Unsupervised Discovery of Functional Neural Circuits in Visual System

Zhang Liu [1] (Mentors: Steven W. Zucker [2] Luciano Dyballa[2])

[1]Yale-NUS College [2]Department of Computer Science, Yale University

Research Question: Neural Data Analysis

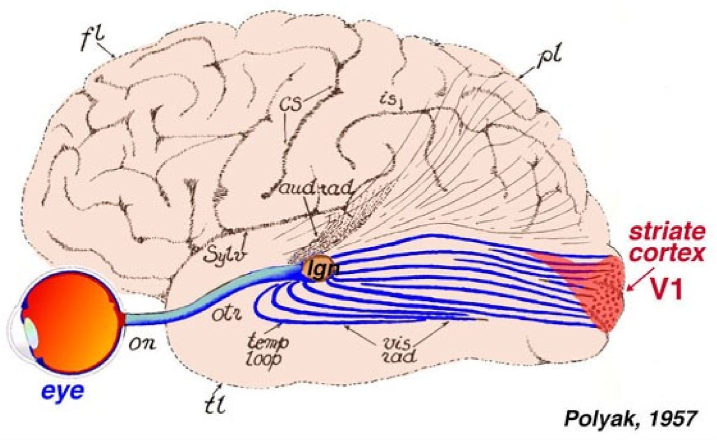


Figure 1. Visual input goes from the eye to LGN and then to primary visual cortex (V1). Adapted from Polyak (1957).

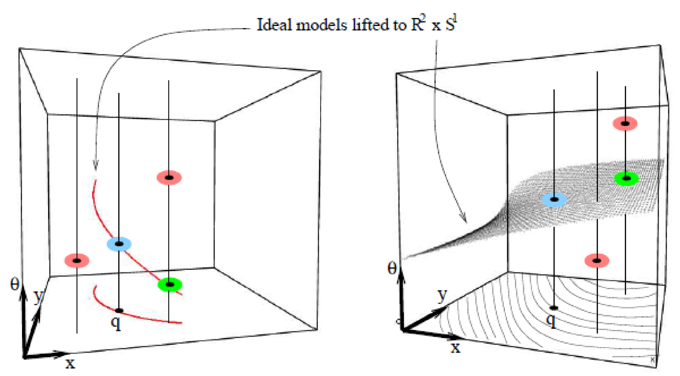


Figure 2. Mathematical model of the cortical columns in V1.

- How do humans process visual information?
- How are the neurons organised and connected in the primary visual cortex?
- Experimental data: Neural spiking data were recorded and encoded in (PSTH) diagrams, each of which shows the firing rate of some neuron over time, in response to some visual stimuli of artificial gratings.

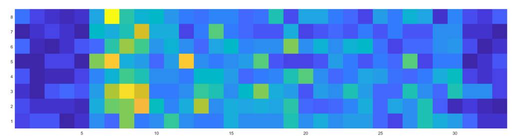


Figure 3. Visualising PSTH diagram.

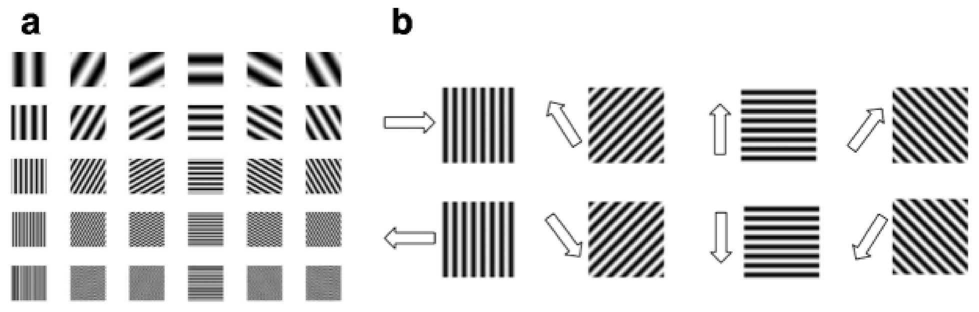


Figure 4. Visual stimuli: artificial gratings.

Background and Significance

- **Challenges in neural data analysis:** limited experimental results, very high-dimensional, complex geometric structure...
- **What we see is not what is actually out there.** It is instead an incredibly cleverly processed version of the actual world. Understanding how visual perception works gives us insights to the most powerful visual machine: our own visual system.
- **Current deep neural networks for computer vision are very different from the actual process in the brain.** More accurate mathematical models of the visual system can inspire more powerful algorithms (reverse engineering).

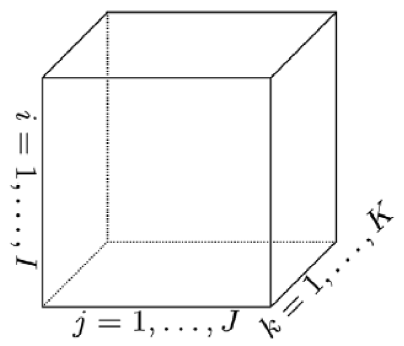
Methods

Neural spiking data
CNN: Transfer Learning
→ Artificial neuron tensors
→ Neuron tensors
→ Tensor factorisation
→ Neurons grouped by similar responses
→ Diffusion map (non-linear dimensionality reduction)
→ Manifolds of neurons
→ Analysing the coordinates and neighborhoods
→ Local functional circuits

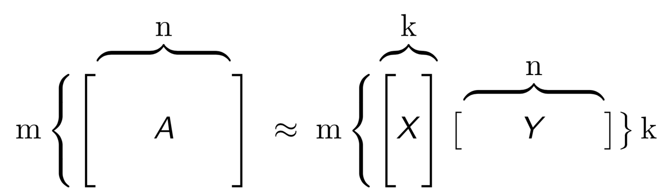
Linear Dimensionality Reduction: Tensor Factorisation

From matrices to tensors: N-way tensor is an element of the tensor product of N vector spaces.

- 1-way tensor = vector, $v = [v_1 \quad v_2 \quad \dots \quad v_n]^T$.
- 2-way tensor = matrix, $A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}$.
- 3-way tensor of size I-by-J-by-K:

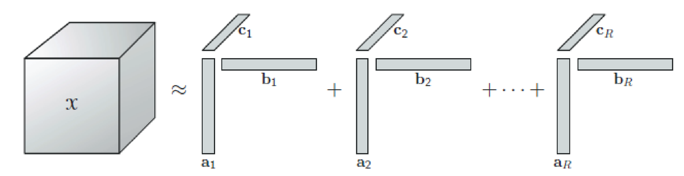


From matrix factorisation to tensor factorisation:



(a)

(b) Intuition for matrix factorisation.



(c)

(d) Intuition for tensor factorisation. (Adapted from [1].)

Demonstration with Face Recognition

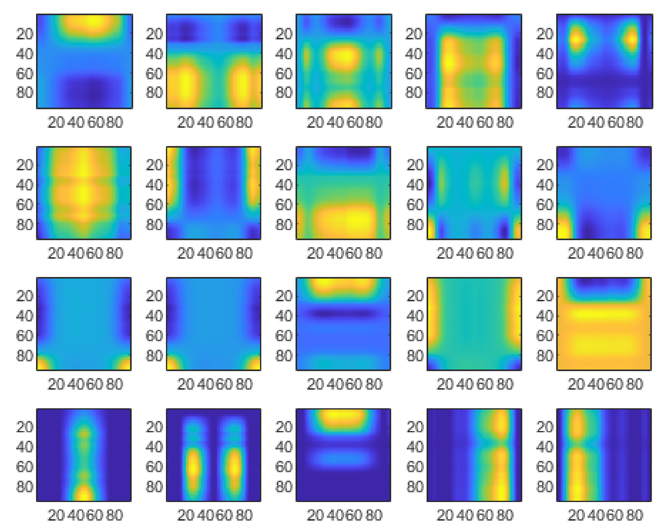


Figure 6. First 20 tensor factors for face image data.

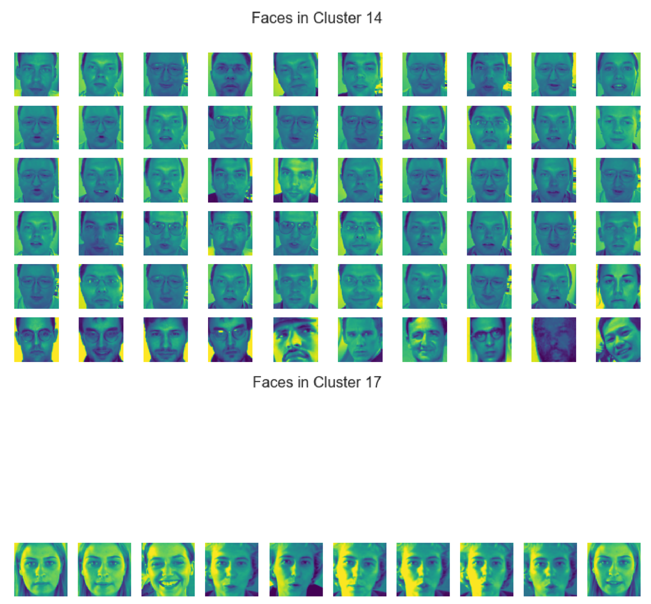


Figure 7. Two arbitrary clusters.

Application to Neural Tensors

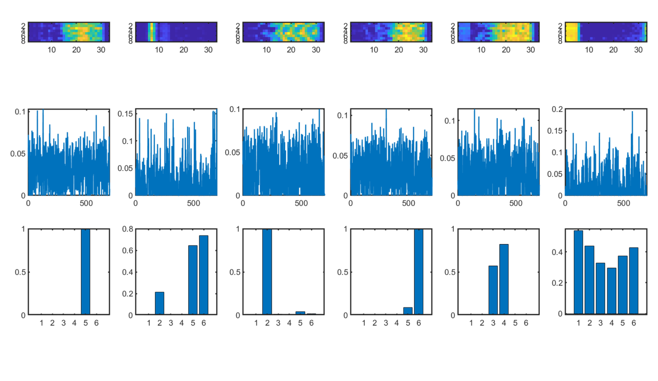
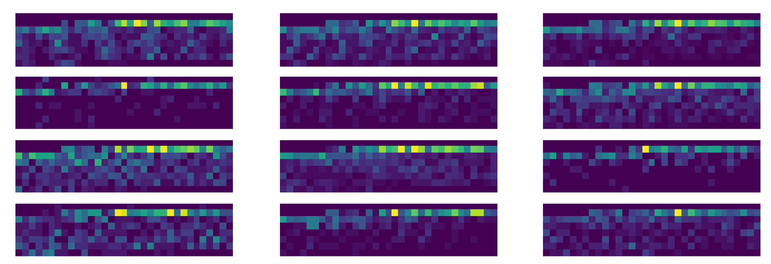
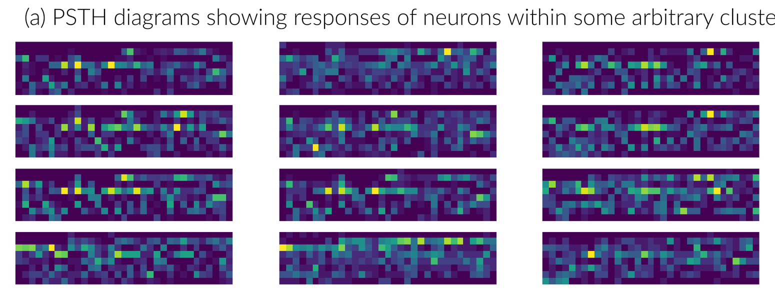


Figure 8. Visualising the first six tensor factors for retina data.



(a) PSTH diagrams showing responses of neurons within some arbitrary cluster.



(b) PSTH diagrams showing responses of neurons within a different cluster.

Non-Linear Dimensionality Reduction: Diffusion Map

When the coordinate function from the data points to their respective positions is highly non-linear, tensor factorisation might not work well. Thus, we introduce diffusion map, a non-linear dimensionality reduction technique. The main idea behind diffusion map is to define a new set of coordinates such that we "redistribute" the points on some lower-dimensional manifold while preserving the "closeness information" in the original high-dimensional space. The formal steps to derive such set of coordinates (called "diffusion maps") are outlined as follows.

- *Similarity Matrix, W*: $W_{ij} = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2})$.
- *Random walk matrix, also Markov matrix, M*: $M = D^{-1}W$, where $D_{ii} = \sum_k W_{ik}$.
- *Probability of reaching node j from node i after t steps*: $p(t, j \mid i) = e_i^T M^t e_j$.
- *Eigendecomposition* of the Markov matrix M: $M = D^{-1/2}\Omega\Lambda\Omega^T D^{1/2}$, where Ω contains the eigenvectors of $M_s = D^{1/2}MD^{1/2}$ and Λ contains the eigenvalues of M_s.
- *Diffusion maps*: the columns of $\gamma = D^{-1/2}\Omega$.
- *Diffusion distance*: $\|e_i^T\gamma - e_j^\gamma\| = \sum_k (p(t,k \mid i) - (p(t,k \mid j))^2 w(k)$, where $w(k) = \frac{1}{d}(k) = D_{kk}^{-1}$.

Demonstrations with High-dimensional Data

Applying diffusion map to synthetic spiral data and MNIST handwritten digits images data[2]:

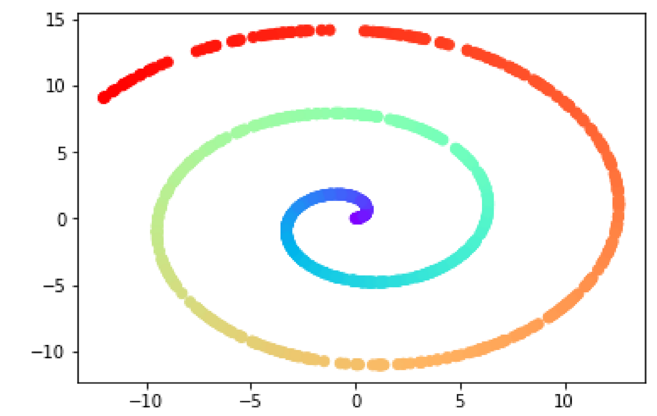


Figure 10. Visualising the original spiral data.

Figure 12. Sample data from the MNIST.

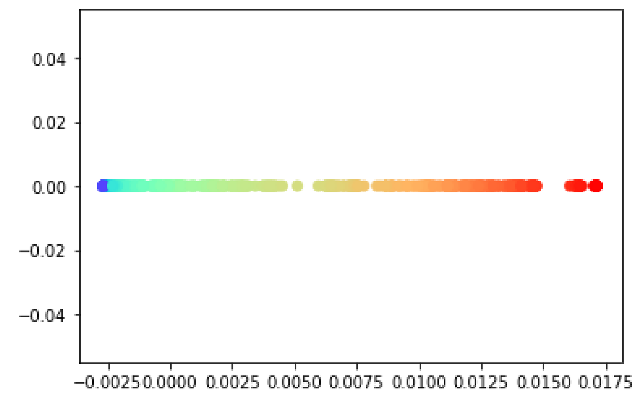


Figure 11. First non-trivial coordinate function.

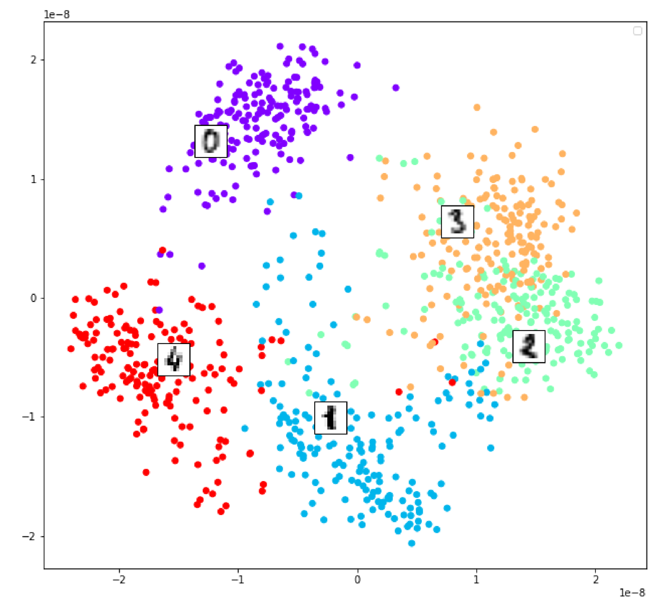


Figure 13. First two non-trivial coordinate functions.

Conclusion and Future Work

In this project, we demonstrated how linear and non-linear dimensionality reduction techniques can be used to discover the structure of high-dimensional data in an unsupervised setting. We would like to develop these techniques further to infer and compare the manifold structure in both biological and artificial neural networks. Eventually, we hope to come up with a more accurate mathematical model of visual perception, which is a significant task not only for understanding how perception works but also for inspiring better computer vision algorithms.

References

[1] Tamara G. Kolda and Brett W. Bader. Tensor decompositions and applications. *SIAM Rev.*, 51(3):455–500, August 2009.

[2] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.