Segmentation and Binding in Oscillatory Neural Systems

David HORN and Irit OPHER

School of Physics and Astronomy
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University, Tel Aviv 69978, Israel
`horn@vm.tau.ac.il` `irito@ccsg.tau.ac.il`

September 15, 1995

Abstract

Segmentation and binding are cognitive operations which underlie the process of perception. They can be understood as taking place in the temporal domain, i.e. relying on features like simultaneity of neuronal firing. We demonstrate them in a system of oscillatory networks, consisting of Hebbian cell assemblies of excitatory neurons and inhibitory interneurons in which the oscillations are implemented by dynamical thresholds. We emphasize the importance of fluctuating input signals in producing binding and in enabling segmentation of a large set of common inputs. Segmentation properties can be studied by investigating the cyclic attractors of a neural system. Employing this method we show that for constant inputs full segmentation is limited to a small set of excited memories, however fluctuating inputs can lead to approximate segmentation of a large set of memories.

1 Introduction

Segmentation is a first processing step of various modalities. We define it as the task of parallel retrieval of individual memorized patterns which appear together in an input. Using the visual analysis of a scene as a convenient example (von der Malsburg and Buhmann, 1992), it is clear that a complicated picture is built out of elements which can be recognized separately. Before identification of the different elements in the scene takes place, an earlier step is required, namely the segmentation process, that separates the different elements, thus making them accessible for further processing. In addition to vision we find that it should play a role in auditory signal separation such as in the cocktail party effect (von der Malsburg and Schneider, 1984) and odor separation in the olfactory bulb (Hopfield and Gelperin, 1989).

One natural mechanism for segmentation is temporal tagging, which can be easily implemented in oscillatory systems. The realization of segmentation in oscillatory networks was demonstrated by Wang, Buhmann and von der Malsburg (1990) and by Horn and Usher (1991). The way it works is that the activities of the different memory-patterns that are turned on by the input, oscillate in a staggered fashion. In Wang *et al.* the elements of the networks are oscillators by themselves, composed of excitatory-inhibitory pairs of neurons. Horn and Usher have worked with neurons that possess dynamic thresholds that exhibit adaptation: they vary as a function of the activity of the neurons to which they are attached. As such, they introduce time dependence which can turn a neural network from a dissipating system which converges onto fixed-points into one which moves from one center of attraction to another (Horn and Usher 1989).

The use of dynamic thresholds leads to a variety of models which can represent cognitive functions which are more complex than content-addressable memory retrieval. It is quite obvious how this can lead to free associative transitions between memories. It may be less obvious that it allows us to model segmentation and binding. These features were obtained within an excitatory-inhibitory network which will be described in the next section. Based on that we will dwell on how segmentation is achieved and what is needed to obtain binding. We will point out the importance of noise in obtaining these goals.

To understand better the origin of these properties we turn to a symmetric model in which we investigate the segmentation mode as a particular kind of limit cycle (Horn and Opher 1995a, 1995b.) It becomes clear in this model why full segmentation cannot be obtained for many memories when a constant input is used. Noisy input allows us to overcome this limitation.

2 Neural Models with Dynamical Thresholds

To illustrate the method of Horn and Usher consider a neural system defined by the equation of motion

$$S_i(t+1) = F\left(\sum_j J_{ij} S_j(t) - \theta_i(t)\right) \ , \tag{1}$$

where F is a transfer function whose threshold parameters θ_i change with time in a fashion which depends on the history of the neuron S_i at the same location:

$$\theta_i(t+1) = \frac{\theta_i(t)}{c} + bS_i(t+1). \tag{2}$$

This is an effective integration of the S_i over time which saturates at the value $g = bc/(c-1)$ if the S_i stays constant at 1. c is chosen to be slightly bigger than 1. If

g is bigger than the input from all other neurons, this destabilizes the tendency of the system to stay in a fixed-point and leads to oscillations of the network and to movement between different patterns.

Trying to stay closer to biological models we turn to an excitatory-inhibitory network (Horn and Usher, 1990). In this model the memory patterns are assumed to be carried by excitatory neurons only. A pattern is defined by a cell-assembly of excitatory neurons which excite each other. These patterns affect one another through their joint excitation of the inhibitory neurons which, in turn, inhibit all the excitatory neurons. If the same excitatory neuron participates in two different patterns, or cell-assemblies, it plays the role of a two-way pointer between these two patterns. In the absence of any such cases, i.e. when all patterns are disjoint, we can derive differential equations for their activities:

$$dm^\mu/dt = -m^\mu + \sigma(Am^\mu - Bm^I - \theta_0^E - br^\mu + i^\mu) \tag{3}$$

$$dr^\mu/dt = (1/c - 1)r^\mu + m^\mu \tag{4}$$

$$dm^I/dt = -m^I + \sigma(CM - Dm^I - \theta^I) \tag{5}$$

$m^\mu(t)$, the activity of pattern number μ, is the fraction of the cell-assembly number μ which fires at time t. Similarly, m^I defines the fraction of all inhibitory neurons which fire. M represents the total excitatory activity, $M = \sum_\mu m^\mu$, and σ is the sigmoid function

$$\sigma(x) = (1 + e^{-x/T})^{-1}. \tag{6}$$

θ_0^E and θ^I are the (static) thresholds of all excitatory and inhibitory neurons correspondingly. The excitatory thresholds are assumed to have dynamical components which are represented by their average br^μ within the cell assembly μ.

In equations (3)-(5) we have included external inputs i^μ affecting memory number

μ. In other words, we operate our memory sytem under the influence of continuous external stimuli, and test its functioning under these conditions.

3 Segmentation and Binding in an Excitatory-Inhibitory Network

The problematics of segmentation and binding may be illustrated by the following example: consider the case studied in psychophysical experiments (Treisman and Schmidt, 1982) in which an observer is presented with a display consisting of three colored shapes, e.g., a red diamond, a blue square and a green circle. If we suppose that shapes and colors are stored in different cortical modules (networks), we are faced with the double problem of segmentation and binding. That is, the 'shape' module should recognize and segment the shapes, while the 'color' module should recognize and segment the colors. The binding problem, then, is to provide the correct matching between the shapes and their corresponding colors. Fig. 1 is an illustration of the problem.

Fig. 1. Schematic description of a problem involving segmentation and binding.

This problem was studied (Horn Sagi and Usher, 1992) within a model of two coupled networks, each obeying equations of the type (3)-(5). A coupling was introduced between the inhibitory components of the two networks, in order to drive them with the same frequency, while avoiding any semantic relations between the attributes (shapes and colors) represented by the excitatory cell assemblies.

To study segmentation we let n memories receive a common input. For suitable parameters of the system we find that memory activities which are aroused by the input oscillate in a staggered manner, each one peaking at different times, as displayed in the first two frames of Fig. 2. Thus we obtain temporal segmentation: the system identifies the mixed input by decomposing it dynamically into the different memories contained in it.

Binding is modeled by assuming that patterns corresponding to the related objects should oscillate in phase (e.g., the activity of the pattern representing the shape 'diamond' should oscillate in phase with the activity of the color 'red'). To achieve phase-locking of the attributes corresponding to the same object we assumed that both are affected by some common random activity fluctuation. Different pairs of attributes are driven by different and uncorrelated random noises. Fig. 2 describes results when two pairs of input are used. In this figure we show in addition to the activities of the two different cell-assemblies in the two networks also the two inputs. Note that the time scale of phase-locked oscillations is much larger than that of the autocorrelations of the fluctuating noise.

Fig. 2. Segmentation and binding (Horn Sagi and Usher, 1992). The first two frames exhibit activities of memory patterns (excitatory cell assemblies) of two coupled networks (see Fig. 1). The first network has five memory patterns and the second has three. The activities of the two memories excited by the input are shown by the full line and the dashed line. The dot-dashed curve represents an activated memory which does not receive an input. We observe both segmentation and binding. Segmentation means that the two different patterns in the two networks oscillate in a staggered fashion, and binding means that the patterns which are associated with one another oscillate in phase. This phase-locking is brought about by the common noise of the two attributes belonging to the same object. The inputs of the two objects are shown in the third frame.

In the case of three objects it takes longer time to achieve correct binding. Moreover, we have noticed that the system can move out of correct binding into erroneous phase correlations. This could correspond to the phenomenon of illusory conjunctions observed in psychophysical experiments (Treisman and Schmidt, 1982).

4 Segmentation viewed as a Limit-Cycle of a Dynamical System

To gain understanding of segmentation we have recently studied a system of equations which is similar in nature to (3)-(5), using methods of the type employed in dynamical systems. We consider n excitatory neurons interacting with one inhibitory neuron and receiving external inputs (Horn and Opher, 1995):

$$du_i/dt = -u_i + m_i - am - br_i + I_i \tag{7}$$

$$dr_i/dt = m_i - cr_i \tag{8}$$

$$dv/dt = -gv - em + f\sum_i m_i \tag{9}$$

u_i denote post-synaptic currents of excitatory neurons, whose average firing rates are

$$m_i = (1 + e^{-\beta u_i})^{-1} \tag{10}$$

while v and m are analogous quantities for an inhibitory neuron which induces competition between all excitatory ones. r_i are dynamical thresholds which rise when their corresponding neurons i fire. They quench the active neurons and lead to oscillatory behavior. $a, \cdots, g$ and β are fixed parameters. To study segmentation we choose $I_i = I$ as a common external input. Here each neuron plays the role of a Hebbian cell assembly in the model considered before. Note that this system is now fully symmetric under the interchange of any two neurons $i \leftrightarrow j$.

Fig. 3. Limit cycles of the $n = 3$ system (Horn and Opher, 1995). Parameters were the following: $a = 0.5$ $b = 0.4$ $c = 0.2$ $g = 0.1$ $e = 1.1$ $f = 0.5$ $\beta = 9$. The different m_i are plotted *vs* time after the system has reached stability. The time scale is arbitrary but is chosen to be the same in all figures. Each m_i is represented by a different symbol. The limit cycles are: a. Fully synchronous. $I = 0.8$. b. Partial synchronous waveform. $I = 0.4$. c. Full segmentation. $I = 0.4$.

In general, this dynamical system flows into a set of dynamic attractors. Thus, for 3 excitatory elements and constant input, one finds the following types of attractors: 1. Common fixed point or common oscillatory mode. 2. Two of the elements oscillate in phase and a third out of phase. (There are three possibilities for this type of attractor corresponding to the three possibilities of choosing 2 oscillators out of 3). 3. Staggered oscillations of all elements. (There are 2 possibilities for this attractor corresponding to the 2 possible arrangements of 3 oscillators in a ring). The last type fits our understanding of temporal segmentation. Examples of all limit cycles are shown in Fig. 3, which displays different solutions of $m_i(t)$ obtained for some fixed values of parameters in the system of equations, but different initial conditions. The segmentation mode competes here with only two other kinds of limit-cycles. The choice of one or the other depends on the initial conditions of the system. It is clearly also a function of the parameters which are being chosen.

In order to estimate the sizes of the basins of attraction of all waveforms we have chosen random initial conditions over the whole seven dimensional space of initial conditions of the $n = 3$ problem, and checked to which waveforms the system converges. Our interest is focused on the segmentation limit cycles. We found that the probability of converging onto the two waveforms of type c is 0.45 for the set of parameters specified in Fig. 1c.

As n is being increased the situation gets much more complicated. The phenomenon of full segmentation becomes less abundant and, in our simulations, stops at $n = 5$. However, partial segmentation is often obtained. This can happen in one of two ways (or a combination of both): 1. Formation of clusters of amplitudes which move in unison in a segmentation pattern. 2. Appearance of leading and non-leading amplitudes, in which the leading ones display segmentation behavior. Examples of these two types of waveforms are shown in Figs. 4 and 5.

Fig. 4. Partial segmentation in the $n = 6$ problem, through the formation of clusters of pairs of patterns which vary synchronously.

Fig. 5. A quasiperiodic solution of the $n = 6$ problem which displays partial segmentation. The three large amplitudes form a segmented pattern, while the low amplitudes display very different periodicities.

5 The Limitation on Temporal Segmentation

The fact that pure segmentation cannot be obtained in our model for $n > 5$ ties in with previous observations that temporal segmentation works only as long as n is small if the input is constant (Wang *et al.*, 1990, Horn and Usher, 1991) . It seems to be a general conclusion which is valid in many variants of these models: above $n =$ 4 to 6 the system looses its ability to perform temporal segmentation. It is tempting to speculate that this feature could provide an explanation for the known limits on attention and short term memory capacity (Miller 1956). This is the case if attention can be understood in the context of segmentation, as in the cocktail-party effect (von der Malsburg and Schneider, 1982). The connection with short term memory was made by Horn and Usher (1992) who have added a potentiation term to the fatigue term of the neuron's threshold, thus allowing for continued reverberations after the stimulus is disconnected. As expected, the mechanism works for only small numbers of memories which are jointly activated.

In this section we will look for an explanation of the limitation within the symmetric dynamical system which we study. In the system of equations (7)-(9), the interaction between all elements is provided by the inhibitory unit (9). The individual excitatory unit i, described by (7) and (8), is influenced by all other units through the amplitude m of the inhibitory unit in (9). The behavior of each m_i in any given waveform can therefore be also viewed as the response of equations (7) and (8) to a driving term $am(t)$. All the waveforms which we encounter in our numerical study have an overall period τ which is roughly the same as that of the free oscillator ($a = 0$). In a segmentation mode $m(t)$ oscillates with a period of τ/n. This can be seen in Fig. 6, where we show $m_1(t)$ and $m(t)$ for the $n = 5$ segmentation. m_1 has a waveform of period τ whose local peaks have widths of τ/n. If we think of m as the driving term then the phenomenon observed here is that of subharmonic oscillation which is known to exist in non-linear oscillating systems (Mandelstam and Papalexi, 1932, Hayashi, 1964).

Fig. 6. The variation of m_1 (solid line) and m (the inhibitory amplitude - dotted line) in the $n = 5$ fully segmented limit cycle.

A stable linear system follows the frequency of the driving term. Only nonlinear systems exhibit different periodic solutions, including the subharmonic ones which are of interest to us. The nonlinear characteristics of the system determine the possible orders of the subharmonic oscillations. In particular, a subharmonic solution of order $1/v$ is likely to occur when one of the terms of the nonlinear function has the power v. Its realization depends however on stability conditions which have to be met. Note that any such solution has a v-fold degeneracy determined by the phase which, in turn, strongly depends on the initial conditions. Full segmentation is a $1/n$-ordered subharmonic solution where each oscillator makes a different phase choice. In our case the nonlinearity is that of a sigmoid function which, in principle, contains all powers.

To test directly this property, we ran the system of (7)-(8) with a constant+sinusoidal driving term replacing m. In other words, we investigated solutions of the set of equations

$$du_1/dt = -u_1 + m_1 - h - br_1 + I_1 \tag{11}$$

$$dr_1/dt = m_1 - cr_1 \tag{12}$$

in which $h(t)$ is chosen to be similar to $am(t)$ of Fig. 6 but with a tunable frequency. We were able to generate narrow subharmonic solutions of 1/2 to 1/5. From 1/6 onwards the subharmonic solutions were no longer narrow, i.e. the m_1 amplitude generated by such a driving term has a width which is considerably larger than τ/n. This explains why full segmentation is limited in our system to $n \leq 5$. Higher n values cannot sustain the narrow subharmonic solution needed to build up segmentation.

It is interesting to find the stability of the subharmonic oscillations. We tested it (Horn and Opher, 1995b) in two ways. First we ran the system with variations of the frequency ω of the pure sinusoidal driving term and measured the window $\Delta\omega/\omega$ for which subharmonic oscillations were obtained. The results show large ranges for the subharmonics 1/2 and 1/3, for which the relevant values are 0.32 and 0.11 respectively. The other subharmonic solutions were obtained for considerably smaller frequency windows. The results for 1/4, 1/5 and 1/6 are 0.03, 0.015 and 0.013 respectively. Then we tested stability of the subharmonic solution against the mixing of a lower frequency $\omega - \delta\omega$ with the frequency ω which is the driving term of the nth subharmonic. Surprisingly the 1/2 solution turns out to be unstable, whereas the higher solutions of 1/3 1/4 etc have a range of stability of the order $\delta\omega/\omega \approx 0.2$. These results explain why in partial segmentation solutions, observed for high n values, three segments of leading amplitudes are dominant. The small amplitudes add a varying background to the driving term created by the large amplitudes, the ones responsible for the subharmonic solution. The fact that the 1/3 subharmonic solution has a large frequency window and is stable against admixture of several frequencies is the reason for the dominance of structures like the one displayed in Fig.

5.

The stability of an effective 3-segmentation pattern is particularly evident when we break the symmetry and use graded inputs, i.e. I_i with different values. The characteristic result, shown in Fig. 7, is that of three leading amplitudes (with different values reflecting the inputs) and a background of small amplitudes of the oscillators which have low inputs.

Fig. 7. Waveforms of an $n = 8$ system with graded inputs. I_i were chosen as 0.6, 0.575, 0.55, 0.525, 0.5, 0.475, 0.45, 0.425

6 Facilitation of Segmentation by Noisy Inputs

The effective segmentation obtained for the case of graded inputs leads us to investigate the possibility of unlimited effective segmentation for a noisy input. Looking for regularity in the oscillation patterns becomes more difficult when I_i contain stochastic elements. To facilitate the study and interpretation let us consider the correlation matrix

$$C_{ij}(\theta) = \frac{1}{T}\int_{t_1}^{t_1+T} m_i(t)m_j(t+\theta)dt \tag{13}$$

where the integral is carried out over a time period $T > \tau$ after some transient time t_1 has elapsed, ensuring that the limit cycle has been reached. Regularity may emerge only if we allow for an integration time T which is much larger than the characteristic period. Integrating over long time scales we may hope that the fact that the system of differential equations is symmetric on average will reflect itself in some effective symmetry of the waveforms which the correlation matrix elements may reveal.

Fig. 8. Correlations C_{12} (solid line) and C_{13} (dotted line) *vs* θ for noisy input in the $n = 3$ case display almost perfect segmentation.

An example of $n = 3$ is shown in Fig. 8. Here we choose an input of $I_i = 0.4 + 0.1\eta_i$, where η_i is a random variable between 0 and 1, that changes rapidly and obtain a regular structure which reflects full segmentation. The symmetry is obtained in spite of the random component in the input. Moreover, full segmentation is the only limit cycle observed. Increasing n to 4 and more, we find that the symmetry is broken. The general pattern is one of approximate segmentation, as demonstrated in Fig. 9:

the correlation structure displays distinct peaks, as is the case in all segmentation modes, but the regularity of relative phases is broken, even to the extent that the order of different peaks can change with time.

Fig. 9. Correlation structures for $n = 4$ of the pairs 12 (solid line), 13 (dotted line) and 14 (dashed line), show that noisy inputs lead to approximate segmentation.

Another possible form of introducing noise into the system is by giving a sudden increase (a random kick) to the constant input of one of the oscillators. This kick may activate this oscillator, and, since there is competition between all the excitatory units (provided by the inhibitory unit) other oscillators are prevented from rising at that instance. The result is a segmented mode, where the order of oscillation is not fixed, as can be seen in Fig. 10, showing eight separate oscillators. Fig 11 shows that the rise of each oscillator is preceded by an increase of the input to this oscillator. However, not every kick activates the relevant oscillator. This is due to the dynamics of the system, that could prevent an oscillator from rising if at that very moment another oscillator is at its peak.

Fig. 10. Staggered oscillation of the $n = 8$ problem. Activities of all oscillators are displayed as function of time.

Fig. 11. Timing of the random kicks and the resulting activation of oscillators. The ordinate is the number of the excitatory unit. In the upper diagram x denotes the time when the particular oscillator was kicked, and the horizontal line denotes the time span when the oscillator was active. The lower diagram displays only the dominant oscillators, and is therefore an alternative representation of Fig. 10.

This mechanism of effective segmentation through noise activation is quite similar to the noise driven activation in the segmentation and binding problem displayed in Fig. 2.

7 Summary: The Importance of Noise

We have seen the importance of a random component in the input for both the binding problem and the segmentation problem. For binding, the way it was formulated here, it was important to find some way to introduce information about the non-semantic correlation between the color and shape types. This was achieved with the random noise applied to each object and reflected in the two attributes which have to be bound. So far there exists no experimental foundation for such a realization of binding. Keele *et al.* (1988) have shown that temporal variations on the scale of 50ms do not facilitate binding in the visual domain. Hence they conclude that temporal variation which is relevant to binding may have to be internally generated.

We have seen that segmentation is a feature which is due to the non-linear nature of the system. In the symmetric model which was described here, pure segmentation corresponds to a limit-cycle whose basin of attraction disappears for $n > 5$. We have an understanding of the origin of this limitation in terms of subharmonic oscillations. We saw that it can be overcome when noisy inputs are used. In that case, approximate segmentation is the general result.

Acknowledgement

It is a pleasure to acknowledge the help of Marius Usher in the work summarized here.

References

Hayashi,C. (1964). *Nonlinear oscillations in physical systems* Princeton University Press.

Hopfield,J.F. & Gelperin,A. (1989). Differential conditioning to a compound stimulus and its components in the terrestrial mollusc Limax maximus. *Behav. Neurosci.* 103, 329-333.

Horn,D. & Opher,I. (1995a). Dynamical symmetries and temporal segmentation, to be published in *Jour. of Nonlinear Science.*

Horn,D. & Opher,I. (1995b). Temporal segmentation in a neural dynamical system, to be published in *Neural Computation.*

Horn,D., Sagi,D, & Usher,M. (1992). Segmentation, binding and illusory conjunctions. *Neural Comp.* 3, 509-524.

Horn,D.,& Usher,M. (1989). Neural networks with dynamical thresholds, *Phys. Rev. A* 40, 1036–1044 .

Horn,D. & Usher,M. (1990). Excitatory-inhibitory networks with dynamical thresholds.*Int. J. of Neural Systems* 1, 249–257.

Horn,D. & Usher,M. (1991). Parallel activation of memories in an oscillatory neural network. *Neural Computation* 3, 31–43.

Horn,D.,& Usher,M. (1992). Oscillatory model of short term memory, in *Advances in Neural Infromation and Processing Systems 4*, J.E.Moody S.J.Hanson R.P.Lippmann eds,(Morgan Kaufmann Pub.) p. 125-132.

Keele,S.W., Cohen,A., Ivry,R., Liotti,M. & Yee,P. (1988). Tests of a temporal theory of attentional binding. *Jour. of Exp. Psychology: Human Perception and Performance* 14, 444–452.

von der Malsburg,C. & Buhmann,J. (1992). Sensory segmentation with coupled neural oscillators. *Biol. Cybern.* 67, 233–242.

von der Malsburg,C. & Schneider,W. (1986). A neural cocktail party processor. *Biol. Cybern.* 54, 29–40.

Mandelstam,L. & Papalexi,N. (1932). Über resonanzerscheinungen bei frequenzteilung. *Z. Physik* 73, 223–248.

Miller,G.A., 1956. The magical number seven, plus or minus two: Some limits on our capacity of processing information. *Psychological Rev.* 63, 81–97.

Treisman,A.M. & Schmidt,H. (1982). Illusory conjunctions in the perception of objects. *Cognit. Psych.* 14, 107–141.

Wang,D.,Buhmann,J. & von der Malsburg,C. (1990). Pattern segmentation in associative memory.*Neural Computation* 2, 94-106.