Oscillatory Dynamic Link Matching for Pattern Recognition

Ramin Pichevar and Jean Rouat
Département GEI, Université de Sherbrooke, J1K 2R1, QC, Canada
ERMETIS, Université du Québec, Chicoutimi, G7H 2B1, QC, Canada
pichevar@hermes.usherb.ca, jean.rouat@ieee.org

ABSTRACT

1 Introduction

The "dynamic link matching" (DLM) has been first proposed by Konen et al. [1] to solve the visual correspondence problem. The approach consists of two layers of neurons connected to each other through synaptic connections constrained to some normalization. The reference 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 of two *blobs*, each belonging to a different layer, are sufficiently similar, then weight strengthening between the *blobs* and activity similarity will be observed between sets of similar *blobs*. The size of the *blobs* remains fixed during all the simulation. In the original DLM network, the behavior is based on rate coding (averaged neuron activity over time is encoded).
Here, we propose the Oscillatory Dynamic Link Matching algorithm (ODLM) that uses models of conventional spiking neurons and for which the coding is based on phase (place coding). We observe that the network is capable of performing motion analysis without optical flow computation and no additional signal processing should be made between layers unlike in [2] (translation, rotation, etc. between the patterns in the first and second layers can be seen as motion). More generally, our proposed network can solve the correspondence problem, and at the same time, performs the segmentation of the scene, which is in accordance with the Gestalt theory of perception.

2 The oscillatory dynamic link matcher

The building blocks of this network are oscillatory neurons [3]. 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). There is an active phase when the neuron spikes and a relaxation phase when the neuron is silent.
A neighborhood of 4 is chosen in each layer for the connections (Fig. 1). A global controller is connected to all neurons in the first and second layers as in [4]. During the first processing stage, the two layers are not connected and image segmentation is performed in each layer. Then, during the pattern matching stage, each neuron in the first layer is connected to all neurons in the second layer and vice-versa (extra-layer connections). The intra and extra-layer connection weights $w_{i,j,k,m}(t)$ are defined as follows:

$$w_{i,j,k,m}(t) = \frac{0.25}{Card\{N^{int}(i,j) \cup N^{ext}(i,j)\} e^{\lambda |p(i,j;t)-p(k,m;t)|}} \tag{1}$$

where $p(i,j;t)$ is the external input to $neuron_{i,j}$. $Card\{N^{int}(i,j)\}$ is a normalization factor and is equal to the cardinal number (number of elements) of the set $N^{int}(i,j)$ that comprises the neighbor neurons connected to $neuron_{i,j}$. It is equal to 4, 3 or 2 depending on the location of the neuron on the map (i.e. center, corner, etc.) and on the number of active connections. A connection is active when $H(w_{i,j,k,m} - 0.01) = 1$, and is true both for intra-layer and extra-layer connections. $Card\{N^{ext}(i,j)\}$ 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.

3 Behavioral description of the network

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

- **Segmentation:** During 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 will force the segments on the two layers to have different phases. At the end of this stage, no two objects (on either layers) have the same synchronization phase.

  The coupling influence $S_{i,j}$ for each layer is computed by :

$$S_{i,j}(t) = \sum_{k,m \in N^{int}(i,j)} w^{int}_{i,j,k,m}(t) H(x^{int}(k,m;t)) - \eta G(t) \quad (2)$$

  $H(.)$ is the Heaviside function, $G(t)$ is the coupling of the global controller [3]. η is a constant and $x^{int}(k,m;t)$ is the output potential of the $neuron_{k,m}$.

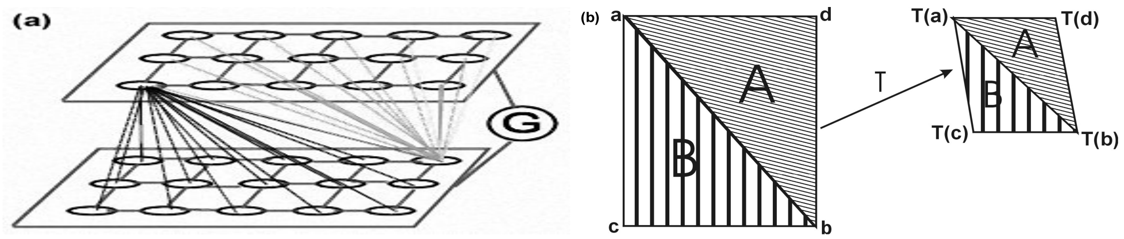

Figure 1: **(a)**: The architecture of the oscillatory dynamic link matcher. The number of neurons in the figure does not correspond to the real number of neurons. The global controller has bidirectional connections to all neurons in the two layers. **(b)**: An affine transformation T for a four-corner object.

- **Dynamic Matching:** During the matching phase, extra-layer connections (Eq. 1) are established. If there are similar objects in the two layers, these extra-layer connections will help the synchronization of neurons between layers. In other words, these two segments are bound together through these extra-layer connections . This stage may be seen as a folded oscillatory texture segmentation device as the one proposed in [3]. The coupling $S_{i,j}$ for each layer during the matching phase is defined as follows :

$$S_{i,j}(t) = \sum_{k,m \in N^{ext}(i,j)} w^{ext}_{i,j,k,m}(t) H(x^{ext}(k,m;t)) + \sum_{k,m \in N^{int}(i,j)} w^{int}_{i,j,k,m}(t) H(x^{int}(k,m;t)) \quad (3)$$

  Note that since no segregation (phase discrepancy) is required in the second phase, there is no global controller activity.

Rate coding vs. Place coding: The aim in this paragraph is to show that the original DLM is an averaged coding approximation of the ODLM. Aoinishi et al. [5] 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 [1] is as follows:

$$\frac{dx}{dt} = -\alpha x + (k * \sigma(x)) + I_x \quad (4)$$

Where $k(.)$ is a neighborhood function, I_x is the summed value of extra-layer couplings, σ is the sigmoidal function, and x is the output of the rate coded neuron. On the other side, the Integrate-and-Fire equivalent of our network (knowing that our relaxation oscillator has the same behavioral dynamics as the integrate-and-fire neurons), we can write for a single neuron of our network:

$$\frac{dV_{i,j}}{dt} = -V_{i,j} + \Sigma_{k,m} w^{int}_{i,j,k,m}(V_{ref} - V_{k,m}) + I_V \tag{5}$$

Where V is the potential (output) of the neuron and is reset when it crosses a threshold, and I_V is the summed value of extra-layer couplings. We know that there is a correspondence between the dependence of the output value $\sigma(\Sigma_{k,m} w^{int}_{i,j,k,m} x_{k,m})$ of a sigmoidal neuron on its input values $x_{k,m}$ on one hand, and the dependence of the firing rate of a spiking neuron on the firing rates of presynaptic neurons on the other hand. Therefore, if we replace $V_{i,j}$ in Eq. 5 by the averaged rate code x and use the equivalence explained above to replace the intra-layer coupling by $\sigma(x)$, the equations 5 and 4 will become equivalent. Therefore the dynamics of the original Dynamic Link Matcher is the rate-coding approximation of our place-coding network.

Affine Transformations: 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. 1b):

$$\frac{\Delta_{abc}}{\Delta_{abd}} = \frac{\Delta_{T(abc)}}{\Delta_{T(abd)}} \tag{6}$$

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. 1 is equivalent to (neglecting the effect of intra-layer connections, since $N^{ext} \gg N^{int}$):

$$w^{ext}_{i,j,k,m}(t) = \frac{f(p(i,j;t) - p(k,m;t))}{\Delta_{T(abc)} + \Delta_{T(abd)}}, \text{with} \quad f(x-y) = \frac{0.25}{e^{\lambda|x-y|}} \quad \forall x, y \tag{7}$$

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 influence for $neuron_{i,j}$ from all $neuron_{k,m}$ becomes :

$$S_{i,j}(t) = \frac{\Delta_{T(abc)} f(0)\psi(t, \phi_1) + \Delta_{T(abd)} f(A-B)\psi(t, \phi_2)}{\Delta_{T(abc)} + \Delta_{T(abd)}}, \text{with} \quad \psi(t, \phi) = H(x^{ext}_{k,m}(t)) \tag{8}$$

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. 6 we obtain:

$$S_{i,j}(t) = \frac{f(0)\psi(t, \phi_1)}{1 + \frac{\Delta_{abd}}{\Delta_{abc}}} + \frac{f(A-B)\psi(t, \phi_2)}{1 + \frac{\Delta_{abc}}{\Delta_{abd}}} \tag{9}$$

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 if 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.

4 Preliminary Results

In this paper, we use prototypes to illustrate the capacity of the network.
There are 5x5 neurons in each layer. A vertical bar in a background is presented to the first layer. The second layer receives the same object transformed by an affine transformation (rotation, translation, etc.). A vertical bar is applied to the first layer and a translated horizontal bar to the second. After the segmentation phase, different segments on different layers are desynchronized. During the dynamic matching stage, similar objects among different layers are synchronized. The thresholded sum of the activity of all neurons ($\sum_{i,j} H(x(i,j) - 0.5)$)) is shown in Figure 2 for the segmentation phase and for the dynamic matching phase. Since there are four different regions in the two layers with different phases in the end of the segmentation phase, four different synchronization regions can be seen in Figure 2. During the dynamic matching phase, the similar objects (and the backgrounds) merge with each other producing only two distinct regions. In a second stage a pseudo-triangle stimulus is applied to the first layer while the bar is maintained on the second. Since there isn't any affine transformation that can map the bar to the triangle no matching occurs, although segmentation is achieved in the segmentation phase.

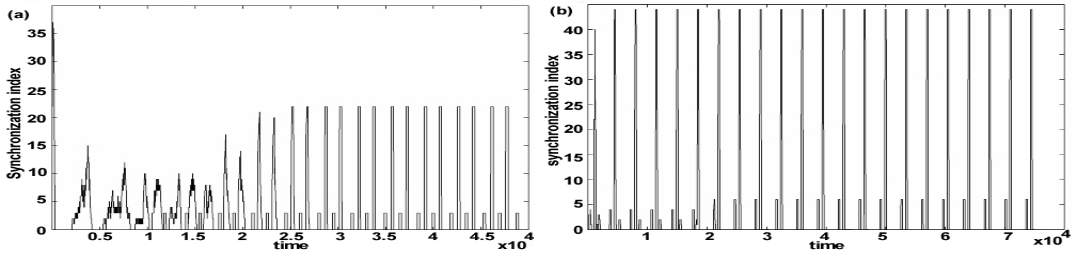

Figure 2: **(a)**: The evolution of the thresholded activity of network in the segmentation phase. Each vertical rod represents a synchronized ensemble of neurons ($\sum_{i,j} H(x(i,j) - 0.5)$). The network synchronizes at time step 3×10^4. **(b)**: The evolution of the thresholded activity of the network in the dynamic matching phase.

5 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 affine transforms. Generalization to real-world images is under investigation and will be presented in further works.We are investigating the possibility of the insertion of this architecture in our bottom-up sound segregator [6] as a top-down processor.

Keywords: Oscillatory Dynamic Link Matching, Relaxation Neurons, Pattern Recognition.

References

[1] W. Konen, T. Maurer, and C. Von der Malsburg. A fast dynamic link matching algorithm for invariant pattern recognition. *Neural Networks*, 7(6/7):1019–1030, 1994.

[2] X. Zhang and A. Minai. Detecting corresponding segments across images using synchronizable pulse-coupled nerual networks. In *IJCNN2001*, 2001.

[3] D.L. Wang and D. Terman. Image segmentation based on oscillatory correlation. *Neural Computation*, pages 805–836, 1997.

[4] E. Cesmeli and D. Wang. Motion segmentation based on motion/brightness inegration and oscillatory correlation. *IEEE Trans. on neural networks*, 11(4):935–947, 2000.

[5] 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.

[6] R. Pichevar and J. Rouat. Double-vowel segregation through temporal correlation: A bio-inspired neural network paradigm. In *NOLISP'2003*, 2003.