

IMAGE RETRIEVAL AND CLASSIFICATION USING ASSOCIATIVE RECIPROCAL-IMAGE ATTRACTORS

Douglas S. Greer¹ and Mihran Tuceryan²

General Manifolds¹, Indiana University Purdue University Indianapolis²

ABSTRACT

In this paper, image processing and symbol processing are bridged with a common framework. A new computational architecture allows arbitrary fixed images to be used as attractors in a general-purpose association processor that can be used for the retrieval and recognition of images. Direct image-to-image associations eliminate the need to extract edges or other features. The creation of attractor basins around the reciprocal-image pairs permits the construction of stable implementations.

The algorithms, developed as a neurophysiological model, can form global image associations using only local, recurrent connections. A powerful composite structure can be created with an array of interconnected image processors.

We show the results of using this framework successfully and the convergence of partial images to nearby reciprocal-image attractors.

Index Terms— Image retrieval, pattern recognition, associative memory, image restoration, signal detection, non-linear dynamical systems.

1. INTRODUCTION

The basic unit of memory, the Set-Reset (SR) flip-flop, 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 extend this fundamental principle by replacing the bits in the SR flip-flop with images and the logic gates with local association processors. Each attractor corresponds to a pair of arbitrarily chosen images. Rather than having only two attractors for “0” and “1,” as is the case for an SR flip-flop, the dynamical system can have any number of attractors. Each “image attractor” is constructed from two arbitrarily chosen *reciprocal* images.

This document contains material that is patent pending by General Manifolds LLC, <http://www.gmanif.com/ip>.

When a new image is presented to the system, the association model converges to the closest stable state corresponding to an image pair, thus resulting in the recalled and reconstructed image. These fixed-point image attractors can be viewed as a basic representation of continuous symbols.

The bridge between image processing and neurophysiology, created by modeling each Brodmann area as a pair or reciprocally associated images, implies that research into a large class of image processing algorithms is cognitive science research and vice-versa [1][2]. The ability to associate images is the basis for learning cause-and-effect relationships involving vision, hearing, tactile sensation, and kinetic motion. Consequently, the model described can serve as part of a scientific explanation of cognition. Moreover, concrete implementations can serve as a basis for *in silico* experiments in neuroscience.

The algorithms described use only *direct* image-to-image associations and consequently differ from much of the previous work in image retrieval and classification which rely on the extraction of edges, corners or discrete features as an intermediate step. Recognition or classification can take the form of retrieving an associated image which contains a gray-scale barcode that can be quickly identified. The justification for the direct image-to-image approach is that discrete bit patterns do not occur in nature and therefore the central nervous system evolved to form associations directly between functions of time, space and frequency without an intermediate reduction in dimension and the concomitant loss of information.

2. COMPUTATIONAL MODEL

The computations of neurotransmitter fields and psymap arrays are described in [1][3] and only briefly reviewed here.

2.1 Neurotransmitter Fields

In neurotransmitter field models, the state variables are the concentration of neurotransmitters in the extracellular space rather than the neuron action potentials. Since there are effectively two summations, one intracellular and one extracellular, it has mathematical characteristics that are

fundamentally different than a traditional neural network [3]. Neurotransmitter fields are also simpler to train, since no Hebbian learning or back-propagation is required.

The computation with three topologically aligned input images, A, B, and C, and an output image G, is illustrated in Fig. 1.

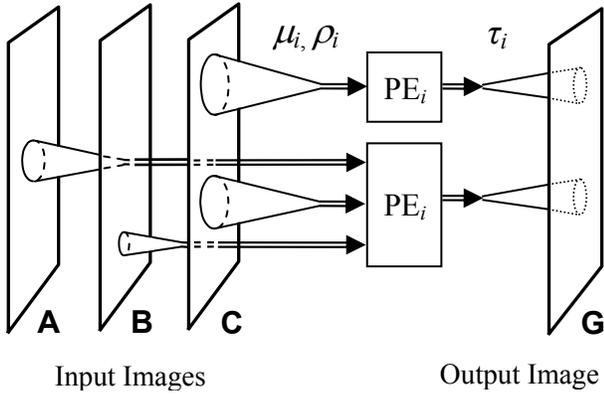


Fig. 1. The internal structure of a single Λ -map. Each Processing Element (PE_i) represents a single neuron i . The image data values correspond to the concentration of neurotransmitters in the extracellular space. There are effectively two summations, one inside the neuron and one outside the neuron in the space surrounding multisynaptic dendritic spines. Consequently, the model differs mathematically from a standard neural network.

Each Processing Element (PE_i) corresponds to a single neuron i and is characterized by the functions $\rho_i(u,v)$, $\mu_i(u,v)$ and $\tau_i(x,y)$ where the input images are parameterized by the variables (u,v) and the output image is parameterized by (x,y) . Each separate PE_i is “tuned” to a specific input pattern $\rho_i(u,v)$. When that pattern of neurotransmitter concentration is detected in the input image, the neuron fires, causing the release of neurotransmitter into the output image. The amount of neurotransmitter released by the neuron axons is specified by the transmitter function $\tau_i(x,y)$. The $\mu_i(u,v)$ are two-dimensional (product) measures, which in the results demonstrated, were simply set to a uniform positive value inside the cones on the left-hand side of the PEs in Fig. 1 and set to zero outside of the cones.

Let $\theta(\cdot)$ be a real-valued radial basis function, for example a Gaussian $\exp(-x^2/(\text{stddev})^2)$, that has its maximum at the origin and the properties $\theta(x) > 0$ and $\theta(x) \rightarrow 0$ as $|x| \rightarrow \infty$. If $h(u,v)$ is an arbitrary input image, then $\theta(h-\rho_i)$ is a two-dimensional image which equals one where the input, h , equals the pattern, ρ_i , and falls to zero elsewhere.

By chance, a few pixels in the input will match the pattern generating values of $\theta(h-\rho_i)$ near one. However, when the image $\theta(h-\rho_i)$ is integrated over a local region, and the result mapped by a non-linear sigmoidal transfer function $\sigma(\cdot)$, (similar to the transfer function used in neural

networks) this value can be made arbitrarily small. This also allows us to characterize some fractional portion of an image region as a near perfect match.

Multiplying this value by the transmitter function and summing over all of the neurons i results in the output g :

$$g(x,y) = \sum_i \sigma \left(\int_H \theta(h-\rho_i) d\mu_i \right) \cdot \tau_i(x,y) \quad (1)$$

For brevity, the variables of integration (u,v) which parameterize the functions h , ρ_i and μ_i on the input image H have been omitted from the equation.

Multiple input images can be combined in several ways, for example by forming a vector or direct sum of the multiple measures, μ , and patterns, ρ . Figure 1 also shows how the input regions can vary in size and shape. These have been posited to correspond to the receptive fields of varying sizes such as those found in the retina [1].

2.2 Associative Reciprocal-Image Attractors

The basic building block is the image association processor *Logic-map* or Λ -map shown in Fig. 1. Since a Λ -map can accept multiple input images, two Λ -maps can be combined and connected recursively as shown in Fig. 2. Because of the biological analogy to the exterior and interior lamina of one of the Brodmann areas in the cerebral cortex, the two Λ -maps are labeled Λ_E and Λ_I and the circuit as a whole is referred to as a psymap.

Just as stateless NAND gates in an SR flip-flop form a dynamical system with two stable attractors, so too can feed-forward Λ -maps form stable image attractors in a psymap. Even though the PEs in each Λ -map have only *local* connections in a small neighborhood, they can create overall *global* image attractors. This is demonstrated in the results section where initially, many of the PEs have no local input values that match the overall global image pattern.

A psymap dynamical system with any number of reciprocal-image attractors can be created in the following manner. Let $q = \Lambda_E(p, s)$ and $p = \Lambda_I(q, r)$ denote the exterior and interior Λ -maps shown in Fig 2 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 a new set of PEs to both Λ -maps so that $\Lambda_E(b_i, \text{Null}) = a_i$ and $\Lambda_I(a_i, \text{Null}) = b_i$. 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. Since each PE is tuned to a specific input pattern, any number of attractors can be added to the system and all of the computations can be done in parallel.

The partial differential equations which describe the behavior of reciprocal-image attractors over time form an

infinite-dimensional dynamical system [4]. However, we can try to gain some insight into the capacity and training of the system by imagining a large geographic area in which attractors are created by the force of gravity where a frictionless ball would roll into valleys or “energy wells.” These valleys are attractor basins in the corresponding dynamical system. The ability to specify the $\rho_i(u,v)$, $\mu_i(u,v)$ and $\tau_i(u,v)$ functions in (1) in addition to the θ and σ characteristics provides a great deal of flexibility in controlling the potential energy landscape. By analogy, the patterns ρ_i for a reciprocal-image pair control the longitude and latitude of the bottom of the wells or valleys; the θ and σ functions control the width and steepness of the sides. Decreasing the width or “size” of the energy wells allows the number of stable wells in a large area to be increased without limit. Thus the image storage capacity can be made arbitrarily large by decreasing the standard deviation of the θ of the functions. The smaller attractor size may make the associations more difficult to “find”, but once located, all of the attractors will be stable.

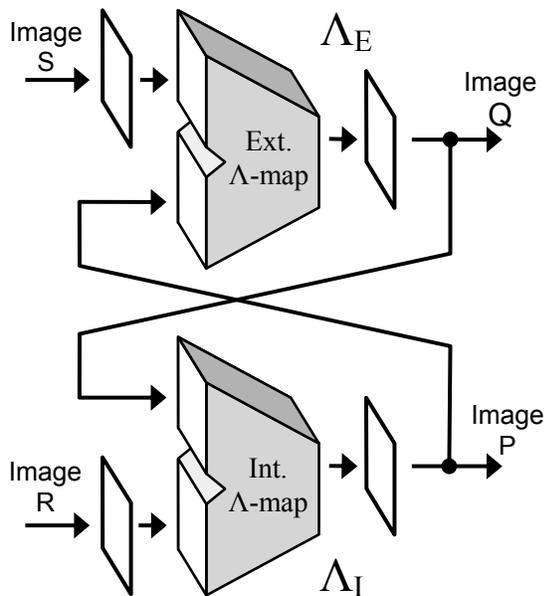


Fig. 2. A single psymap composed of the recursively connected external and internal Λ -maps labeled Λ_E and Λ_I . Both the circuit and the dynamic behavior are analogous to the SR flip-flop where the NAND gates have been replaced by Λ -maps and bits have been replaced by images.

As shown in Fig. 1, using topological alignment of the images, a Λ -map can receive any number of inputs. Consequently, a composite retrieval and recognition system can be created with multiple psymaps by connecting the output of one psymap to one or more inputs. An array of psymaps whose connections are illustrated with an I/O Bus notation is shown in Fig. 3.

Using multiple psymaps is a potentially very powerful technique since it can serve as a mechanism to trigger a “recognition cascade” where convergence in one or more psymaps causes a cascade of convergence in several others. Again using the metaphor of gravity and geographical landscapes, each psymap has its own landscape. While we may be searching for a relatively tiny reciprocal-image attractor in one psymap, another psymap may have a large expansive valley where convergence toward the bottom will cause us to move closer to the precise solution in the first psymap. This phenomenon is well known and well studied in psychology, where verbal “hints” about what an image contains help us locate it. Crossword puzzles are an example of multi-sensory integration, where phonetic sounds are combined with semantic meaning to help find the correct association.

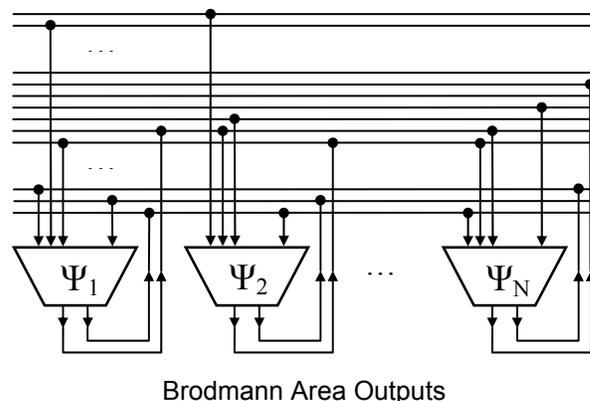


Fig. 3. The Psymap Array Model. Each psymap, Ψ_i , corresponds to a separate reciprocal-image attractor. Lines in the diagram correspond to images, and dots represent specific input and output connections. The array as whole, forms a composite image recognition and retrieval system.

3. EXPERIMENTAL RESULTS

In the experimental setup, we attempt to show the power and utility of this computational framework by implementing a simple recognizer which performs well even with input images that have been degraded with noise. The computational framework described above was implemented in Java and a single psymap was created with the image associations shown in Fig. 4. Associations A4 and A5 are the barcodes with their corresponding images. The test inputs and experimental results are shown in Figures 5, 6 and 7. After three iterations, the system converges to an image attractor where it will remain indefinitely. Notice in these results that (i) the barcode and the associated image are properly retrieved/recognized, and (ii) the results are robust under extreme degradation due to noise (Figs. 6 & 7).

All of the images used were 256×256 pixels. Each of the associations was created from a grid of (16+1)×(16+1) or

289 PEs with local inputs taken from a 64×64 pixel area. Since the ρ_i and τ_i functions are computed directly, and do not require conventional neural network training methods, the set-up time was less than 10s and the run time for each of the tests was approximately 30s on a 2 GHz personal computer.

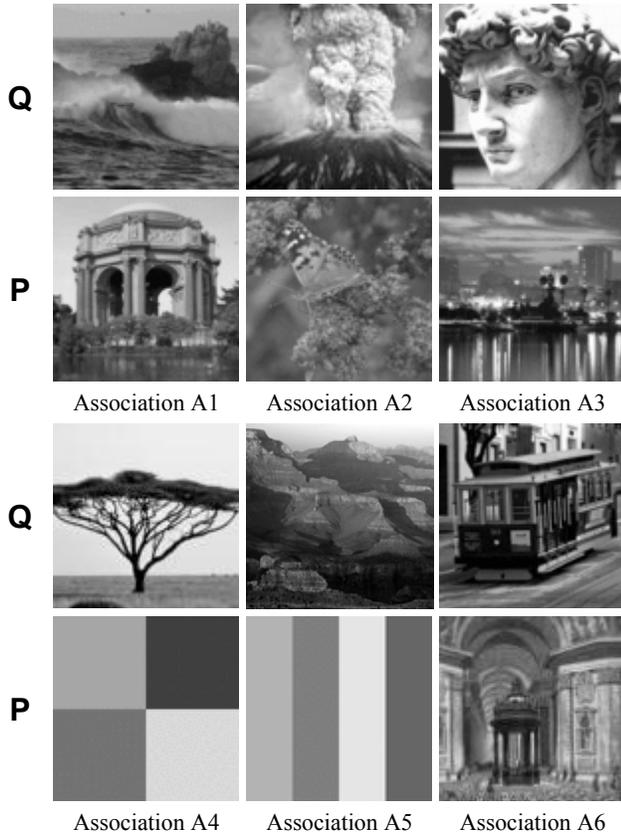


Fig. 4. The six pairs of associated images used to construct a single psymap that contains six reciprocal-image attractors.

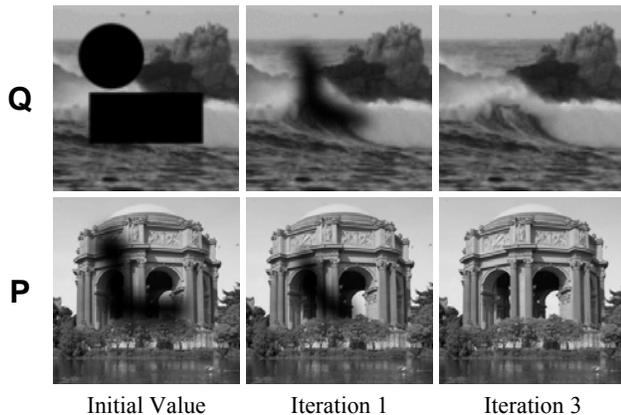


Fig. 5. An initial value, which is equal to one of the predefined attractor images with a circle and rectangle removed, converges to the association A1.

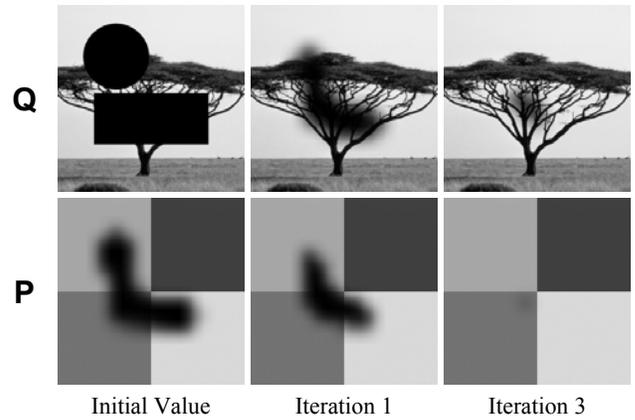


Fig. 6. An initial value that converges to the association A4. Creating an association with a square “barcode” image pattern permits straightforward classification and retrieval.

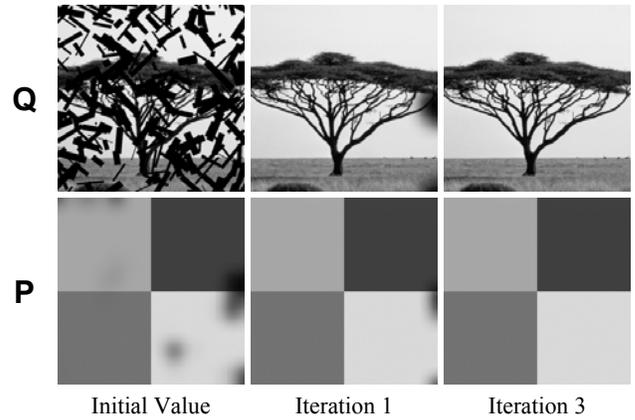


Fig. 7. An initial value with “noise” in the form of random line segments that converges to the association A4.

5. REFERENCES

- [1] D. S. Greer, “An Image Association Model of the Brodmann Areas,” Proc. 6th IEEE International Conf. on Cognitive Informatics, 2007
- [2] D. S. Greer, “A unified system of computational manifolds,” Tech. Rep. TR-CIS-0602-03, Dept. of Comp. and Info. Sci., IUPUI, Indianapolis, IN, 2003.
- [3] D. S. Greer, “Neurotransmitter fields,” Proc. of International Conf. on Artificial Neural Networks, (ICANN’07), Porto, Portugal, 2007.
- [4] J. C. Robinson, Infinite-Dimensional Dynamical Systems: an Introduction to Dissipative Parabolic PDEs and the Theory of Global Attractors. Cambridge University Press, 2001.