Network 1 (1990) 41–59. Printed in the UK

Application of the elastic net algorithm to the formation of ocular dominance stripes

G J Goodhill†§ and D J Willshaw‡

† Department of Artificial Intelligence, University of Edinburgh, Edinburgh, UK
‡ Centre for Cognitive Science, University of Edinburgh, Edinburgh, UK

Received 4 September 1989

Abstract. The elastic net algorithm, an iterative technique for the solution of combinatorial optimisation problems that have a geometric interpretation, was applied to the problem of explaining the development of ocular dominance stripes in the vertebrate visual system. Simulations show that this algorithm produces stripes under certain conditions. Analysis is presented that predicts the moment at which stripes form and an expression is derived for how stripe width depends on the parameters of the system. In contrast to most other models for stripe formation, the elastic net algorithm provides a common explanatory framework for the development of stripes and of retinotopically ordered projections.

1. Introduction

A general property of the wiring of the vertebrate brain is that axons projecting from sensory structures to more central brain structures preserve neighbourhood relations. Such mappings are topographic, i.e. minimally discontinuous. The neighbourhood relations preserved can be both spatial and featural, and thus the problem may involve mapping between structures of different dimensionalities.

The best studied example of a topographic mapping in biological systems is the retinotopic mapping between the retina and the visual centres of the brain. In the course of ontogenesis, the ganglion cells of the retina send axons through the eye stalk to the brain, where they project to cells of central structures. In mammals the retinae each project to the visual cortex via the lateral geniculate nucleus; in lower vertebrates there are direct projections to the optic tectum. Several mechanisms have been proposed that produce such neighbourhood-preserving mappings (see, for example, Willshaw and von der Malsburg 1979).

A more complex pattern of connections is that formed in the binocularly innervated part of the visual cortex. This is where fibres from both retinae map to the same part of the brain. In the early stages of development, the retinotopic projections from the two eyes to the common target structure are uniformly intermingled. As development proceeds, each part of the target gradually becomes more densely innervated by one eye and less densely innervated by the other. Eventually, a striped pattern of innervation is observed that is reminiscent of the pattern of zebra stripes: the fibres from the two eyes divide the projection area up between them to form distinct regions (figure 1).

§ Present address: School of Cognitive Sciences, University of Sussex, UK.

0954-898X/90/010041+19$03.50

The individual pattern of innervation from each eye remains retinotopic. In animals such as cats and monkeys the thickness of these stripes is approximately 400μm. In lower vertebrates, similar striped projections do not occur naturally but can be induced experimentally (Constantine-Paton and Law 1978, Fawcett and Willshaw 1982). At present, there is no satisfactory explanatory model for how retinotopically ordered yet striped neural mappings are produced.

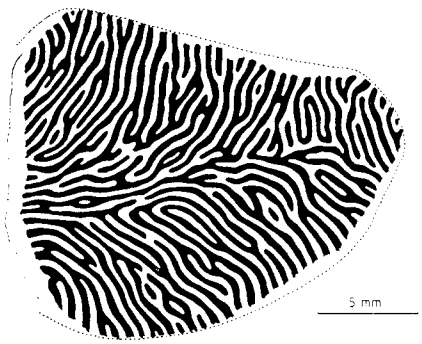


Figure 1. A reconstruction of the pattern of ocular dominance stripes in area 17 of the Macaque monkey (Hubel and Wiesel 1977).

In this paper the elastic net algorithm (Durbin and Willshaw 1987), a method for combinatorial optimisation that developed from the Tea Trade model for the establishment of monocular retinotopic mappings (von der Malsburg and Willshaw 1977), is applied to the development of binocular striped projections. For a fuller account see Goodhill (1988).

1.1. Models for the development of nerve connections

1.1.1. Monocular mappings. von der Malsburg and Willshaw (1977) proposed a mechanism for the formation of topographically ordered mappings between two groups of nerve cells. It was originally explained by analogy with a mythical commercial situation involving the import of tea from India to Britain—hence the name 'Tea Trade model'. They assumed that the neighbourhood relations within the pre-synaptic sheet are expressed by a set of chemical markers. These markers are induced by the optic nerve fibres into the post-synaptic sheet, where they are used to guide the fibres to their places of termination. It is assumed that (*a*) there exist sources of chemical markers fixed at certain positions in the pre-synaptic sheet and (*b*) there are no such sources in the post-synaptic sheet. Due to processes of diffusion, eventually each pre-synaptic cell attains a unique collection of markers, which encodes the position of the cell relative to its neighbours. The pre-synaptic sheet then develops connections with the post-synaptic sheet. Initially these connections are arranged almost entirely at random. Through each synaptic contact, the collection of markers characteristic of the presynaptic cell of origin is induced into the post-synaptic sheet, where the markers diffuse as before. The strength of each connection is continually changed in accordance with the instantaneous similarity between the collection of markers that the fibre carries and the collection of markers already existing at that point in the post-synaptic sheet. Weak connections are abolished, while strong connections sprout to produce further

connections with neighbouring post-synaptic cells. To prevent synaptic strengths growing without bounds, a pre-synaptic sum rule operates: all synapses associated with each pre-synaptic cell are normalised so that their sum is a constant value. Computer simulations with this model successfully produced neighbourhood-preserving mappings and provided an account of various biological experiments investigating monocular retinotectal projections in lower vertebrates.

The Tea Trade model inspired Kohonen (1982) to develop a general method for establishing neighbourhood preserving mappings between spaces of different dimensionalities. Given an initially arbitrary mapping from input units to output units, input patterns are presented at random. For a given input, the output unit with the maximal response is found, either directly or by lateral inhibition between output units. The connections to this unit and its neighbours are then adjusted so that they will give an even greater response to that input pattern in the future. Kohonen (1984) presented simulations that demonstrate that this process produces neighbourhood-preserving mappings. Mathematical analysis of the conditions for the stability of topographically ordered projections between two sets of elements has been carried out for the Tea Trade model (Takeuchi and Amari 1979, Häussler and von der Malsburg 1983) and for the Kohonen algorithm (Ritter and Schulten 1986).

1.1.2. Striped projections. von der Malsburg (1979) extended the Tea Trade model to apply to the formation of ocular dominance stripes. He postulated the existence of two further markers to indicate ocularity, one for each eye. For the mechanism to produce satisfactory mappings, these new markers had to be introduced once the initial topographic mapping had been established.

A completely different theory was proposed by Swindale (1980), in terms of the effect exerted by synapses on the growth of other synapses. He assumed mutually reinforcing interactions, of a circularly symmetrical form, between synapses from the same eye, and similar inhibitory interactions between the synapses from different eyes. Computer simulations of the development of this system showed that a wide variety of conditions incorporating these assumptions leads to the formation of striped patterns. These patterns exhibited many of the morphological features of striped projections in cat and monkey binocular visual cortex. However, the main drawback of this model is that the origin of retinotopicity was not considered. In a recent extension of Swindale's work, Miller *et al* (1989) give a more detailed and thorough treatment in which they assume a pre-existing retinotopicity.

1.2. Combinatorial optimisation and the travelling salesman problem

One way of viewing the formation of topographic mappings is as a form of combinatorial optimisation. A cost function may be defined in terms of the relative positions of fibres in the post-synaptic sheet that were neighbouring in the pre-synaptic sheet which is optimised for a neighbourhood-preserving mapping. Any cost function with this property will do if a continuous map is possible; however, if a continuous map is not possible then the precise form of the function will determine the type of discontinuities introduced (Stone 1989).

The method to be described here is based on a logical extension to the Tea Trade model, called the elastic net (Durbin and Willshaw 1987), which was originally developed for the well known optimisation problem, the travelling salesman problem (TSP). The problem is to find the shortest closed path that visits each of N cities just once (Lawler *et al* 1986). In the general case, each city–city distance can be

arbitrary, but in the so-called geometrical case each city–city distance corresponds to the Euclidean distance between the coordinates of the two cities in the plane. An exact solution is computationally intractable for distributions of more than a few cities, and so much effort has been devoted to techniques for finding good solutions. Perhaps the best-known algorithm is that due to Lin and Kernighan (1973). The TSP is representative of the class of problems called NP-complete, which means that the amount of computation required to solve them increases faster than any power of the scale of the problem. All NP-complete problems are interconvertible. Since the TSP is easy to state and conceptually simple, it is often used as a benchmark against which to test optimisation techniques.

Durbin and Willshaw formulated the elastic net algorithm by noting that the geometrical TSP involves the construction of a mapping from a circle to a plane such that neighbouring points on the circle map to points that are close together in the plane. Thus, an algorithm by which a plane is mapped into a plane, preserving neighbourhood relationships (such as the Tea Trade model), could be adapted to this case. In the case of the geometrical TSP, it is imagined that a closed ring of points placed in the plane of the cities moves under the influence of two types of force: forces of attraction from the cities and an elastic force from neighbouring points on the ring. Eventually the ring comes to pass through the cities, thus defining a tour.

The rule for the change $\Delta \boldsymbol{y}_j$ in the position $\boldsymbol{y}_j$ of point j at each iteration is

$$\Delta \boldsymbol{y}_j = \alpha \sum_i w_{ij}(\boldsymbol{x}_i - \boldsymbol{y}_j) + \beta k(\boldsymbol{y}_{j+1} - 2\boldsymbol{y}_j + \boldsymbol{y}_{j-1}) \tag{1}$$

where $\boldsymbol{x}_i$ is the position of city i, and α and β are scaling constants for the contributions from the cities and from neighbouring points on the ring respectively. w_{ij} represents the normalised force of city i on path point j:

$$w_{ij} = \frac{\Phi(|\boldsymbol{x}_i - \boldsymbol{y}_j|, k)}{\sum_p \Phi(|\boldsymbol{x}_i - \boldsymbol{y}_p|, k)} \tag{2}$$

where

$$\Phi(|\boldsymbol{x}_i - \boldsymbol{y}_j|, k) = \exp\left(\frac{-|\boldsymbol{x}_i - \boldsymbol{y}_j|^2}{2k^2}\right). \tag{3}$$

Note that the influence of each city on each point is modulated by the scale parameter k, which defines the effective range of the interaction. Initially, k is large and is then gradually reduced during the execution of the algorithm. When k is large, each city has a large domain of attraction and shares its force between many points. Likewise, each point experiences a substantial force from many cities. The force field over the positions of the rope points is fairly uniform and only gross disparities in the distribution of the cities cause deformations in the shape of the path. The forces become more specific as k decreases, until eventually each city effectively has an attraction for only the closest point to it, thus completing the tour. This is a different approach from conventional TSP algorithms since all intermediate states of the system are now in a continuous space, rather than the discrete space of valid tours. In this respect, the elastic net resembles Hopfield and Tank's algorithm (Hopfield and Tank 1985).

Durbin and Willshaw defined an energy function E:

$$E = -\alpha k \sum_i \log \sum_j \Phi(|\boldsymbol{x}_i - \boldsymbol{y}_j|, k) + \frac{\beta}{2} \sum_j |\boldsymbol{y}_{j+1} - \boldsymbol{y}_j|^2. \tag{4}$$

E has the property that

$$\Delta \boldsymbol{y}_j = -k \frac{\partial E}{\partial \boldsymbol{y}_j} \tag{5}$$

and thus changing $\boldsymbol{y}_j$ according to equation (1) always has the effect of reducing E. In the limit as $|\boldsymbol{x}_i - \boldsymbol{y}_j|$ tends to zero for small k, the first term of E vanishes. The second term is just the length of the tour, and hence E is minimised for the shortest tour.

It should be noted that E is a function of the scale parameter k. Durbin *et al* (1989) have shown that, for large enough k, there is just one minimum of E, and at some value of k the energy surface bifurcates. The algorithm tracks the minimum as k decreases. In effect, k determines the amount of detail seen in the structure of the energy surface.

Since the scale parameter k plays a somewhat analogous role to the temperature in simulated annealing (Kirkpatrick *et al* 1983), it seems reasonable to expect that the slower k is reduced, the better will be the solution obtained. In Durbin and Willshaw's original application to the TSP, k was reduced at a constant rate of 1% every 25 iterations. Tours were obtained within 1–2% of the best tours obtained by more conventional algorithms. Unlike Hopfield and Tank's method, the elastic net scales well with the number of cities, and good tours have been obtained for up to 100 cities (Durbin and Willshaw 1987).

1.3. Other recent algorithms for the TSP

The self-organising algorithm due to Kohonen (1984) for mapping between spaces of different dimensionalities has recently been applied to the TSP (Angeniol *et al* 1988, Fort 1988, Stone 1989). In these algorithms, each city attracts only the points on the path closest to it, and they generally require less computation than the elastic net. It is difficult to compare these algorithms with the elastic net since the elastic net acts synchronously and deterministically while the others act asynchronously and non-deterministically. Since the elastic net uses more global information than the Kohonen algorithm, it might be expected to obtain shorter tours, but in longer time. Results (Angeniol *et al* 1988, Stone 1989) indicate that, comparing several runs of the Kohonen algorithm with one elastic net run, the best tour obtained is shorter but the *average* length is longer.

Burr (1988) has recently attempted to refine the elastic net algorithm in two ways. Firstly, the rate at which k is reduced is determined adaptively in terms of the current average (RMS) city–point distance. Secondly, in determining the forces to be exerted at each point from a given city, the point is replaced by the point nearest to the city. With these additions, he claimed tours as good as those produced by the elastic net algorithm, but found in far fewer iterations.

2. Application of the elastic net algorithm to the formation of striped projections

We formulate the problem in terms of stretching an elastic rope, representing the cortex, between points, representing pre-synaptic cells, in an abstract space. This space

is analogous to the geometric space for the TSP, and is also analogous to the space defined by the marker concentrations in the Tea Trade model. The essential difference between the two models is that in the Tea Trade model, distances between points represent the similarities between pre- and post-synaptic cells, rather than the weights of the connections themselves as in the elastic net algorithm.

Two cases were investigated: a simplified case of two one-dimensional retinae mapping to a one-dimensional region of brain, and the general case of two two-dimensional retinae mapping to a two-dimensional region of brain (here called the cortex).

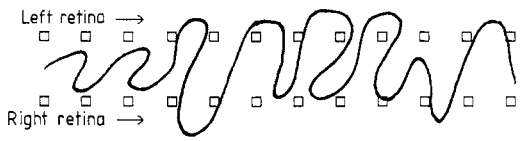


Figure 2. The cells are shown as squares, the rope as a line.

For the one-dimensional case the representation used is as follows: the cells in the two retinae occupy two parallel rows in the unit square, the rows running in the horizontal direction and separated by a certain specified distance. Position along the horizontal axis represents position within one retina and the vertical separation of the two rows indicates ocularity. The cortex, represented as a rope, is free to move between the two rows in the plane of the cells (figure 2). The distance of a rope point from a cell represents the amount of innervation of that part of the cortex by that retinal cell. The closer the point to the cell, the greater the amount of innervation. Thus if a point lies half way between two cells, then the cortex at tnat point is innervated equally by the two retinal cells. Note that in this model *the topology of the rope represents the topology of the cortex.* Neighbouring regions of the cortex are represented by adjacent points on the rope.

To apply the elastic net algorithm, we view the cells as cities and the rope as a tour, by analogy with the TSP. The position of the rope is then adjusted according to equation (1), with the difference that the rope now has a break in it. The tension at the endpoints of the rope is therefore taken to be proportional to the difference in position between the end point and its neighbour, i.e. $y_{N-1} - y_N$ for point N.

The 2D case is a straightforward generalisation of the 1D case. The two retinae are represented as two planes of cells, lying on top of one another and separated by a small gap. The cortex becomes a 2D elastic sheet, and again the topology of the sheet represents the topology of the cortex. The definition of tension in the cortical sheet has to be generalised to take account of the fact that it is now a sheet and not a rope. It was assumed that there is tension between a point and its four nearest neighbours, and suitable boundary conditions were applied.

The initial position of the rope in the one-dimensional case was essentially random, except for a small amount of bias applied to the horizontal components of the position of each point to ensure that the rope comes to have a specific orientation. For example, endpoint 0 of the rope is to be mapped to the left-hand end of the two chains of cells rather than the right-hand end. This reflects the fact that the biological mapping has a specific orientation, and it was the method of specifying orientation used by Willshaw and von der Malsburg (1979). The vertical component of the initial position of each rope point was random, with no correlation between the positions of neighbouring

points. A similar method was used for the two-dimensional case.

3. Results

Preliminary results showed that stripes are formed. The simulations and analysis concentrated on:

(1) whether the patterns obtained were striped and retinotopic (as in figure 3);

(2) the time course of the development;

(3) the nature of the patterns produced;

(4) the role of the parameters in determining the development of the system and the width of the stripes.

The parameters that control the behaviour of the system are α, β, the initial value of k, the rate at which k is reduced, the number of points in the rope, the separation $2l$ between the two layers of cells and the separation $2d$ between cells within a layer. The parameters investigated most thoroughly were l and d. The values of α, β and the initial value of k were fixed at 0.2, 2.0 and 0.2 respectively for all simulations (Durbin and Willshaw 1987).

There are several different broad domains in the parameter space which exhibit qualitatively different behaviours. If the cell separation between layers is much greater than the separation within a layer, i.e. $l \gg d$, or if k is reduced too rapidly, the points in the rope are *captured* by cells as soon as the simulation starts, and an entirely random pattern governed solely by the initial conditions soon results. By *capturing* we mean that, on the scale set by the value of k, a given point p becomes so close to a cell c that c is effectively the only cell that influences p. This domain was not examined, and the results described below are for the region where $l < 10d$ and the factor by which k is reduced at each iteration is less than 10%.

The programs which performed the calculations were written in C and run on a colour Sun 3/60.

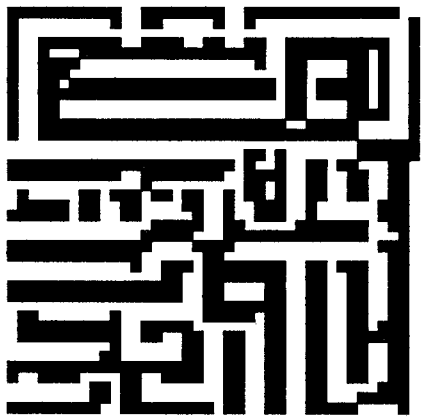
3.1. The one-dimensional case

The system investigated most thoroughly had 20 cells in each layer and between 40 and 120 points in the rope. Larger simulations indicated that the results obtained do not depend on the number of cells.

The rope initially flattens out to lie equidistant between the two layers. Small discrepancies from the midway position remain, which become the asymmetry needed for stripe formation (see section 4). The rope also shortens. Note that this represents a retinotopic mapping whereby there is no specialisation of the cortex to one retina or the other, as is observed in the early stages of biological development.

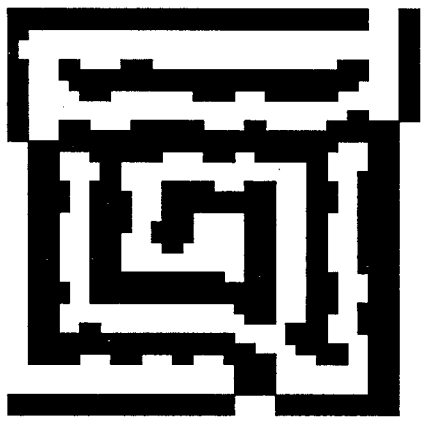
The points in the rope are distributed densely at the ends of the rope and widely spaced in the middle. As k is reduced further, the rope gradually expands horizontally. Points move from the ends towards the middle of the rope, where they start to become more evenly spaced (figure 4). Stripes (resulting from rope points moving vertically and being captured by cells) may start to form during this time, depending on the separation between the layers. The wider the separation, the earlier the vertical movement occurs.

Eventually, the rope extends to the full length of the rows of the cells. The points are now evenly distributed over its length. If stripes have not yet begun to form, a new phenomenon occurs that we call *clustering*. This is where the points on the rope move horizontally and form into groups halfway between pairs of cells in opposite layers.



(a)

Separation of retinae is $2l = 0.1$



(b)

Separation of retinae is $2l = 0.15$

(c)

Separation of retinae is $2l = 0.2$

Figure 3. Stripe patterns obtained in the 2D case. Black areas represent the projection from one retina, and white areas the projection from the other. The number of cells in each retina is 20×20, the number of points in the sheet is 40×40, the separation of cells within a retina is $2d = 0.05$.

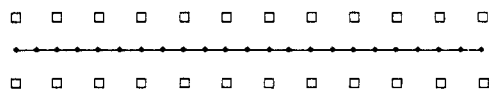


Figure 4. Even spacing of points on the rope.

In the case when the number of rope points is divisible by half the number of cells, each cluster consists of the same number of points. Provided that stripes do not form during this process, the points in each cluster become tightly bunched. In biological terms, clustering corresponds to further topographic refinement of connections.

At some point, stripes begin to form. At least one, and usually several, of the points on the rope simultaneously develop a small vertical fluctuation from the other points on the rope. This quickly becomes magnified, and soon all points in the rope have moved up or down. If clustering has occurred by this time, each cell captures at least one point on the rope. If clustering has not occurred, then some cells may not initially capture a point (*incomplete* mappings). If the system in this state is left to develop further, points captured by cells slowly move the small remaining distance to lie on top of the particular cell, and points that have not been captured remain in approximately the position they occupied when stripes had first formed.

In order to explore the possibility that the properties of the system are simply boundary effects due to the break in the rope, another one-dimensional case was briefly examined. The ends of the rows of cells and the ends of the rope were wrapped round horizontally to form a cylinder, and the rope was constrained to move on the surface of this cylinder. The forces are now slightly more complicated, since there are two forces between each cell and each point in the rope, corresponding to the two possible vectors between a cell and a point. However, no fundamental change in behaviour from the original case was observed, and this system will not be considered further here.

3.2. Quantitative considerations

3.2.1. Rate at which k is reduced. The faster k is reduced, the more disordered the final pattern of stripes. However, the rate of reduction was found to have no effect on the value of k at which clustering and stripe formation occurred. These values will henceforth be referred to as k_c and k_s respectively. For most simulations, the rate at which k was reduced was fixed at 0.3% per iteration, which seemed to provide an acceptable compromise between speed of simulation and degree of disorder. Reducing k more slowly than this made little difference to the pattern of stripes produced.

3.2.2. Number of rope points. The case of one rope point per cell was investigated first, for speed of simulation. Adding more points to the rope does not fundamentally alter the development of the final pattern of stripes (figure 5). One minor effect is that if the ratio of points in the rope to cells is greater than about 2.5 (i.e. five points per cluster), the outer points in each cluster remain slightly separated and do not coincide with the other points in the cluster.

3.2.3. Separation of cells. k_c increases with d, the intralayer separation parameter (figure 6), which has virtually no effect on k_s. Conversely, k_s increases with l, the interlayer separation parameter (figure 7), which has virtually no effect on k_c. The values of l

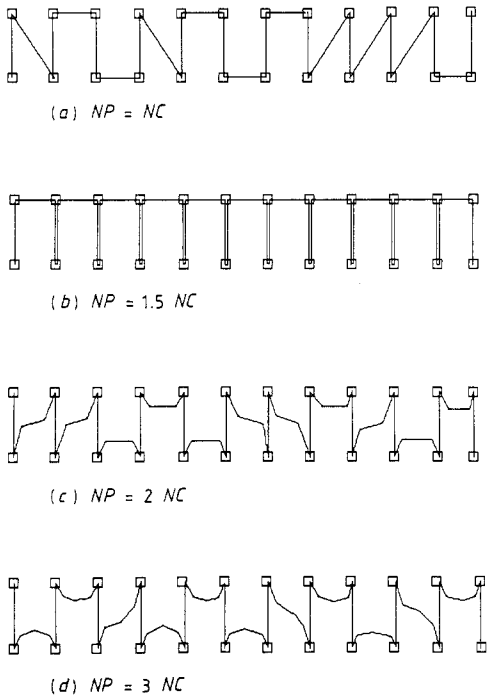


Figure 5. Patterns obtained as the ratio of rope points to cells is increased. Number of cells $= NC$, number of rope points $= NP$.

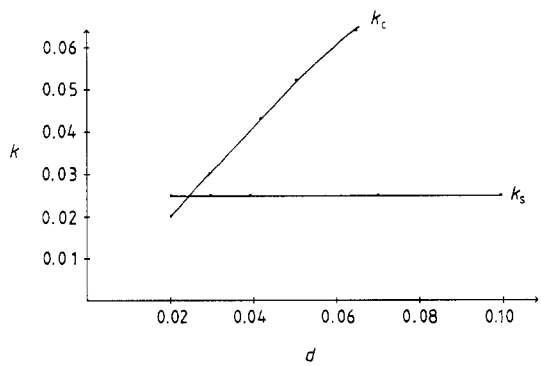


Figure 6. Dependence of k_s and k_c on the intralayer separation parameter, d. $l = 0.05$.

and d interact to determine the pattern of stripes formed. If $k_s > k_c$, the stripes form before clustering can occur and the resulting pattern is one of broad, rather irregular stripes. If $k_s < k_c$, the pattern tends to be more homogeneous.

3.2.4. Energy function. The behaviour of the energy function E as defined in equation (4) was investigated in the one-dimensional case. Equation (5) guarantees that each rearrangement of the point positions will reduce E at constant k, but it is not clear how E will be affected by the fact that k is reduced at each iteration. It was in fact found that E increases rapidly as k decreases, and that this increase far outweighs

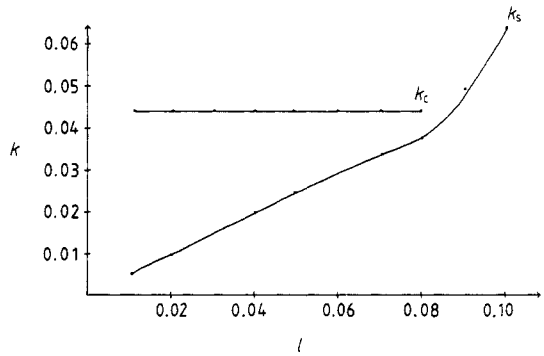


Figure 7. Dependence of k_s and k_c on the interlayer separation parameter, l. $d = 0.025$.

the decrease in E due to the rearrangement of the points on the path. In effect the algorithm moves towards the minimum of E at each shelf of the energy surface of constant k. The form of E for a typical simulation is shown in figure 8.

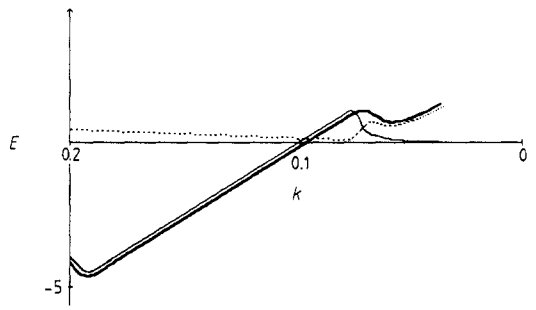


Figure 8. The form of E for a typical simulation. The thin line represents the contribution from the first term in the energy equation, the broken curve represents the second term and the thick curve is the total.

3.3. The two-dimensional case

Behaviour analogous to the one-dimensional case was observed. The sheet of rope points initially flattens out and contracts. At this stage the points are widely spaced in the middle of the sheet and dense at the edges. The sheet now expands uniformly in the horizontal direction until the points in the sheet are evenly spaced. Clustering of rope points between pairs of cells may occur as before. All pairs of cells will have the same number of rope points in their cluster if the number of retinal points is an even multiple of the number of cells in the target sheet. At some point, stripes form in a way analogous to the previous cases. The variation in the form of the resulting patterns is similar to that observed in the one-dimensional case. Some examples of the striped patterns obtained are shown in figure 3.

4. Mathematical analysis

Qualitatively, the task is to provide an explanation of the various stages the system

goes through as described in the previous section. Quantitatively, the task is to predict the values of k at which clustering occurs and at which stripes form, and to explain the resulting pattern of stripes in terms of the parameters of the system.

4.1. Development of the pattern of stripes

Stage 1

Behaviour. The rope flattens out and settles between the two rows of cells.

Explanation. Consider a point p displaced from the middle of the gap between the two layers of cells, as in figure 9. Provided that k is large compared with l, cells $c1$ and $c2$ exert almost the same total influence F on p. Then the point p will be pulled down to the middle since the downward component $F\sin\theta$ is greater than the upward component $F\sin\phi$.

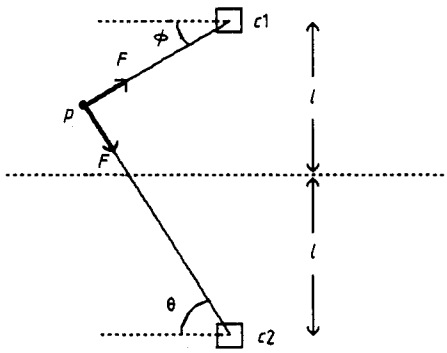


Figure 9. Point p experiences approximately equal forces from cells $c1$ and $c2$ initially.

The rope will therefore settle into the middle even without any tension, and this has been confirmed by simulations with zero tension. Other simulations have shown that tension alone will not pull the rope flat. The tension does, however, have the effect of contracting the rope horizontally.

Stage 2

Behaviour. The rope gradually expands until it reaches to both ends of the distribution of cells. Points are initially clustered at the ends of the rope and sparsely distributed in the middle.

Explanation. The horizontal distribution of the points in the rope follows from the fact that each cell has a finite amount of influence that it must share out between all the points in its domain of attraction. A cell at the end of a row has only half as many points to share its influence over compared to a cell in the middle of a row, and hence attracts each point twice as strongly. Thus initially, even though the rope contracts due to the tension, the points move to the ends of the rope.

As k decreases, the tension decreases, allowing the rope to be pulled out by the end cells. The tension force then acts to space the points evenly, since for a given length of rope this distribution gives the least contribution to the tension component of the energy E.

Stage 3

Behaviour. The points in the rope begin to cluster between pairs of cells.

Explanation. Suppose that the points in the rope are evenly spaced except for one, which has a small horizontal displacement. For k sufficiently small, it can be assumed

that each cell influences only the points in its immediate neighbourhood. The cases of three points and five points only being influenced were analysed (see the appendix). It was found in both cases that the net horizontal force acts to restore even spacing until a value of k is reached at which the displacement increases, and this is taken to be the value of k_c.

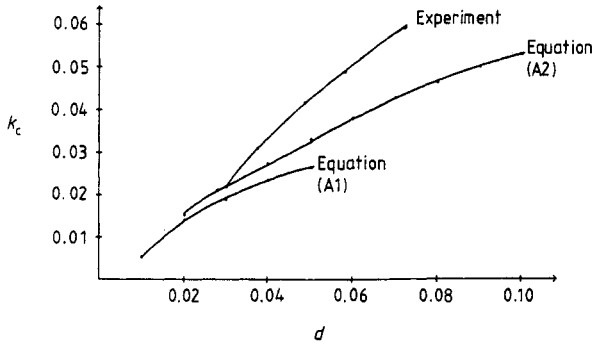


Figure 10. The values of k_c calculated from the two approximations.

In figure 10 the values obtained with these approximations are compared with the experimental results. The second approximation yields higher values of k_c than the first. This is to be expected, since the second approximation is taking into account the influence of the cells on more points, while the tension remains the same as in the first approximation. However, it can be seen that these predictions underestimate the experimental results.

Stage 4

Behaviour. Stripes begin to form.

Explanation. As in the case of clustering described above, the forces on each point from the cells are balanced against the tension in the rope. Simplifications were made in order to obtain a relation between k_s and the parameters of the system (see the appendix). Three cases were investigated.

(1) Only one cell in each layer influences each point.

(2) Three cells in each layer influence each point.

(3) An infinite number of cells in each layer influence each point.

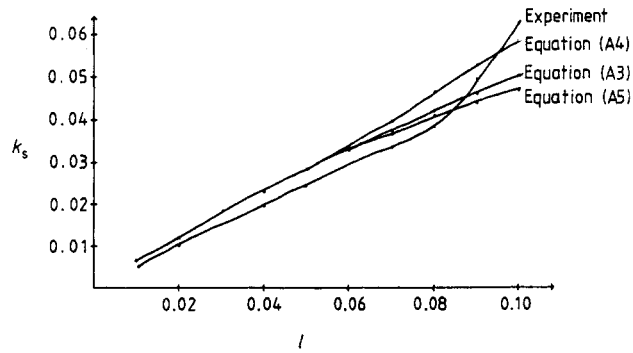


Figure 11. The values of k_s calculated from the three approximations.

These theoretical predictions agree closely with the experimental results (figure 11). The slight overestimate may be due to a time lag between the change in sign of the force and when the points were observed to move vertically.

4.2. The pattern of stripes—path length analysis

The form of the pattern of stripes resulting for given parameters can be predicted by regarding the problem as a TSP; that is, find the shortest length of rope that joins all the cells. First, we assume that all cells are visited. There are then three basic shapes that the rope can take on (figure 12). The relative lengths of these different paths depend only on the ratio of the interlayer and intralayer separations l/d. If it is assumed that each pattern repeats regularly along the rows of cells, and that there are a large number of cells, it is straightforward to calculate the relative lengths of the paths. The following conditions hold (see figure 12 for notation).

(1) P_1 is the shortest path for $l > d$.

(2) P_2 is the shortest path for $l < d$.

(3) P_3 is a generalisation of P_2, and can be shown to have minimum length if the number n of cells per stripe is $n = l^2/d^2$.

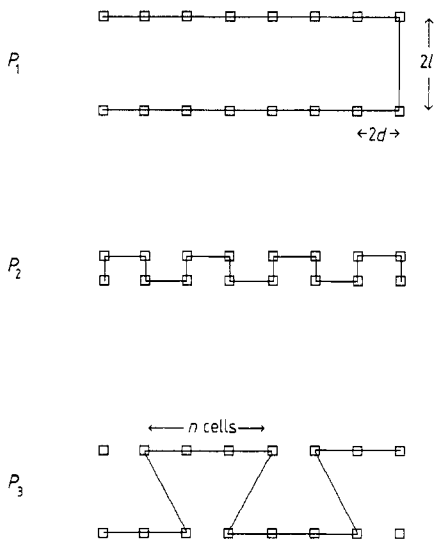


Figure 12. Three possible arrangements of the rope.

In practice, P_1 is not observed since for this pattern the topographic condition that the two ends of the rope map to different ends of the two rows of cells is not satisfied. These results explain why the stripes get broader as l increases: there are fewer pieces of rope between the two rows of cells.

The case when not all cells have been visited (incomplete mappings) has also been analysed (Goodhill 1988). A value of k can be derived at which it is energetically favourable for this type of mapping to form, and it is a function of the ratio l^2/d^2. Although the intralayer separation does not affect the moment at which stripes form, it contributes to the pattern of stripes formed via its role in determining the point at which incomplete mappings become favourable.

5. Discussion

The elastic net algorithm was applied to the formation of binocular projections in the mammalian visual cortex and the optic tectum in amphibia and fishes. Two cases were examined: two one-dimensional retinae mapping to a one-dimensional cortex, and two two-dimensional retinae mapping to a two-dimensional cortex. In both cases, the system exhibits qualitatively the stages of development as observed in biological systems: for certain values of the parameters, overlaid topographically ordered projections are formed initially, which are then modified to give regularly striped projections. Quantitatively, it was found that k_c, the value of the space parameter k at which the points in the rope (the representation of the cortex in the algorithm) begin to cluster, depends almost exclusively on d, the intralayer parameter, and that k_s, the value of k at which stripes begin to form, depends almost exclusively on l, the interlayer parameter. The pattern of stripes formed was found to depend largely on the ratio l/d. In biological terms, d represents the similarity between adjacent retinal cells and l the corresponding similarity between the two retinae. The relative values of the parameters α and β specify the balance between the retino-cortical and the cortico-cortical interactions.

Certain critical points of the system were investigated theoretically, and equations derived to predict the dependence of k_c and k_s on the parameters of the system. These equations provide more accurate predictions in the case of stripe formation than clustering. Predictions relating the form of the stripes to the parameters of the system were derived by considering the formation of stripes as a special case of the travelling salesman problem. Whilst the value of the energy function E decreases with respect to the rearrangement of the positions of the rope points specified at each iteration, overall it *increases* due to its dependence on k.

It is instructive to compare our model (hereafter called GW) with that proposed recently by Miller, Keller and Stryker (1989) (hereafter called MKS), itself an elaboration of earlier work by Swindale (1980). In MKS, four factors are held to determine the development of ocular dominance stripes.

(1) The initial pattern of connectivity between afferents and cortex.

(2) The correlations between the activity in the various types of afferent at the cortex.

(3) Intracortical interactions, which are a property of the cortex itself.

(4) Constraints on the total synaptic strength available to each presynaptic cell.

The following correspondences can be made between MKS and GW.

(a) The arbor function in MKS is not modelled explicitly in GW, but is related to the value of the parameter k, which determines the effective distribution of cortical cells contacted by a given presynaptic cell.

(b) In GW, the parameters d (specifying the distance between adjacent retinal cells) and l (specifying the interretinal distance) play similar roles to the functions used in MKS (which we call here $C_{\text{same}}(= C_{CC}, C_{RR})$ and $C_{\text{diff}}(= C_{CR}, C_{RC})$), that specify the correlations between afferents from the same eye and different eyes respectively. MKS allows anticorrelations, whereas GW does not.

(c) The two models use effectively similar cortico-cortical interaction functions and synaptic normalisation rules.

In GW, the ratio l/d determines the width of the stripes formed. The stripes widen as l is increased or d is decreased. In MKS terms, this would correspond to a reduction in the size of the function C_{diff} or an increase in C_{same}. Miller *et al* (1989) confirmed

Swindale's original analysis (1980), according to which the width of the stripes is related to $C_{same} - C_{diff}$, and thus this model and the elastic net make broadly similar predictions concerning the dependence of stripe width on the parameters of the system.

The essential difference between the two models is that GW offers an explanatory model for the development of the retinotopic projection *and* the emergence of stripes, starting from an almost entirely random set of connections. In MKS the retinotopic mapping is supposed to exist from the start, as specified by the fixed arbor function. The continuous change in the geometrical scale of the computation employed in GW, through the value of k, seems to avoid the problems encountered by von der Malsburg (1979) in his direct extension of the Tea Trade model to stripe formation.

The usual, informal, explanation for the occurrence of 'retinotopic' stripes in the retinotectal projection, where the interplay of retinotopy and ocularity has been extensively investigated, is that it is the result of two processes: one that guides each retinal fibre to its 'normal' retinotopic target, and one that clusters together the sites of termination of the fibres from the same eye (Constantine-Paton 1981, see also Levay *et al* 1975). This method would, however, seem to be too rigid to deal with cases when the normal retinotopic target is lacking, such as during the development of stripes (Constantine-Paton and Law 1978) and in the case of stripe formation in *Xenopus* compound eyes (Fawcett and Willshaw 1982).

In summary, the elastic net algorithm provides a method by which striped projections can be produced over a wide range of conditions which might be expected to occur in the biological situation, such as for differing numbers of target cells. The attribute of striped double projections that can be most easily assessed experimentally is that of stripe width. The elastic net method predicts that stripe width is a function of the retinal parameters. A direct way of testing this would be to introduce a controlled imbalance between the two retinae, such as that produced when forcing a double nasal eye and a normal eye to develop a double projection on a *Xenopus* optic tectum (Straznicky and Tay 1982).

Acknowledgments

We are grateful to John Hallam for many helpful comments and suggestions.

Appendix. Analysis of clustering

A simplified picture of the situation arising once points have become evenly spaced is shown in figure 4. It is assumed that the number of points in the rope is equal to the number of cells. Imagine rope point p_0 is displaced to the left by a small amount ε from its position midway between its two neighbouring rope points (figure A1). Cells $c1$ and $c2$ now attract p_0 more strongly, while cells $c3$ and $c4$ attract p_0 less strongly. However, the tension in the rope contributes a counterbalancing force. As k decreases, the tension becomes weaker and the difference between the attractions by the two pairs of cells becomes greater. Thus points will start to cluster once the imbalance of the forces from the cells outweighs the extra tension.

The balance of forces on point p_0 in figure A1 was calculated to first order in ε, assuming that k is small enough so that the effective range of attraction of each cell is $\sqrt{l^2 + d^2}$, that is, cells $c1$ and $c2$ influence points p_{-2}, p_{-1} and p_0 only. (The bracketing

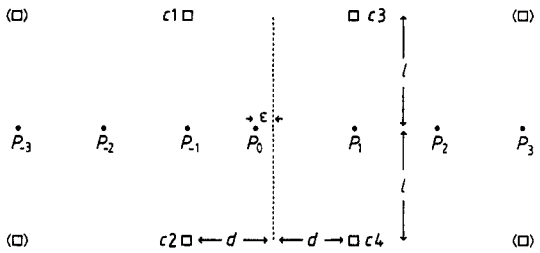


Figure A1. Analysis of clustering.

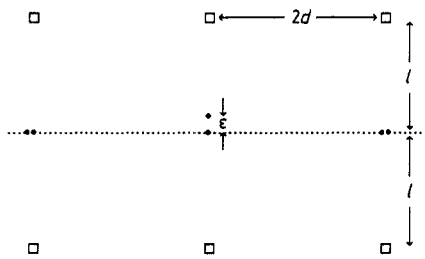


Figure A2. Analysis of stripe formation.

of some cells indicates that they are drawn for completeness, but are not taken into account in the calculation). When $k < 0.8d$ the attraction from the cells that have been ignored is less than 10% of the attraction from the cells considered. The net force F_1 on p_0 to the left in figure A1 was calculated to be

$$\frac{F_1}{\varepsilon} = \frac{2\alpha}{(2+e^y)^2}[e^y(2y-1)+2(y-1)] - 2\beta k \tag{A1}$$

where

$$y = d^2/2k^2.$$

Thus, if F_1 is negative, this acts as a restoring force towards the midway position, which is then stable. However, if F_1 is positive, any small deviation will be magnified at the next iteration, indicating the onset of clustering. Thus the value of k_c is given by the zero of equation (A1). Note that there is no dependence on l in equation (A1), in keeping with the simulation results.

In the second approximation it was assumed that $c1$ and $c2$ attract only $p_{-3}, p_{-2}, p_{-1}, p_0$ and p_1. This assumes that the effective range of the force from each cell is $\sqrt{l^2+2d^2}$. The 10% condition described above now requires $k < d$. The force F_2 was found to be

$$\frac{F_2}{\varepsilon} = \frac{\alpha}{1+e^y+e^{-2y}+2e^{-3y}+e^{-6y}+e^{2y}/4} \times \left(\frac{d^2}{k^2}(1+e^y+2e^{-3y}) - (2+e^y+2e^{-3y})\right) - 2\beta k \tag{A2}$$

where y is as given above.

The values of k_c as given implicitly in equations (A1) and (A2) cannot be found analytically, but can be easily obtained numerically.

For the analysis of stripe formation, three cases were investigated, each corresponding to different assumptions about the effective domain of attraction of the cells. In each case, it is assumed that clustering has already occurred (figure A2) and there are N points in each cluster. The three cases, with their respective upward forces, are as follows.

(1) Only one cell in each layer influences each point:

$$\frac{F_1}{\varepsilon} = 2\frac{\alpha}{N^2}\left(\frac{l^2}{k^2}(N-1) - N\right) - 2\beta k. \tag{A3}$$

(2) Three cells in each layer influence each point:

$$\frac{F_2}{\varepsilon} = 2\alpha\left(\frac{l^2}{N^2k^2(2e^{-y}+1)^2}\left(2(2N-1)e^{-2y} + 4Ne^{-y} + N - 1\right) - \frac{1}{N}\right) - 2\beta k \tag{A4}$$

where

$$y = \frac{2d^2}{k^2}.$$

(3) An infinite number of cells in each layer influence each point:

$$\frac{F_3}{\varepsilon} = \frac{2\alpha}{N^2}\left(\frac{l^2 a}{k^2}(Na - b) - Na^2\right) - 2\beta k \tag{A5}$$

where

$$a = \sum_{n=-\infty}^{n=\infty} e^{-n^2h^2/2k^2}$$

and

$$b = \sum_{n=-\infty}^{n=\infty} e^{-n^2h^2/k^2}$$

with

$$h = 2d.$$

In all three cases, the zeros of the expressions for the force can be regarded as the point at which stripes begin to form, and thus values for k_s can be calculated. These were found numerically for the three cases. For equations (A4) and (A5), d was set to 0.025, the value used in most of the simulations. Note that when $d \gg k$, both equations (A4) and (A5) reduce to the form of equation (A3), as expected. The three approximations all give similar results when l is small. This is to be expected, since for l small the effect of neighbouring cells is much less than that of the cells directly above and below the point under consideration.

An important point is that equation (A3) does not involve the intralayer separation d. This seems desirable, since the values of k_s obtained from the simulations show little dependence on d. Equations (A4) and (A5) do involve d, but here k_s is only a weak function of d.

References

Angeniol B, de la Croix Vaubois G and le Texier J-Y 1988 Self-organising feature maps and the travelling salesman problem *Neural Networks* **1** 289–93

Burr D J 1988 An improved elastic net method for the travelling salesman problem *Proc. IEEE Int. Conf. on Neural Networks (San Diego, CA, July 1988)*

Constantine-Paton M 1983 Position and proximity in the development of maps and stripes *TINS* **6** 32–6

Constantine–Paton M and Law M I 1978 Eye-specific termination bands in tecta of three-eyed frogs *Science* **202** 639–41

Durbin R, Szeliski R and Yuille A 1989 An analysis of the elastic net approach to the traveling salesman problem *Neural Computation* **1** 348–58

Durbin R and Willshaw D J 1987 An analogue approach to the travelling salesman problem using an elastic net method *Nature* **326** 689–91

Fawcett J W and Willshaw D J 1982 Compound eyes project stripes on the optic tectum in Xenopus *Nature* **296** 350–2

Fort J C 1988 Solving a combinatorial problem via self-organising process: an application of the Kohonen algorithm to the travelling salesman problem *Biol.Cybern.* **59** 33–40

Goodhill G J 1988 Application of the elastic net algorithm to the formation of ocular dominance stripes *MSc thesis* University of Edinburgh (unpublished)

Häussler A F and von der Malsburg C 1983 Development of retinotopic projections: an analytical treatment *J. Theoret. Neurobiol.* **2** 47–73

Hopfield J J and Tank D W 1985 Neural computation of decisions in optimisation problems *Biol.Cybern.* **52** 141–52

Hubel D H and Wiesel T N 1977 Functional architecture of the macaque monkey visual cortex *Proc. R. Soc.* B **198** 1–59

Kirkpatrick S, Gelatt C D and Vecchi M P 1983 Optimisation by simulated annealing *Science* **220** 671–80

Kohonen T 1982 Self-organised formation of topologically correct feature maps *Biol. Cybern.* **43** 59–69

—— 1984 *Self Organization and Associative Memory* (Berlin:Springer)

Lawler E L, Lenstra J K, Rinnooy Kan A H G and Schmoys D B 1986 *The Travelling Salesman Problem* (New York:Wiley)

Levay S, Hubel D H and Wiesel T N 1975 The pattern of ocular dominance columns in the macaque visual cortex revealed by a reduced silver stain *J. Comp. Neurol.* **159** 559–76

Lin S and Kernighan B W 1973 An effective heuristic algorithm for the travelling salesman problem *Oper. Res.* **21** 498–516

Miller K D, Keller J B and Stryker M P 1989 Ocular dominance column development: analysis and simulation *Science* **245** 605–15

Ritter H and Schulten K 1986 On the stationary state of Kohonen's self–organising mapping *Biol. Cybern.* **54** 99–106

Stone J 1989 In preparation

Straznicky C and Tay D 1982 Retino-tectal map formation in dually innervated tecta: a regeneration study in *Xenopus* with one compound eye following bilateral optic nerve section *J. Comp. Neurol.* **206** 119–30

Swindale N V 1980 A model for the formation of ocular dominance stripes *Proc. R. Soc.* B **208** 243–64

Takeuchi A and Amari S 1979 Formation of topographic maps and columnar microstructures in nerve fields *Biol. Cybern.* **35** 63–72

von der Malsburg C 1979 Development of ocularity domains and growth behaviour of axon terminals *Biol. Cybern.* **32** 49–62

von der Malsburg C and Willshaw J 1977 How to label nerve cells so that they can interconnect in an ordered fashion *Proc. Natl Acad. Sci., U.S.A.* **74** 5176–8

Willshaw D J and von der Malsburg C 1979 A marker induction mechanism for the establishment of ordered neural mappings: its application to the retinotectal problem *Phil. Trans. R. Soc.* B **287** 203–43