Thesis Proposal

Coupled Dictionary Learning for Image Analysis

Tian Cao

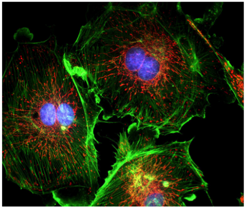
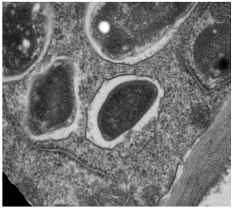
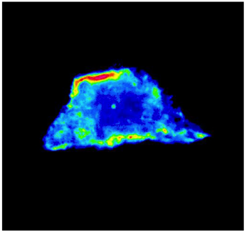
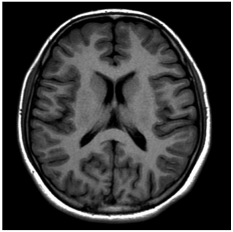
Department of Computer Science
University of North Carolina at Chapel Hill

Motivations

- Modern imaging technologies allow us to visualize various objects

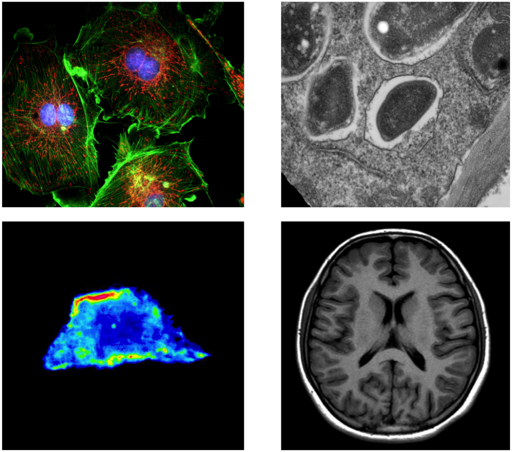
Motivations

- Modern imaging technologies allow us to visualize various objects

Motivations

- Modern imaging technologies allow us to visualize various objects
- Common problem in image analysis is how to relate the information from different modalities/sources



Multi-modal Image Registration

Correlative Microscopy

- combination of different microscopy technologies
  - Confocal
  - Transmission Electron Microscopy (TEM)

Requires image registration for joint analysis

Relating image appearances from different modalities

Confocal TEM

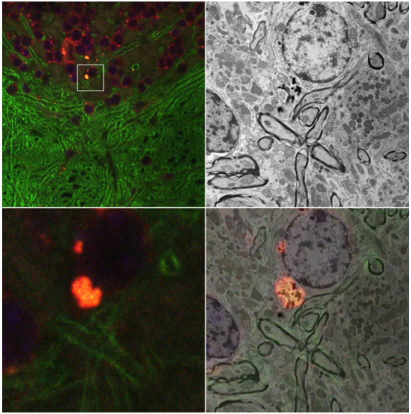
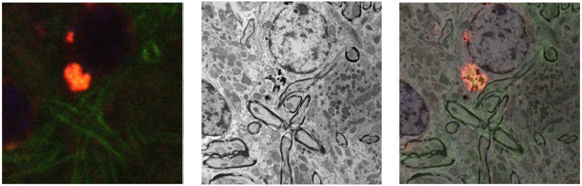


Image Registration

Image Registration is the process of estimating a spatial transformation between two images



(a) Source image (b) Target image (c) Registration result

Figure: Image Registration Example

Deformation Estimation

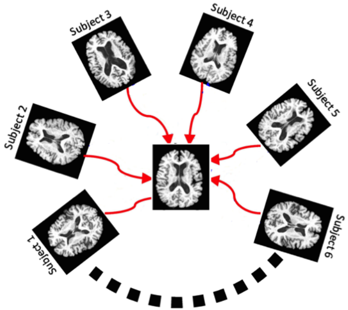
Estimating deformations to register subject images to a common reference image (atlas)

- Brain Images Analysis

Learning based approach

- estimating deformations from image appearances

Relating deformations with image appearances



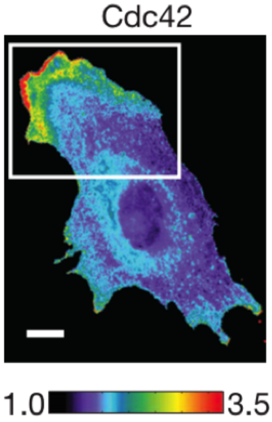
Coordination between GTPase Activations and Cell Movements in Mouse Embryonic Fibroblasts

GTPases are a family of enzymes that can bind Guanosine triphosphate(GTP)

Establishing correspondences between GTPase activations and cell movements

- crucial for understand cell dynamics

Relating spatiotemporal information



Similarities between The Three Examples

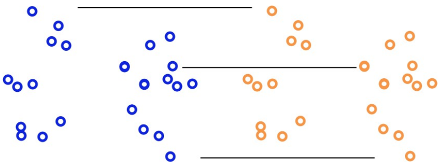
Examples:

- multi-modal registration for correlative microscopy
- deformation estimation from image appearances
- coordination between GTPase activations and cell movements

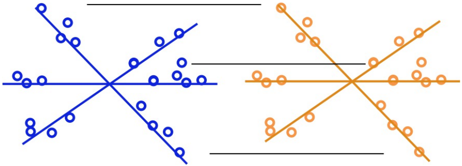
Similarities:

- relating data from two different spaces
  - spatial data: appearances
  - spatial data: appearances and deformations
  - spatiotemporal data: GTP activations and cell movements
- can be solved by modeling the relationship between the data in two spaces

My Solution

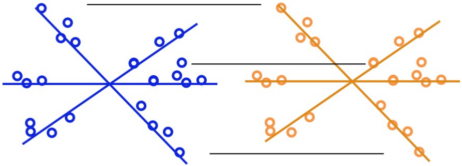


My Solution



- Learning a joint basis for compact representation of two spaces

My Solution



- Learning a joint basis for compact representation of two spaces
- Joint basis is learned by coupled dictionary learning

Thesis

- Learning a coupled basis for the compact representation of two spaces can be achieved by **coupled dictionary learning** (CDL).

Thesis

- Learning a coupled basis for the compact representation of two spaces can be achieved by **coupled dictionary learning** (CDL).
- Such dictionaries can be learned to capture **appearance differences** of different imaging modalities, dependencies between **image appearances and deformations** as well as the **spatiotemporal patterns** for cell signaling and boundary protrusions and retractions.

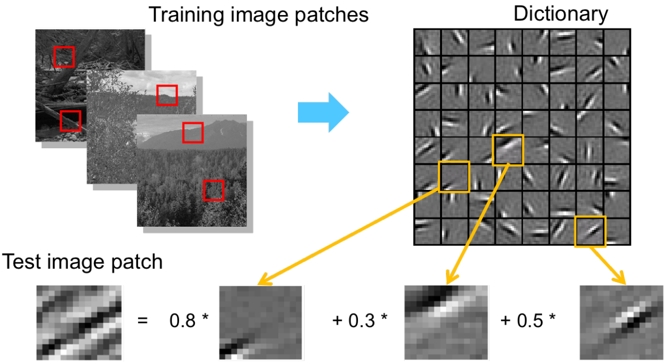
Thesis

- Learning a coupled basis for the compact representation of two spaces can be achieved by **coupled dictionary learning** (CDL).
- Such dictionaries can be learned to capture **appearance differences** of different imaging modalities, dependencies between **image appearances and deformations** as well as the **spatiotemporal patterns** for cell signaling and boundary protrusions and retractions.
- To account for data inconsistencies, a robust coupled dictionary can be obtained based on a **probabilistic dictionary model**.

Outline

- Background
- CDL for Multi-modal Image registration
- CDL for Deformation Estimation
- CDL for GTPase Activations and Cell Protrusions
- Robust CDL
- Summary and Future Work

Sparse Representation Models (SRM)



- Dictionary is a basis to represent the data in a space
- Each coefficient have few non-zero values (sparsity)

slide credit: Andrew Ng

Sparse Representation Model (SRM)

Suppose D is pre-defined, sparse representation model includes solving the optimization problem

$$\hat{\alpha} = \arg\min_{\alpha} \frac{1}{2}\|x - D\alpha\|_2^2 + \lambda\|\alpha\|_p$$

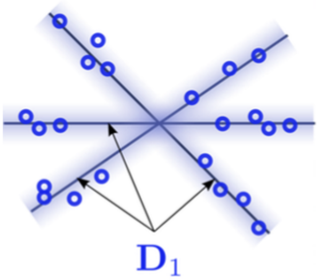
Usual sparsity-inducing regularizers

- $p = 0$: ℓ_0 norm, number of non-zero elements, non-convex, greedy algorithm
- $p = 1$: ℓ_1 norm, relaxation of ℓ_0 norm, convex optimization (lasso)

Dictionary Learning

$$(\hat{\alpha}, \hat{D}) = \arg\min_{\alpha, D} \sum_{i=1}^{m} \frac{1}{2} \|x_i - D\alpha_i\|_2^2 + \lambda \|\alpha_i\|_p$$

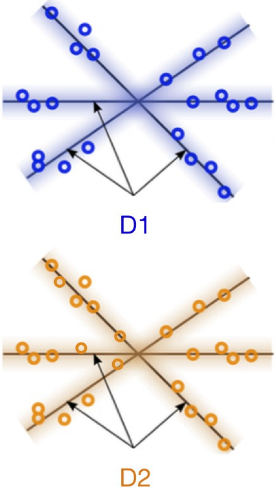
- Atoms satisfy $\|d_i\|_2 \leq 1$
- Usual Sparsity-inducing regularizes
  - ℓ_0 norm: $\|a_i\|_0$
  - ℓ_1 norm: $\|a_i\|_1$



Coupled Dictionary Learning

$$(\hat{\alpha}, \hat{D}^*) = \arg\min_{\alpha, D^*} \sum_{i=1}^{m} \frac{1}{2} \|x_i^* - D^* a_i\|_2^2 + \lambda \|a_i\|_1$$

- $x_i^* = [x_i^{1\,T}, x_i^{2\,T}]^T$,
- $D_i^* = [D_i^{1\,T}, D_i^{2\,T}]^T$
- Atoms satisfy $\|d_i^*\|_2 \leq 1$
- Use a single α to enforce the correspondences between data in two spaces



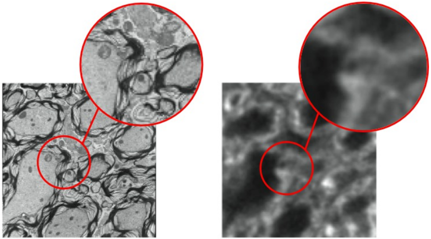
Outline

- Background
- CDL for Multi-modal Image registration
- CDL for Deformation Estimation
- CDL for GTPase Activations and Cell Protrusions
- Robust CDL
- Summary and Future Work

Multi-modal Image Registration

Challenges

- different appearances between different image modalities



Previous Methods

- distance measure: mutual information
- change image appearance [S. Yang 2011]
- convert image modality [W. Wein 2008]
- MR image synthesis [S. Roy 2011]

Multi-modal Image Registration

Challenges

- different appearances between different image modalities

My method

- image analogies (IA)
- simplify a multi-modal registration problem to a mono-modal one

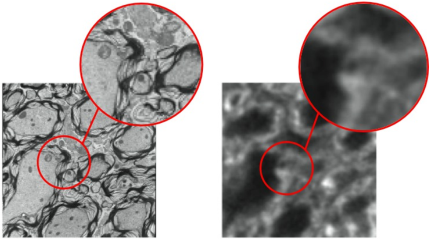


Image Analogies Problem

Training images

- image pairs A and A'
- A' is a "filtered" version of A

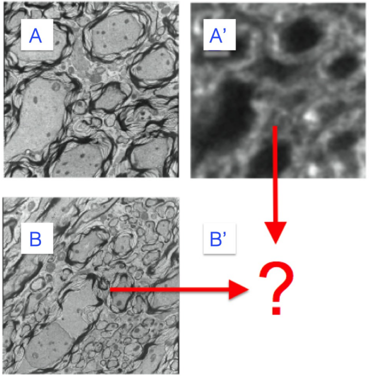
Input image

- source image B

Output image

- predicted "filtered" image B', based on the relationship between A and A'

Training Images

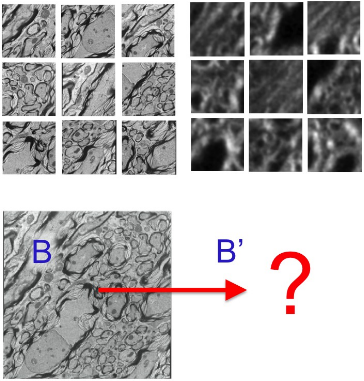


Input Image Predicted Image

Standard Image Analogies (IA)

- Generate image patch pairs from A and A'

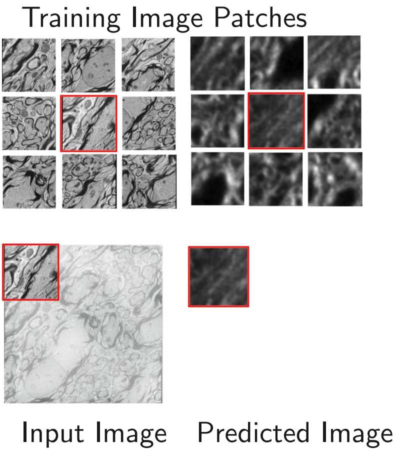
Training Image Patches



Input Image Predicted Image

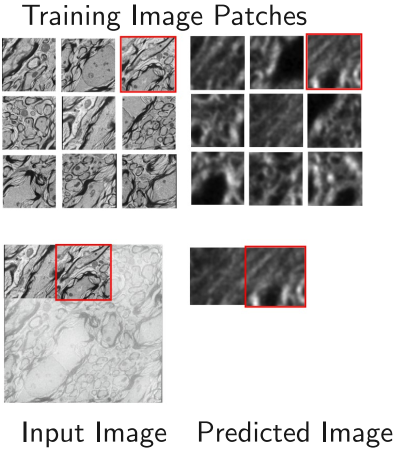
Standard Image Analogies (IA)

- Generate image patch pairs from A and A'
- Reconstruct source image B by searching the closest patches in training set
- Predict B' based on the found patches



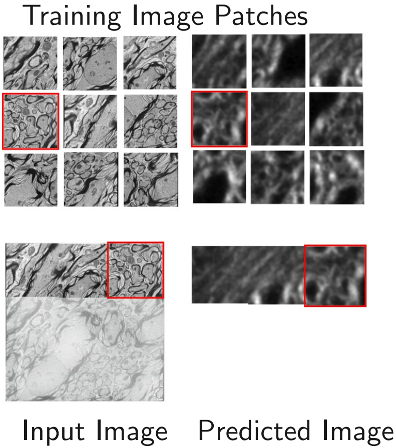
Standard Image Analogies (IA)

- Generate image patch pairs from A and A'
- Reconstruct source image B by searching the closest patches in training set
- Predict B' based on the found patches



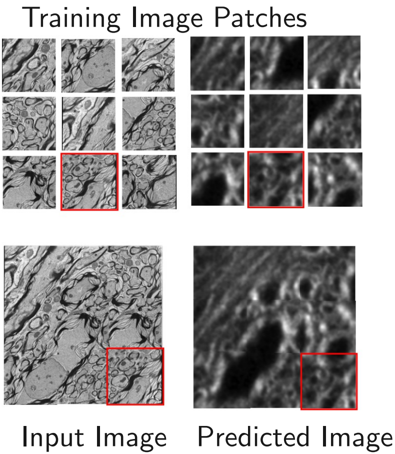
Standard Image Analogies (IA)

- Generate image patch pairs from A and A'
- Reconstruct source image B by searching the closest patches in training set
- Predict B' based on the found patches

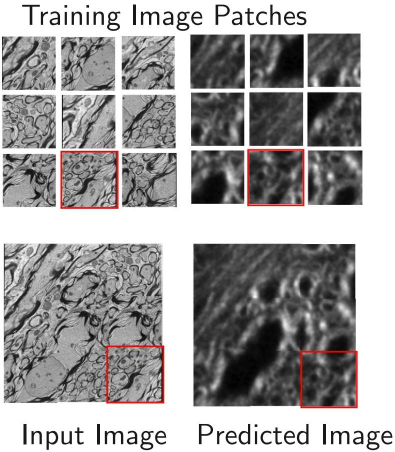


Standard Image Analogies (IA)

- Generate image patch pairs from A and A'
- Reconstruct source image B by searching the closest patches in training set
- Predict B' based on the found patches



Standard Image Analogies (IA)



Limitation:

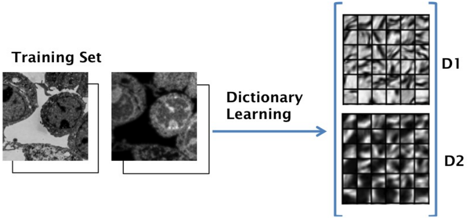
- nearest neighbor search in a large training set is computationally expensive

Proposed Method

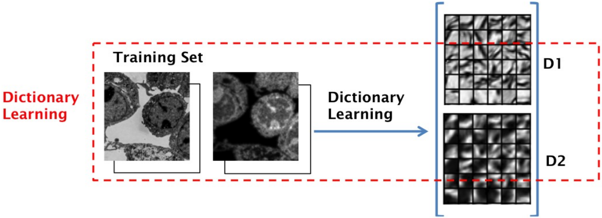
A general framework for image analogies based on

- coupled dictionary learning for image appearances from different modalities
- sparse representation model

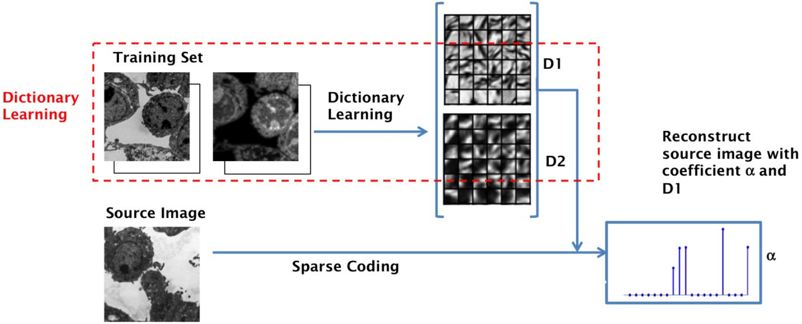
Framework of Proposed Method



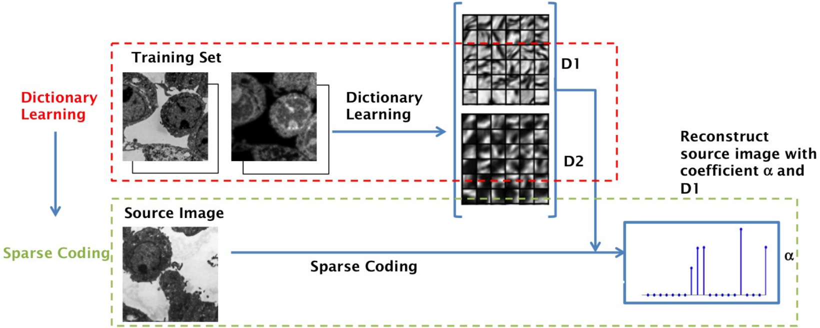
Framework of Proposed Method



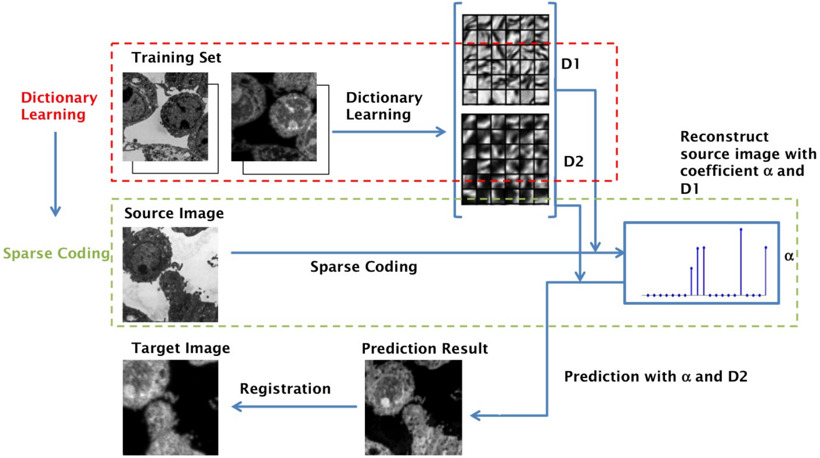
Framework of Proposed Method



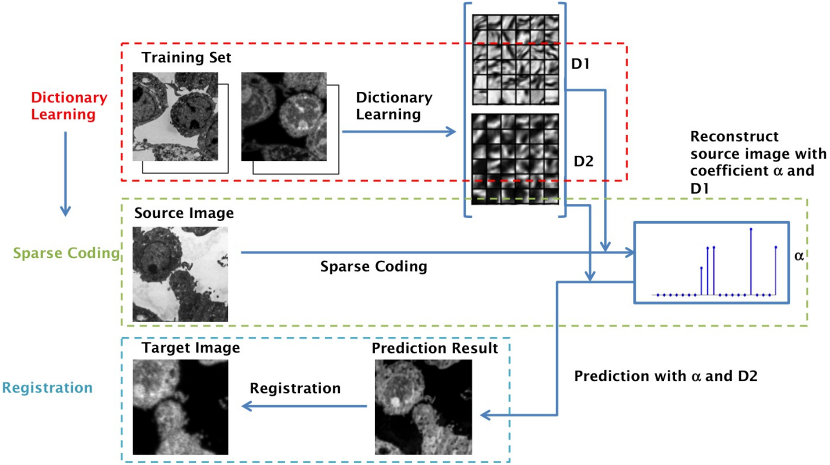
Framework of Proposed Method



Framework of Proposed Method



Framework of Proposed Method



Formulation of Image Analogies Problem with SRM

- f_1 is the input image B, f_2 is the corresponding image B'
- (D_1, D_2) are learned coupled dictionary
- u_1, u_2 are reconstructed images of f_1, f_2 respectively

Formulation of Image Analogies Problem with SRM

- f_1 is the input image B, f_2 is the corresponding image B'
- (D_1, D_2) are learned coupled dictionary
- u_1, u_2 are reconstructed images of f_1, f_2 respectively
- to estimate u_1, u_2 together without the knowledge of f_2

Formulation of Image Analogies Problem with SRM

- f_1 is the input image B, f_2 is the corresponding image B'
- (D_1, D_2) are learned coupled dictionary
- u_1, u_2 are reconstructed images of f_1, f_2 respectively
- to estimate u_1, u_2 together without the knowledge of f_2
- $R_i u$ selects ith image patch in u

$$E(u_1, u_2, \{\alpha_i\}) =$$

$$\boxed{\frac{\gamma}{2}\|u_1 - f_1\|_2^2} + \frac{1}{N}\left(\boxed{\frac{1}{2}\sum_{i=1}^{N}\left\|R_i \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} - \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \alpha_i\right\|_2^2} + \boxed{\lambda\|\alpha_i\|_1} \right)$$

- Similarity

Formulation of Image Analogies Problem with SRM

- f_1 is the input image B, f_2 is the corresponding image B'
- (D_1, D_2) are learned coupled dictionary
- u_1, u_2 are reconstructed images of f_1, f_2 respectively
- to estimate u_1, u_2 together without the knowledge of f_2
- $R_i u$ selects ith image patch in u

$$E(u_1, u_2, \{\alpha_i\}) =$$

$$\frac{\gamma}{2}\|u_1 - f_1\|_2^2 + \frac{1}{N}\left(\frac{1}{2}\sum_{i=1}^{N} \left\|R_i \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} - \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \alpha_i\right\|_2^2 + \lambda\|\alpha_i\|_1 \right)$$

- Similarity
- Reconstruction error

Formulation of Image Analogies Problem with SRM

- f_1 is the input image B, f_2 is the corresponding image B'
- (D_1, D_2) are learned coupled dictionary
- u_1, u_2 are reconstructed images of f_1, f_2 respectively
- to estimate u_1, u_2 together without the knowledge of f_2
- $R_i u$ selects ith image patch in u

$$E(u_1, u_2, \{\alpha_i\}) =$$

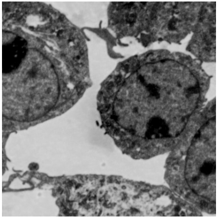
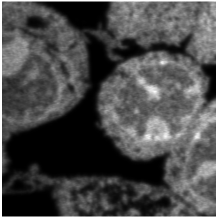
$$\boxed{\frac{\gamma}{2}\|u_1 - f_1\|_2^2} + \frac{1}{N}\left(\boxed{\frac{1}{2}\sum_{i=1}^{N} \left\|R_i \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} - \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \alpha_i\right\|_2^2} + \boxed{\lambda\|\alpha_i\|_1} \right)$$

- Similarity
- Reconstruction error
- Regularization

Results

Eight pairs of 2D correlative Scanning Electron Microscopy (SEM)/Confocal images

- with gold bead fiducials
- use landmark based registration for validation

Six pairs of 2D correlative TEM/Confocal images

- without fiducials
- manually choose landmarks for validation

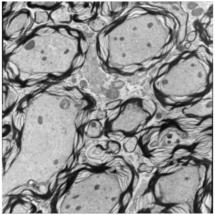
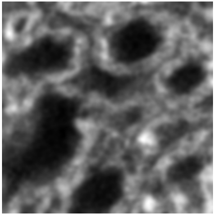
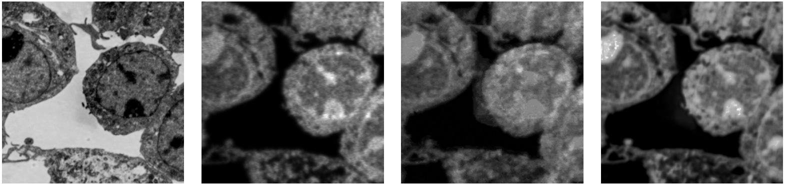



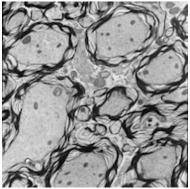
Image Analogies Results



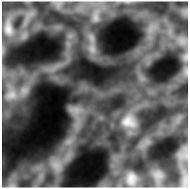
SEM Image Confocal Image Standard IA Proposed IA

Method	mean	std	p-value
Proposed IA	$\mathbf{1.52 \times 10^5}$	5.79×10^4	**0.0002**
Standard IA	2.83×10^5	7.11×10^4	

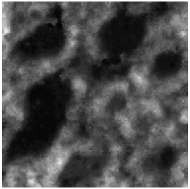
Image Analogies Results



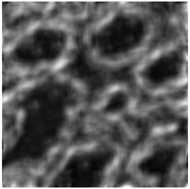
TEM Image



Confocal Image



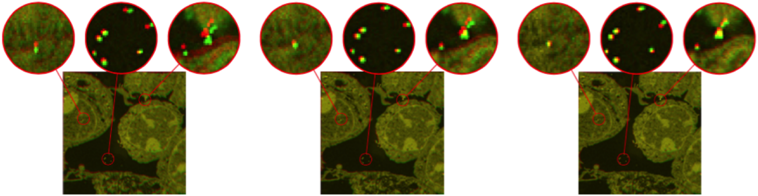
Standard IA



Proposed IA

Method	mean	std	p-value
Proposed IA	$\mathbf{7.43 \times 10^4}$	4.72×10^3	**0.0015**
Standard IA	8.62×10^4	6.37×10^3	

Registration Results



(a) Direct registration (b) Standard IA (c) Proposed IA

Method	mean	std
Proposed IA	**83.30**	54.31
Standard IA	140.40	135.07
Direct registration	176.20	116.29

unit is nm, pixel size is 40nm

Tian Cao, Marc Niethammer and etc. Multi-modal registration for correlative microscopy using image analogies. Medical image analysis, 2013

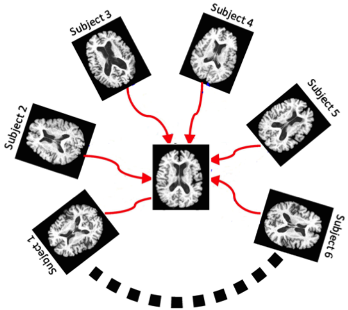
Outline

- Background
- CDL for Multi-modal Image registration
- CDL for Deformation Estimation
- CDL for GTPase Activations and Cell Protrusions
- Robust CDL
- Summary and Future Work

Deformation Estimation

Challenges

- difficult to model the relationship between deformations and appearances
- whether a basis for deformations and appearances exist



Previous work

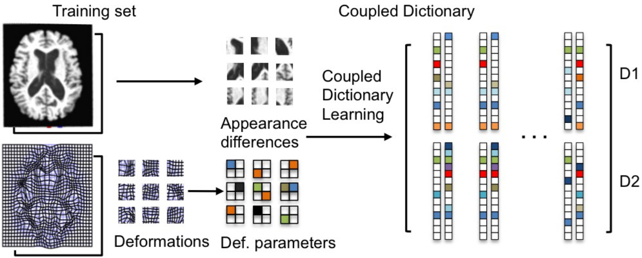
- Global linear regression [C.Chou2010]
- Estimate deformations of a set of key points [Q.Wang2012]

Proposed Method

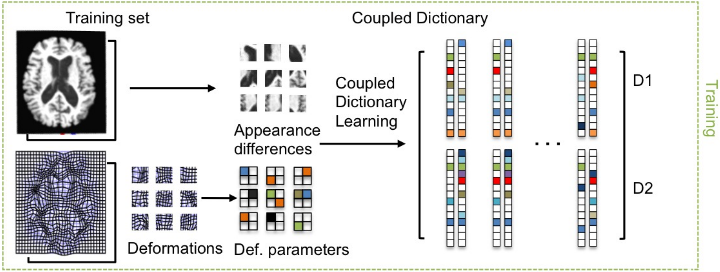
A general framework for deformation estimation based on

- semi-coupled dictionary learning on deformation parameters and image appearances
- investigation of different deformation parametrization methods
  - b-spline
  - initial momentum

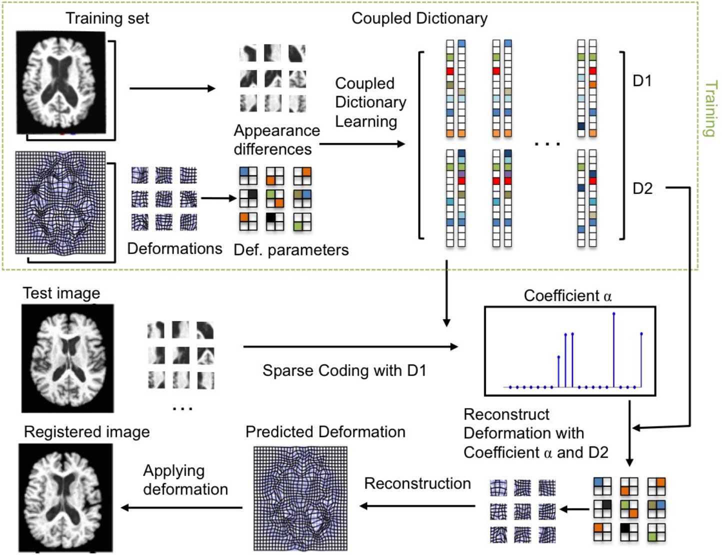
Framework of Proposed Method



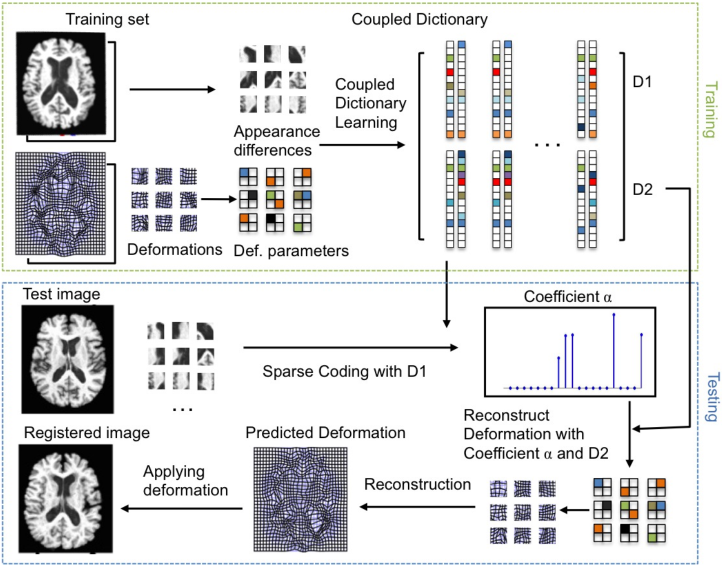
Framework of Proposed Method



Framework of Proposed Method



Framework of Proposed Method



Semi-coupled Dictionary Learning

- Dictionary learning for D^1

$$\underset{D^1, D^2, \alpha_i^1, \alpha_i^2, W}{\text{minimize}} \sum_{i=1}^{N} \boxed{\frac{1}{2}\|x_i^1 - D^1\alpha_i^1\|_2^2 + \lambda_1\|\alpha_i^1\|_1} +$$

$$\boxed{\frac{1}{2}\|x_i^2 - D^2\alpha_i^2\|_2^2 + \lambda_2\|\alpha_i^2\|_1} + \boxed{\gamma_1\|\alpha_i^2 - W\alpha_i^1\|_2^2 + \gamma_2\|W\|_F^2}$$

Semi-coupled Dictionary Learning

- Dictionary learning for D^1

$$\underset{D^1, D^2, \alpha_i^1, \alpha_i^2, W}{\text{minimize}} \sum_{i=1}^{N} \frac{1}{2}\|x_i^1 - D^1\alpha_i^1\|_2^2 + \lambda_1\|\alpha_i^1\|_1 +$$

$$\frac{1}{2}\|x_i^2 - D^2\alpha_i^2\|_2^2 + \lambda_2\|\alpha_i^2\|_1 + \gamma_1\|\alpha_i^2 - W\alpha_i^1\|_2^2 + \gamma_2\|W\|_F^2$$

- Dictionary learning for D^2

Semi-coupled Dictionary Learning

- Dictionary learning for D^1

$$\underset{D^1,D^2,\alpha_i^1,\alpha_i^2,W}{\text{minimize}} \sum_{i=1}^{N} \frac{1}{2}\|x_i^1 - D^1\alpha_i^1\|_2^2 + \lambda_1\|\alpha_i^1\|_1 +$$

$$\frac{1}{2}\|x_i^2 - D^2\alpha_i^2\|_2^2 + \lambda_2\|\alpha_i^2\|_1 + \gamma_1\|\alpha_i^2 - W\alpha_i^1\|_2^2 + \gamma_2\|W\|_F^2$$

- Dictionary learning for D^2
- Linear mapping between α^1 and α^2

Semi-coupled Dictionary Learning

- Dictionary learning for D^1

$$\underset{D^1, D^2, \alpha_i^1, \alpha_i^2, W}{\text{minimize}} \sum_{i=1}^{N} \frac{1}{2}\|x_i^1 - D^1\alpha_i^1\|_2^2 + \lambda_1\|\alpha_i^1\|_1 +$$

$$\frac{1}{2}\|x_i^2 - D^2\alpha_i^2\|_2^2 + \lambda_2\|\alpha_i^2\|_1 + \gamma_1\|\alpha_i^2 - W\alpha_i^1\|_2^2 + \gamma_2\|W\|_F^2$$

- Dictionary learning for D^2
- Linear mapping between α^1 and α^2

- Atoms satisfy $\|d_i^1\|_2 \leq 1$ and $\|d_i^2\|_2 \leq 1$

Semi-coupled Dictionary Learning

- Dictionary learning for D^1

$$\underset{D^1,D^2,\alpha_i^1,\alpha_i^2,W}{\text{minimize}} \sum_{i=1}^{N} \frac{1}{2}\|x_i^1 - D^1\alpha_i^1\|_2^2 + \lambda_1\|\alpha_i^1\|_1 +$$

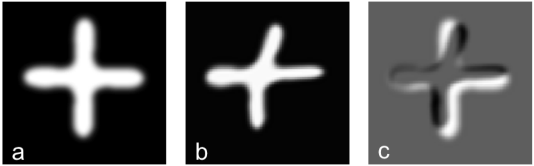
$$\frac{1}{2}\|x_i^2 - D^2\alpha_i^2\|_2^2 + \lambda_2\|\alpha_i^2\|_1 + \gamma_1\|\alpha_i^2 - W\alpha_i^1\|_2^2 + \gamma_2\|W\|_F^2$$

- Dictionary learning for D^2
- Linear mapping between α^1 and α^2

- Atoms satisfy $\|d_i^1\|_2 \leq 1$ and $\|d_i^2\|_2 \leq 1$
- Use a linear mapping W to relating α^1 and α^2 to relax the strong coupling

Results

- synthesize cross image as atlas
- randomly generate 3×3 b-spline coefficients to transform atlas as training subject images
- test on 100 images with random b-spline transformations



Results

Method	median	mean	min	max	std
RAW	17.9776	19.2912	6.4132	45.1841	9.1304
NN	6.4063	7.1153	1.1399	12.2234	4.0948
GR	4.9261	5.2245	2.6279	10.2138	3.7925
CDL	3.6325	3.7531	0.724	8.5439	2.1158
SCDL	**3.0679**	**2.9254**	**0.3532**	**5.0559**	**1.4815**

Unit: pixel
RAW: direct registration
NN: nearest neighbor search
GR: global regression
CDL: coupled dictionary learning
SCDL: semi-coupled dictionary learning

mean absolute errors (MAE) in pixels

Tian Cao, Nikhil Singh, Vladimir Jojic, and Marc Niethammer. Semi-coupled dictionary learning for deformation prediction. Accepted by ISBI 2015.

Outline

- Background
- CDL for Multi-modal Image registration
- CDL for Deformation Estimation
- CDL for GTPase Activations and Cell Protrusions
- Robust CDL
- Summary and Future Work

GTPase Activations and Cell Protrusions Modeling

Challenges

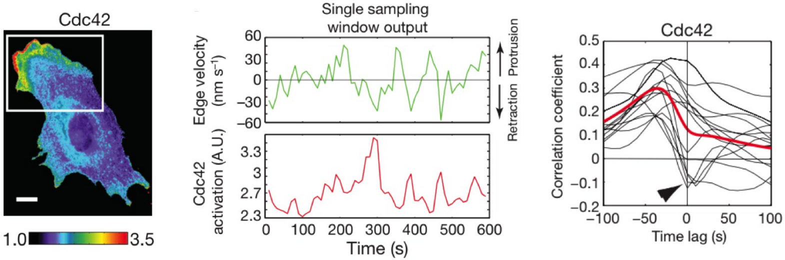
- how to represent the data
- how to model the global and local relationships

Previous work

- compute correlation between activations and protrusions on boundary points [M. Machacek 2009]

Previous Work

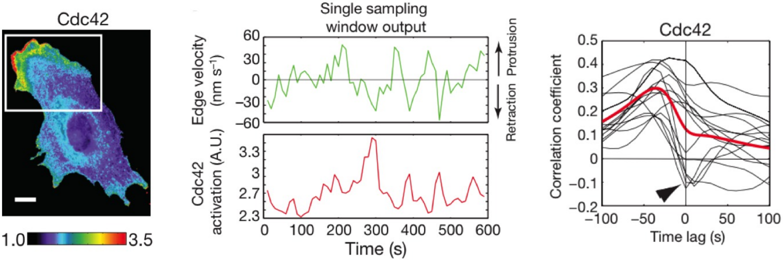
- Track boundary points in specific regions
- Record intensity values for small windows centered at boundary points
- Compute cross-correlation between GTPase activations and cell movements on boundary points



[Matthias Machacek and etc, Nature letter, 2009]

Limitations of Previous Work

- regions of interest are hand-picked
- missing spatial interaction between neighboring boundary points



[Matthias Machacek and etc, Nature letter, 2009]

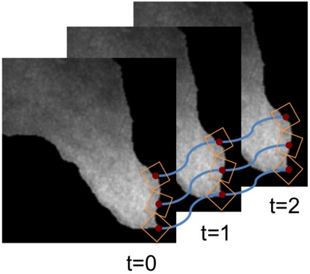
Proposed Method

A framework to model the GTPase activations and cell movements based on

- coupled dictionary learning for spatiotemporal data
  - GTPase activations and cell movements on boundary points
- represent data with sparse coefficients
- clustering boundary points based on sparse coefficients

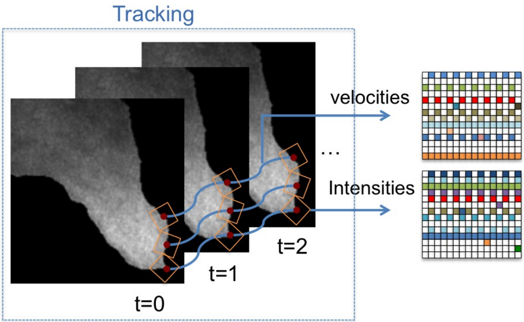
Proposed Method

- Learn coupled dictionary on protrusion and activation for single boundary points



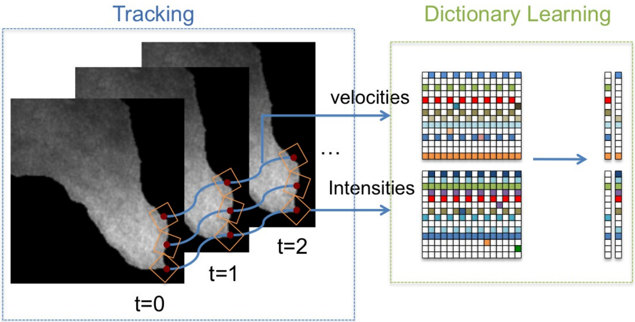
Proposed Method

- Learn coupled dictionary on protrusion and activation for single boundary points



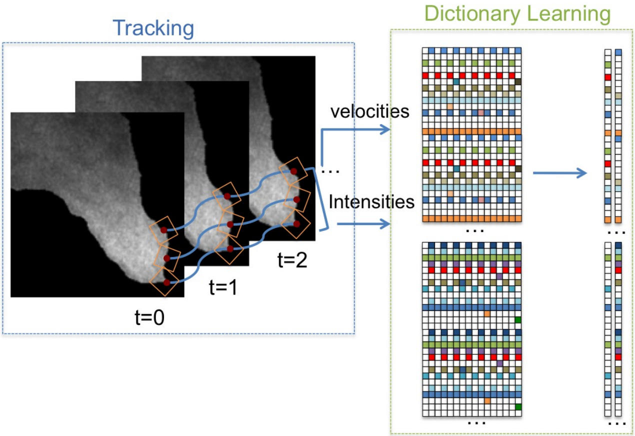
Proposed Method

- Learn coupled dictionary on protrusion and activation for single boundary points



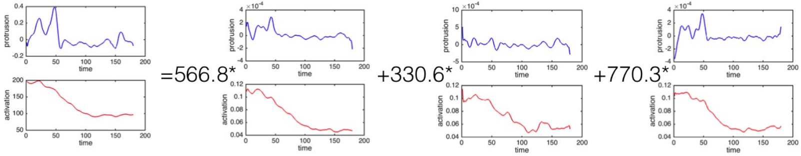
Proposed Method

- Learn coupled dictionary on protrusion and activation for single boundary points
- Learn coupled dictionary on protrusion and activation for boundary point "patches"



Proposed Method

- Use coupled dictioanry to reconstruct the data
- Represent data with compact coefficients based on learned dictionary

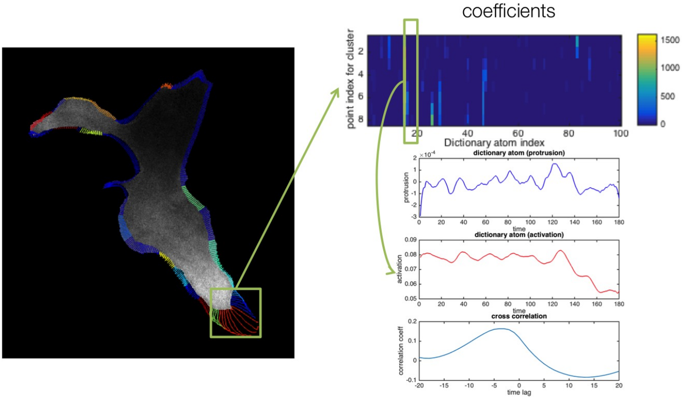


Coefficients = [0, 0, …, 566.8, 0, 0,…, 330.6, 0, 0, …, 770.3]

Preliminary Results

To study spatial interactions between neighboring points or "patches"

- Represent data with compact coefficients based on learned dictionary
- k-means clustering on coefficients of points or "patches"



Hierarchical clustering

Limitations of K-means clustering

- need assign number of clusters in k-means clustering
- find optimal k is difficult

Hierarchical clustering

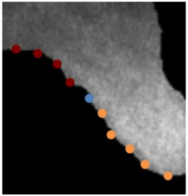
- arrange the data in a tree structure
- provide hierarchical relationship between different data points

Discussion

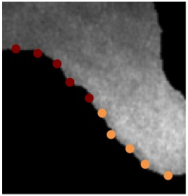
Why do clustering on coefficients instead of raw data?

- coefficients can provide information about the patterns used to represent the data

Why do clustering on "point patches"?



clustering on single points



clustering on point "patches"

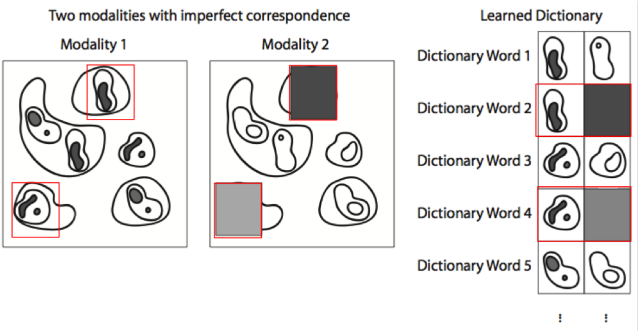
Outline

- Background
- CDL for Multi-modal Image registration
- CDL for Deformation Estimation
- CDL for GTPase Activations and Cell Protrusions
- Robust CDL
- Summary and Future Work

Robust Coupled Dictionary Learning

Challenges for Coupled dictionary learning

- may fail without sufficient correspondences between training data in two spaces



Our Method

A probabilistic model for coupled dictionary learning

- confidence measure
  - a conditional probability $p(h|x_i)$
  - $h = 0$ indicates x_i reconstructed by "clean" dictionary atoms
  - $h = 1$ indicated x_i reconstructed by "noise" dictionary atoms
- EM algorithm

Confidence Measure

- Applying Bayes' rules, $p(h = 1|x_i)$ can be represented as

$$P(h = 1|x_i) = \frac{P(x_i|h = 1)P(h = 1)}{P(x_i|h = 1)P(h = 1) + P(x_i|h = 0)P(h = 0)}$$

- Assuming independence of x_i and the Gaussian distribution of pixels in x_i

$$p(x_i|h = 1, \theta_1) = \mathcal{N}(x_i; \mu_1, \sigma_1^2),$$
$$p(x_i|h = 0, \theta_0; D, \alpha_i) = \mathcal{N}(x_i - D\alpha_i; 0, \sigma_0^2).$$

EM Algorithm

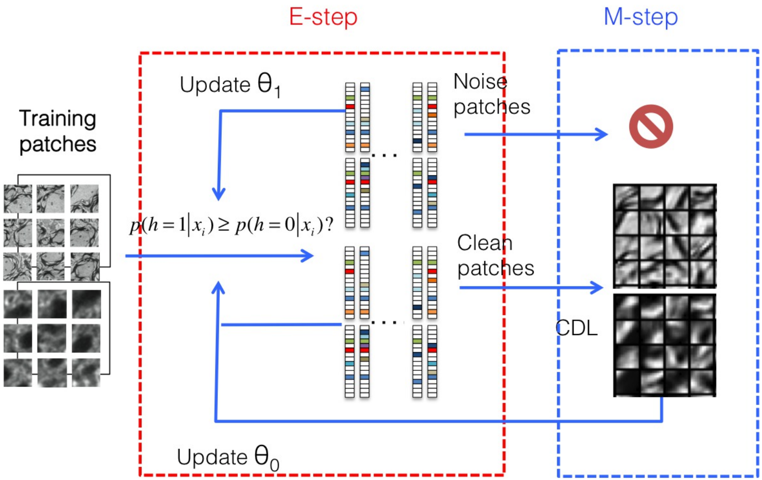
The parameters to estimate are

- $\theta_1 = \{\mu_i, \sigma_1\}$
- $\theta_o = \sigma_0$

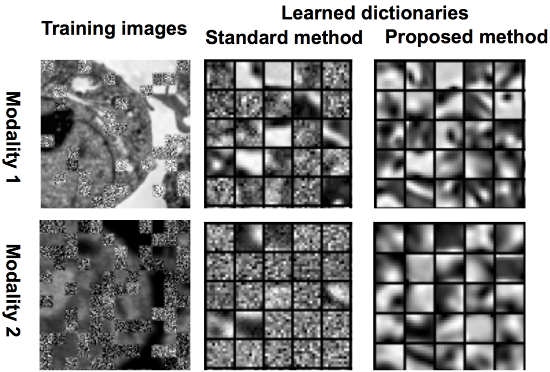
Maximum likelihood (ML) estimation

- classify non-corresponding data and update coupled dictionary iteratively

Framework of Proposed Method



Results



Tian Cao, Vladimir Jojic, Marc Niethammer and etc. Robust multi-modal dictionary learning. MICCAI, 2013.

Outline

- Background
- CDL for Multi-modal Image registration
- CDL for Deformation Estimation
- CDL for GTPase Activations and Cell Protrusions
- Robust CDL
- Summary and Future Work

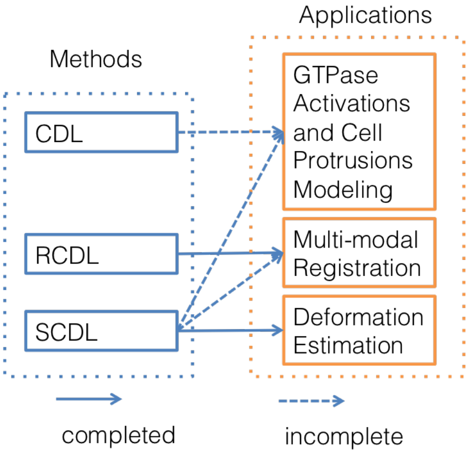
Summary

Coupled dictionary learning method to relate data from two spaces:

- spatial data: appearances for different modalities
- spatial data: appearances and deformations
- spatiotemporal data: GTPase activations and cell movements

A probabilistic model for robust coupled dictionary learning to account for data non-corresponding training data

Summary and Future Work



Research Status and Tentative Timeline

CDL for multi-modal registration	Completed([1, 4])
CDL for deformation prediction	Completed([3])
CDL for GTPase activation and cell protrusion modeling	May 2015
Robust CDL	Completed([2])
Future Work	June 2015
Proposal	January 2015
Oral Exam	February 2015
Defense	August 2015

Table: Research plan

List of publications

- T. Cao, C. Zach, S. Modla, D. Powell, K. Czymmek, and M. Niethammer. Registration for correlative microscopy using image analogies. *Biomedical Image Registration*, pages 296–306, 2012.
- Tian Cao, Vladimir Jojic, Shannon Modla, Debbie Powell, Kirk Czymmek, and Marc Niethammer. Robust multimodal dictionary learning. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 259–266. Springer Berlin Heidelberg, 2013.
- Tian Cao, Nikhil Singh, Vladimir Jojic, and Marc Niethammer. Semi-coupled dictionary learning for deformation prediction. In *Accepted by ISBI 2015*.
- Tian Cao, Christopher Zach, Shannon Modla, Debbie Powell, Kirk Czymmek, and Marc Niethammer. Multi-modal registration for correlative microscopy using image analogies. *Medical image analysis*, 2013.

Thesis

- Learning a coupled basis for the compact representation of two spaces can be achieved by **coupled dictionary learning** (CDL).
- Such dictionaries can be learned to capture appearance differences of different imaging modalities, dependencies between image appearance and deformation as well as the spatiotemporal patterns for cell signaling and boundary protrusions and retractions.
- To account for data inconsistencies, a robust coupled dictionary can be obtained based on a probabilistic dictionary model.