual image, is a transformation from a function defined on the retina, a two-dimensional manifold, to an internal representation of matter in the surrounding space, which is a function defined on a three-dimensional manifold.

Tonotopic maps, which map audio frequencies to locations in the brain, have been documented and studied for some time [10]. Some of these resemble audio spectrograms which plot time in the horizontal direction and frequency in the vertical direction. Analogous auditory maps inherit a product topology [11] from the two manifolds, time and frequency. In the computational manifold approach, a child learning new nouns is learning associations between visual images and audio images.

Kinematic models in robotics often use a few real-valued variables to specify the articulated joint angles and control signals for the actuators. In contrast, a myology diagram of human anatomy shows how muscle fibers connect along considerable sections of the bone. Fibers within the same muscle may pull the limb in different directions. This is illustrated in Fig. 2, where two nearby fibers (*a* and *b*) exert forces that are approximately equal (F_a and F_b), while a third fiber (*c*) pulls the limb in a separate direction (F_c). Consequently, the two-dimensional image of the muscle cross-section smoothly maps to a manifold in a tangent bundle that describes the forces exerted on the limb. Images on the surface of the cross-section can be used to describe the efferent nerve signals that control the contraction of the muscle fibers as well as the afferent signals that convey the fiber length and tension [12].

Complex and coordinated motor control is possible without the presence of skeletal joints. Common, well known examples might include the tentacles of an octopus or the tongue of a reptile. Arthrology, the physiological study of articulated joints, reveals a wide variety of bone attachment mechanisms and even the synovial joints, commonly modeled as simple hinges or ball-in-socket connections, usually involve

surfaces with complex sliding motions [13]. The standard models employed in robotics are often too simplistic to describe these motions. However, all types and varieties of articulated and non-articulated movements can be uniformly described in a computational manifold framework. The model of a finger shown in Fig. 3 illustrates the corresponding transformations.

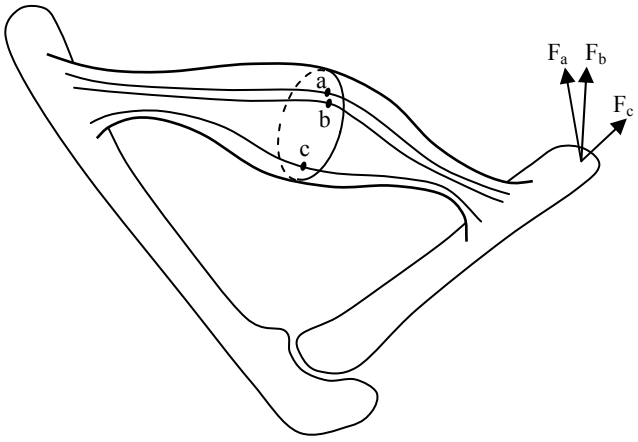

Fig. 2. Neighboring points in the muscle cross-section map to neighboring force vectors in a tangent bundle. Consequently, there exists a differentiable map between the images representing the muscle control and feedback signals and the manifolds describing the resulting forces and the motion of the limb in space.

The outside of a finger is shown in Fig. 3(a) with a path along the surface of the skin between two points, labeled p_1 and p_2. The interphalangeal joints $\mathbf{J_1}$ and $\mathbf{J_2}$ represent the idealized center of rotation. The surface of the skin as it might appear from a viewpoint inside the finger is shown in Fig 3(b) with the same path from p_1 to p_2. If the articulation was a "smoothly bending" cartilaginous joint we could calculate the spatial position of any point on the path using integrals of the principal curvatures along the surface. However, in the case of idealized joints the changes in surface orientation would become sudden jumps and where the curvature is undefined. These discontinuous regions are represented in Fig 3(b) by the two rings labeled $\mathbf{J_1}$ and $\mathbf{J_2}$. Using the more general Lebesgue integrals rather than Riemann integrals, the same equations apply in all cases by allowing the inclusion of measures similar to Dirac delta functions in the curvature [12].

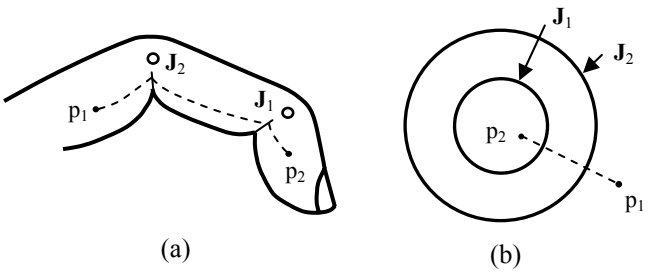

Fig. 3. The two-dimensional image defining the surface of the skin viewed from "inside" of the finger (*b*). The spatial position can be calculated in a uniform manner for both idealized and biologically realistic joints using Lebesgue integrals.

The olfactory bulb responds to odors with a spatial map generated by a distributed assembly of specific and general molecular receptors. As a result, different mixtures of odor molecules produce unique odor maps [14]. Moreover, there are general similarities between the retina and olfactory bulb in the cellular circuits producing lateral inhibition and receptive fields, which indicate conserved mechanisms of neural processing in vision and olfaction [15]. This is consistent with an image association model of sensory information processing in a frequency space such as one generated by the continuous wavelet transform [16].

III. EMBRYOLOGY

The computational manifold approach to describing natural intelligence is consistent with the relationships established during embryonic development. The topological structure of the germ layers can be visualized in an idealized diagram such as Fig. 4. During the first few days of development, the notochord induces the formation of the neural tube from the ectoderm, and the neural crest cells split off creating the innervation by the peripheral nervous systems [17].

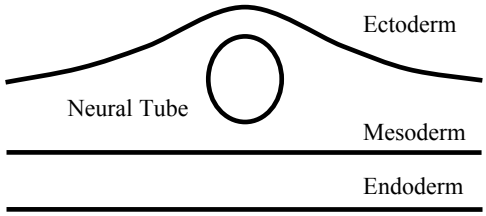

Fig. 4. The topological relationships created during the embryonic development of vertebrates are shown in a transverse section. The precursors of the intestines (*endoderm*), bones and musculature (*mesoderm*) and skin (*ectoderm*) have a manifold structure which is connected to the central nervous system (*neural tube*) by the peripheral nervous system.

During early development, the interface between the neural tube and the mesoderm can be represented as a mapping between two connected spaces. As the distinct muscles form, the space representing the mesoderm becomes disconnected but is still a manifold since every point is still locally homeomorphic to $\mathbb{R}^n$. Each muscle cross-section is topologically connected internally, but is disconnected from the other cross-sections. This can be conceptualized as a surface covered with "spots" where each spot is the cross-section of a single muscle. Consequently, the mapping between forces and muscle cross-sections illustrated in Fig. 2 can be extended to the entire musculature and viewed as a single transformation between two images. During development, the peripheral nervous system maintains the relationship between the two surfaces.

The ectoderm shown in Fig. 4 is the precursor of the skin surface and is also mapped to the neural tube by the peripheral nervous system. For the portions that are controllable (in terms of engineering control theory), there is a unique isomorphism that characterizes the mapping between the spatial position of the body surface and the lengths of the muscle fibers. The neurological representation of the position of the body surface can be expressed as a two-dimensional manifold embedded in the internal representation of three-dimensional space shown on the right side of Fig. 1. Moreover, once the position of this surface is calculated, the

forces on the surface of the skin can be mapped to the tension and forces on the combined cross-sections of the muscles.

The product topologies of these surfaces over time provide a framework for motion planning and collision avoidance. Consequently, "hand-eye" coordination, and all forms of perception and movement can be completely specified as transformations between images.

The endoderm layer shown in Fig. 4 gives rise to the gastrointestinal tract and major portions of the respiratory system. The mapping between this manifold and the neural tube defines an image representing the physiological awareness of the viscera, chemical properties of ingested material, and internal sensation.

IV. Computational Manifolds

Knowledge of the density and momentum of matter in the external world at the present moment is sufficient to predict its location a short time in the future using Newtonian physics. Differential geometry removes the need to specify a particular coordinate system in the expression of the intrinsic properties of the relationships. The correct result of this prediction is also independent of the sampling resolution or the data representation used in the calculations. The same can be said about many psychological experiments that measure spatial reasoning and the ability to transform mental images [18]. This suggests the need for an abstract description that does not require discretization of the data. This formalism can also be applied to help separate the state, which is a collection of images, from the symbols, which are inextricably linked to the system behavior.

A metric automata is a type of topological automata [19] where the set of possible states is a metric space. However, in computational manifold automata, the state is not a point in the manifold but rather a function defined over the entire manifold. That is, the metric which defines the automata is the L^2 metric defined on the function space rather than the metric of the underlying manifold.

As an informal metaphor, we can imagine a room with a collection of flat-screen displays which are labeled 1 through *N*. Each display screen is a flat surface, i.e. a rectangular two-dimensional manifold, which we label $\mathbb{M}_i$. The visual image currently being displayed on a screen can be represented by a function $f_i(x,y)$ where x and y parameterize the horizontal and vertical dimensions of the screen. If the square of this function integrated over the surface of the screen is less than infinity, we have $f_i \in L^2(\mathbb{M}_i)$. The entire set of images $\{f_i : 1 \le i \le N\}$ currently being displayed on all screens is the system *state*.

Additional screens labeled $\mathbb{J}_i$ may serve as inputs where the data is fed in from external video cameras. Images on screens labeled $\mathbb{K}_i$ serve as outputs. Except for the externally generated video inputs, we assume the images are synthetically generated using arbitrary computer graphics or image processing algorithms.

At the core of an automata is a state transition function which determines the next state based on the current state and the inputs. The *current state* is the collection of images currently being displayed on the flat-panel screens. The *next state* is also a set of images which are uniquely defined by the current-state images and the input images. Since the inputs, outputs and current state are functions, we refer to the mapping between the current state and the next state as the *next state transformation* $\mathcal{Z}$. Similarly we define the mapping that generates the functions defined on the output manifolds as the *output transformation* $\mathcal{T}$.

Asynchronous state transformation functions are possible, but to simplify the discussion, we will limit ourselves to synchronous transitions analogous to those produced by a digital system clock. In the case of flat-panel display screens, this is the frame refresh rate. This rate may be considered independently of any discretization within the data images.

Formally, let:

$$\begin{aligned} &\mathbb{M}_i : \ 1 \le i \le N_1 \ \text{ be the internal state manifolds,} \\ &\mathbb{J}_i \ : \ 1 \le i \le N_2 \ \text{ be the input manifolds, and} \\ &\mathbb{K}_i \ : \ 1 \le i \le N_3 \ \text{ be the output manifolds.} \end{aligned} \tag{1}$$

The images or functions defined on these manifolds are:

$$\begin{aligned} &f_i \in \mathbf{L^2}(\mathbb{M}_i) : \ 1 \le i \le N_1, \quad \text{the state images,} \\ &g_i \in \mathbf{L^2}(\mathbb{J}_i) \ : \ 1 \le i \le N_2, \quad \text{the input images, and} \\ &h_i \in \mathbf{L^2}(\mathbb{K}_i) \ : \ 1 \le i \le N_3, \quad \text{the output images.} \end{aligned} \tag{2}$$

At time *t*, the complete input, output, and state are formed by creating a vector of the combined images:

$$\begin{aligned} \mathbf{f}(t) &= (f_1, f_2, \ldots, f_{N1}) \ \in \mathbf{F} \ = \prod_{i=1}^{N_1} \mathbf{L^2}(\mathbb{M}_i) \\ \mathbf{g}(t) &= (g_1, g_2, \ldots, g_{N2}) \ \in \mathbf{G} \ = \prod_{i=1}^{N_2} \mathbf{L^2}(\mathbb{J}_i) \\ \mathbf{h}(t) &= (h_1, h_2, \ldots, h_{N3}) \ \in \mathbf{H} \ = \prod_{i=1}^{N_3} \mathbf{L^2}(\mathbb{K}_i) \end{aligned} \tag{3}$$

Let Δt be the time interval for one cycle of a synchronous system clock. The next state transformation $\mathcal{Z}$ is a function of the inputs and the current state:

$$\mathbf{f}(t + \Delta t) = \mathcal{Z}\big(\mathbf{f}(t), \mathbf{g}(t)\big) \tag{4}$$

In the case of a discrete automata, the outputs of a *Mealy* machine may depend on the current inputs whereas those of a *Moore* machine depend only on the current state [2]. We generalize the more general Mealy machine with an output transform $\mathcal{T}$:

$$\mathbf{h}(t) = \mathcal{T}\big(\mathbf{f}(t), \mathbf{g}(t)\big) \tag{5}$$

Let the initial or starting state be denoted as $\mathbf{f}_I \in \mathbf{F}$.

We can now define a *Computational Manifold Automata* $\mathbf{\Omega}$ as the 6-tuple:

$$\mathbf{\Omega} = (\mathbf{f}_I, \mathbf{F}, \mathbf{G}, \mathbf{H}, \mathcal{Z}, \mathcal{T}) \tag{6}$$

The functionality of any discrete automata can be replicated using the state variables in a CMA or in terms of the symbols described below. However, compared to the computational capabilities of a discrete state machine, $\mathbf{\Omega}$ has a more powerful mathematical structure. Each component of the function spaces $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{H}$ is defined on a manifold which locally has continuous bijections to neighborhoods of $\mathbb{R}^n$. These manifold

substrates have a metric topology [11], which in the case of $\mathbb{R}^1$ is the one induced by the open sets $(a,b) = \{x: a < x < b\}$. The images f_i, g_i, and h_i are themselves elements of an $\mathbf{L}^2$ normed linear space [20], which has an inner product defined by the integral over the manifold of the product of two functions. This enables the construction of generalized functions and spectral operators such as the continuous wavelet transform. Consequently, **F**, **G** and **H** are equipped with robust mathematical mechanisms that may be used to construct the next-state and output transformations $\mathcal{Z}$ and $\mathcal{T}$. For example, spherically symmetric wavelets, analogous to the receptive fields in the retina, can be used to construct reciprocal-image association operators [16].

We do not place restrictions on the transformations $\mathcal{Z}$ and $\mathcal{T}$ in terms of their continuity or differentiability. The discontinuities may be similar to the branching in a processor controller, where the automata may undergo arbitrarily defined, deterministic, state transitions. Various input or state values may cause sudden jumps or "resets" similar to sharp "cuts" in a video sequence.

A process whose characteristics change is time is sometimes referred to as *non-stationary* or *time-varying* [21]. A deterministic, *stationary* system generates the same outputs when presented with the same inputs; that is, the system does not learn. Similar to neural networks, we can define *long-term memory* as changes in the state-transition transformation $\mathcal{Z}$ and the output transformation $\mathcal{T}$. In contrast, *short-term memory* is defined as the current state **f**. Non-stationary behavior, that is changes in long-term memory, can result from the formation of symbols and symbolic associations.

V. Symbolic Processing

In many computer science courses, symbols are either taken for granted, or assumed to exist in axioms. Like addition and multiplication, binary digits are abstract mathematical concepts. If someone were to take apart a typical desktop computer searching for ones and zeros, they might be disappointed to find only analog resistors and transistors. The digital circuit specifications themselves refer only to acceptable ranges of voltages on the inputs and outputs. These specifications are, in effect, a contract with the digital circuit designer that defines a *guaranteed behavior*. This concept of equating symbols with behavior can be extended to computational manifold automata.

The basic unit of memory, the Set-Reset (SR) flip-flop can be constructed from two NAND gates as shown if Fig 5(a). This circuit acts as a dynamical system with two attractors corresponding to "0" and "1". The two binary digits are, in effect, voltage ranges where the circuit conceptually falls into one of two "energy wells". These fixed-point attractors create the stability required for an actual physical realization to operate in the presence of the inevitable noise and transient errors in the inputs.

We define a Λ-*map* (Λ from the Greek word *Logikos*) as a feed-forward process that accepts one or more images as inputs and produces a single associated output image. We can create a circuit analogous to an SR flip-flop by replacing the bits with images and replacing the NAND gates with Λ-maps. The recurrent connections are shown in Fig. 5(b) where the *exterior* and *interior* Λ-maps are labeled Λ_E and Λ_I. We refer to the two outputs as *reciprocal images* and use the term *psymap* to refer to the structure containing two recursively-connected Λ-maps.

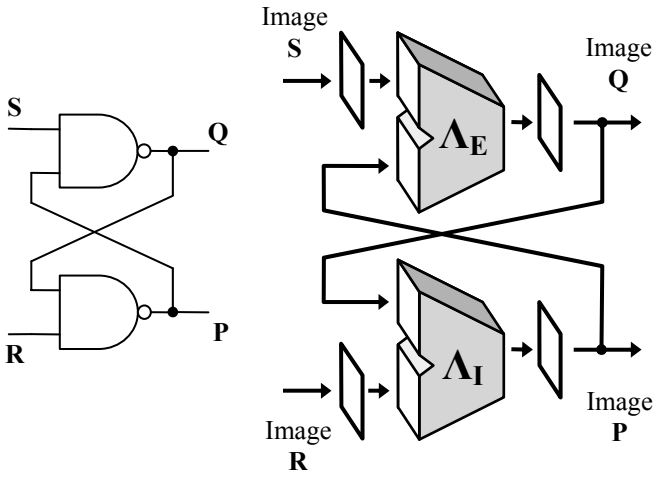

Fig. 5. The fundamental unit of discrete memory, the set-reset flip-flop, is formed from two NAND gates by creating a simple feedback loop (a). The abstract binary values, "Zero" and "One", are actually attractor basins in a dynamical system. A *psymap* (b) is constructed from two Λ-maps (Λ_E and Λ_I). The psymap structure is analogous to an SR flip-flop, where the bits have been replaced by images, and permits the realization of symbols as reciprocal-image attractors.

Rather than having only two attractors for "0" and "1", as is the case for an SR flip-flop, a psymap can have any number of attractors. Each *reciprocal-image attractor* can be constructed from two arbitrarily chosen reciprocal images. Let $q = \Lambda_E(p, s)$ and $p = \Lambda_I(q, r)$ denote the exterior and interior Λ-Maps shown in Fig 5(b) and let *Null* denote a predefined "blank" image. For an arbitrary collection of image pairs (a_i, b_i) we can create a new attractor by adding associations to both Λ-Maps so that $\Lambda_E(b_i, Null) = a_i$ and $\Lambda_I(a_i, Null) = b_i$. In addition, we can define associations for the *S* and *R* inputs – for example $\Lambda_E(X, s_i) = a_i$ where *X* is any image – that allow us to force the psymap to the (a_i, b_i) state.

VI. Brodmann Areas

The human cerebral cortex is divided into approximately fifty Brodmann areas that are shown in Fig. 6 in a diagram first published one hundred years ago [22]. Many of the areas have well documented maps with various sensory modalities. [15][17].

The division into areas is based on the thickness or prominence of six cellular layers, labeled I through VI, which are present throughout all parts of the cortex. Within a single Brodmann area, the layers have the same thickness, but at the boundaries between areas, there is a sudden change in the relative thickness of the six layers. Most of the cellular layers occur in *exterior* / *interior* pairs. Layers II, III, IV and V are known as the *exterior* granule, the *exterior* pyramidal, the *interior* granule and *interior* pyramidal layers respectively. If we group together the internal and external cortical layers and

connect them recursively, the result closely resembles a psymap, where the external layers form the external Λ-map, Λ_E, and the internal layers form the internal Λ-map, Λ_I.

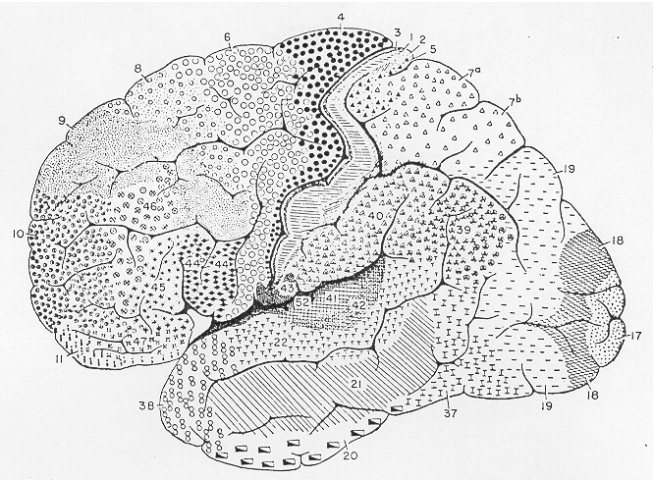

Fig. 6. Lateral view of the human brain divided into areas from the original work by Brodmann. The boundaries between the Brodmann areas are determined by the thickness of the six cellular layers.

VII. The Psymap Array

Fiber tracts that connect different areas of the central nervous system are similar to smooth laminar flows, where points that are close to each other at the source project to points that are close to each other at the destination. Mathematically, the axon fibers preserve the intrinsic topological structure. By constructing "composite overlays" with multiple images as the components, a Λ-map can accept any number of input images [16].

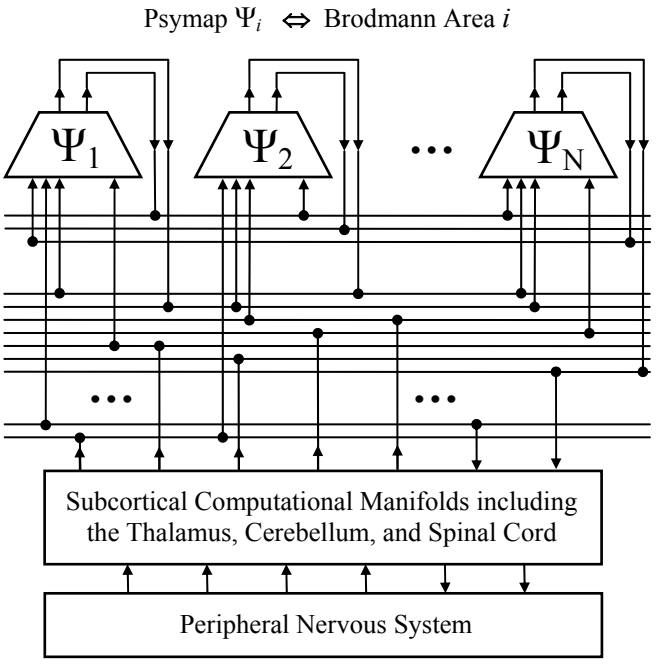

Fig. 7. The psymap array model of the cerebral cortex models each Brodmann area as one psymap Ψ_i. Lines in the diagram correspond to images, and dots represent specific input and output connections between areas. The images in the cortex are connected to other computational manifolds in subcortical regions.

If we identify each Brodmann area with a single psymap, the cerebral cortex can be modeled as a *psymap array* such as the one shown in Fig. 7. Using topological alignment, each psymap Ψ_i can accept input images from several different areas. However, each Brodmann area has two output images (which arise from the internal and external pyramidal layers) and these are shown connected to the other psymaps in Fig. 7 using an "I/O Bus" notation. The lines and dots in the diagram are used to represent the transfer of entire images between the Brodmann areas. Operations in the subcortical areas are also modeled as two or three-dimensional computational manifolds.

The psymap array image association model predicts that the cellular layers within each Brodmann area should be of constant thickness. Moreover, since distinct Brodmann areas have different inputs and store a different number of image associations, there should be a sharp transition in cellular thickness from one area to the next.

VIII. Neurotransmitter Fields

Neurotransmitter fields differ from many neuron models in the assumption that the data values are not encoded in the neuron action potentials, but rather by the chemical density of neurotransmitter molecules in the extracellular space. This model has advantages in terms of size, energy and abstract computing power [8]. The general theory can be developed from the following conjecture.

The Neurotransmitter Cloud Hypothesis: When multicellular fauna first appeared, organisms began to represent quantities such as mass, force, energy and position by the chemical concentration of identifiable molecules in the extracellular space. The basic principles of operation developed during this period still govern the central nervous system today.

In this model, synapses are seen as an incremental evolutionary adaptation that reduces chemical diffusion and the amount of neurotransmitter that must be released and absorbed by the axons. The dendritic trees of neurons perceive the current state of "input" clouds and release neurotransmitters into an "output" cloud via their axonal trees. When each cloud is viewed as a continuum rather than as discrete molecules, the result is a natural implementation of an abstract computational manifold automata.

If we model the external and internal layers of a single Brodmann area as two recurrently connected neurotransmitter fields, the result is the psymap implementation illustrated in Fig. 8. Each Processing Element (PE) in the diagram represents the action of a single neuron. The large majority of synaptic connections in the cortex are between neighboring neurons. The recurrent connections between the layers create cyclical structures analogous to the cortical columns.

The rightmost image, $Q_{+\Delta t}$, is equal to the leftmost image following one iteration of the recurrent computations. Consequently, for this simple case of a computational manifold automata, the composite operation ($\Lambda_E \circ \Lambda_I$) is the next state transformation $\mathcal{Z}$.

The transformation of the images over time can be

described by equations that form an infinite-dimensional dynamical system [23]. Pairs of images are the "points" in this dynamical system, that is, elements in the Cartesian product of two $\mathbf{L}^2$ function spaces. Stable *fixed-point attractors* are the result of creating neighborhoods around the image pairs where similar images converge toward the reciprocal-image attractor.

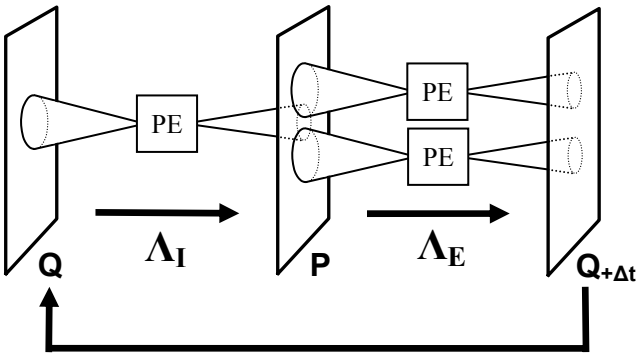

Fig. 8. The Λ-map computations are performed by Processing Elements (PEs) with overlapping support regions. The image data is the concentration of neurotransmitter in the extracellular space of the neuron synapses. Even though the dendritic and axonal trees cover only local limited areas, the recurrence allows their effect to spread over the entire image. The pattern of interconnection closely resembles the cytoarchitecture of the cortical columns.

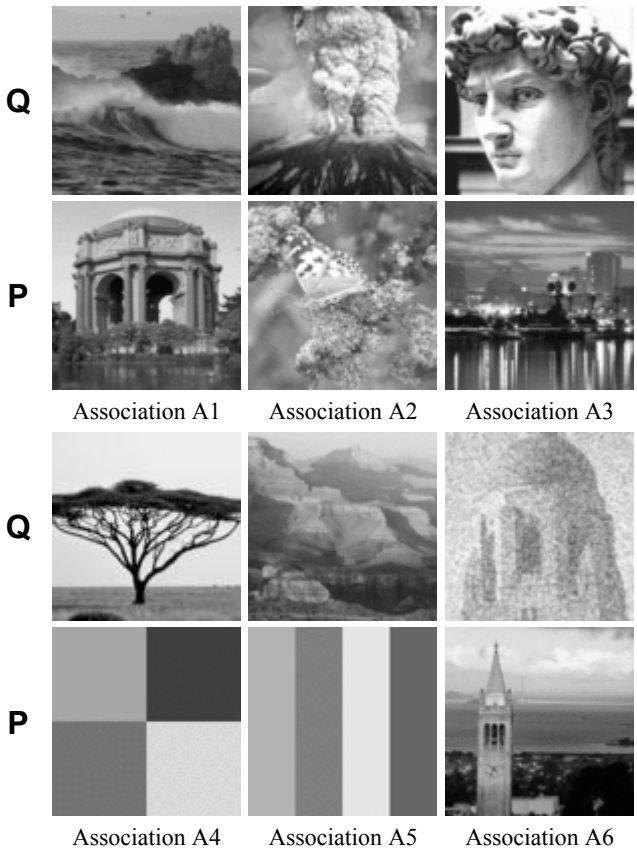

Fig. 9. The six pairs of associated images used to construct a single psymap with six reciprocal-image attractors.

To demonstrate the stability of a psymap constructed from two Λ-map neurotransmitter-fields as shown in Fig 8, a software simulation was written in Java and tested. The system was trained with the six pairs of images illustrated in Fig. 9. The images can be selected arbitrarily, and could be used to store associations representing any subjective cause-and-effect relationship.

The two examples shown in Fig. 10 and Fig. 11 illustrate the convergence of the same psymap to two different stable reciprocal-image attractors. In each example, the initial value is an image a short distance in function space from one of the attractors (A5 or A6 in Fig. 9). The iterated sequence of images shows the image pair moving toward the stable fixed-point. These examples illustrate the distinction between state and symbol. Each image pair represents the current state of the CMA. However, symbols are interpreted as behavior, in particular, convergence towards the image attractors.

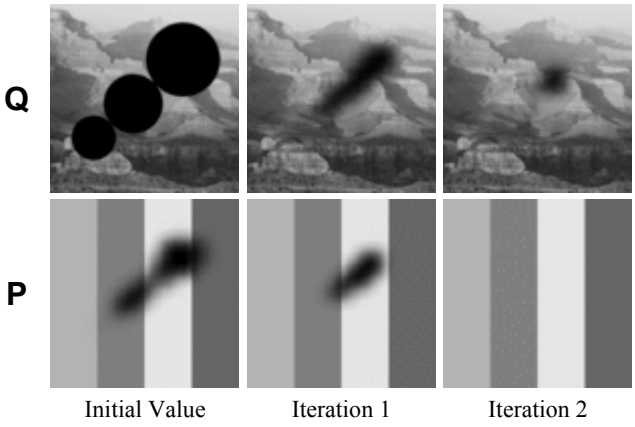

Fig. 10. Results showing convergence to one of the stable attractors (A5). The initial value is the first image P which represents a displacement away from the image, but within its attractor basin. The images constitute the system state, whereas symbols are defined by the system behavior. Since all six reciprocal-image attractors are stable, they can serve as symbols in a cognitive system.

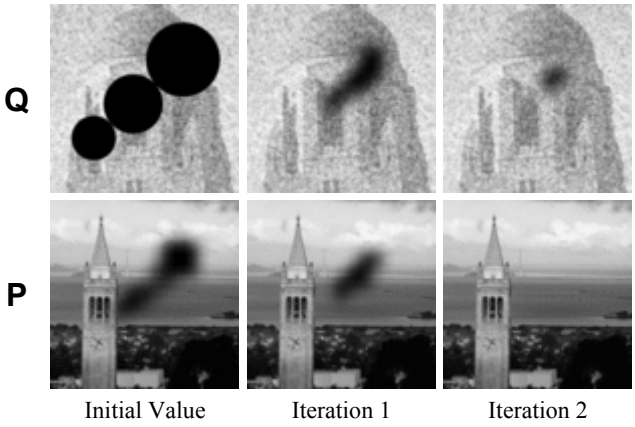

Fig. 11. Starting from a different initial value, the system converges to the stable attractor A6.

The input regions of the processing elements, corresponding to the dendritic trees of the neurons may vary in size and shape. For example, the use of two-dimensional wavelets, similar to the receptive fields in the retina may improve association accuracy and reduce noise [16].

The example images in Fig. 10 are two-dimensional fields representing visual data. However, the same computational principles can be applied to fields of any dimension and the two halves of a reciprocal-image pair may be of heterogeneous types. It is possible that subcortical fields may represent the three-dimensional shape of individuals and objects in the external world.

The CMA approach to recognition or classification does not reduce the dimensionality of the data. Instead, the process is one of moving toward a reciprocal-image attractor. For example, we can view the recognition of unique individuals in the external world as the process of the psymap dynamical system moving toward a unique attractor.

A Λ-map can accept multiple input images by aligning them topologically as shown in Fig. 12. One or more of the input images man serve as a control mechanism by regulating how the other images are processed. Image masks that overlay the multiple input images can be used to control the association formation process and to focus attention on specific regions.

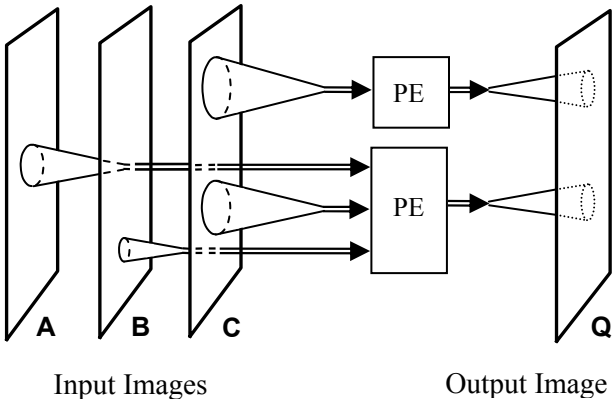

Fig. 12. The internal structure of a single Λ-Map which shows the topological alignment of multiple images. One of the images may serve as a "control mask" that regulates how the other images are combined.

If one of the two reciprocal images in a psymap association were an audio spectrogram, the resulting symbol may be a word.

IX. Language

Words spoken by many different people with a variety of vocal characteristics have the same spelling and meaning. These similar but distinct sounds are all mapped to the same symbol. This process is analogous to the convergence demonstrated in figures 10 and 11.

The symbols can be defined as either *sets* or *attractors*. These definitions are related, but while using sets may be easier initially, it is more difficult to maintain in practice. An attractor basin is like a geographic *catchment region*, the set of all points on the surface of the earth where the water flows into a particular lake or river. This is similar to defining voltage ranges as the ones and zeros in digital logic, that is, the set of all voltages between the minimum and maximum value. While this approach works well for binary values, for manifold automata it can lead to complications.

As the dimension of the state space is increased, chaotic attractors are more likely to be created [5]. The boundaries between these attractors will be fractal, so while the sets corresponding to the attractor basins can be defined in theory, in practice they are impossible to precisely locate.

Imagine a high-resolution photograph of a room containing assorted items. If several people were asked to describe the room, their narratives would have significant differences in the words used and the grammatical constructions. Consequently, it is difficult to define the photograph in terms of sets. However, if we consider words to be CMA attractors in a dynamical system, then we can describe the process as each person having a "visual mask" that changes in shape and moves around the photograph of the room, bringing various items into "focus". The resulting image portion may come close to a reciprocal-image attractor that is associated with a distinct word. The images that form the system state then move toward that attractor, eventually causing that word to be spoken. In this approach, there is no need to define the actual sets. The symbols are defined by the system behavior.

It is possible to convert any grammar from a discrete formal system to a continuous formal system [24]. The psymap array shown in Fig. 7 allows several heterogeneous types of symbols to exist simultaneously. The diversity of the sensory images, such as the computational maps of olfaction or the viscera formed from the endoderm, suggests that the complete range of qualia can be expressed by the value of functions defied on manifolds. Like multiple convoluted valleys in a geographical landscape, each of the image attractors may have complex relationships with each other. Computation manifold automata provide a general framework in which these relationships can be articulated and parsed.

Letters are visual glyphs that are associated with sounds. When we read, we not only understand the meaning of the words, but we instinctively know when the words rhyme. The words may evoke mental images that are then transformed back into other sounds. These cycles are consistent with the psymap array structure shown in Fig 7. In a similar fashion, the flow of information through the cortical and subcortical network of computational manifold processors forms the basis of cognition.

X. Consciousness

The framework of computational manifold automata provides a mathematical model for unifying psychology and neurophysiology. For example the description of short-term memory as the state of a psymap array, or multimodal sensory integration using image alignment as shown in Fig. 12.

Every area of the cerebral cortex has reciprocal connections with the thalamus [25], which is known to play an important role in consciousness. We can hypothesize the involvement of these connections as "control masks" that overlay the psymap images and regulate their operation. Control masks can be used to direct our attention to a small region in the visual field, a single sound in a word, a particular part of the body surface, or the motion of a single muscle.

In the psymap shown in Fig 5(b), both the external and internal Λ-maps, Λ_E and Λ_I, have two input images. The first

latch image is fed back from the reciprocal Λ-map output (*Q* or *P*) and the second *extrinsic image* (*S* or *R*) is fed in from outside. A control mask can be used to focus attention on specific regions of the extrinsic image or ignore it altogether. In the later case, the psymap images would tend to move toward a nearby reciprocal-image attractor and remain there, until the control mask forced it to again focus on outside stimuli. During this time, the psymap output images *Q* and *P* would remain fixed. The control mask may be viewed as similar to an alpha channel, such as those used in computer graphics to control the merging or blending of multiple images [26].

The collection of control masks are themselves images that coordinate the overall psymap array operation. If all of the psymap masks were such that the *Q* and *P* output images were based solely on the *S* and *R* extrinsic input images, the internal state would be a function only of the current sensory inputs. If the control mask gradually or suddenly changed so that the latch feedback images began to be used in calculating the psymap outputs, the presence of various reciprocal-image attractors would then alter the next-state transitions.

In this way, the psymap control masks serve as a mechanism for switching between the immediate awareness of the surrounding environment and internal reflection and thoughtful deliberation. Moreover, the switch need not be an "all-or-nothing" phenomena. Some of the psymaps may change between external awareness and internal reflection while others do not. Various portions of each image may be partially or completely removed from consideration.

Time varying control masks can be used to spatially scan a visual image, or temporally scan through the sounds of a sentence. CMA control image sequences can serve as programs analogous to the instruction sequences in discrete automata. In addition, images from one component of a psymap array can be used to modify the control image sequence in other parts.

The presence of reciprocal-image attractors in connected psymaps can generate a "recognition cascade". In one psymap, a small portion of an image moving near an attractor can generate a sequence of images moving toward that attractor. Its outputs may then cause the state of other psymaps to move toward related attractors. In this way, the slightest "hint" may evoke an intricate and detailed recollection.

Traditional artificial intelligence often uses methodologies, such as first-order predicate calculus, to model the world, and knowledge about it, in ontological graphs. Objects, properties and relationships are defined as discrete symbols. This approach contrasts with computational manifold automata where the current state is a collection of images, and where various portions may at any moment be "in the neighborhood of" an unlimited number of symbols, or none at all. However, many discrete problem-solving techniques can be redefined in a computational manifold framework. For example, we could imagine a simple psymap array with only two psymaps storing two collections of symbolic attractors; the first psymap could identify nodes and the second psymap could identify relationships. The size of the graph would be limited only by the association capacity of the psymaps, which in a neurotransmitter field implementation is proportional to the number of neurons [8]. Thus, a relatively small number of interconnected psymaps, each having a large capacity, can encode a large complex graph of interconnected ideas. Algorithms that operate on these graphs can then be structured in terms of images and image attractors.

Manifold automata allow us to unify diverse cognitive processing mechanisms, including verbal and non-verbal reasoning, in a straightforward manner than is consistent with observed behavior and intuitive experience. Sounds, physical objects, visual images and muscular movements are all functions defined on manifolds that can be mentally shifted, transformed and manipulated. The formation of composite images, by overlaying multiple components topologically, allows their content to be combined and processed in the context of past experience.

Associating images in computational manifold automata with multiple feedback cycles provides a framework for language and cognition. The mathematics that describes these surfaces and volumes of computation and the functional relationships between them provides a unified theory which can be used to model natural intelligence.

ACKNOWLEDGMENT

The author thanks Professors M. Tuceryan and B. Thomsen for their comments and suggestions.

REFERENCES

[1] M. Spivak, *A comprehensive introduction to differential geometry* (3rd ed.), Wilmington, DE: Publish or Perish, 1979.
[2] Z. Kohavi, *Switching and Finite Automata Theory*. New York: McGraw-Hill, 1970.
[3] M. A. Harrison, *Introduction to Formal Language Theory*. Reading, MA: Addison-Wesley, 1978.
[4] A. Turing, "On computable numbers, with an application to the entscheidungsproblem," *Proc. London Mathematical Society*, 1936.
[5] J. M. T. Thompson and H. B. Stewart, *Nonlinear Dynamics and Chaos*. Hoboken, NJ: Wiley, 2002.
[6] J. Monad, *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology* (Trans. A. Wainhouse). New York: Knopf, 1971.
[7] J. B. Marion, *Classical Dynamics of Particles and Systems* (2nd Ed.). New York: Academic Press. 1970, pp. 50-51.
[8] D. S. Greer, "Neurotransmitter fields," in *Proc. of International Conf. on Artificial Neural Networks*, Porto, Portugal, 2007.
[9] I. Rock, *Perception*. New York: Scientific American Books. 1984.
[10] E. I. Knudsen, S. du Lac, and S. D. Esterly, "Computational maps in the brain," *Ann. Rev. of Neuroscience*, vol. 10, 1987. pp. 41-65.
[11] J. R. Munkres, *Topology* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall, 2000.
[12] D. S. Greer, "A unified system of computational manifolds," Dept. of Comp. and Info. Sci., IUPUI, Indianapolis, IN, Tech. Rep. TR-CIS-0602-03, 2003.
[13] P. L. Williams, and R. Warwick, *Gray's Anatomy* (36th Ed.). Philadelphia, PA: Saunders. 1980.
[14] F. Q. Xu, C. A. Greer, and G. M. Shepherd, "Odor maps in the olfactory bulb," *J. Comparative Neurology*, vol. 422, pp. 489-495. July 2000.
[15] L. R. Squire, F. E. Bloom, S. K. McConnell, J. L. Roberts, N. C. Spitzer and M. J. Zigmond, *Fundamental Neuroscience*. New York: Academic Press, 2003. pp. 657-663.
[16] D. S. Greer, "An image association model of the Brodmann areas," in *Proc. 6th IEEE International Conf. on Cognitive Informatics*, Lake Tahoe, NV, 2007
[17] R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science* (4th ed.). New York: McGraw-Hill, 2000.
[18] R. N. Shepard, and L. A. Cooper, *Mental Images and their Transformations*. Cambridge MA: MIT Press. 1982.

[19] E. Jeandel, "Topological Automata," in 22*nd* *Annual Symposium on Theoretical Computer Science*, Heidelberg: Springer 2005.
[20] A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*. New York: Dover. 1970.
[21] L. H. Koopmans, *The Spectral Analysis of Time Series*. New York: Academic Press, 1974.
[22] K. Brodmann, *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: Barth, 1909.
[23] J. C. Robinson, *Infinite-Dimensional Dynamical Systems: an Introduction to Dissipative Parabolic PDEs and the Theory of Global Attractors*. Cambridge University Press, 2001.
[24] B. MacLennan, "Continuous formal systems: A unifying model in language and cognition, In *Proc. IEEE Workshop on Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems*. Monterey, CA. 1995.
[25] T. E. J. Behrens, H. Johansen-Berg, M. W. Woolrich, S. M. Smith, C. A. M. Wheeler-Kingshott, P. A. Boulby, et al. "Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging," Nature Neuroscience, vol. 6 (7), pp. 750-757, 2003.
[26] K. Thompson, "Alpha blending," In A. Glassner (Eds.), Graphics Gems, pp 210-211, Cambridge: Academic Press, 1990.