

LETTER TO THE EDITOR

Physiological Interpretation of the Self-Organizing Map Algorithm

Stephen Grossberg
Boston University
Department of Cognitive and Neural Systems
111 Cummington Street
Boston MA 02215 USA

Submitted: February 1994

Revised: July 1994

Supported in part by the Air Force Office of Scientific Research (AFOSR F49620-92-J-0499), the Advanced Research Projects Agency (ONR N00014-92-J-4015), and the Office of Naval Research (ONR N00014-91-J-4100).

Kohonen (1993) provides an interesting discussion of how a self-organizing feature map (SOFM) may be biologically realized. He notes that such a map “has been used as an idealized theoretical principle and effective practical computing method, without explicit reference to any physiological processes...It is now argued in the present article that there may exist a physiological implementation of this principle” (p.895). Kohonen’s goal is to physiologically interpret a winner-take-all dot product rule and a self-normalizing synaptic weight change rule, where weights change only if they are in the neighborhood of the winner, and do so by an amount that varies with their distance from the winner. He builds upon his own contributions to the SOFM model, dating from 1982.

As with other key neural network models, however, substantive model discovery and development were carried out in the 1970’s and even the 1960’s. That is why the renewed interest in neural networks in the 1980’s led to such explosive development of the field. Often key models were substantially known, ready to be presented to a new audience for further development. In this heady atmosphere, the original discoveries were often overlooked, much as backpropagation is often attributed to Rumelhart, Hinton and Williams (1986), rather than to Amari (1967; see also Amari, 1993), Werbos (1974), and Parker (1982).

As Kohonen (1993) notes on p.895, a number of investigators studied the SOFM model before 1982. Many of these investigators considered how maps self-organize in the cerebral cortex. In addition, in Grossberg (1976a) and Grossberg (1978a), neurobiological modeling rules were articulated and restated in the familiar SOFM formalism as an algebraic winner-take-all dot product rule, and a self-normalizing synaptic weight change rule whose weights change only if they are in the neighborhood of the winner. This transition was made possible by a decade of prior mathematical and simulation work that analysed how neural networks carry out associative learning (e.g., Grossberg, 1969a, 1972a) and how recurrent on-center off-surround networks carry out competitive decision making, including winner-take-all competition (e.g., Grossberg, 1973; Ellias and Grossberg, 1975; Grossberg and Levine, 1975; Levine and Grossberg, 1976). The mathematical properties of these associative and competitive networks guided the selection of idealized rules for SOFM computation.

These rules, schematized in Figure 1 from Grossberg (1976a), are used in all contemporary SOFM models. The 1976 article also contained the first mathematical proof of the key SOFM properties that synaptic weights self-normalize, track the statistics of environmental inputs, and stabilize in an input environment that is sufficiently “sparse”. The 1976 paper developed the L1 norm model that was subsequently applied by Rumelhart and Zipser (1985). The 1976 proof of weight normalization led to the observation in Grossberg (1978a) of how the L2 norm case that is used by Kohonen (1984) carries out unbiased learning. The relationship of these results to the seminal work of Malsburg (1973) is reviewed in Grossberg (1987) and Carpenter and Grossberg (1991). In brief, Malsburg (1973) modified the additive model for adaptive instar classification that was developed in Grossberg (1972b) to introduce the first biologically motivated SOFM. His model, however, used a non-local adaptive weight equation and did not operate in real-time. The Grossberg (1976a) article showed how to modify the Malsburg SOFM rules to introduce the first local real-time SOFM model.

These research streams on both biological and technological SOFM models have continued

unabated to this day. Contributions have been made to all the areas that Kohonen (1993) mentions, and seem to exhibit computational and data-predictive properties that have not yet been attained in Kohonen (1993), while providing some evidence against one of his key hypotheses. Since the primary focus in Kohonen (1993) was to discuss physiological mechanisms for winner-take-all (WTA) competition, that will also be the focus below.

To better appreciate this discussion, the reader should keep in mind that one of the results proved in Grossberg (1976a) led to a bifurcation of research on neural networks after the mid-1970's. This result proved that learning in the SOFM model is unstable in response to an arbitrary nonstationary input stream. One large group of researchers, with Kohonen (1984) as one of its most visible stimuli, applied SOFM models to problems in sparse or stationary environments, and externally controlled learning at large times to ensure good convergence. The second group of researchers focused on how to design an SOFM that could *self-stabilize* its learning in response to *arbitrary* nonstationary inputs. To accomplish this goal, I introduced Adaptive Resonance Theory in Part II of the original SOFM paper (Grossberg, 1976b). Thus ART models were introduced as examples of self-controlling SOFM models. One theme of considerable research interest still today is whether, and how, external or internal controllers are used to stabilize SOFM learning. All such approaches can be viewed as variants or elaborations of the basic SOFM model in Figure 1.

Kohonen (1984, 1989) regulated convergence by using a topological neighborhood, or bubble, whose size was shrunk through time by an external controller. One reason why external control as needed is that Kohonen, like Malsburg (1973) and Grossberg (1972b) before him, used an additive model as his state equation for neuron interactions. In Grossberg (1976a), a shunting model of neuron interactions was used. A shunting model realizes the membrane equations of neurophysiology (Hodgkin, 1964), in addition to having important computational properties. For example, shunting models tend to self-normalize their total activity. Thus, as learning proceeds, the topological neighborhoods, or bubbles, in a shunting SOFM model shrink automatically, due to internal self-control, as winning nodes get relatively more active due to the map learning process. This difference between additive and shunting models is also relevant to evaluating Kohonen's proposed WTA model.

Kohonen correctly notes that "one of the original WTA networks" was presented in Grossberg (1973). This 1973 article introduced and analysed networks with transient inputs that instate an activity pattern before they are shut off. Then nonlinear internal feedback selects a winning population or populations. Kohonen incorrectly states that Cohen and Grossberg (1983) analysed a more general class of nets with transient inputs. He then asserts that his goal is to analyse "a WTA function with respect to persistent external inputs". Actually, the Cohen-Grossberg model and theorem were developed to handle a large class of biological networks with persistent inputs that were introduced subsequent to 1973, with Ellias and Grossberg (1975), Grossberg and Levine (1975), and Levine and Grossberg (1976) presenting some of the earliest. All of these networks are built up from cells undergoing shunting equations that interact through recurrent on-center off-surround interactions (Kuffler, 1953; Ratliff, 1965), as in Kohonen's Figure 2. They thus provided the type of WTA "physiological implementation" that Kohonen is now seeking, as well as a mathematical characterization of when "bubbles", or partial contrast-enhancement, will occur.

Kohonen (1993) does not use a shunting equation to describe WTA of model activities η_i . Instead, he uses an additive equation with variable gain $\gamma(\eta_i)$; see his equation (6). He also uses slow inhibitory feedback interneurons ζ_i to periodically inhibit, and thus reset, the WTA choice; see his Figure 4. Periodic inhibition using slow feedback inhibitory interneurons in additive networks was mathematically analysed in Grossberg (1970, Section 3); see also Ögmen and Gagné (1990).

Kohonen adds this well-known device to his WTA network. If the feedback inhibition gets large, however, then the activities η_i could potentially become negative and thereby destabilize network dynamics. To prevent this from happening, Kohonen imposes unphysiological mathematical conditions on equation (6). In contrast to these unphysiological properties of additive models, shunting networks can automatically preserve the sign of their activities (Grossberg, 1973). In summary, Kohonen uses an additive network with slow inhibitory interneurons to make and reset a WTA choice. Because the interneuron varies more slowly than the WTA choice, he ignores it in his mathematical treatment of the WTA property in Lemma 2.

This difference between the additive and shunting equations having been noted, it can be seen that Kohonen (1993) uses the same mathematical method in his Lemma 2 to prove the WTA property for an additive network as was developed for the shunting case in Grossberg (1973). In particular, fixed signs of terms and faster-than-linear signals (condition $\gamma'' > 0$ in Lemma 2) are used to prove that activity differences grow, thereby leading to selection of a winning population. Kohonen (1973, p.903) goes on to note that “a connectivity of $O(N)$ can be achieved if the inhibitory effects are first summed up by a common interneuron”. The proofs in Grossberg (1973) and Ellias and Grossberg (1975) showed this by rewriting network interactions as an “adaptation level” $\sum_k \sigma(\eta_k)$ that can be computed by an inhibitory interneuron. Kohonen (1993) also uses the adaptation level formulation; see his equation (12).

This “adaptation level” observation led to another burst of research activity. It was generalized in a series of articles in the mid 1970’s, leading to the characterization and mathematical analysis of a general class of “adaptation level” recurrent networks, all of whose trajectories approach equilibria (Grossberg, 1978b). This analysis introduced global Liapunov methods into the theory of neural networks (Grossberg, 1980). The effort to develop a unified method for proving convergence in both additive and shunting models also led to the Cohen-Grossberg model and theorem (Cohen and Grossberg, 1983; Grossberg, 1982), which anticipated and subsumes the additive model and Liapunov function that is sometimes called the Hopfield model after their use in Hopfield (1984). Because Hopfield did not first describe either the model or the Liapunov function, the classical name of additive model is more accurate.

As noted above, one of Kohonen’s WTA design goals is to “reset [the network] before a new input becomes active” (p.898). He uses a slow negative feedback loop that generates a reset at a *fixed* delay after input *onset*. However, in both biological and artificial networks, inputs typically need to be reset aperiodically, in a stimulus-dependent way, at a *variable* delay after input *offset*, much as the information stored in a working memory is dependent more on the order of events than on their input timing (Atkinson and Shiffrin, 1968, 1971). In principle, reset could follow an input at a fixed period and thus be insensitive to subsequent environmental feedback; or it could

be driven nonselectively by changes in inputs; or it could be driven to selectively reset only those nodes that mismatch present or subsequent input data. All of these types of reset have been studied before in the literature. An example of the third, most selective type of reset, is found in the ART3 model of Carpenter and Grossberg (1990). This model uses a chemically modulated on-center off-surround recurrent network whose competitive decisions are reset selectively in response to mismatches with the changing inputs themselves, whether they be periodic or aperiodic in time.

Another design goal of Kohonen is to determine the cut-off that decides how many winners there are. He notes that one option is to vary a selection threshold. Grossberg (1973) showed that shunting competitive nets compute a “quenching threshold” that can be used as a control parameter to select any prescribed number of winners. In the more advanced masking field model of Cohen and Grossberg (1986, 1987), the informational ambiguity of an input pattern can automatically determine how many approximate winners are stored, with fewer winners being stored as the input becomes more predictive. In addition, as noted above, learning automatically “shrinks the bubble” as the information became more sure. Bubble shrinking is often done by an external controller in the SOFM applications of Kohonen.

Kohonen emphasizes the possible role of “chemical substances rapidly diffusing” (p.900) to achieve fast and even signaling for control of learning at synaptic weights. In this regard, he reviews a stochastic learning law discussed in Kohonen (1984) that has self-normalizing adaptive weights; see equations (7) and (8) in Kohonen (1993), where chemical signals subtract from synaptic strength to cause a type of “forgetting” that normalizes synaptic strength. A similar self-normalizing learning law was described in Grossberg (1984) and used in Grossberg and Kuperstein (1986, Chapter 10; see also Grossberg, 1988, Sections 20 and 21) to model stochastic learning of invariant self-regulating maps. This learning law also uses chemical signals to control the forgetting that leads to synaptic normalization. Similar chemical signals were used to regulate the reset and learning operations in the ART systems of Grossberg (1980) and Carpenter and Grossberg (1990). These results are consistent with Kohonen’s hypothesis about the importance of chemical signals.

On the other hand, diffusive chemical signaling has not been needed to explain the fast and even spread of a postsynaptic learning factor to multiple sites, in contrast to Kohonen’s central claim that extracellular chemical signaling is needed because intracellular electrical “signal attenuation, especially in large ramified cells, is a severe problem” (p.899). In Grossberg (1975, Section 24) it was suggested that a massive action potential, triggered by the cell body of hippocampal pyramidal cells, could broadly activate cell dendrites and dendritic spines to provide such an evenly distributed learning signal. This prediction was applied again in masking field models of cortical processing (Cohen and Grossberg, 1987, Figure 16) and in models of classical conditioning (Grossberg and Schmajuk, 1987, Figure 9), both of which include SOFM variants. Jaffe et al. (1992) have summarized supportive neurophysiological data of their own and related experiments that suggest how voltage-dependent Na⁺ spikes, that are initiated in the cell body of a hippocampal pyramidal cell, actively propagate along hippocampal dendrites and trigger local Ca⁺⁺ currents that may help to initiate synaptic learning. A role for such local Ca⁺⁺ currents in synaptic learning was suggested in some of the earliest real-time network models of associative learning (e.g., Grossberg, 1968, 1969b). Experimental results do not yet seem, however, to have

tested another property of this prediction; namely, that the global action potential helps to dissociate the local read-out of synaptic weights from the global read-in to synaptic weights after the cell body experiences a competitive decision process.

In summary, the problems which Kohonen (1993) raises have been actively studied for many years. In particular, his WTA analysis uses familiar equations, mechanisms, and methods, but his use of the additive model seems less compelling to me as a physiological model than the earlier shunting models and generalizations thereof. In fact, it was precisely these sorts of deficiencies of the additive model that led me to increasingly use shunting models over the years.

Many researchers may not be aware of these earlier results because they have often been told that the 1960's and 1970's were a dark age in neural network history. This is true from a political perspective. From a scientific perspective, however, this era was a golden age of neural network discovery during which many key models, including symmetric autoassociators, back propagation, self-organizing feature maps, and WTA networks, were discovered and substantially developed. Now that the neural network renaissance is well underway, it is time to expand our scholarship and to get our history right. Only then can our young researchers reliably plan their own research programs with confidence.

REFERENCES

- Amari, S.-I. (1967). Theory of adaptive pattern classifiers. *IEEE Transactions on Electronic Computers*, **EC-16**, 299-307.
- Amari, S.-I. (1993). Backpropagation and stochastic gradient descent method. *Neurocomputing*, **5**, 185-196.
- Atkinson, R.C. and Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence and J.T. Spence (Eds.), **Advances in the psychology of learning and motivation research and theory (Volume 2)**. New York: Academic Press.
- Atkinson, R.C. and Shiffrin, R.M. (1971). The control of short term memory. *Scientific American*, August.
- Carpenter, G.A. and Grossberg, S. (1990). ART3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks*, **3**, 129-152.
- Carpenter, G.A. and Grossberg, S. (Eds.) (1991). **Pattern recognition by self-organizing neural networks**. Cambridge, MA: MIT Press.
- Cohen, M.A. and Grossberg, S. (1983). Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-13**, 815-826.
- Cohen, M.A. and Grossberg, S. (1986). Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short term memory. *Human Neurobiology*, **5**, 1-22.
- Cohen, M.A. and Grossberg, S. (1987). Masking fields: A massively parallel architecture for learning, recognizing, and predicting multiple groupings of patterned data. *Applied Optics*, **26**, 1866-1891.
- Ellias, S.A. and Grossberg, S. (1975). Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks. *Biological Cybernetics*, **20**, 69-98.
- Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences*, **60**, 758-765.
- Grossberg, S. (1969a). On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *Journal of Statistical Physics*, **1**, 319-350.
- Grossberg, S. (1969b). On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, **22**, 325-364.

Grossberg, S. (1970). Neural pattern discrimination. *Journal of Theoretical Biology*, **27**, 291-337. Reprinted in Carpenter, G.A. and Grossberg, S. (Eds.) (1991), **Pattern recognition by self-organizing neural networks**. Cambridge, MA: MIT Press.

Grossberg, S. (1972a). Pattern learning by functional-differential neural networks with arbitrary path weights. In K. Schmitt (Ed.), **Delay and functional-differential equations and their applications**. New York: Academic Press. Reprinted in Grossberg, S. (1982), **Studies of mind and brain**. Amsterdam: Kluwer Academic.

Grossberg, S. (1972b). Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik*, **10**, 49-57. Reprinted in Carpenter, G.A. and Grossberg, S. (Eds.) (1991), **Pattern recognition by self-organizing neural networks**. Cambridge, MA: MIT Press.

Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, **52**, 217-257. Reprinted in Grossberg, S. (1982), **Studies of mind and brain**. Amsterdam: Kluwer Academic.

Grossberg, S. (1975). A neural model of attention, reinforcement, and discrimination learning. *International Review of Neurobiology*, **18**, 263-327. Reprinted in Grossberg, S. (1982), **Studies of mind and brain**. Amsterdam: Kluwer Academic.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, **23**, 121-134.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, **23**, 187-202.

Grossberg, S. (1978a). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen and F. Snell (Eds.), **Progress in theoretical biology**, Vol. 5. New York: Academic Press. Reprinted in Grossberg, S. (1982), **Studies of mind and brain**. Amsterdam: Kluwer Academic.

Grossberg, S. (1978b). Competition, decision, and consensus. *Journal of Mathematical Analysis and Applications*, **66**, 470-493. Reprinted in Grossberg, S. (1982), **Studies of mind and brain**. Amsterdam: Kluwer Academic.

Grossberg, S. (1980). Biological competition: Decision rules, pattern formation, and oscillations. *Proceedings of the National Academy of Sciences*, **77**, 2338-2342. Reprinted in Grossberg, S. (1982), **Studies of mind and brain**. Amsterdam: Kluwer Academic.

Grossberg, S. (1982). Associative and competitive principles of learning and development: The temporal unfolding and stability of STM and LTM patterns. In S.I. Amari and M. Arbib (Eds.), **Competition and cooperation in neural networks**. New York: Springer-Verlag.

Grossberg, S. (1984). Some psychophysiological and pharmacological correlates of a developmental, cognitive, and motivational theory. In R. Karrer, J. Cohen, and P. Tueting (Eds.), **Brain and information: Event related potentials**. New York: New York Academy of Sciences. Reprinted in Grossberg, S. (1987), **The adaptive brain, I: Cognition, learning, reinforcement, and rhythm**. Amsterdam: Elsevier/North-Holland.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, **11**, 23-63.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, **1**, 17-61.

Grossberg, S. and Kuperstein, M. (1986). **Neural dynamics of adaptive sensory-motor control: Ballistic eye movements**. Amsterdam: Elsevier. (1989, Expanded edition. New York: Pergamon Press.)

Grossberg, S. and Levine, D. (1975). Some developmental and attentional biases in the contrast enhancement and short term memory of recurrent neural networks. *Journal of Theoretical Biology*, **53**, 341-380.

Grossberg, S. and Schmajuk, N.A. (1987). Neural dynamics of attentionally-modulated Pavlovian conditioning: Conditioned reinforcement, inhibition, and opponent processing. *Psychobiology*, **15**, 195-240.

Hodgkin, A.L. (1964). **The conduction of the nervous impulse**. Liverpool: Liverpool University Press.

Hopfield, J.J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences*, **81**, 3058-3092.

Jaffe, D.B., Johnston, D., Lasser-Ross, N., Lisman, J.E., Miyakawa, H., and Ross, W.N. (1992). The spread of Na⁺⁺ spikes determines the pattern of dendritic Ca⁺⁺ entry into hippocampal neurons. *Nature*, **357**, 244-246.

Kohonen, T. (1984). **Self-organization and associative memory**. New York: Springer-Verlag.

Kohonen, T. (1989). **Self-organization and associative memory** (third edition). Berlin: Springer-Verlag.

Kohonen, T. (1993). Physiological interpretation of the self-organizing map algorithm. *Neural Networks*, **6**, 895-905.

Kuffler, S. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, **16**, 37-68.

- Levine, D. and Grossberg, S. (1976). On visual illusions in neural networks: Line neutralization, tilt aftereffect, and angle expansion. *Journal of Theoretical Biology*, **61**, 477-504.
- Malsburg, C. von der (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, **14**, 85-100.
- Ögmen, H. and Gagné, S. (1990). Neural network architectures for motion perception and elementary motion detection in the fly visual system. *Neural Networks*, **3**, 487-505.
- Parker, D.B. (1982). Learning-logic. Invention Report 581-64, File 1. Office of Technology Licensing, Stanford University.
- Ratliff, F. (1965). **Mach bands: Quantitative studies on neural networks in the retina**. New York: Holden Day.
- Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart and J.L. McClelland (Eds.), **Parallel distributed processing**. Cambridge, MA: MIT Press.
- Rumelhart, D.E. and Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science*, **9**, 75-112.
- Werbos, P. (1974). **Beyond regression: New tools for prediction and analysis in the behavioral sciences**. Doctoral thesis, Harvard University, Cambridge, Massachusetts.

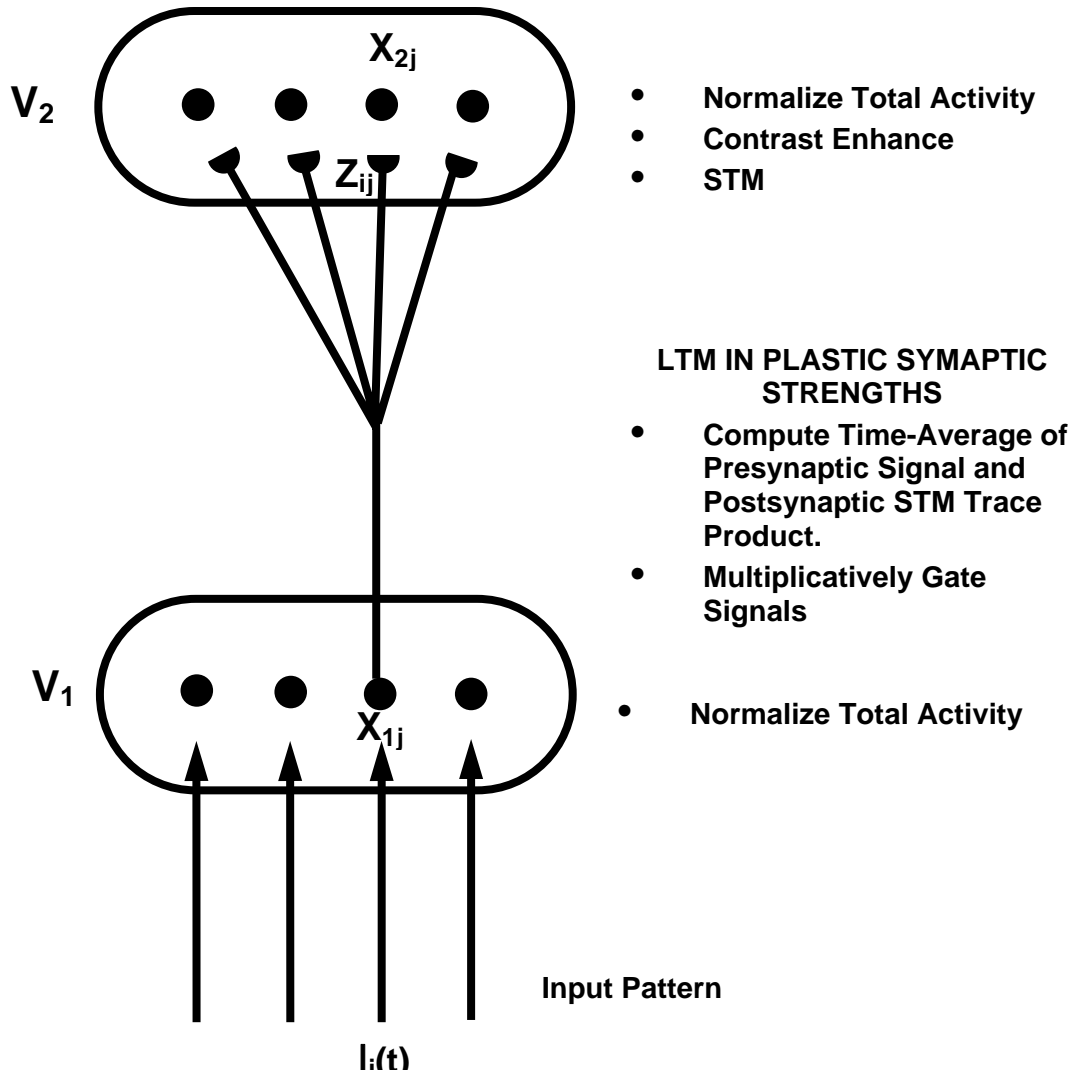


Figure 1. SOFM rules as described in Grossberg (1976). [Reproduced with permission.]