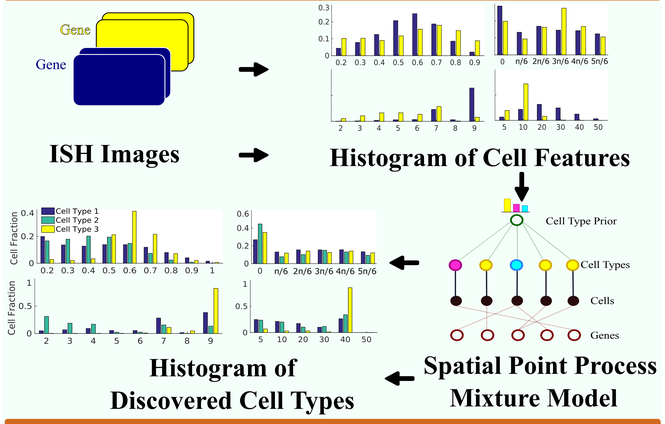
Discovering Neuronal Cell Types and Their Gene Expression Profiles Using a Spatial Point Process Mixture Model

*F. Huang, A. Anandkumar, C. Borgs, J. Chayes,
E. Fraenkel, M. Hawrylycz, E. Lein, A. Ingrosso, S. Turaga

Summary

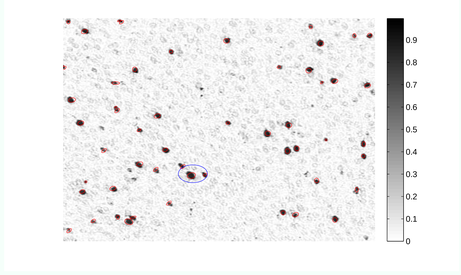
- **Goal**: Cataloging the neuronal cell types that comprise circuitry of individual brain regions
  - Major goal of modern neuroscience
  - Number 1 task of the Obama BRAIN initiative
- **Method** to infer cell types and their gene expression profiles
  1. computational analysis of brain-wide cell resolution in situ hybridization (ISH) imagery.
  2. measure spatial distribution of neurons in the ISH image for each gene
  3. model it as a spatial point process mixture
     - mixture weights = the distribution of cell types which express that gene
- **Validation** using single cell RNA sequencing data
  - Our results show superiority (better perplexity) over an *average expression level* model

Overall Flow



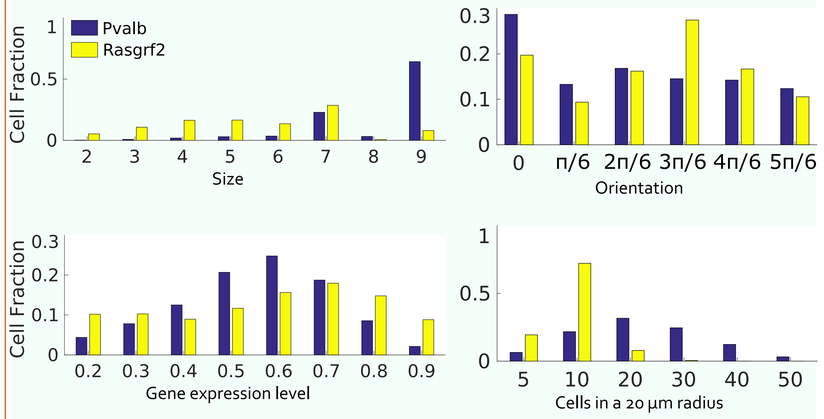
Extracting Cell Features

A cellular resolution *in situ* hybridization image for gene Pvalb along with detected cells:



Histogram Features

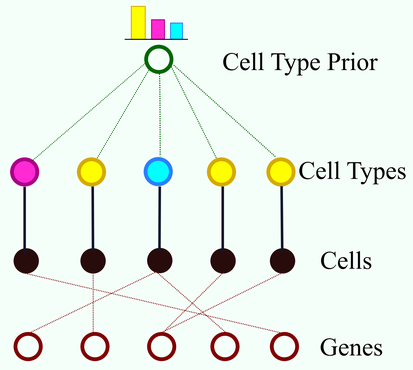
Marginalized point process feature histograms for genes *Pvalb* and *Rasgrf2*:



Pvalb and *Rasgrf2* are well-known markers for a specific class of *inhibitory* and *excitatory cortical* neuronal cell-types respectively.

Spatial Point Process Mixture Model

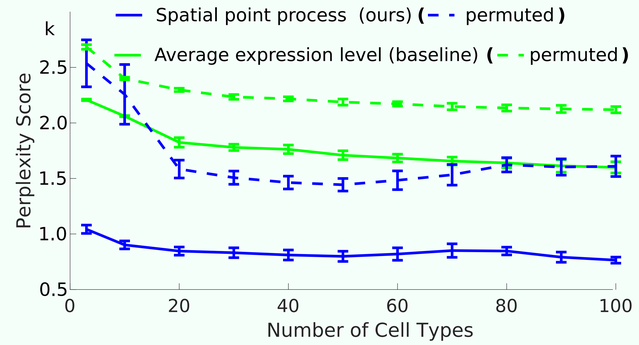
Method and Model



1. ISH image (for each gene) → point process features (for each gene)
2. model the point process features as a mixture of point processes belonging to individual cell-types
3. Motivation: cells allocated in the same joint histogram bin have similar
   - sizes, orientations, gene expression profiles, spatial distributions
4. why LDA model with a Dirichlet prior?
   - encourage this sparse membership of neuron types

Validation of Our Model

Comparison of gene expression profiles recovered for cell-types using **spatial point process features (ours)** vs **standard average gene expression level feature (baseline)**.



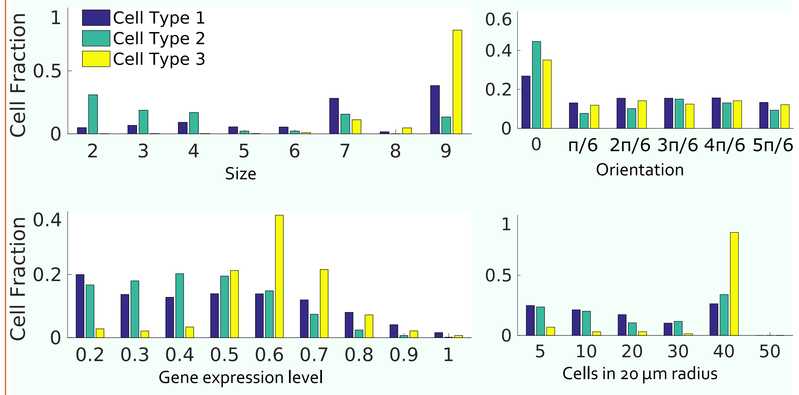
Discovered Neuronal Cell Types

Cell types and Top expressed genes within

- Three discovered cell types overlap of gene expression profile with three broad cell-type categories
  1. *astrocytes*
     - Putative astrocyte (cell type 1) expresses *Gfap* as expected.
  2. *interneurons*
     - Putative interneuron (cell type 2) expresses classic interneuron markers: *Gad1*, *Gad2* and *Sst* which is a marker for the somatostatin expressing interneuron sub-type.
  3. *oligodendrocytes*

Type	Top Genes			
Cell-type 1	***Gfap***	*Itga3*	*Asb4*	*Actr2*
Cell-type 2	***Sst***	***Gad2***	***Gad1***	*Ifnar1*
Cell-type 3	*Bhlhe22*	*Ptgds*	*Srgap3*	*Cry2*

Estimated histograms for the cell types:



*Email: furongh@uci.edu

NIPS workshop 2015