

A NEURAL NETWORK APPROACH FOR ANALYZING THE ILLUSION OF MOVEMENT IN STATIC IMAGES

JAVIER ROPERO PELÁEZ⁽¹⁾, DIEGO ANDINA⁽²⁾, SILVIO SZAFIR,
FRANCISCO JAVIER RAMÍREZ FERNÁNDEZ⁽¹⁾

¹DEPARTAMENTO DE ENGENHARIA DE SISTEMAS ELETRÔNICOS,
UNIVERSITY OF SÃO PAULO, BRAZIL

²UNIVERSIDAD POLITÉCNICA DE MADRID, SPAIN

fjavier@usp.br, d.andina@gc.ssr.upm.es, silviosz@yahoo.com.br, jramirez@lme.usp.br

ABSTRACT— The purpose of this work is to analyze the illusion of movement that appears when seeing certain static images. This analysis is accomplished by using a biologically plausible neural network that learned (in a unsupervised manner) to identify the movement direction of shifting training patterns. Some of the biological features that characterizes this neural network are: intrinsic plasticity to adapt firing probability, metaplasticity to regulate synaptic weights and firing adaptation of simulated pyramidal networks. After analyzing the results, we hypothesize that the illusion is due to cinematographic perception mechanisms in the brain due to which each visual frame is renewed approximately each 100 msec. Blurring of moving object in visual frames might be interpreted by the brain as movement, the same as if we present a static blurred object.

Key Words: movement illusion, plasticity, metaplasticity, cinematographical perception, Kitaoka's designs.

1. INTRODUCTION

Perception of illusory movement in static images is commonplace in Kitaoka's famous designs [1], such as his "rotating snakes" shown in Fig.1. Although the explanation of this effect is not known, Kitaoka and Ashida (2003) [2] describe the characteristics in the figure that intensify the effect. An example of these characteristics is the transition from black to light gray and then to white in each of the elements of Fig.1. We have focused on this characteristic to design Fig. 2, in which each triangle is colored with a smooth transition from white to gray. This type of transition occurs, for example, when shooting a photo of an object, and either the camera or the object moves, thereby producing a blurred image with a meteor-like vanishing tail. If the object

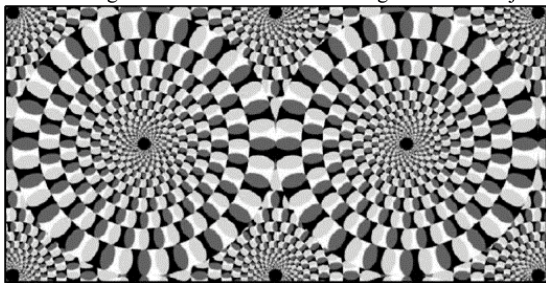


Fig. 1. Detail of "Rotating snakes", by Akiyoshi Kitaoka

is moving to the right there is a vanishing tail at the left and vice versa. This effect is also related to the concept of "cinematographic perception" [3], [4] according to which the brain is able to capture reality as in cinematography, frame by frame, one frame each 100 ms, approximately. If our brain shoots a "photo" each 100 ms, our "cerebral photograph" would be also moved or blurred when the perceived object is moving quickly. We propose that, as these blurred photographs appears associated with movement, movement detectors in our brain would be activated by these static images as actually happens in the experiments of Conway and collaborators [5]. Conway and collaborators discovered that the neural circuits involved in recognizing movement, i.e. the movement detectors of the middle temporal area (MT area), are also active when Kitaoka's-like images are presented. Although the neural circuits in the MT area still are not unraveled, we have tried to understand their function with a biologically plausible neural network. Being trained with several moving patterns, our un-supervised network learns to detect the direction of movement of moving patterns [6]. When the pattern moves in a certain direction one of the output neurons fires. When the object moves in the opposite direction the neuron stops firing and another neuron fires instead, thus codifying that the direction of movement is reversed.

In this paper we re-examine this network and apply it to a static pattern with a vanishing tail. As will be shown, the network interprets the static pattern as a moving pattern. The network also interprets the direction of the apparent movement according to the direction of the vanishing tail, activating the correct neuron in its output layer.

This paper is organized as follows: in section two we analyze how neural plasticity was implemented in each

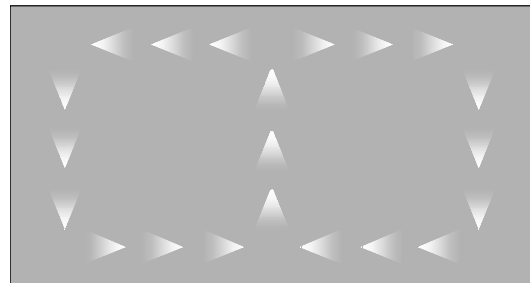


Fig. 2. "Rotating triangles". We have simplified Figure 1 and considered only the essential elements responsible for the illusion of movement.

neuron, in section three we study the network topology and its implementation in Matlab. Results are shown in section four and are commented in section five.

2. MECHANISMS OF NEURONAL PLASTICITY.

Neurons display complex adaptable responses for regulating input stimuli in synapses and the output generated in their soma.

In our network two relevant regulatory or homeostatic mechanisms of neuronal activity were used: metaplasticity and intrinsic plasticity [7]. For a correct understanding of these two mechanisms we will present each in turn.

2.1. Synaptic Plasticity

The transmission of information between neurons is mediated by neurotransmitters that act on synapses, and interact with receptors residing in the synaptic membrane, thereby allowing the inflow or outflow of positive or negative ions such as potassium, sodium, chlorine and calcium. Synaptic plasticity, refers to the modulation of the efficacy of information transmission between neurons, being related to the regulation of the number of ionic channels in synapses.

The first model of synaptic plasticity was postulated by Hebb and is known as the Hebb rule [8], that may be stated as follows: when two neurons fire together they wire together or, in other words, the synaptic strength between neurons with correlated firing tends to increase. Mathematically the change in the synaptic strength (synaptic weight) between neurons i and j is calculated by the product of the output of neuron i , O_i , and the input I_j (which corresponds to the output of neuron j) multiplied by a learning constant.

$$\Delta w_{ij} = \varepsilon O_i I_j \quad (1)$$

Some authors proposed revised versions of Hebb's rule [9], taking into account more recent biological studies. The formulation that was adopted for our simulation of synaptic plasticity, due to its biological plausibility, is Grossberg's presynaptic learning rule [10] either in its incremental or in its probabilistic version.

The incremental version of the pre-synaptic rule is as follows:

$$\Delta w_{ij} = \varepsilon I_j (O_i - w_{ij}) \quad (2)$$

According to Minai [11], this incremental version of the pre-synaptic rule is asymptotically equivalent to the following probabilistic version, where the synaptic weight between two neurons is the conditional probability of the output neuron's firing, given that the input neuron has fired:

$$w_{ij} = P(O_i / I_j) \quad (3)$$

2.2. Synaptic Metaplasticity

One of the important biological characteristics of the presynaptic rule is that it exhibits metaplasticity, which is an important homeostatic mechanism of neurons [7] [12]. Metaplasticity slows down the process of weight increment

or decrement, making it more difficult for the neuron to become either saturated or inactive.

The property of metaplasticity is shown in Fig. 3, in which each curve shows the variation of weight as a function of the neuron's activation. The parameter that defines which curve must be used is the value of the synaptic weight. According to Fig.3, for higher values of the synaptic weight the curves are more elongated to the right.

This means that in synapses with higher weights, the interval in which the variation of weight is negative is broader, thereby favoring synaptic depression. The opposite holds in the lower weight curves.

Synaptic metaplasticity is a homeostatic mechanism because it regulates weight variation, down-regulating weight increment in synapses with high initial weights and up-regulating weight increment in synapses with low initial weights.

Metaplasticity is usually associated with the BCM

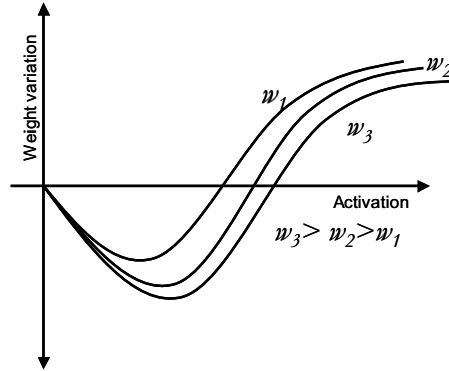


Fig. 3. This family of plasticity curves represents both synaptic plasticity and metaplasticity. They show the increment in weight, Δw , in terms of the neuron's activation. For higher initial synaptic weights, w_i , plasticity curves are more elongated to the right.

model of Bienenstock, Cooper and Munro [13]. In the BCM model, the parameter that defines the curve to be used is the average activation of the neuron, and this value is the same for all of the neuron's synapses. Instead of this, metaplasticity depends on the current synaptic weight in each synapse, so that each different synapse is governed by a different curve. The plasticity curves obtained with the presynaptic rule, either in its incremental or in its probabilistic version, captures well this fact as demonstrated elsewhere [14]. Regarding to this, Mockett and colleagues [15] emphasize that metaplasticity is a homosynaptic phenomenon (restricted to the synapse under study) in contrast to the heterosynaptic nature of the BCM rule, in which even non active synapses are modified according to the average activation of the neuron.

2.3. Intrinsic Plasticity

Although synaptic metaplasticity makes it difficult for synaptic weights to become either null or saturated, it does not totally preclude either of these two extreme situations. To eliminate the possibility of either weight annihilation or saturation, another important homeostatic property of real neurons should be taken into account: the so-called intrinsic plasticity [16]. Intrinsic plasticity regulates the position

(shift) of the neuron's activation function according to the past average level of activity in the neuron. The neuron's activation function is usually modeled as a sigmoidal function:

$$P(O) = \frac{1}{1 + e^{-25(a - shift)}} \quad (4)$$

In which $P(O)$ is the output probability of the neuron and a is the activation given by the sum of synaptic contributions. Intrinsic plasticity was modeled according to the following equation, that yields the position of the sigmoid in terms of the previous position $shift_{t-1}$ and the previous sigmoid output O_{t-1} .

$$shift_t = \frac{\xi O_{t-1} + shift_{t-1}}{1 + \xi} \quad (5)$$

Parameter ξ is a small arbitrary factor ranging in our case from 0.1 to 0.0001. This equation means that the more the neuron keeps firing, the higher will be the rightward shift of the activation function, leading to a moderation of the neurons' firing probability in the future. Conversely, if the firing probability is low, the sigmoid will move leftwards, thereby increasing the probability of the neuron's firing in the future. Parameter ξ works as a learning factor. If we want quick convergence of the shifts with little interest in accuracy, we select a higher ξ . However if we are not worried about the duration of learning and we are mainly interested in accuracy, ξ must be set to a very small value.

3. COMPUTATIONAL SIMULATIONS

In this section we study the neural network architecture

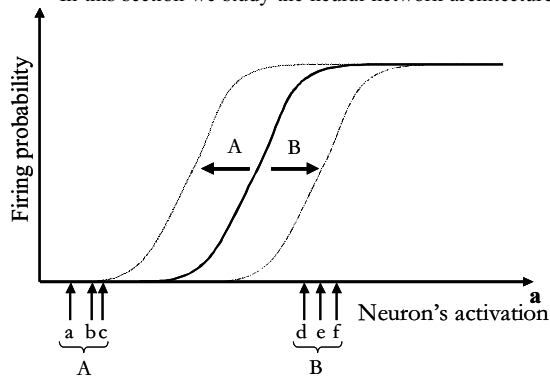


Fig. 4. According to intrinsic plasticity, low neuron's activation (cases a,b,c) contributes to a leftward shift of the sigmoidal activation function. On the contrary high neuron activation (cases d,e and f) contributes to a rightward shift of this sigmoidal function.

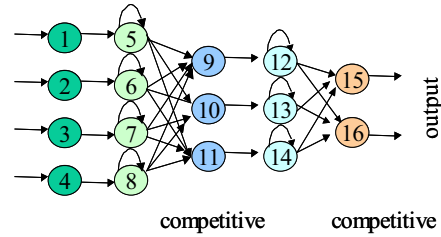


Figure 5. Movement direction identification – Architecture of artificial neural network with Hebbian pre-synaptic learning.

that allows movement direction recognition. Matlab (a commercial software for simulation purposes) was used for the development of the aforementioned architecture and the whole simulation was done on a simple Pentium 4 home desktop computer.

3.1 Network Topology

Figure 5 represents the architecture of the neural network developed in this paper. It is composed by one input layer with 4 neurons, which maps input data into a memory layer. The memory layer's main function is to store past states by a fading mechanism similar to that of pyramidal neurons in the cerebral cortex. Pyramidal neurons fire in a "comet tail" manner: when triggering a burst of action potentials, the initial action potentials are close-together but, at the end of the burst, the action potentials are wide apart [17]. We simulate this behavior by connecting each neuron of the memory layer to itself with a fixed weight value (less than one). At the right of the memory layer, there is a third layer with 3 neurons. This layer performs dimensionality reduction through a competitive process. Therefore, we have two memory layers and two competitive layers, the latter with only 2 neurons. The first competitive layer is intended to encode the trajectories presented in the first layer so they can help in the categorization tasks of the second competitive layer. In each competitive layer only one neuron wins: the one whose output probability is the highest. According to the presynaptic rule, in either the probabilistic or incremental version, the neuron that wins increases its synaptic weights in its active synapses. However non-winning neurons reduce the weights in their active synapses. This process of favoring winning neurons at expense of non-winning ones would lead winning neurons to saturation and non-winning neurons to weight annihilation. The two homeostatic properties for regulating neuron activity, i.e. metaplasticity and intrinsic plasticity collaborate to avoid this deleterious process: metaplasticity, by favoring potentiation of the low synaptic weights and favoring depression of high weights, and intrinsic plasticity by allowing non-winning neuron to win the competition in the future. With this architecture we expect that when the movement of input patterns is in one direction, neuron 15 will fire, whereas when the movement is in the opposite direction, neuron 16 will fire. The architecture of the ANN presented in figure 2 can be represented in matrix notation as shown below. The matrix M is a topological representation where each line corresponds to a neuron (so we have 16 lines) and each column represents whether the neuron A (defined by the line number) receives input from another

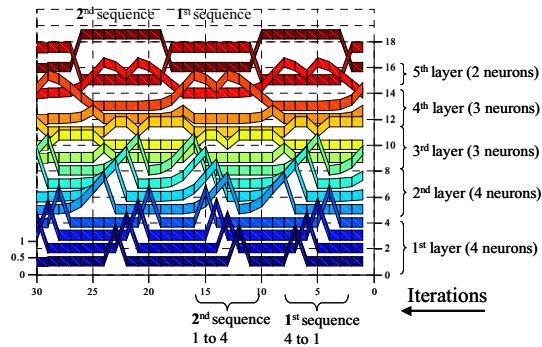


Fig. 6. Each ribbon represents the evolution over time of each of the 16 neurons of the network when neurons in the first layer are activated from right to left (first sequence) and from left to right (second sequence). There is a delay in the transmission of information through the network so that the response to the first sequence appears in the 5th layer 4 iterations later. In this case each time the first sequence is input to the network, the output of neuron 15 is activated.

```

Initialize Ot, with zeros
While condition is true
  Adjust Ot
  update epsilon, according to equation 8
  initialize W, according to equation 3
  calculate the net input, according to equation 9
  update shifting, according to equation 5
  calculate and adjust Ot+1, according to equation 4
End while

```

The initial weights of the competitive neurons were initialized by setting the numerators of the conditional probabilities (see Equation 3) to a random number between 0 and 1 and all the denominators to one. The shifts of the activation functions of all the neurons in the network were initialized to 0.5. Only competitive neurons change their shifts according to Equation 8 in which ξ_i was set to 0.3 and ξ_r to 0.001. The number of iterations was set to $n=10000$. Similar simulations were done by using the incremental version of the presynaptic rule (with a learning

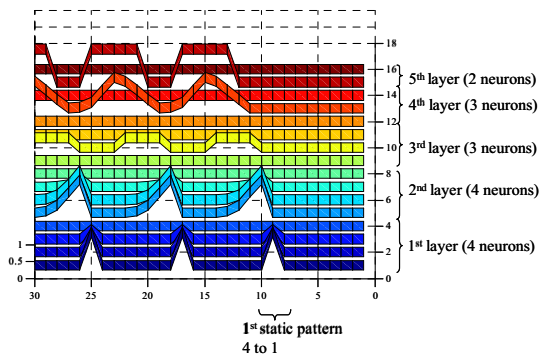


Fig. 7. When the static pattern $[1 \ 0.6 \ 0.3 \ 0]$ is input to the network in the first layer, neuron 15 that served to recognize the first sequence is also activated.

constant set to 0.01). Both formulations yielded consistent results that were almost indistinguishable after a higher number of iterations (above $n=2000$).

Once the neural network is trained with the input patterns described above, we present two types of static patterns for studying the response of the network. The two static patterns are:

First static pattern: $[1 \ 0.6 \ 0.3 \ 0]$. This pattern represents a fading static pattern with a vanishing tail at the right as if a “photo” of a leftward moving pattern were taken.

Second static pattern: $[0 \ 0.3 \ 0.6 \ 1]$. This pattern represents an static pattern with a vanishing tail at the left as if a “photo” of a rightward moving pattern were taken.

The network also worked well varying the speed of the input dot used for training although for the slowest dots we needed more iterations for the network to converge.

4. RESULTS

The results of the simulation confirmed the expected behavior, either with shifting or with static patterns. The case of shifting patterns is depicted in Figure 6. In this figure each ribbon represents the evolution of the outputs of each neuron when either a sequence from right to left or from left to right is presented to the first layer. We can see that the outputs of the neurons in the second and fourth layers have a slower decay because they are self-connected. These neurons mimic the firing adaptation property of pyramidal neurons[17].

The third and fifth layers are competitive, and only one neuron in each layer fires at any one time.

It is important to notice that the response to a certain pattern in the first layer is delayed in the following layers. In this way the response of the neurons in the fifth layer to the

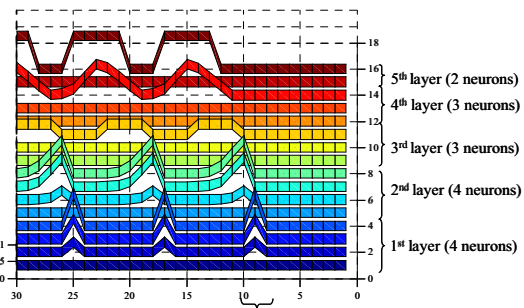


Fig. 8 When the static pattern $[0 \ 0.3 \ 0.6 \ 1]$ is input to the network, neuron 16 that served to recognize the rightward sequence was also activated.

sequences presented in the first layer is delayed by four iterations. Taking this into account, neuron 15 is the one that, after learning, fires every time in the fifth layer when the leftward sequence (the first sequence) is presented. Conversely, neuron 16 fires every time when the rightward sequence is presented. This occurs after an unsupervised learning process that takes place in the network along 10.000 iterations, so that, at the end, leftward and rightward sequences are clearly separated in the last layer.

If a static pattern with a vanishing tail is presented, the network reacts as if a sequence were input to the network. In

Figure 7, the static pattern with a “vanishing tail” [1 0.6 0.3 0] was presented. In this case, neuron 15 that was activated for a leftward sequence was also activated each time the pattern [1 0.6 0.3 0] was presented. Figure 8 corresponds to the case in which the pattern [0 0.3 0.6 1] is presented to the network.

Notice that the orientation of the vanishing tail indicates an apparent rightward movement, and that the neuron that was activated for a rightward movement, neuron 16, is also activated every time this static pattern is presented to the network.

5. CONCLUSIONS

In this work we have implemented a neural network architecture that was able to auto-organize in order to capture the information of a rightward or leftward movement presented as an input to the network. Instead of using a conventional auto-organizing neural network, our network learns by using recently discovered properties of neurons like metaplasticity and intrinsic plasticity. Without these properties, no learning at all would take place in this network. Once the network is trained with moving patterns, the network is deceived by a static pattern with a vanishing tail, the same that occurs with humans when seeing such type of pattern. We hypothesized that this type of illusion takes place due to cinematographic perception by which perception is produced at a sampling rate of around 10 Hz. According to this paradigm, quickly moving objects would be blurred or moved in each perceptual frame, and these blurred patterns could be correlated with moving patterns through learning. We expect that our network would help future studies to elucidate the dynamic of the networks of the middle temporal area (MT area) that, not only are capable of recognizing the direction of movement but are also deceived by static patterns like Kitaoka’s famous designs.

Acknowledgment: This research has been partially supported by the National Spanish Institution “Ministerio de Educación Y Ciencia” as part of the project AGL2006-12689/AGR.

REFERENCES

1. <http://www.ritsumei.ac.jp/~akitaoka/index-e.html>
2. Kitaoka A, Ashida H (2003) Phenomenal characteristics of the peripheral drift illusion. *Vision* 15:261–262.
3. VanReullen, R. & Koch, C. (2003) Is perception discrete or continuous? *Trends in Cognitive Sciences*. 7(5):207-213.
4. VanReullen, R., Reddy, L. & Koch, C. (2005) Attention-driven discrete sampling of motor perception. *Proceedings of the National Academy of Sciences*. 102(14):5292-5296
5. Conway BR, Kitaoka A, Yazdanbakhsh A, Pack CC, and Livingstone MS. (2005) Neural basis for a powerful static motion illusion. *Neuroscience* 25: 5651–5656.
6. Kinto, E.A. & Peláez F.JR, (2007) A Preliminary Neural Model for Movement Direction Recognition Based on Biologically Plausible Plasticity Rules. In *Inspired Problem-Solving Methods in Knowledge Engineering. Lecture Notes in Computer Science*. Vol. 4528/2007, pp. 628-636.
7. Peláez, J.R. & Piqueira, J. R. C. (2006). Biological clues for up-to-date artificial neurons”. In: Diego Andina; Duc Truong Phan. (Org.). *Computational Intelligence: for Engineering and Manufacturing*. Berlin: Springer-Verlag, 2006, v. 1, p. 1-19.
8. Hebb D.O., *The organization of behavior*. New York: Wiley, 1949.
9. Brown T.H., Kairiss E.W., Keenan C.L., *Hebbian Synapses: Biophysical Mechanisms and Algorithms*, *Annual Review in Neurosciences*, Vol. 13, 1990, pp. 475-511
10. Grossberg S, *Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors*. *Biological Cybernetics*, Vol. 23, 1976, pp.121-134.
11. Minai AA, Levy WB, *The Dynamics of Sparse Random Networks*, *Biological Cybernetics*, Vol.70(2), 1993, pp.177-187
12. Desai, N. (2003) Homeostatic plasticity in the CNS: synaptic and intrinsic forms. *Journal of Physiology*. 97:391-402.
13. Bienenstock, E.L., Cooper, L.N.& Munro,P.W. (1982). Theory for the development of neuron selectivity : orientation specificity and binocular interaction in visual cortex. *J Neurosci.*, 2(1), 32-48.
14. Peláez, F. J. R., Simões, M. G. (1999). A computational model of synaptic metaplasticity. In: *Proceedings of the IJCNN 99 International Joint Conference of Artificial Neural Networks*. Washington D.C.
15. Mockett,B., Coussens C., Abraham, W. C. (2002). NMDA receptor-mediated metaplasticity during the induction of long-term depression by low-frequency stimulation. *European Journal of Neuroscience*, 15, 1819-1826.
16. Desai, N. S., Rutherford, L. C. and Turrigiano, G. (1999) Plasticity in the intrinsic excitability of cortical pyramidal neurons. *Nature Neuroscience*, 2(6):515-520.
17. Shepherd, G.M.(1998). *The synaptic organization of the brain*. Oxford University Press. New York. Pag.437.