Assignment 2: Investigations in Self-Ordered Mapping

Methodology

I began this assignment by editing the given code to make it functional. This entailed randomly seeding the cluster vectors, adding code to find the most similar cluster (I did this my maximizing the dot product), and adding code to update each cluster's feature vector (see equation 1 and 2 below). At this point, the code was functional, but did not perform self-ordered mapping (the radius of influence and alpha term were both still constants.)

$$X' = \hat{X} + \alpha * \left(\hat{Y} - \hat{X}\right)$$

Equation 1

X' is the updated cluster feature vector. $\hat{X}$ is the old cluster feature vector (normalized), and $\hat{Y}$ is the pattern feature vector. α is the influence coefficient.

$$\hat{X}' = \frac{X'}{\|X'\|}$$

Equation 2

X' is the updated cluster vector (normalized). X is the cluster vector calculated in Equation 1.

Initially, I experimented with simple linear functions for radius and alpha terms. These seemed somewhat effective, but I wanted to improve them. After some rather unsuccessful experiments with logistic functions, I consulted the text, which suggested an exponential decay function for the radius and a Gaussian function multiplied by an exponential decay function for alpha.

$$r = r_{init} * e^{-\frac{t}{\tau_r}}$$

Equation 3

r is the radius of influence, as a function of an initial radius r_{init}, a time constant τ, and time t.

$$\alpha = \alpha_{init} * e^{-\frac{t}{\tau_\alpha}} * e^{-\frac{d^2}{2*r^2}}$$

Equation 4

α is the influence coefficient, as a function of initial coefficient α_{init}, a time constant τ_α, the distance of the cluster from the best match cluster, and the radius of influence r, which is calculated in equation 3.

Using equations 3 and 4 from the text showed some promise. I started tuning from the values suggested in the text: initial radius of 3 (half the radius of the output image), radius decay constant of about 2000 (so the radius decays to 1 in 1000 iterations), and an alpha decay constant of VALUE, so that alpha decays to VALUE in 1000 iterations. AUTHOR recommended these values to that the first 1000 iterations order the map and the remaining iterations refine it.

I began testing these methods by doing 50000-iteration runs of the code (hooray for quad-core i5 processors). In practice, I found that these values were not very well suited to the task, so I experimented with them. I found that reducing the initial radius improved the results up to a point. When the initial radius is reduced below about 2, the responses don't cluster properly. Instead, isolated "islands" form in the test pattern responses, as shown in figure 1.

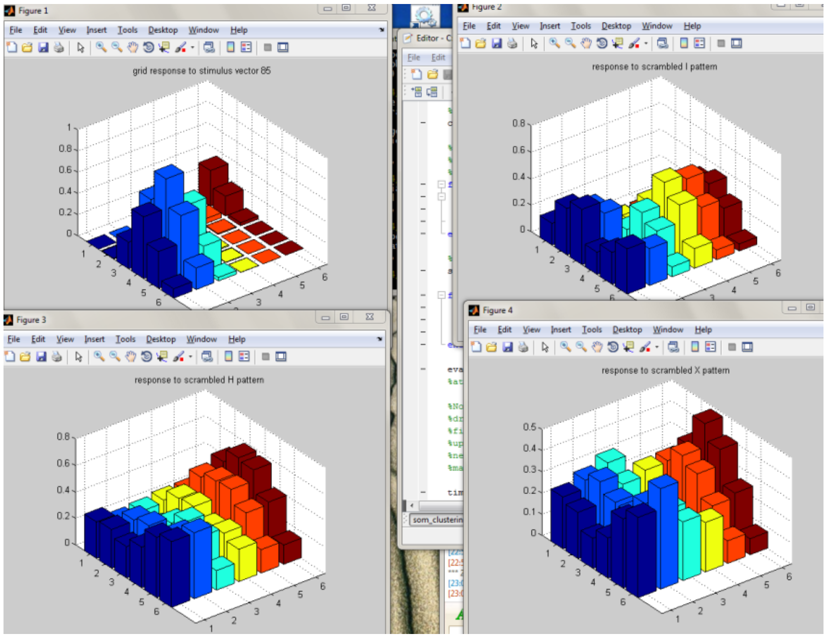


Figure 1
File Edit View Insert Tools Desktop Window Help
grid response to stimulus vector 85
Figure 2
response to scrambled I pattern
Figure 3
response to scrambled H pattern
Figure 4
response to scrambled X pattern
som_clusterin

Figure 1: Initial radius too low (r=1).

Having decided to set my initial radius to about 2, I started varying the other parameters. I found that increasing the time constants, made the algorithm more effective. Essentially, this lengthens the "ordering" period of the algorithm, before it proceeds to refining the clusters. I ultimately decided on 16000 for the radius time constant and 50000 for the alpha time constant. I also varied the initial value of alpha, increasing from .2 to .3 and .4. I found that the higher initial alpha resulted in poor performance, particularly on the H pattern.

I did some limited exploration into the effects of changing the number of clusters. I found that there was little if any advantage to increasing the number of clusters from 36 to 64, despite the fact that the input vectors were of length 64.

Results and Conclusions

I had two sets of values that yielded promising results after trial runs of 50000 iterations.

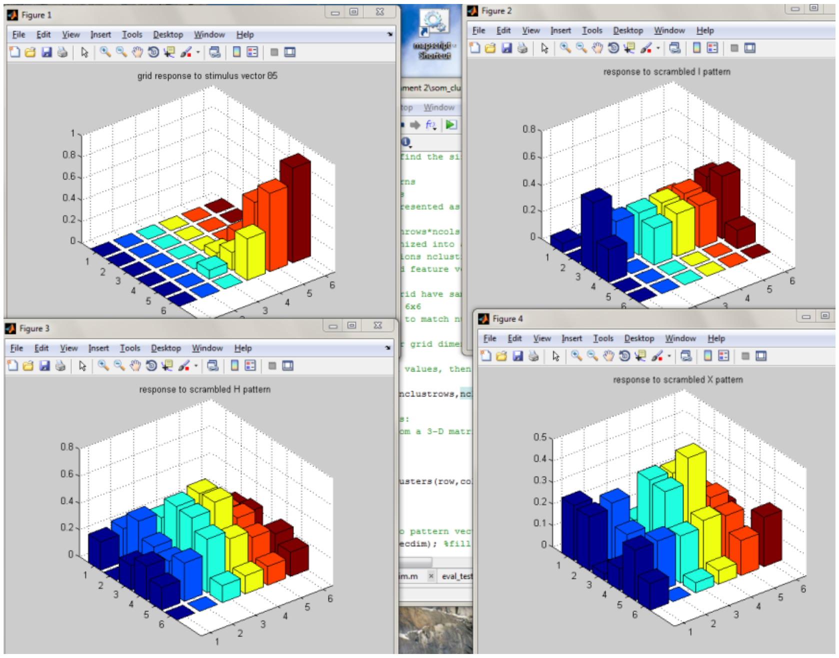


Figure 1
grid response to stimulus vector 85
Figure 2
response to scrambled I pattern
Figure 3
response to scrambled H pattern
Figure 4
response to scrambled X pattern

Figure 2: Clustering results after 50,000 iterations using r_{init}=2, t_r= 16000, α_{init}= .3, τ_α=50000

As you can see in figure 2 above, the H pattern was easily the worst of the three with every combination of parameters I tried. Even with the best values I could come up with, the H is mostly garbled. The X fared slightly better. The center is clearly defined, and the empty spaces on the sides are also well-defined. The I has the characteristic bar down the center, but the top and bottom bars are less clear.

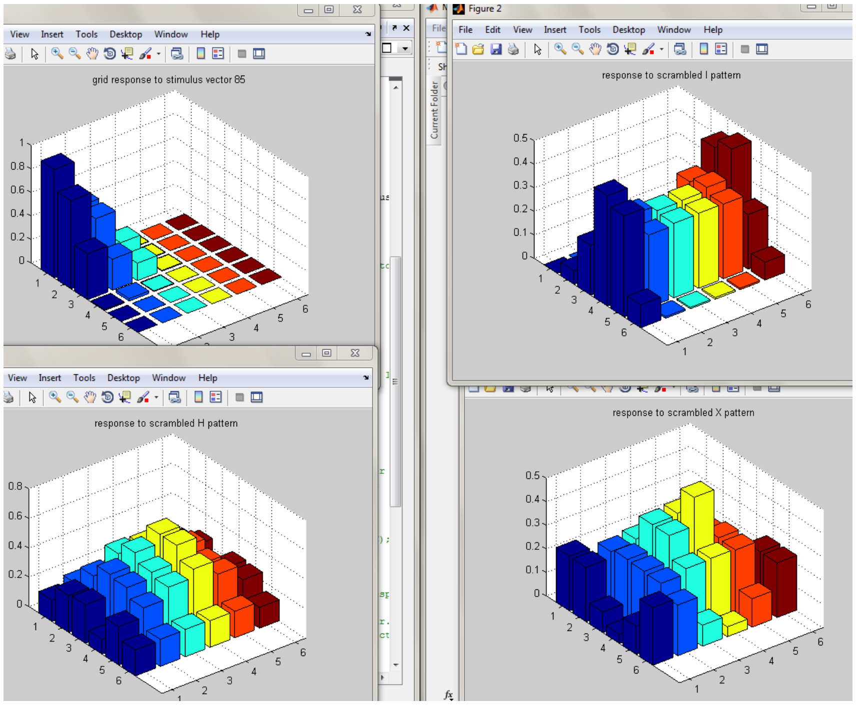

grid response to stimulus vector 85
Figure 2
response to scrambled I pattern
response to scrambled H pattern
response to scrambled X pattern

Figure 3: Clustering results after 50,000 iterations using r_{init}=3, t_r= 16600, α_{init}= .2, τ_α=10000

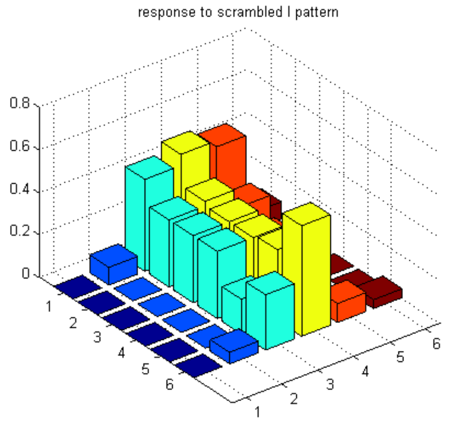
In figure 2 above, I used a smaller alpha that decays faster, but I used a larger initial radius of influence. This seems to result in better output resolution on the I and X patterns, but the H pattern is still difficult to make out. It is possible that running the algorithm longer would resolve the H, but I think that the H may be impossible to resolve without very carefully chosen parameters.

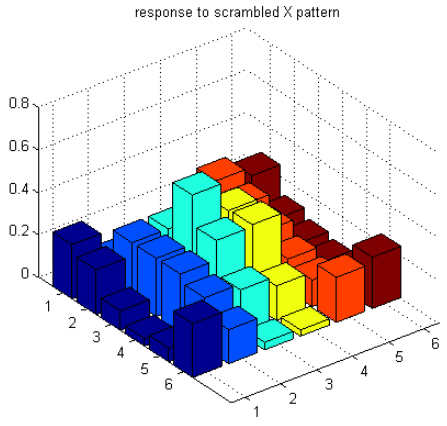
Two interesting effects that I noticed were that the output was arbitrarily rotated from the original. The clustering algorithm has no concept of how to orient the clusters, as long as the shapes are the same, so the orientation is entirely decided by the random seeding.

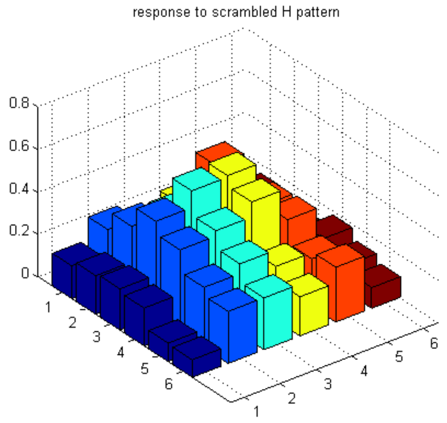
Also interesting was the fact that changing the number of output clusters from 6x6 to 8x8 had little to no effect on the quality of the results. Although we are using the algorithm here for ordering, it is in fact a clustering algorithm, so the number of features in the pattern need not be the same as the number of clusters.

Appendix: Results of Running Algorithm for 100,000 Iterations

Parameter	Value
Clusters	6x6
Radius_init	3
Tau_r	16600
Alpha_init	.2
Tau_a	50000


response to scrambled I pattern


response to scrambled X pattern


response to scrambled H pattern