Alpha Divergence for Feature Pruning and Indexing of Large Biological Databases

John D. Gorman and Alfred O. Hero

Cytoprint, Inc., Santa Fe, NM

and

Departments of Electrical Engineering and Computer Science, Statistics, and Biomedical Engineering, University of Michigan, Ann Arbor, MI

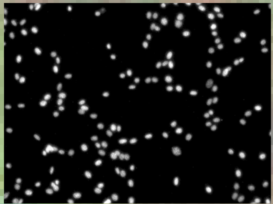
URSI 2002 National Radio Science Meeting, January 9-12, 2002, University of Colorado, Boulder, CO.

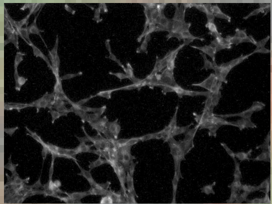
Outline

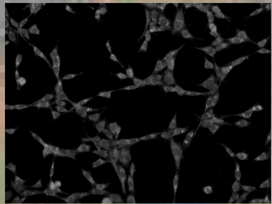
1. Indexing Application: Large-Scale Cellular Imagery Database
2. Statistical Framework: Alpha-Divergence for Indexing
3. Application Cellular Imagery
4. Summary, Open Issues

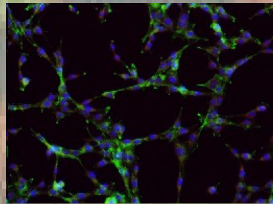
Cellular Imaging with Fluorescence Microscopy

- Can label and image several distinct cell components

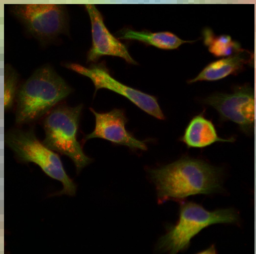

Nucleii

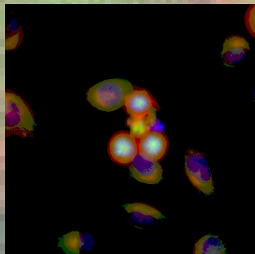

Actin


Mitochondria


Composite

- Can measure biological effect of drugs by tracking changes vs. dose


Before


After

Large-Scale Cellular Imagery Database

- Combinatorial Biology: large biological effects database

  ~ 30 cell lines (heart, liver, cervix, prostate, lung, etc.)

  ~ 30 labels/markers (nucleus, membranes, cytoskeleton)

  ~ 10000 compounds (pharmacopia)

  ~ 100 images (4 replicates $\times$ 5 doses $\times$ 5 timepoints)

  $\Rightarrow\ 10^{10}$ images (300M multi-color image sets)

- Objectives:
  - Catalog biological effects of known compounds
  - Infer properties of novel compounds by comparison to known drugs

Methodology

- Combinatorial Biology $\Rightarrow$ Imaging Experiments
- High-Throughput Cellular Imaging $\Rightarrow$ Image Responses

$$\left\{\left(X^k, Y^k\right), k = 1, \ldots, K\right\}$$

X^k are images, $Y^k \in \mathcal{Y}$ are labels assigned by experiment protocol

- Image Responses $\Rightarrow$ Response Signatures (feature vectors)

$$Z^k = Z(X^k)$$

- Response Signatures $\Rightarrow$ Biological Effects

$$Z^k = Z(X^k) \sim f(z|Y^k)$$

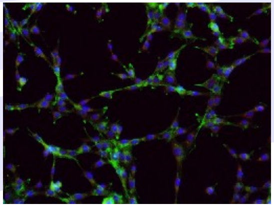
f a feature density (likelihood) conditioned on label Y^k

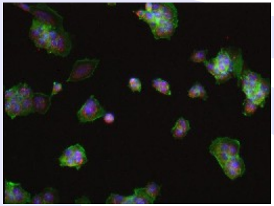
$\Rightarrow$ Biological Similarity? $X^1 \sim X^2$ if $f(z^1|Y^1) \sim f(z^2|Y^2)$

Comparison of Biological Effects

- Define a Cell Imaging Experiment
  $\mathcal{E}_k = \{\, Y^k \sim$ treatment k :
  n cell lines $\times$ m labels $\times$ p time points $\times$ q replicates $\}$
- Can we infer biological effects by comparing experiments?

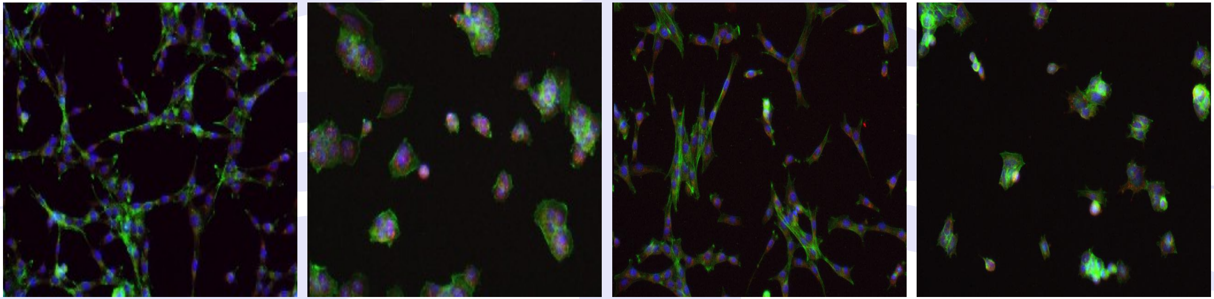
$Y^k \in \mathcal{Y} \stackrel{\text{def}}{=} \{$ Apoptosic (cell death), Normal (blank)$\}$
$\mathcal{E}_{\text{Blank}} = \{$ “Blank” : 400 imgs of untreated 3T3 cells $\}$
$\mathcal{E}_{\text{Apop}} = \{$ “Pos. Control” : 90 imgs of 3T3 w/ Paclitaxel$\}$


Blank


Pos. Control

Apoptosis Experiment

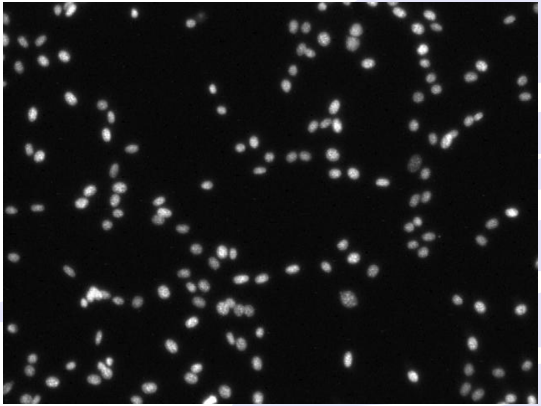
- Collected 1000 images of 10 different compounds, $\mathcal{E}_k, k = 1, \ldots, 10$
  - 3 known apoptosis inducers, 7 known non-apoptotic agents
  - Imaged 3T3 cells treated with various compound doses
- Objective: Classify compounds according to effect (apoptosis inducer?) (classify: $\mathcal{E}_k \sim \mathcal{E}_{\text{Blank}}$ or $\mathcal{E}_k \sim \mathcal{E}_{\text{Apop}}$)



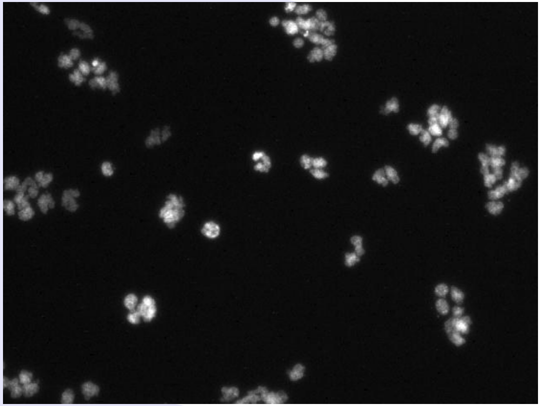
Blank | Pos. Control | Penicillin G | Vinblastine

Apoptosis Effects: Nucleus

- Nuclear Fragmentation (“blebbing”) and Condensation



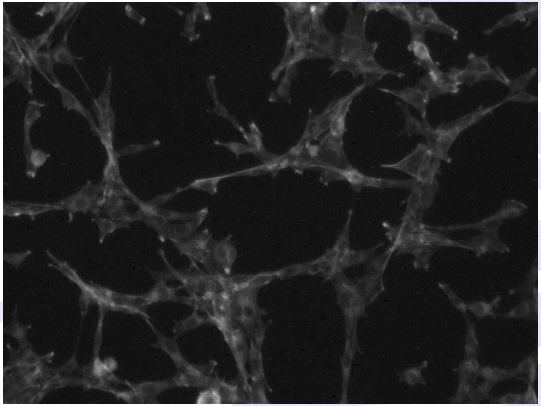
Normal



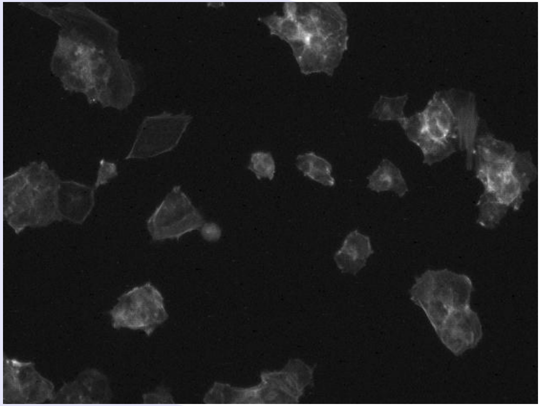
Apoptosis

Apoptosis Effects: Actin Filaments

- Actin Reorganization



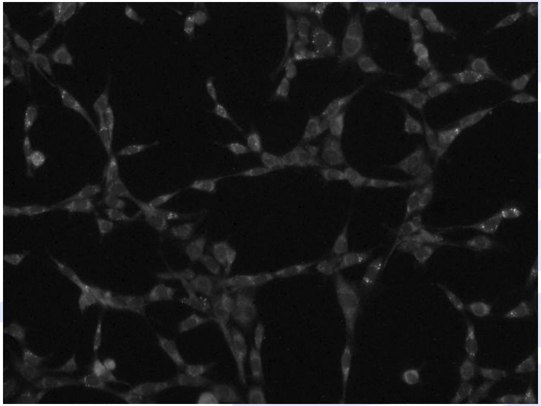
Normal



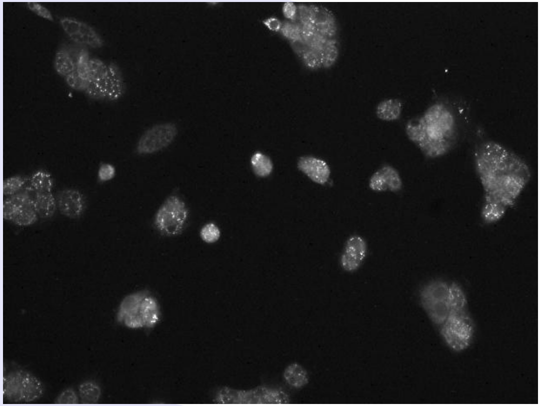
Apoptosis

Apoptosis Effects: Mitochondria

- Mitochondrial Disintegration



Normal



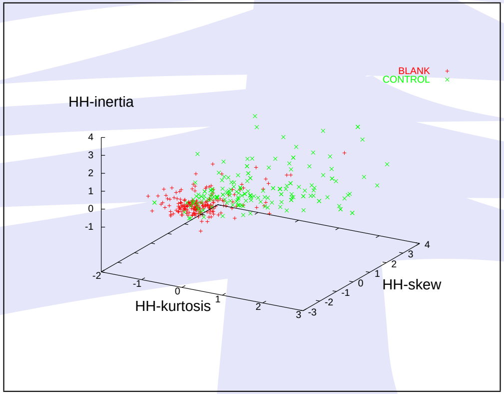
Apoptosis

Feature Extraction

- Top 12 Features:

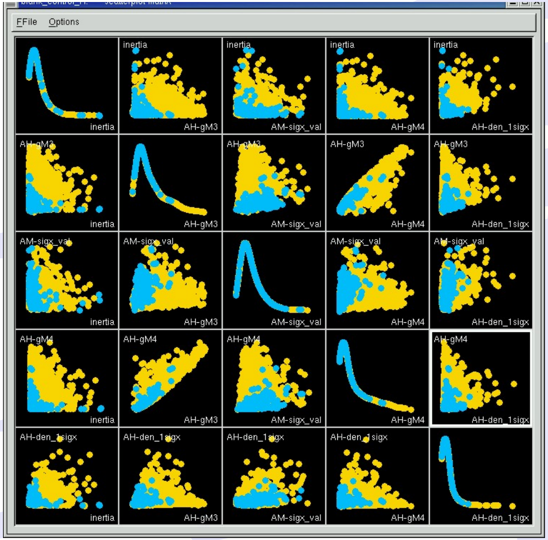
Nucleus	Actin	Mitochondria
H-Mass(Size in pix)	A-Mass(Size in pix)	HM-Coef.Var.
H-Coef.Var.	A-Rot.Inert.	M-2nd X-Mom. Inert.
H-Kurt.	AH-3rd Jt. Cumulant	M-Rot.Inert.
H-Skew	AH-4th Jt. Cumulant	
	AH-2nd Y-Mom. Inert.	

- Graylevel features from Nuclear channel:

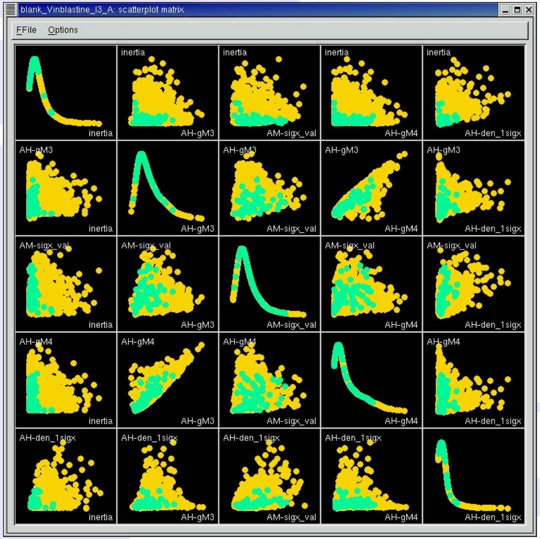


Feature Extraction

- Lattice plot of five selected features
  - Blank vs. Control and Blank vs. Vinblastine scatterplots visually similar



Blank vs. Control



Blank vs. Vinblastine

Candidate Feature Similarity Metrics

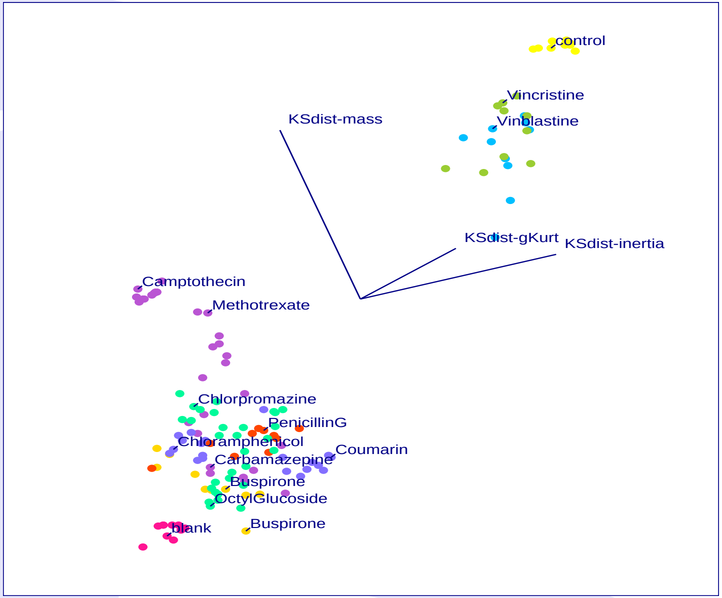
- Univariate: Kolmogorov-Smirnov Distance:
  - Product density approximation: combine several 1D feature scores
- Multivariate: Rényi α-Divergence
  - Computation via second-order approximation of densities
  - Direct computation via minimal graphs

Univariate Distance Measures

Score computed from 1D Kolmogoroff-Smirnoff distance

$$D_b(z) = D_{KS}(f_1(z) \| f_{\text{blank}}(z))$$
$$D_c(z) = D_{KS}(f_1(z) \| f_{\text{control}}(z))$$
$$D_{bc}(z) = D_{KS}(f_{\text{blank}}(z) \| f_{\text{control}}(z))$$

$$\text{Score}(z) = (D_b(z) - D_c(z))/D_{bc}(z)$$



Renyi α-Divergence

Define: $f_i = f(z|Y^i), f_0 = f(z|Y^0)$

The Rényi α-divergence of fractional order $\alpha \in [0,1]$ [Renyi:61, 70]

$$\begin{aligned} D_\alpha(f_i \| f_0) &= \frac{1}{\alpha - 1} \ln \int f_0 \left(\frac{f_i(z)}{f_0(z)} \right)^\alpha dz \\ &= \frac{1}{\alpha - 1} \ln \int f_i^\alpha(z) f_0^{1-\alpha}(z) dz \end{aligned}$$

Note: $D_\alpha(f_i \| f_0) = D_\alpha(Y^i \| Y^0)$ is indexed by Y^i and Y^0.

Renyi α-Divergence: Special Cases

α-Divergence vs. Batthacharrya-Hellinger distance

$$D_{\frac{1}{2}}(f_i \,\|\, f_0) = \ln\left(\int \sqrt{f_i(z)\, f_0(z)}\, dz\right)^2$$

$$D_{BH}(f_i \| f_0) = \int \left(\sqrt{f_i(z)} - \sqrt{f_0(z)}\right)^2 dz$$

$$= 2\left(1 - \int \sqrt{f_i f_0}\, dz\right)$$

α-Divergence vs. Kullback-Leibler divergence

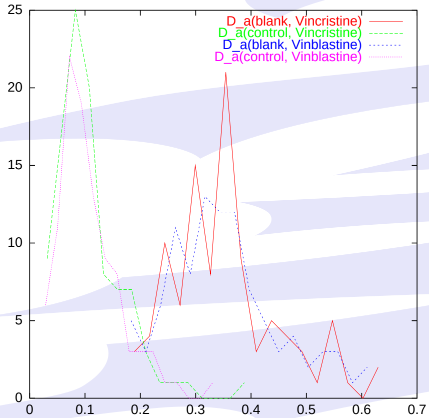
$$\lim_{\alpha \to 1} D_\alpha(f_i, f_0) = \int f_0(z) \ln \frac{f_0(z)}{f_i(z)}\, dz \tag{1}$$

Renyi α-Divergence via Approximation

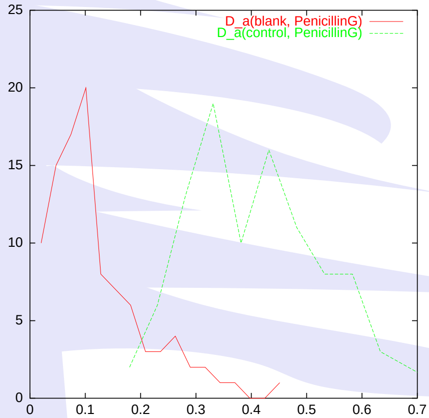
- Compute D_α using second-order approximation:

$$D^u_\alpha(f(x;\mu_1,\Lambda_1)\|f(x;\mu_0,\Lambda_0)) = \underbrace{-\frac{1}{2}\ln\frac{|\Lambda_0|^\alpha|\Lambda_1|^{1-\alpha}}{|\alpha\Lambda_0+(1-\alpha)\Lambda_1|}}_{\text{Term A}} + \underbrace{\frac{\alpha(1-\alpha)}{2}\Delta\mu^T(\alpha\Lambda_0+(1-\alpha)\Lambda_1)^{-1}\Delta\mu}_{\text{Term B}}$$

where $\Delta\mu = \mu_1 - \mu_0$.



$D_\alpha(\text{Vin}*\|\text{Blank})$ vs. $D_\alpha(\text{Vin}*\|\text{Control})$



$D_\alpha(\text{Penicillin G}\|\text{Blank})$ vs. $D_\alpha(\text{Penicillin G}\|\text{Control})$

Experiment Results

- Compounds used in Experiment

Training Data	Test Data
Negative Blank, Penicillin G	**Non-Apoptotic** Chloramphenicol, Buspirone, Sodium Azide, Carbamezapine, Octyl Glucoside, Carbon Tetrachloride
Positive Paxlitaxel (pos. control)	**Apoptotic** Methotrexate, Vincristine, Vinblastine

- Correctly classified Vincristine and Vinblastine as apoptosis inducers
- Methotrexate did not penetrate the 3T3 cells with the experiment protocol we used → did not induce apoptosis.

Summary and Ongoing Issues

- Successful Comparison of Biological Experiments
  - Similarity of cell image changes $\Leftrightarrow$ similarity of biological effect
- Rényi Divergence via 2nd Order Approximation
  - Reasonable fit for nuclear features (unimodal, symmetric)
  - Will be an issue for other feature densities (skew, multimodal)
- Open Questions:
  - What is the appropriate feature set for cellular imagery analysis?
  - Does image similarity imply biological similarity?