

Note that filter groups are the horizontal neighbors in all plots. All pooling was with no overlap.

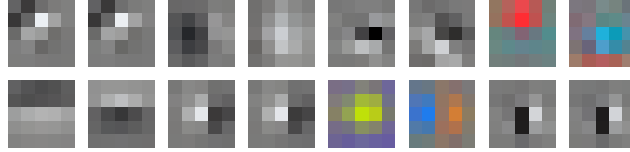


Figure 1: L_2 pooling across features only (groups of 2)

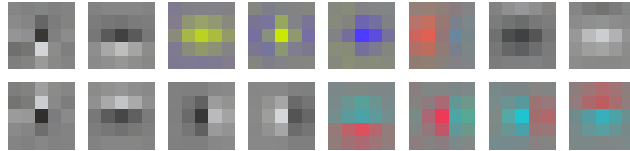


Figure 2: Max pooling across features and space (groups of 2, 4 spatial neighbors). Weak slowness penalty.

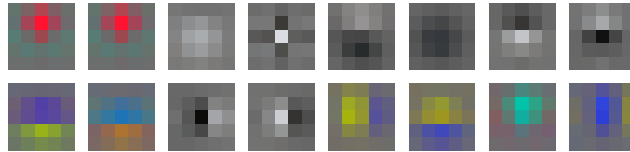


Figure 3: Max pooling across features and space (groups of 2, 4 spatial neighbors). Strong slowness penalty.

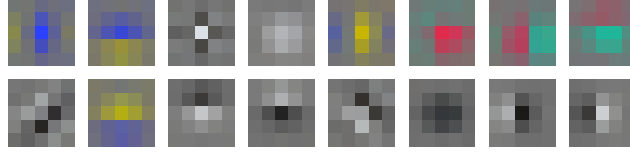


Figure 4: Max pooling across features and space (groups of 2, 16 spatial neighbors). Note the weak group association when the spatial neighborhood is increased.

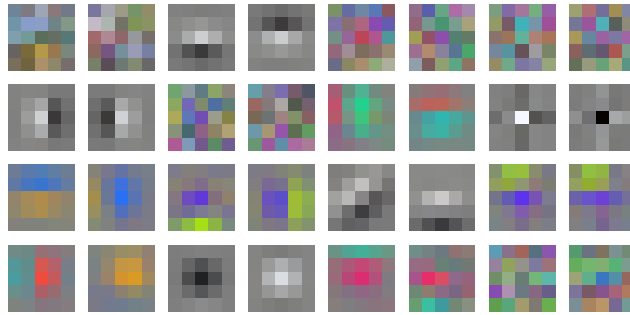


Figure 5: *Training in progress, has not converged yet.* Max pooling across features and space (groups of 2, 4 spatial neighbors).

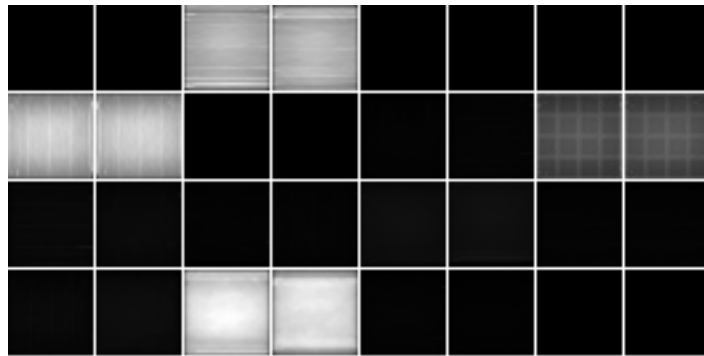


Figure 6: *Training in progress, has not converged yet.* Activations averaged over the dataset for the filters above. I think 'plaid patterns' are due to compression artifacts in the images that it picks up on. Note that, by far the most popular features are local low-pass followed by horizontal and vertical edges.