

Oscillatory Dynamic Link Matcher: A Bio-Inspired Neural Network for Pattern Recognition

Ramin Pichevar and Jean Rouat

Department of Electrical and Computer Engineering, Université de Sherbrooke, QC, Canada
 emails: Ramin.Pichevar@usherbrooke.ca, Jean.Rouat@usherbrooke.ca

Abstract—In this paper we show that an unsupervised two-layered oscillatory neural network with intra-layer connections, and a learning rule based on stimulus difference can behave as a Dynamic Link Matching Machine for invariant pattern recognition. We show that this architecture is robust to affine transformations. We call this architecture Oscillatory Dynamic Link Matching (ODLM).

I. INTRODUCTION

Pattern recognition robust to noise, symmetry, homothety (size change with angle preservation), etc. has long been a challenging problem in artificial intelligence. Many solutions or partial solutions to this problem have been proposed using expert systems or neural networks. In general three different approaches are used to perform invariant pattern recognition:

- **Normalization:** In this approach the analyzed object is normalized to a standard position and size by an internal transformation. The advantages of this approach are: The coordinate information (the "where" information) is retrievable at any stage of the processing and there is a minimum loss of information. The disadvantage of this approach is that the network should find the object in the scene and then normalize it. This task is not as obvious as it can appear [1] [2].
- **Invariant Features:** In this approach some features that are invariant to the location and the size of an object are extracted. The disadvantages of this approach is that the position of the object may be difficult to extract after recognition and information is lost during the process. The advantage is that the technique doesn't require to know where the object is and unlike normalization in which other techniques should be used after this stage to recognize patterns, the invariant features approach already does some pattern recognition by finding important features [3].
- **Invariance Learning from temporal input sequences:** The assumption is that primary sensory

signals, which in general code for local properties, vary quickly while the perceived environment changes slowly. If one succeeds in extracting slow features from the quickly varying sensory signal, he/she is likely to obtain an invariant representation of the environment [4] [5].

Based on the *Normalization* approach, the "dynamic link matching" (DLM) has been first proposed by Konen et al. [2]. This approach consists of two layers of neurons connected to each other through synaptic connections constrained to some normalization. The saved pattern is applied to one of the layers and the pattern to be recognized to the other. The dynamics of the neurons are chosen in such a way that "blobs" are formed randomly in the layers. If the features in these two blobs are similar enough, some weight strengthening and activity similarity will be observed between the two layers, which can be detected by correlation computation [2] [6]. These blobs can or cannot correspond to a segmented region of the visual scene, since their size is fixed in the whole simulation period and is chosen by some parameters in the dynamics of the network [2]. The apparition of blobs in the network has been linked to the attention process present in the brain by the developers of the architecture. The dynamics of the neurons used in the original DLM network is not the well-known spiking neuron dynamics. In fact, its behavior is based on rate coding (average neuron activity over time, for details see section V) and can be shown to be equivalent to an enhanced dynamic Kohonen Map in its Fast Dynamic Link Matching (FDLM) form [2]. Here, we propose the Oscillatory Dynamic Link Matching algorithm (ODLM), which uses conventional spiking neurons and is based on phase (place) coding. The network is capable of doing motion analysis, but neither it computes optical flow nor it performs additional signal processing between the layers, unlike in [7]. In a more general way, our proposed network can solve the correspondence problem, and at the same time, perform

the segmentation of the scene, which is in accordance with the Gestalt theory of perception [8] and it is very useful when pattern recognition should be done in multiple-object scenes. In other words the network does normalization, segmentation, and pattern recognition at the same time. It is also self-organized. In addition, if only one object is present in the scene the segmentation phase can be bypassed, if the speed of convergence is the only concern (section VI). The application of this network is not limited to visual scene analysis, it can be used in sound source segregation problem and may act as a top-down (schema-driven) processor in the Computational Auditory Scene Analysis (CASA) [9].

II. THE OSCILLATORY DYNAMIC LINK MATCHER

The building blocks of this network are oscillatory neurons [10]. The dynamics of this kind of neurons is governed by a modified version of the Van der Pol relaxation oscillator (called the Wang-Terman oscillator) (for a similar approach with different dynamics see [11]). There is an active phase when the neuron spikes and a relaxation phase when the neuron is silent. The dynamics of the neurons follows the following state-space equations, where x_i is the membrane potential (output) of the neuron and y_i is the state for channel activation or inactivation.

$$\frac{dx_{i,j}}{dt} = 3x_{i,j} - x_{i,j}^3 + 2 - y_{i,j} + \rho + H(p_{i,j}^{input}) + S_{i,j} \quad (1)$$

$$\frac{dy_{i,j}}{dt} = \epsilon[\gamma(1 + \tanh(x_{i,j}/\beta)) - y_{i,j}] \quad (2)$$

ρ denotes the amplitude of a Gaussian noise, $p_{i,j}^{input}$ the external input to the neuron, and $S_{i,j}$ the coupling from other neurons (connections through synaptic weights). ϵ , γ , and β are constants. Initial values are generated by a uniform distribution between the interval $[-2; 2]$ for $x_{i,j}$ and between $[0; 8]$ for $y_{i,j}$ (these values correspond to the whole dynamic range of the equations).

A neighborhood of 4 is chosen in each layer for the connections. Each neuron in the first layer is connected to all neurons in the second layer and vice-versa. A global controller is connected to all neurons in the first and second layers as in [12]. In a first stage, segmentation is done in the two layers independently (with no extra-layer connections) as explained in section III, while dynamic matching is done with both intra-layer and extra-layer couplings. The intra-layer and extra-layer connections are defined as follows:

$$w_{i,j,k,m}^{int}(t) = \frac{w_{max}^{int}}{\text{Card}\{N^{int}(i,j) \cup N^{ext}(i,j)\}} \cdot \frac{1}{e^{\lambda|p(i,j;t)-p(k,m;t)|}} \quad (3)$$

$$w_{i,j,k,m}^{ext}(t) = \frac{w_{max}^{ext}}{\text{Card}\{N^{ext}(i,j) \cup N^{int}(i,j)\}} \cdot \frac{1}{e^{\lambda|p(i,j;t)-p(k,m;t)|}} \quad (4)$$

where $w_{i,j,k,m}^{int}(t)$ are intra-layer connections and $w_{i,j,k,m}^{ext}(t)$ are extra-layer connections (between the two layers) and $w_{max}^{int} = 0.2$ and $w_{max}^{ext} = 0.2$ are constants equal to the maximum value of the synaptic weights. $\text{Card}\{N^{int}(i,j)\}$ is a normalization factor and is equal to the cardinal number (number of elements) of the set $N^{int}(i,j)$ containing neighbors connected to the $neuron(i,j)$ and can be equal to 4, 3 or 2 depending on the location of the neuron on the map, i.e. center, corner, etc., and the number of active connections. A connection is active when $H(w_{i,j,k,m} - 0.01) = 1$, which is true both for intra-layer and extra-layer connections. $\text{Card}\{N^{ext}(i,j)\}$ is the cardinal number for extra-layer connections and is equal to the number of neurons in the second layer with active connection to $neuron_{i,j}$ in the first layer. Note that normalization in Eq. 4 is mandatory if someone wants to correspond similar pictures with different sizes. If the aim is to match objects with exactly the same size the normalization factor should be set to a constant for all neurons. The reason for this is that with normalization even if the size of the picture in the second layer was the double of the same object in the first layer the total influence to the $neuron_{i,j}$ would be the same as if the pattern was of the same size.

The schematic of the network is shown in Fig. 1.

III. BEHAVIORAL DESCRIPTION OF THE NETWORK

The network has two different behavioral mode: segmentation and matching.

- **Segmentation:** In the segmentation stage, there is no connection between the two layers. The two layers act independently (unless for the influence of the global controller) and segment the two images applied to the two layers respectively. The global controller forces the segments on the two layers to have different phases. At the end of this stage, the two images are segmented but no two segments have the same phase (Fig. 4). The results from segmentation are used to create binary masks that select one object in each layer in multi-object scenes. In fact a snapshot like the one shown in

Fig. 10 is used to create the binary mask $m(i, j)$ for one of the objects as follows:

$$m(i, j) = \begin{cases} 1 & \text{for } x_{i,j}(t_{sync}) = x_{sync} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

x_{sync} can be the synchronized value that corresponds to either the cross or the rectangle in Fig. 10 at time t_{sync} .

The coupling strength $S_{i,j}$ for each layer as defined in Eq. 1 is computed by :

$$S_{i,j}(t) = \sum_{k,m \in N^{int}(i,j)} w_{i,j,k,m}^{int}(t) H(x^{int}(k, m; t)) - \eta G(t) \quad (6)$$

$H(\cdot)$ is the Heaviside function, $G(t)$ is the influence of the global controller defined by the following equation. η should be set to a value smaller than the maximum value of synaptic weights, i.e. 0.25 in our case.

$$G(t) = \alpha H(z - \theta) \quad (7)$$

$$\frac{dz}{dt} = \sigma - \xi z \quad (8)$$

σ is equal to 1 if the global activity of the network is greater than a predefined ζ and is zero otherwise.

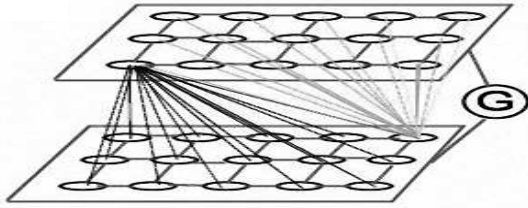


Fig. 1. The architecture of the oscillatory dynamic link matcher. The number of neurons in the figure does not correspond to the real number of neurons. The global controller has bidirectional connections to all neurons in the two layers

- **Dynamic Matching:** In the matching phase, the external input to the layers are defined by the binary masks generated in the segmentation phase. The input to the layers are defined by:

$$p_{i,j}^{matching} = m(i, j) p_{i,j}^{input} \quad (9)$$

Extra-layer connections (Eq.4) are established. If there are similar objects in the two layers, these extra-layer connections will help them synchronize. In other words, these two segments are bound together through these extra-layer connections [13]. In order to detect synchronization double-thresholding can be used [14]. This stage may be seen as a folded oscillatory texture segmentation

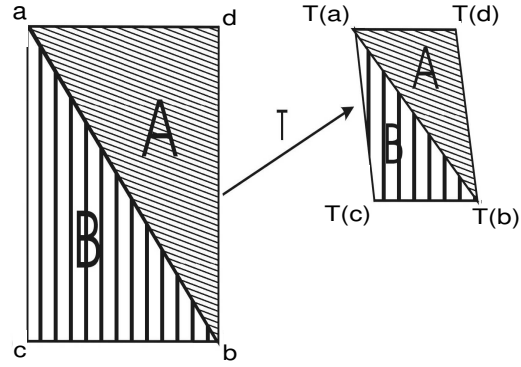


Fig. 2. An affine transform T for a four-corner object.

device as the one proposed in [10]. The coupling strength $S_{i,j}$ for each layer in the matching phase is defined as follows :

$$S_{i,j}(t) = \sum_{k,m \in N^{ext}(i,j)} \{w_{i,j,k,m}^{ext}(t) H(x^{ext}(k, m; t)) + w_{i,j,k,m}^{int}(t) H(x^{int}(k, m; t))\} - \eta G(t) \quad (10)$$

IV. GEOMETRICAL INTERPRETATION OF THE ODLM

We know that an object can be represented by a set of points corresponding to its corners, and any affine transform is a map $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ of these points defined by the following matrix operation

$$\mathbf{p}' = \mathbf{A} * \mathbf{p} + \mathbf{t} \quad (11)$$

Where \mathbf{A} is a 2×2 non-singular matrix, $\mathbf{p} \in \mathbf{R}^2$ is a point in the plane, and \mathbf{p}' is its affine transform. \mathbf{t} is the translation vector. The transform is linear if $\mathbf{t} = \mathbf{0}$. Affine transformation is a combination of several simple mappings such as rotation, scaling, translation, and shearing. The similarity transformation is a special case of affine transformation. It preserves length ratios and angles while the affine transformation, in general does not. In this paragraph we show that the coupling $S_{i,j}$ is independent of the affine transform used. We know that any object can be shattered into its constituent triangles (three corners per triangle). Now suppose that the set $\{a, b, c, d\}$ is mapped to the set $\{T(a), T(b), T(c), T(d)\}$, and that the objects formed by these two sets of points are applied to the two layers of our neural network. Suppose also that points inside the triangle $\{a, b, c\}$ (resp. $\{T(a), T(b), T(c)\}$) have values equal to A (corresponding to the gray-level value of the image at that points) and points inside $\{a, b, d\}$ (resp. $\{T(a), T(b), T(d)\}$) have values equal to B.

We know that for an affine transform (Fig. 2):

$$\frac{\Delta_{abc}}{\Delta_{abd}} = \frac{\Delta_{T(abc)}}{\Delta_{T(abd)}} \quad (12)$$

Where Δ_{abc} is the area of the triangle $\{a, b, c\}$ (expressed in number of neurons). For $neuron_{i,j}$ belonging to $\{a, b, c\}$ and $neuron_{k,m}$ belonging to $\{T(a), T(b), T(c)\}$, Eq.4 is equivalent to (neglecting the effect of intra-layer connections, since $N^{ext} \gg N^{int}$):

$$N^{ext} = \Delta_{T(abc)} + \Delta_{T(abd)} \quad (13)$$

Hence,

$$w_{i,j,k,m}^{ext}(t) = \frac{f(p(i, j; t) - p(k, m; t))}{\Delta_{T(abc)} + \Delta_{T(abd)}},$$

with $f(x - y) = \frac{w_{max}^{ext}}{e^{\lambda|x-y|}} \quad \forall x, y \quad (14)$

There are $\Delta_{T(abc)}$ connections from the region with gray-level value A (triangle $\{T(a), T(b), T(c)\}$) and $\Delta_{T(abd)}$ connections from the region with gray-level value B (triangle $\{T(a), T(b), T(d)\}$) to the $neuron_{i,j}$ belonging to the triangle $\{a, b, c\}$ with gray-level value A. Therefore, the external coupling for $neuron_{i,j}$ from all $neuron_{k,m}$ becomes :

$$S_{i,j}(t) = \frac{\Delta_{T(abc)}f(A - A)\psi(t, \phi_1)}{\Delta_{T(abc)} + \Delta_{T(abd)}} + \frac{\Delta_{T(abd)}f(A - B)\psi(t, \phi_2)}{\Delta_{T(abc)} + \Delta_{T(abd)}},$$

with $\psi(t, \phi) = H(x_{k,m}^{ext}(t)) \quad (15)$

Where $\psi(t, \phi_2)$ and $\psi(t, \phi_1)$ are respectively associated to spikes with phases ϕ_2 and ϕ_1 that appear after segmentation. After factorization and using Eq. 12 we obtain:

$$S_{i,j}(t) = \frac{f(0)\psi(t, \phi_1)}{1 + \frac{\Delta_{abd}}{\Delta_{abc}}} + \frac{f(A - B)\psi(t, \phi_2)}{1 + \frac{\Delta_{abc}}{\Delta_{abd}}} \quad (16)$$

This means that the extra-layer connections are independent of the affine transform that maps the model to the scene (first and second layer objects) and can be extended to more than 4 points.

Note that if there are several objects in the scene and we want to match patterns, we can use the results from the segmentation phase to break the scene into its constituent parts (each synchronized region corresponds to one of the objects in the scene) and apply the objects one by one to the network, until all combinations are tested. This is not possible in the averaged Dynamic Link Matching case where no segmentation occurs.

V. RATE CODING VS. PHASE CODING

The aim in this paragraph is to show that the original DLM is a rate coding approximation of the ODLM. Aoinishi et al. [6] have shown that a canonical form of rate coding dynamic equations solve the matching problem in the mathematical sense. The dynamics of a neuron in one of the layers of the original Dynamic Link Matcher proposed in [2] is as follows:

$$\frac{dx}{dt} = -\alpha x + (k * \sigma(x)) + I_x \quad (17)$$

Where $k(\cdot)$ is a neighborhood function, I_x is the summed value of extra-layer couplings, σ is the sigmoidal function, x is the output of the rate coded neuron, and $*$ is the convolution. On the other hand, we know that the Wang-Terman oscillator can be approximated by the Integrate-and-Fire neuron (for details see [15]), we can write for a single neuron of our network:

$$\begin{aligned} \frac{dx^{two}}{dt} &= -x^{two} + \sum_{k,m \neq i,j} w_{i,j,k,m}^{int} H(x_{k,m}^{two}) \\ &+ \sum_{k,m} w_{i,j,k,m}^{ext} H(x_{k,m}^{one}) + H(p^{input}) \\ x &= 0 \quad x > threshold \end{aligned} \quad (18)$$

Where x^{two} stands for neurons in layer two and x^{one} stands for neurons in layer one. There are synaptic connections (w^{int}) in layer 2 and synaptic connections from layer 1 to layer 2 (w^{ext}).

If we neglect the influence of intra-layer connections, therefore Eq. 18 becomes:

$$\begin{aligned} \frac{dx^{two}}{dt} &= -x^{two} + \sum_{k,m} w_{i,j,k,m}^{ext} H(x_{k,m}^{one}) + H(p^{input}) \\ x &= 0 \quad x > threshold \end{aligned} \quad (19)$$

Note that for an integrate-and-fire neuron the approximation $H(x) = x$ holds, since the output of an integrate-and-fire neuron is either 0 or 1 (it emits spikes or delta functions), therefore Eq. 19 can be further simplified to :

$$\begin{aligned} \frac{dx^{two}}{dt} &= -x^{two} + \sum_{k,m} w_{i,j,k,m}^{ext} x_{k,m}^{one} + H(p^{input}) \\ x &= 0 \quad x > threshold \end{aligned} \quad (20)$$

By averaging the two sides of Eq. 20 we get: ($H(p^{input})$ is considered constant over T) :

$$\begin{aligned} \frac{dx_a^{two}}{dt} &= -x_a^{two} + \sum w^{ext} x_a^{one} \\ &+ H(p^{input}) \\ x_a &= \langle x \rangle_T = \frac{1}{T} \int_0^T x(t) dt \end{aligned} \quad (21)$$

$\langle x \rangle_T$, the averaged version of x over a time window of length T . For the sake of simplicity, the indices are omitted in Eq. 21.

From [16], we know that the averaged output x_a^{two} of an integrate-and-fire neuron is related to the averaged-over-time inputs of a neuron ($\Sigma w^{ext} x_a^{one}$) by a continuous function (sigmoidal, etc.). Let name this function φ (note that β is a proportionality constant):

$$\langle x_{i,j}^{two} \rangle = \beta \varphi(\Sigma w^{ext} \langle x_{k,m}^{one} \rangle) \quad (22)$$

Note that in Eq. 21 we need $\langle x_{i,j}^{one} \rangle$ in function of $\langle x_{k,m}^{two} \rangle$. Note further that Eq. 22 is a set of linear equations in w^{int} and we can deduce $x_{i,j}^{one}$ from that sets of equations:

$$x_{i,j}^{one} = \Sigma_{k,m} \sigma(x_{k,m}^{two}) \quad (23)$$

Replacing the above result in Eq. 21 gives (note that for the sake of simplicity we omitted again the indices):

$$\frac{dx_a^{two}}{dt} = -x_a^{two} + \Sigma \Sigma w^{ext} \sigma(x_a^{two}) + H(p^{input}) \quad (24)$$

Where $\sigma(x) = \varphi^{-1}(x)$.

On the other hand:

$$\Sigma \Sigma w^{int} \sigma(x_a^{two}) = k(x_a^{two}) * \sigma(x_a^{two}) \quad (25)$$

Where $*$ is a 2-D convolution. In our case $k(\cdot)$ is a 2-D rectangular window (in the original DLM $k(\cdot)$ was chosen to be Mexican hat).

Hence, we have proved that the DLM is an averaged-over-time approximation of the ODLM.

VI. RESULTS

As stated earlier, this network can be used to solve the correspondence problem. For example, suppose that in a factory chain, someone wants to check the existence of a component on an electronic circuit board. All he/she has to do is to put an image of the component on the first layer and check for synchronization between the layers. Ideally, any change in the angle or the location of the camera or even the zoom factor should not influence the result. One of the signal processing counterparts of our proposed technique is the morphological processing. Other partial solutions such as the Fourier (resp. Mellin) transform could be used to perform matching robust to translation (resp. scaling). There is no need to train or configure our architecture to the stimulus we want to apply. The network is autonomous and flexible to not previously seen stimuli. This is in contrast with associative memory based architectures in which a stimulus must be applied and saved into memory before retrieval [5]. It doesn't require any pre-configured architecture adapted to the stimulus, like in the hierarchical coding paradigm [17]. DLM can play an important role in structuring memory, e.g. finding structural similarities between stored information during sleep [18].

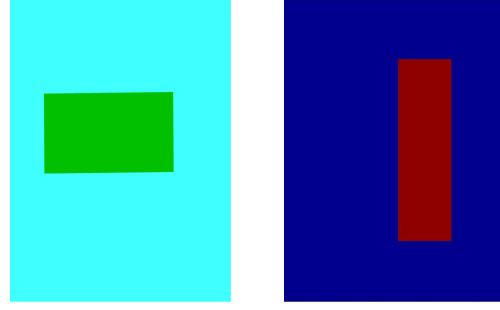


Fig. 3. A snapshot of the activity the first and second layers of the neural map. Colors represent relative phase of oscillations.

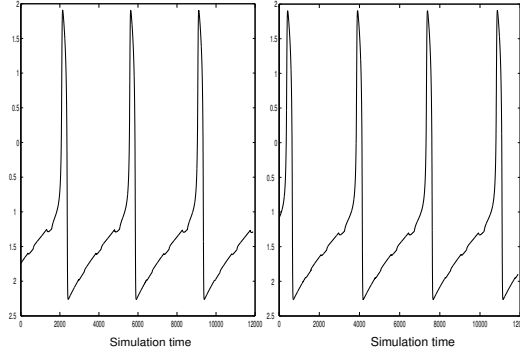


Fig. 4. Left: Activity of one of the neurons associated with the vertical bar in the first layer after segmentation. Right: Activity of one of the neurons associated with the background in the same layer.

In this paper, we show the aforementioned capacities of the network using a prototype that will help us study the dynamics of the network.

A. Segmentation and Matching for Invariant Pattern Recognition

A rectangular neuron map is chosen. There are 5x5 neurons in each layer. A vertical bar in a background is presented in the first layer. The second layer receives the same object transformed by an affine transformation (rotation, translation, etc.). Here are some examples: Figure 3 shows an activity snapshots (instantaneous values of $x(i, j)$) in the two layers after segmentation (first phase). Note that same-colored neurons have similar phases in the figure. On the other hand, different segments on different layers are desynchronized (see Figures 4 and 5). In the dynamic matching stage, similar objects among different layers are synchronized (Figure 7). The thresholded sum (synchronization index) of the activity of all neurons ($\sum_{i,j} H(x(i, j) - 0.5)$) is shown in Figure 6 for the segmentation phase and in Figure 7 for the dynamic matching phase. Since there are four different regions in the two layers with different phases at the end of the segmentation phase,

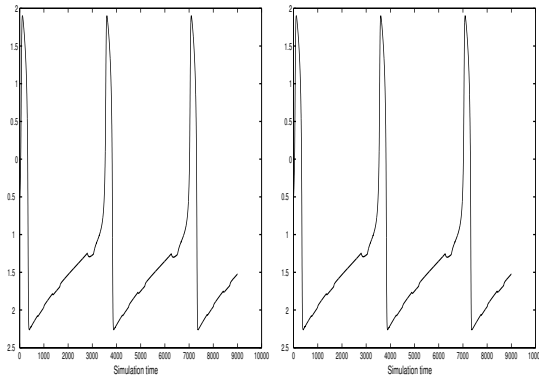


Fig. 5. Left: Activity of one of the neurons associated with the horizontal bar in the first layer after dynamic matching. Right: Activity of one of the neurons associated with the vertical bar after dynamic matching in the second layer.

four different synchronization regions can be seen in Figure 6. In the dynamic matching phase, the similar objects (and the backgrounds) merge with each other producing only two distinct regions. In addition, when a zero-mean Gaussian noise with variance $\sigma^2 = 0.1$ is added to both stimuli ($SNR = 10dB$) the matching results remain unchanged.

B. One-object scenes

Note that if only one object is present in each layer of the scene, then the segmentation phase can be bypassed and the network could function directly in the matching mode. This strategy will help us speed up the pattern recognition process. Fig. 8 and Fig. 9 show the behavior of a 13x5 network when only one object is present in each layer. The synchronization time for the matching-only network is shorter. Note that the matching-only approach cannot be used, if there are multiple objects in the scene. In the latter-mentioned case the segmentation plus matching approach should be used.

VII. CONCLUSION AND FURTHER WORK

We proposed the oscillatory dynamic link matching as a mean to segment images and solve the correspondence problem, as a whole system, using a two-layered oscillatory neural network. We showed that our network is capable of establishing correspondence between images and is robust to translation, rotation, noise and homothetical transforms. More experiments with complex objects and more general transforms like shearing, etc. are under investigation. Pattern recognition of occluded objects is another challenge for this proposed architecture and will be presented in further works.

We are investigating the possibility of the insertion of this architecture in our bottom-up sound segregator [9] [19] as a top-down processor. In

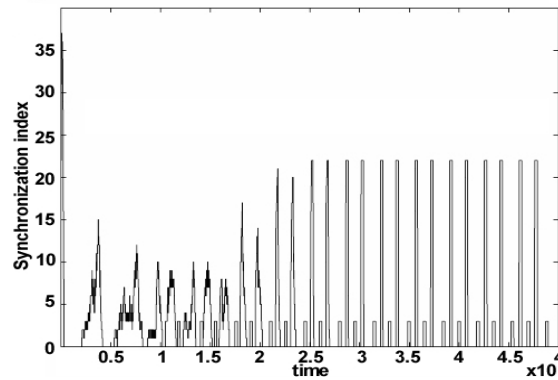


Fig. 6. The evolution of the thresholded activity of network through time in the segmentation phase. Each vertical rod represents a synchronized ensemble of neurons and the number of neurons in that synchronized region is represented on the vertical axis.

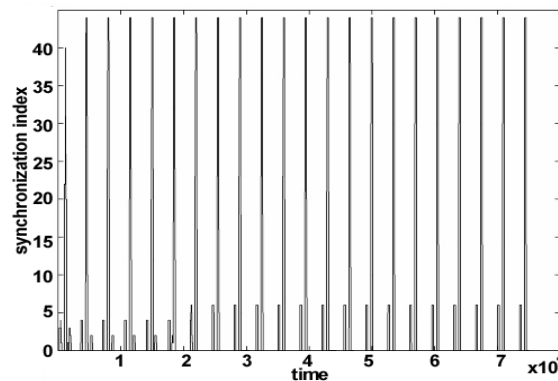


Fig. 7. The evolution of the thresholded activity of the network through time in the dynamic matching phase.

fact, in this application, visual images will be replaced by CAM (Cochleotopic/AMtopic) and CSM (Cochleotopic/Spectrotopic) Maps proposed in [9]. The approach could also be used as a separate discrete-word recognizer.

VIII. ACKNOWLEDGMENTS

Many thanks also to DeLiang Wang for fruitful discussions on the dynamics of the relaxation oscillators. This work was partly supported financially by NSERC, the UQAC foundation, and the University of Sherbrooke.

REFERENCES

- [1] E.O. Postma, H.J. Van der Herik, and P.T. W. Hudson. SCAN: A scalable neural model of covert attention. *Neural Networks*, 10:993–1015, 1997.
- [2] W. Konen, T. Maurer, and C. Von der Malsburg. A fast dynamic link matching algorithm for invariant pattern recognition. *Neural Networks*, pages 1019–1030, 1994.
- [3] K. Fukushima. A neural network model for selective attention in visual pattern recognition. *Biol. Cybernetics*, pages 5–15, 1986.
- [4] L. Wiskott and T. Sejnowski. Slow feature analysis: Unsupervised learning of invariances. *Neural Computation*, pages 715–770, 2002.

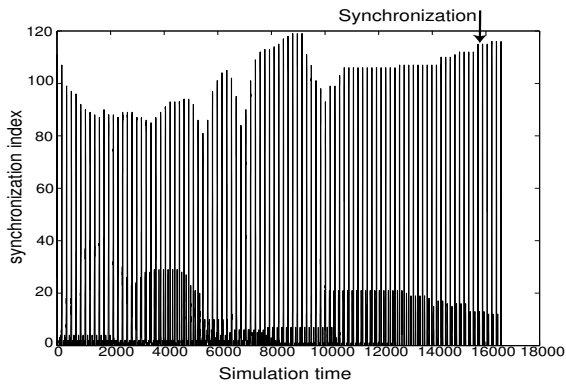


Fig. 8. The Synchronization index of a one-object scene when the segmentation step is bypassed. The synchronization takes 85 oscillations (spikes)

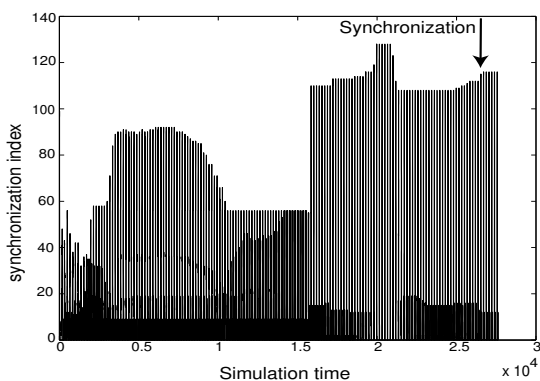


Fig. 9. The synchronization pattern of a one-object scene when the segmentation phase precedes the matching phase. The synchronization takes 155 oscillations (spikes).

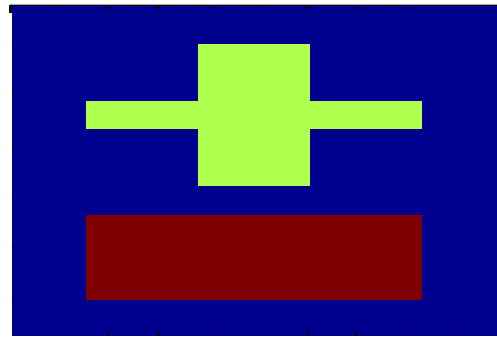


Fig. 10. A scene segmentation done during the segmentation phase of the algorithm. Colors represent synchronization phase. Binary masks are generated by assigning binary values to different oscillation phases (Eq. 5).

nonlinear oscillator network circuit for image segmentation with double-threshold phase detection. In *ICANN 99*, 1999.

- [15] R. Pichevar. *Speech Processing in the Presence of "Cocktail Party" Effect and its Applications in Information Technology*. PhD thesis, University of Sherbrooke (to appear), 2004.
- [16] W. Maass and C. M. Bishop. *Pulsed Neural Networks*. MIT Press, 1998.
- [17] M. Reisenhuber and T. Poggio. Are cortical models really bound by the binding problem? *Neuron*, 24:87–93, 1999.
- [18] L. Wiskott, C. Von der Malsburg, and A. Weitzenfeld. *The Neural Simulation Language: A System for Brain Modeling*, chapter 18, pages 343–372. MIT Press, 2002.
- [19] R. Pichevar and J. Rouat. Binding of audio elements in the sound source segregation problem via a two-layered bio-inspired neural network. In *IEEE CCECE'2003, Montreal, mCanada*.

- [5] T. Vinh Ho and J. Rouat. Novelty detection based on relaxation time of a network of integrate-and-fire neurons. In *IEEE Int'l Joint Conference on Neural Networks, Alaska, USA*, 1998.
- [6] T. Aoinishi, K. Kurata, and T. Mito. A phase locking theory for matching common parts of two images by dynamic link matching. *Biological Cybernetics*, 78(4):253–264, 1998.
- [7] X. Zhang and A. Minai. Detecting corresponding segments across images using synchronizable pulse-coupled neural networks. In *IJCNN2001*, 2001.
- [8] L.E. Gordon. *Theories of Visual Perception*. John Wiley and Sons, 1997.
- [9] R. Pichevar and Jean Rouat. Cochleotopic/AMtopic (CAM) and Cochleotopic/Spectrotopic (CSM) map based sound source separation using relaxation oscillatory neurons. In *IEEE Neural Networks for Signal Processing Workshop, Toulouse, France*, 2003.
- [10] D.L. Wang and D. Terman. Image segmentation based on oscillatory correlation. *Neural Computation*, pages 805–836, 1997.
- [11] R. M. Borisyuk and Y. Kazanovich. Oscillatory neural network model of attention focus formation and control. *Biosystems*, 71:29–36, 2003.
- [12] E. Csemeli and D. Wang. Motion segmentation based on motion/brightness integration and oscillatory correlation. *IEEE Trans. on neural networks*, 11(4):935–947, 2000.
- [13] C. Von der Malsburg. The what and why of binding: The modeler's perspective. *Neuron*, pages 95–104, 1999.
- [14] H. Ando, N. Takashi Morie, M. Nagata, and A. Iwata. A