ESANN'2009 proceedings, European Symposium on Artificial Neural Networks - Advances in Computational Intelligence and Learning. Bruges (Belgium), 22-24 April 2009, d-side publi., ISBN 2-930307-09-9.

# Connection strategy and performance in sparsely connected 2D associative memory models with non-random images

Lee Calcraft, Rod Adams, and Neil Davey

School of Computer Science, University of Hertfordshire
College Lane, Hatfield, Hertfordshire, AL10 9AB, U.K.
{L. Calcraft, R.G. Adams, N. Davey} @ herts.ac.uk

**Abstract.** A sparsely connected associative memory model is tested with different pattern sets, and it is found that pattern recall is highly dependent on the type of patterns used. Performance is also found to depend critically on the connection strategy used to build the networks. Comparisons of topology reveal that connectivity matrices based on Gaussian distributions perform well for all pattern types tested, and that for best pattern recall at low wiring costs, the optimal value of Gaussian σ used in creating the connection matrix is dependent on properties of the pattern set.

## 1 Introduction

It has been established that the performance of sparsely connected associative memory models is strongly influenced by the connection strategy employed in creating the network [1-3]. In this respect, locally-connected networks, in which the input of each node is connected to its *k* nearest neighbours, perform poorly, while randomly-connected networks perform the best [3].

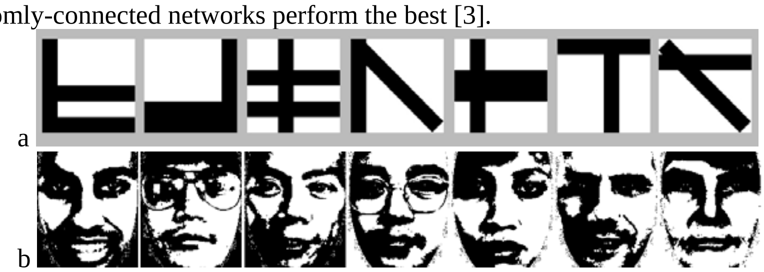


Fig. 1: Samples from two of the image sets used: a. Shapes, b. Faces [4]. The image size used in all cases was 60 x 60 pixels. Black pixels are considered to be *on*, white to be *off*.

These results are based on the use of training sets made from randomly-generated patterns (containing random distributions of *on* and *off* pixels). We find, however, that when real-world images are used (with a more naturalistic distribution of pixels – see Figure 1), the relationship between connection strategy and performance is dramatically altered [5]. Unexpectedly, randomly-connected networks can give relatively poor results with certain types of naturalistic images, with the best performance achieved by networks which show a combination of local and distal connections. In the present study we extend this work to networks built with Gaussian connectivity, and investigate its implications for the optimal connection strategy for sparsely connected associative memory models.

## 2 Network dynamics, training and performance measurement

Our associative memory models consist of a network of perceptrons arranged in a two-dimensional structure with wrap-around at the edges, and the network is trained on sets of patterns of area $N$, where $N$ is the number of nodes in the network. The output of each node is connected to the inputs of a fixed number, $k$, of other nodes. The networks used in the present studies have no symmetric connection requirement [6], and the recall process uses asynchronous random order updates, in which the local field of unit $i$ is given by:

$$h_i = \sum_{j \neq i} w_{ij} S_j$$

where $w_{ij}$ is the weight on the connection from unit $j$ to unit $i$, and $S$ $(= \pm 1)$ is the current state. The dynamics of the network is given by the standard update: $S_i' = \Theta(h_i)$, where $\Theta$ is the Heaviside function. Network training is based on the perceptron training rule [7] chosen for its higher resultant capacity than that of the standard Hopfield model. Further details may be found in [2, 8].

Network performance is determined by measuring Effective Capacity [2]. This is a measure of the number of patterns which a network can restore under a specific set of conditions. The network is first trained on a set of patterns. Once training is complete, the patterns are each randomly degraded with 60% noise, before presenting them to the network. After convergence, a calculation is made of the degree of overlap between the output of the network, and the original learned pattern. The Effective Capacity of the network is the highest pattern loading at which this mean overlap for the pattern reaches a predetermined level. For the present work we use the most stringent version of Effective Capacity (EC-100), requiring 100% correction of the presented patterns.

As in the earlier paper [5], our experiments compare performance using three types of pattern sets: the first is based on purely random patterns (random arrangements of *on* and *off* pixels). The second is a set of 132 hand-generated shapes, illustrated in Figure 1a. These were designed as bold patterns with large contiguous areas of *on* or *off* pixels, and with low correlation between individual patterns across the set. In contrast to this artificially created Shapes set, the third set consisted of 40 digitised faces [4], each of a different individual, as illustrated in Figure 1b. The networks under test all have 3600 nodes, configured as a 60 x 60 two-dimensional associative memory. The networks are sparsely connected, with 40 afferent connections per node.

## 3 Background and motivation

We have previously shown that the way in which pattern recall changes when a locally connected network is progressively rewired depends critically on the properties of the pattern set on which the network is trained [5]. Our results are

ESANN'2009 proceedings, European Symposium on Artificial Neural Networks - Advances in Computational Intelligence and Learning. Bruges (Belgium), 22-24 April 2009, d-side publi., ISBN 2-930307-09-9.

summarised in Figure 2. This illustrates the performance of three different pattern sets as a locally connected network is progressively rewired. The rewiring is to randomly selected nodes, so that once all the connections have been rewired, the network is randomly-connected.

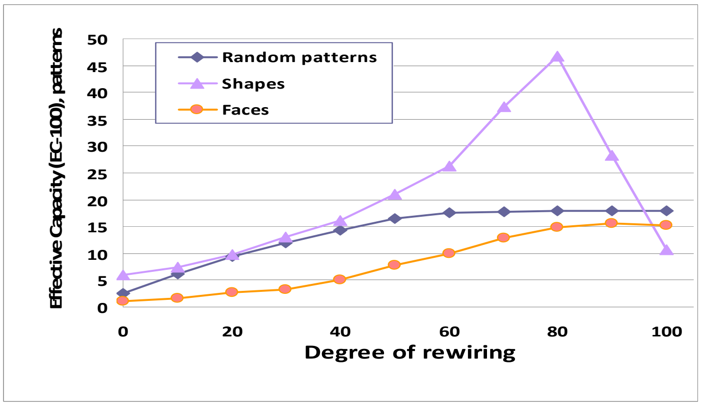


Fig. 2: Effective Capaciy (EC-100) as a function of the degree of rewiring of the network for the Shapes, Faces and random image sets. The network consists of 3600 nodes, with 40 connections per node. Results are averages over 30 runs.

The performance obtained with the random pattern set follows the expected course (broadly echoing the findings of Bholand and Minai for 1D networks [3]), with performance improving as the network is progressively rewired, up to a point at around 50% or 60% rewiring, after which further rewiring gives little improvement. The performance of the Shapes pattern set was quite unexpected, peaking at an Effective Capacity (EC-100) of more than 45 at 80% rewiring. This is more than twice the number of random patterns that could be recalled under similar conditions.

In the case of the Faces pattern set, performance was, by contrast, noticeably poorer than that of the random pattern set. In the paper referred to earlier [5], we argued that the difference in these results could be explained in terms of two factors: image coherence, and correlation between patterns in the set; and other experiments were performed to support this assertion.

In that paper, however, we simply explored the performance of the pattern sets using progressively rewired networks. In the present paper, we use the same pattern sets, but the progressively rewired networks are replaced by more biologically plausible Gaussian networks, where the probability of a connection between any two nodes is a Gaussian function of the distance between them. Comparisons are then made between the performance of the two types of networks. Finally we proceed to examine which connection strategies give rise to the most efficient associative memories for different types of pattern set once wiring length is taken into account.

## 4 Networks built with Gaussian connectivity

The results for the Gaussian networks appear in Figure 3, from where it may be seen that the performance measured for each of the three pattern sets is broadly in line with that obtained with the progressively rewired networks in Figure 2. The performance of the random pattern set increases from a low level as Gaussian σ is increased, and then flattens out just as it did with progressive rewiring at the 60% rewiring point. The Faces pattern set also keeps a similar profile to that seen in Figure 2, flattening out as σ is increased, at a level noticeably below the maximum achieved with the random pattern set. The Shapes pattern set also behaves similarly to the way it performs with progressive rewiring, reaching a slightly higher peak, at an EC-100 of just over 50, from where it drops back, just as in the progressive rewiring case. But the fall is less dramatic: even at a relatively high Gaussian σ of around 50, EC-100 never drops to the levels seen with the progressively rewired network. We would suggest that this is because with a Gaussian distribution there is always a concentration of connections around each node, whereas at 100% rewiring of the progressively rewired network, this is not the case.

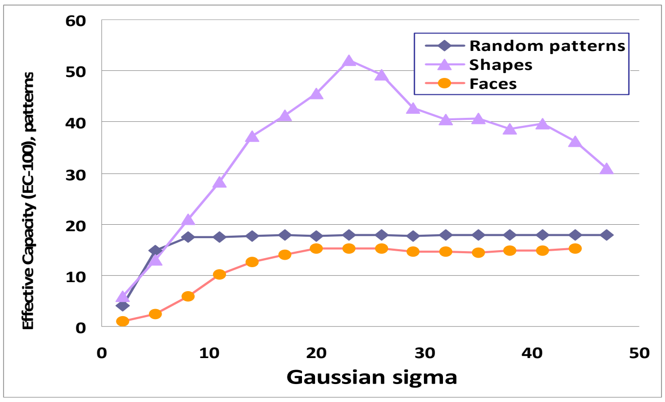


Fig. 3: Effective Capacity (EC-100) as a function of Gaussian σ for the Shapes, Faces and random pattern sets. The network consists of 3600 nodes, with 40 connections per node, using a Gaussian connectivity strategy of varying σ. Results are averages over 30 runs.

## 5 Assessing efficient connection strategies

In any physical implementation of associative memory, the length of wiring involved will be an important factor [9-11], and in seeking an efficient connection strategy

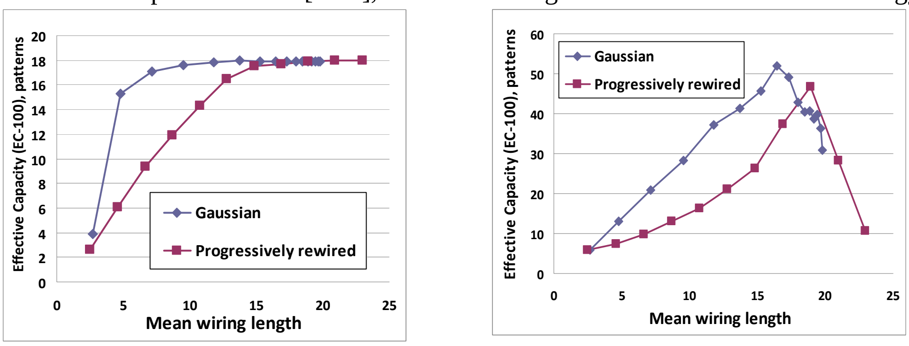


Fig. 4: Effective Capacity (EC-100) as a function of meean wiring length for Gaussian and progressively rewired networks, trained on the random pattern set (left), and trained on the Shapes image set (right). The network consists of 3600 nodes, with 40 connections per node. Results are averages over 30 runs.

here we are concerned to find the topology which yields the best pattern recall for the least expenditure of wiring. To shed light on this we plot EC-100 against wiring length for networks built using the two different connection strategies. Figure 4 gives the results for the random pattern set, revealing that the Gaussian network is the most efficient of the two, and is able to achieve an EC-100 of around 18 patterns at a mean wiring length of 7 units. The equivalent progressively rewired network requires a mean wiring length of twice this value to achieve the same degree of pattern recall.

If we perform a similar plot for the Shapes pattern set (Figure 4 - right), we can see that as well as achieving a slightly higher peak EC-100 value, the Gaussian network does this at a slightly lower mean wiring length of around 16 units compared to 19 units with the rewired network. For the Faces pattern set (Figure 5), the Gaussian network is again the more efficient performer, achieving an EC-100 of around 15 at a mean wiring length of 15. The optimum Gaussian σ for this pattern set would be around 20.

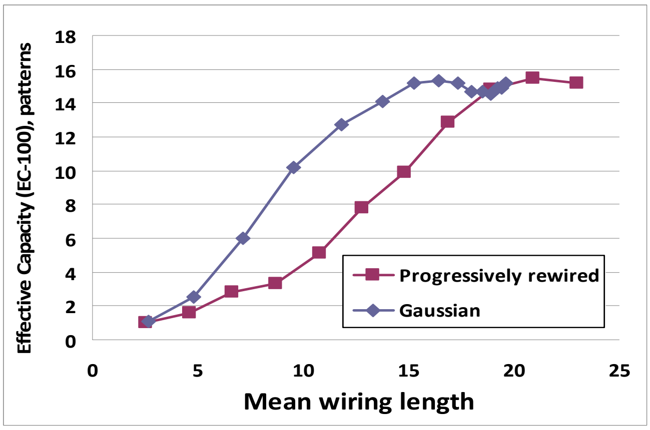


Fig. 5: Effective Capaciy (EC-100) as a function of mean wiring length for Gaussian and progressively rewired networks, trained on the Faces image set. The network consists of 3600 nodes, with 40 connections per node. Results are averages over 30 runs.

Thus the Gaussian network outperforms the progressively rewired network for all three types of pattern (random, shapes and faces) in terms of economy of wiring. Interestingly, the point at which the Gaussian networks achieve greatest efficiency depends on the type of pattern that they are trained on. For random patterns, connectivity based on a Gaussian distribution with a σ of 8 achieves greatest efficiency, but when working with the Faces or the Shapes pattern sets, we need Gaussian distributions with a σ of 20 or 23 respectively for optimal recall. In the case of the shapes set, if σ differs to any extent from the value 23, performance is considerably worsened.

## 6 Conclusion

We have explored the performance of a sparsely connected associative memory model built using a connection strategy based on Gaussian distributions, and compared the results to those obtained with the less biologically plausible progressive rewiring strategy. Simulations with the Shapes, Faces and random pattern sets were performed, and results broadly similar to those using the progressive wiring strategy were obtained. The most significant difference between the two sets of results was that with the Gaussian network, the dramatic drop in performance with the Shapes set, seen in the progressive rewiring network at 100% rewiring, was not apparent even at relatively large values of Gaussian $\sigma$. This is likely to be because in a Gaussian network, even at relatively high values of $\sigma$, there will still be local connections, necessary for correcting faults in clustered groups of pixels. In the progressively rewired network at 100% rewiring, however, there will be virtually no local connections in our 3600 unit network with only 40 connections per node.

Finally we compared the mean wiring length of the different network configurations, and found that for all three pattern types, the Gaussian network was more efficient, recalling a similar number of patterns to the progressively rewired networks, but at a lower wiring cost. This was true for each pattern type tested. Interestingly the point of optimal performance, at which the network is able to recall the maximum number of patterns at the shortest mean wiring length, occurred at different values of Gaussian $\sigma$ for the three different pattern types. Thus when designing an associative memory for maximum pattern recall at a minimum of wiring costs, one needs to take account of the type of patterns that are to be recalled.

## References

[1] L. Calcraft, R. Adams, and N. Davey, "Gaussian and exponential architectures in small-world associative memories," *Proceedings of ESANN 2006: 14th European Symposium on Artificial Neural Networks. Advances in Computational Intelligence and Learning*, pp. 617-622, 2006.

[2] L. Calcraft, R. Adams, and N. Davey, "Efficient architectures for sparsely-connected high capacity associative memory models," *Connection Science,* vol. 19, 2007.

[3] J. Bohland and A. Minai, "Efficient associative memory using small-world architecture," *Neurocomputing,* vol. 38-40, pp. 489-496, 2001.

[4] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms," *Image and Vision Computing,* vol. 16, pp. 295-306, 1998.

[5] L. Calcraft, R. Adams, and N. Davey, "The performance of sparsely-connected 2D associative memory models with non-random images," in *The Proceedings of the Eleventh Neural Computation and Psychology Workshop*, Oxford, 2008.

[6] N. Davey, L. Calcraft, and R. Adams, "High capacity, small-world associative memory models," *Connection Science,* vol. 18, pp. 247-264, 2006.

[7] S. Diederich and M. Opper, "Learning of correlated patterns in spin-glass networks by local learning rules," *Physical Review Letters,* vol. 58, pp. 949-952, 1987.

[8] N. Davey, S. P. Hunt, and R. G. Adams, "High capacity recurrent associative memories," *Neurocomputing,* vol. 62, pp. 459-491, 2004.

[9] D. Attwell and S. Laughlin, "An energy budget for signaling in the grey matter of the brain," *Journal of Cerebral Blood Flow and Metabolism,* vol. 21, pp. 1133–1145, 2001.

[10] D. Chklovskii, "Synaptic connectivity and neuronal morphology: two sides of the same coin," *Neuron,* vol. 43, pp. 609-617, 2004.

[11] G. Mitchison, "Neuronal branching patterns and the economy of cortical wiring," *Proceedings: Biological Sciences,* vol. 245, pp. 151-158, 1991.