# Uniform Image Selection for fMRI studies

Charles Zheng

Stanford University

March 10, 2015

# How should researchers choose images?

- *Relevance*. Using natural images, because we care about how the brain percieves typical scenes in the real world.
- *Control*. Using artificially generated images (like gratings) with known parameters
- *Diversity*. Using a broad set of images to convincingly demonstrate generalizability

In this work we only deal with the question of *diversity*

# Questions about diversity

1. How can we define diversity?
2. How can researchers maximize this measure of diversity in their image set?
3. How does diversity affect metrics such as classification performance?
4. How can diversity improve generalizability?

To explore these questions, we use the dataset from *Kay et al.*

Section 2

# Defining diversity

# Defining diversity

- A first step in defining diversity is *dimensionality reduction* of images to a low-dimensional representation
- Next, we need a *distance metric* on the low-dimensional space
- With coordinates and a distance metric, we can define a measure of diversity

# Dimensionality reduction

- Suppose we have "training data" for $K$ images
- A $K \times p_f$ matrix $X$ of image features (e.g. Gabor filters)
- A $K \times p_v$ matrix $Y$ of mean voxel responses
- Use *sparse canonical correlation analysis* to find a $p_f \times d$ image basis $U$ and a $p_v \times d$ voxel basis $V$
- Use the image basis $U$ to define coordinates

# A measure of diversity

- We have coordinates $z$ for any image, and a distance metric $d(z_1, z_2)$
- Suppose that the coordinates of all images form a continuous, compact support set $S$
- We will define the diversity of a distribution $P$ supported on $S$
- Let $Z_1$, $Z_2$ be random points drawn from $P$
- Let $\kappa(\cdot)$ be a bounded monotone decreasing function which goes to zero at infinity
- Define the diversity of $P$ as

$$-\mathbb{E}[\kappa(d(Z_1, Z_2))]$$

# A measure of diversity

- Let $\kappa(\cdot)$ be a bounded monotone decreasing function which goes to zero at infinity
- Define the diversity of $P$ as

$$-\mathbb{E}[\kappa(d(Z_1, Z_2))]$$

- We did not define the function $\kappa$... but in some sense the choice is unimportant!
- Claim: let $h$ be a bandwidth, and let

$$P_{\kappa,h} = \text{argmin}\mathbb{E}[k(d(Z_1, Z_2)/h)]$$

  Then let $P_\kappa = \lim_{h \to 0} P_{\kappa,h}$. There exist a unique distribution $U$ such that for all $\kappa$ satisfying a *postive definite* condition, $P_\kappa = U$. We call $U$ the *uniform distribution*

Section 3

# Maximizing diversity