Sparsity Based Spectral Embedding: Application to Multi-Atlas Echocardiography Segmentation

Ozan Oktay, Wenzhe Shi, Jose Caballero,

Kevin Keraudren, and Daniel Rueckert

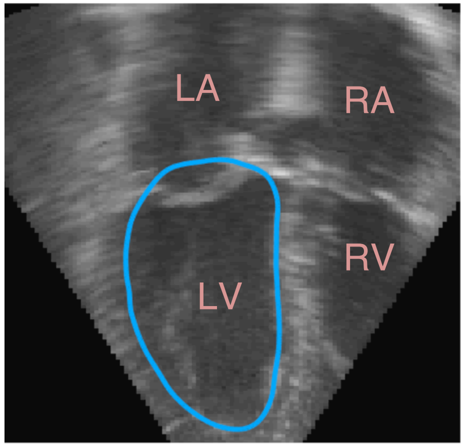
Department of Computing
Imperial College London

Challenge on Endocardial Three-dimensional
Ultrasound Segmentation, MICCAI 2014
14th September 2014

BioMedIA
Imperial College London

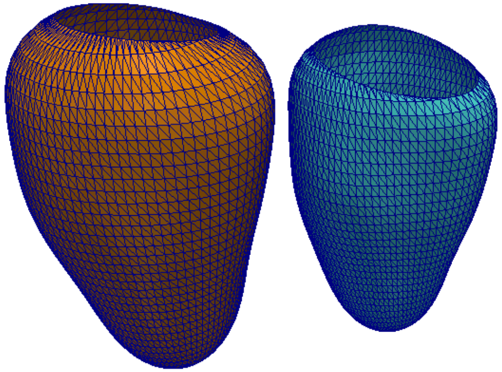
Imperial College
London

Problem Definition and Literature Review



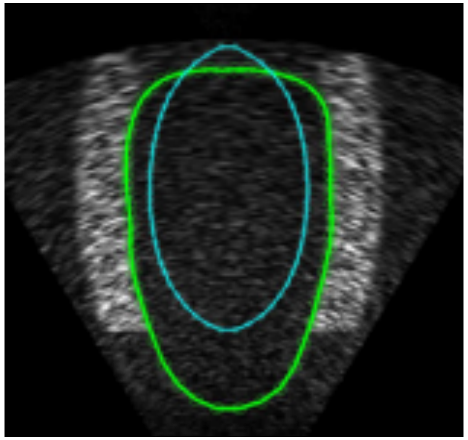
Problem:

Left ventricle (LV) endocardium segmentation in 3D Echo Images



Motivation:

Estimation of clinical indices: (1) ejection fraction, (2) stroke volume, and (3) cardiac motion

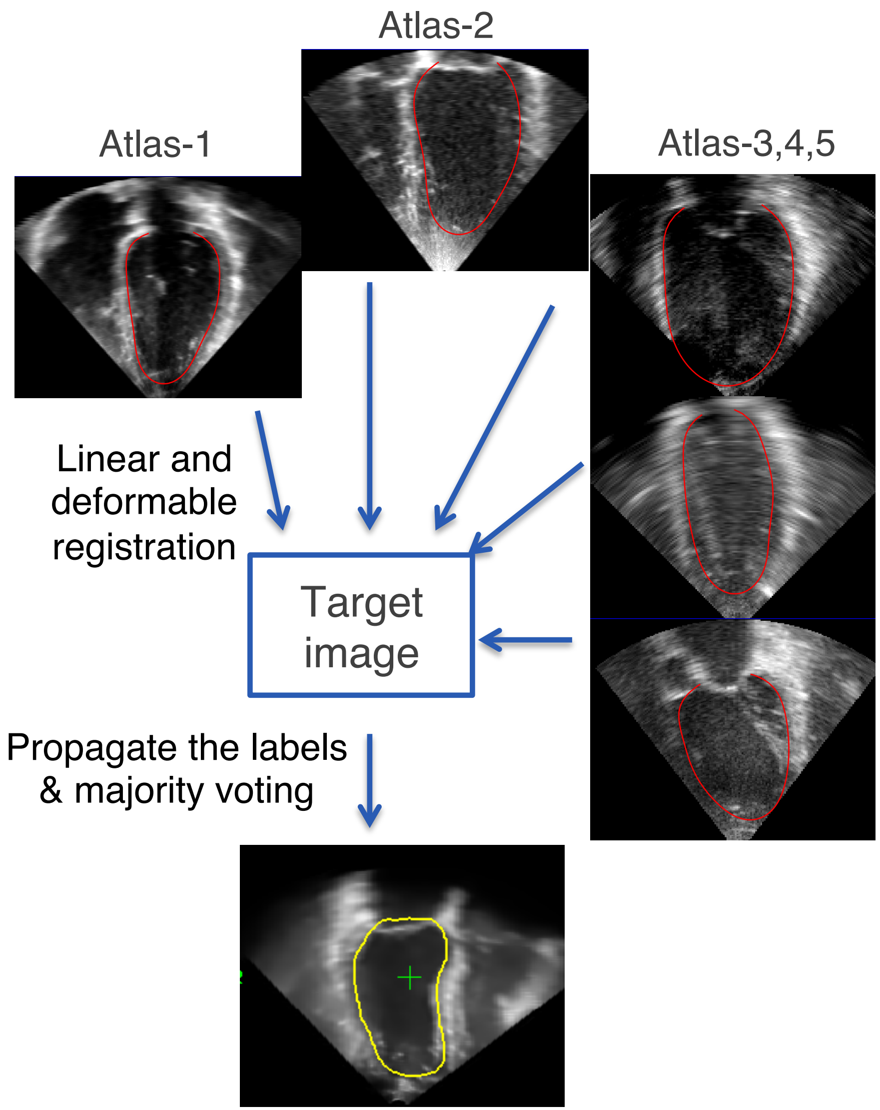


The existing work in the literature:

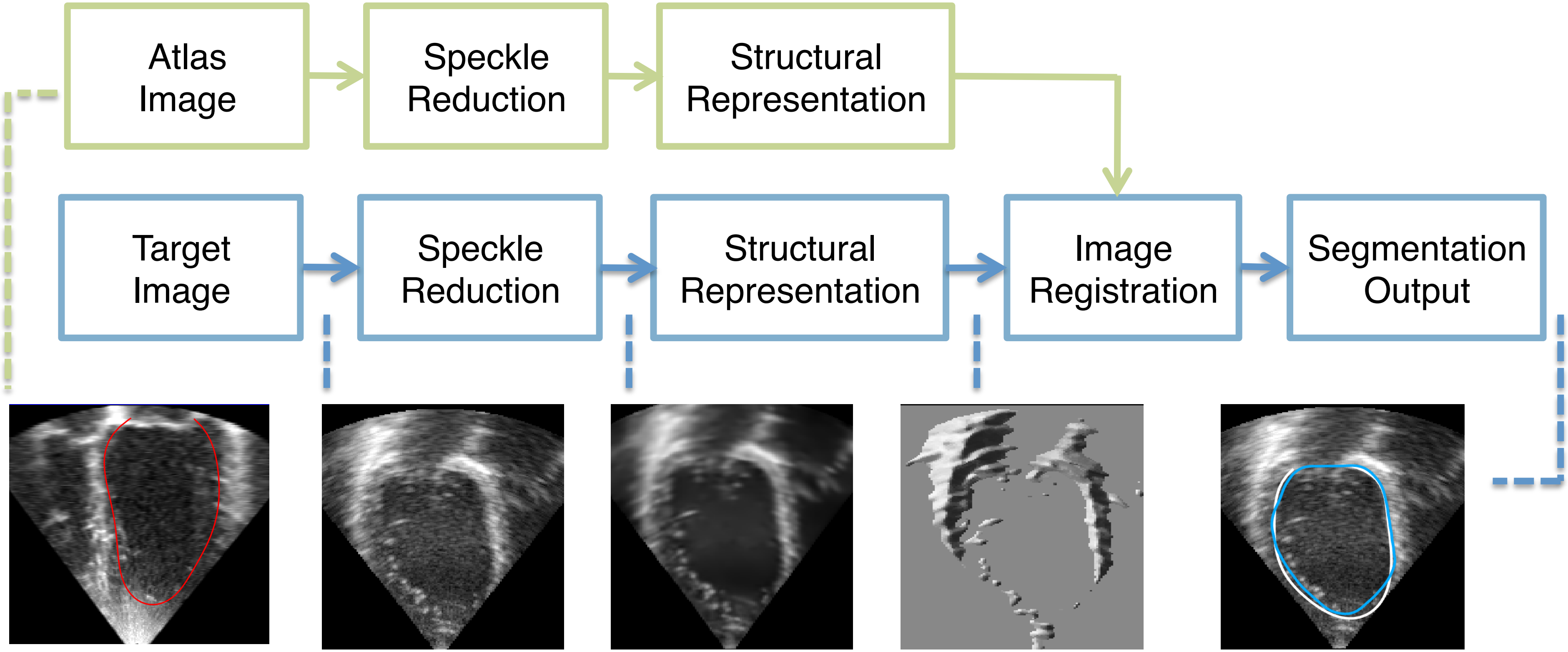
1. B-spline based active surfaces [Barbosa et al., 2013]
2. Statistical-shape models [Butakoff et al., 2011]
3. Edge based level-set segmentation [Rajpoot et al., 2011]

Multi-Atlas Image Segmentation

- It uses the manually labeled atlases to segment the target organ.
- It does not require any training or prior estimation.
- Successfully applied in
  i. Brain MRI Segmentation
     [Aljabar et al. 2009 NeuroImage]
  ii. Cardiac MRI Segmentation
     [Isgum et al. 2009 TMI]

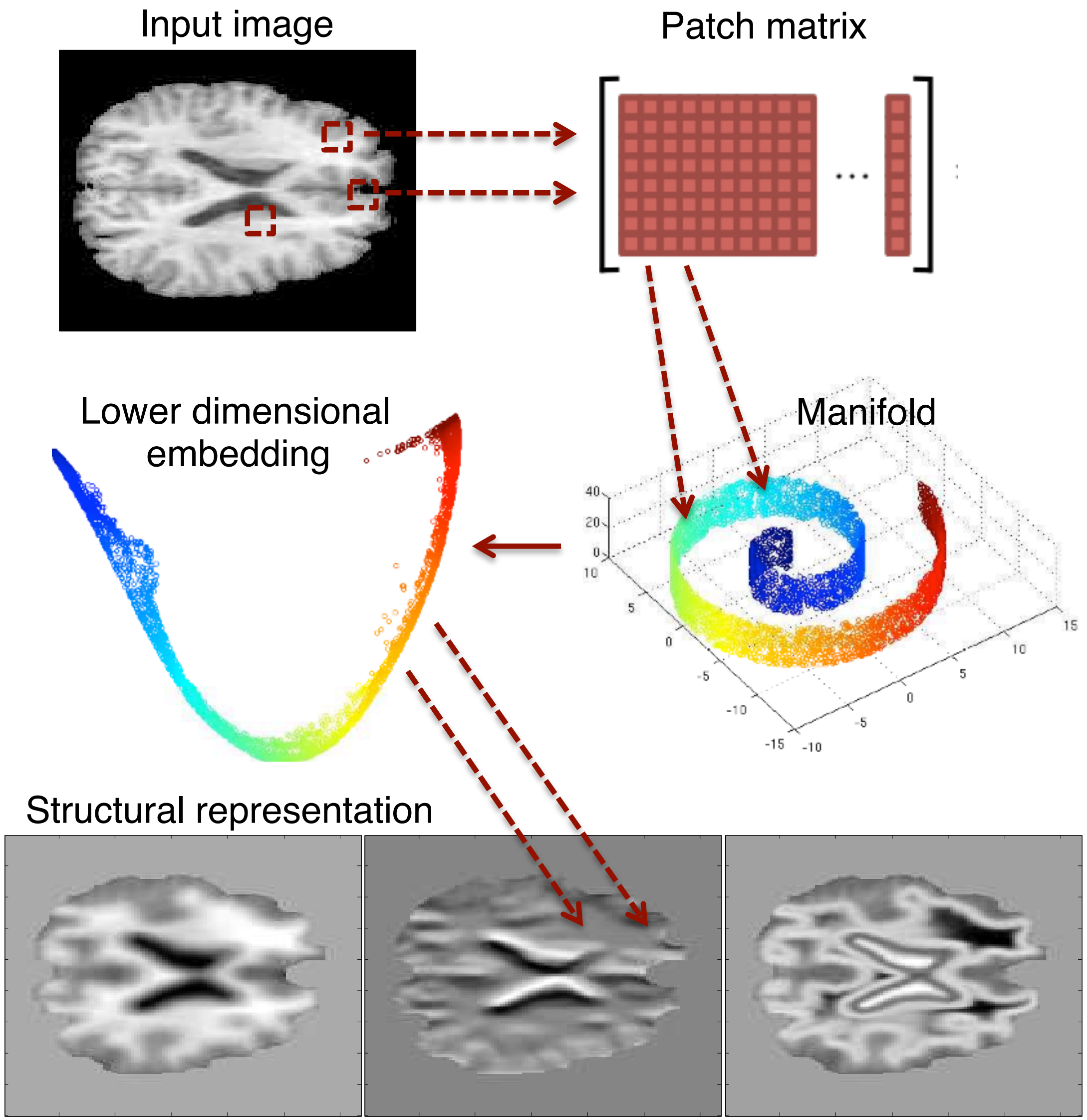


Proposed Segmentation Framework



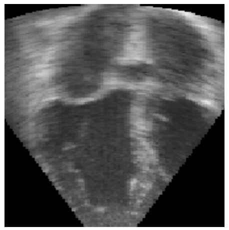
Patch Based Spectral Representation

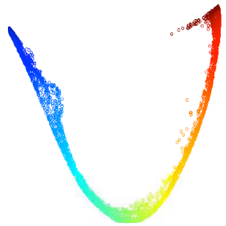
- Structural image representation
- Unsupervised learning of shape and contextual information.
- Laplacian Eigenmaps.
- Useful for echo images since intensity data does not explicitly reveal the structural information.



1. Wachinger C. and Navab N.: "Entropy and Laplacian images: Structural representations for multi-modal registration"
Medical Image Analysis 2012.

Dictionary Based Spectral Representation



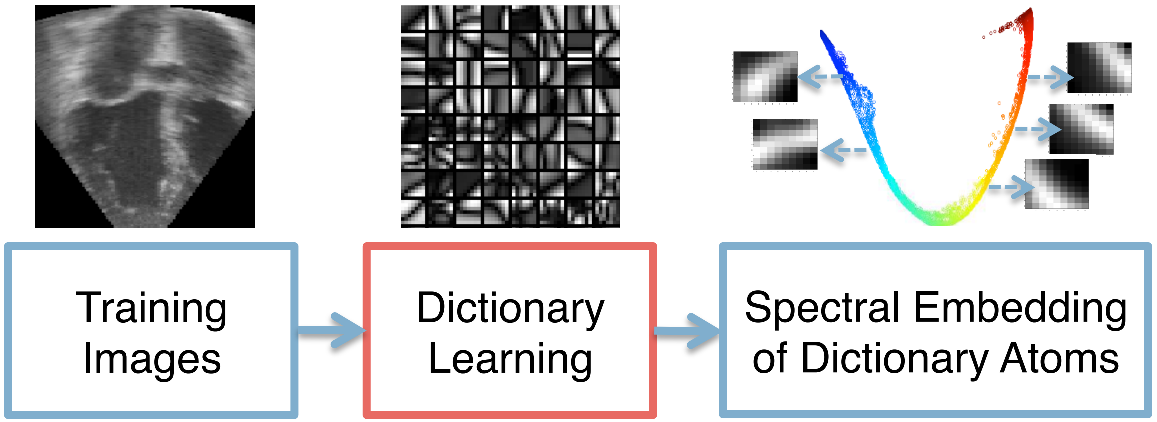


Target image

Computationally prohibitive

Spectral embedding of image patches

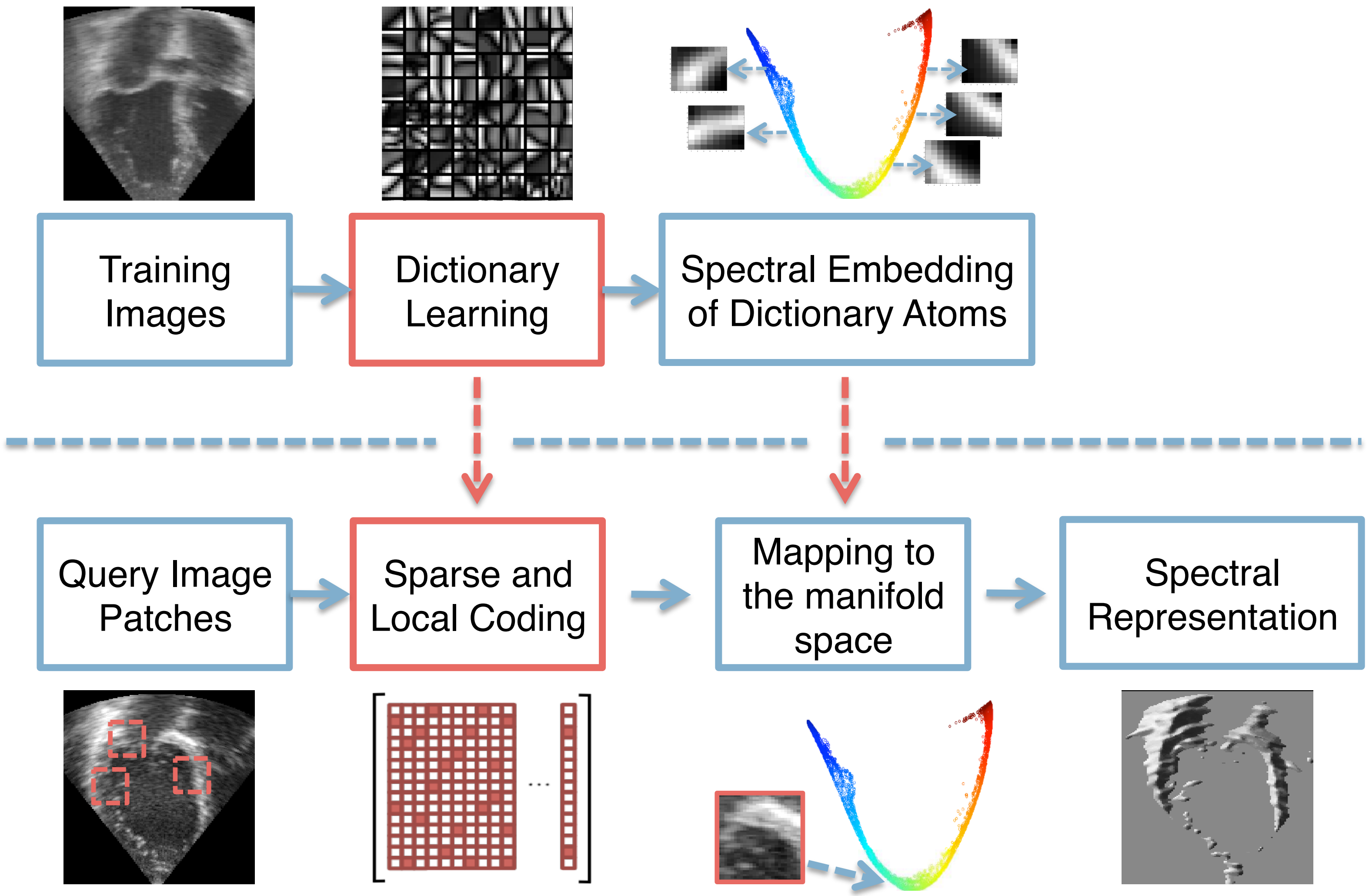
Dictionary Based Spectral Representation



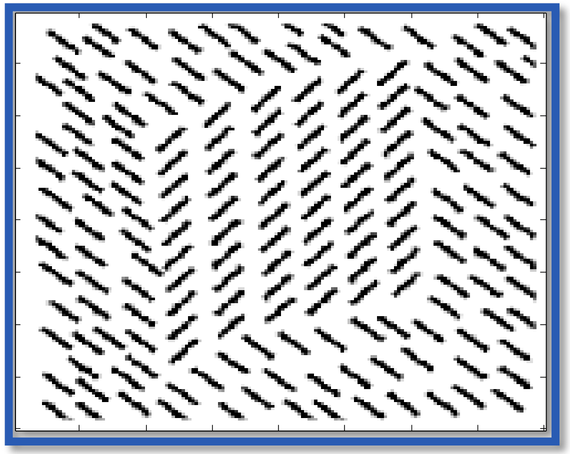
What is the main motivation for the sparsity and dictionary learning ?

- Computationally efficient due to elimination of redundancy
- Large number of images can be mapped to the same embedding space

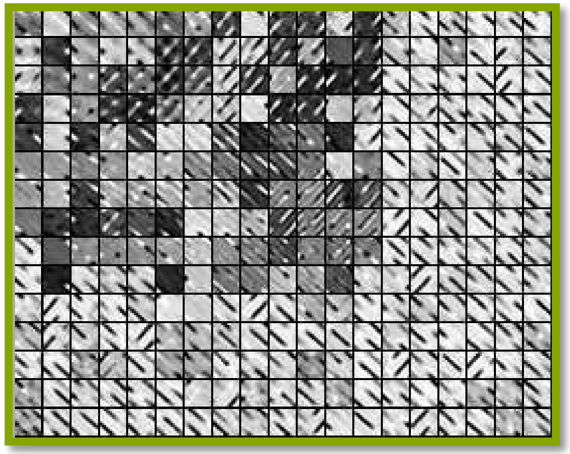
Dictionary Based Spectral Representation



Dictionary Based Spectral Representation



Input Image

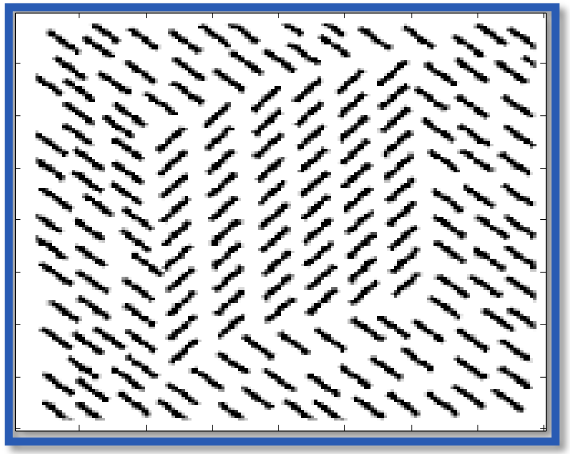


Dictionary Learning

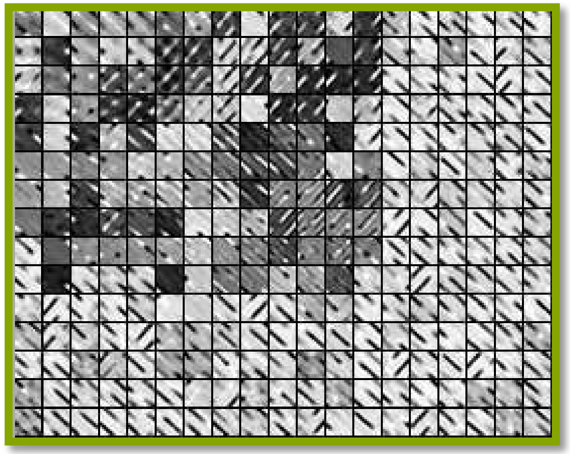
$$\min_{\mathbf{C},\mathbf{X}} \|\mathbf{Y} - \mathbf{CX}\|^2$$

$$\text{s.t.} \quad \forall i \ , \|x_i\|_0 \leq T_0.$$

Dictionary Based Spectral Representation



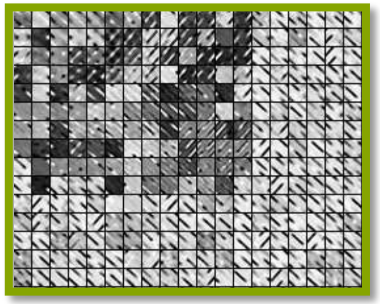
Input Image



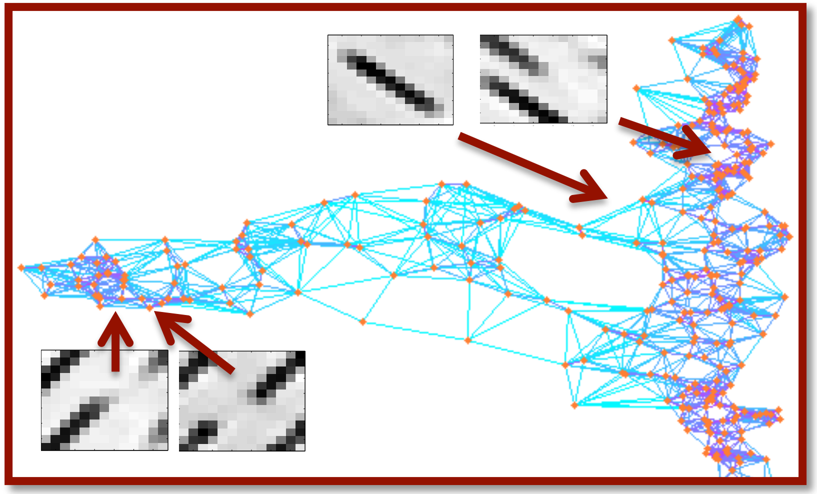
Dictionary Learning

$$\min_{\mathbf{C},\mathbf{X}} \|\mathbf{Y} - \mathbf{CX}\|^2$$

$$\text{s.t.} \quad \forall i\ ,\ \|x_i\|_0 \leq T_0.$$



Dictionary Atoms



Spectral Embedding

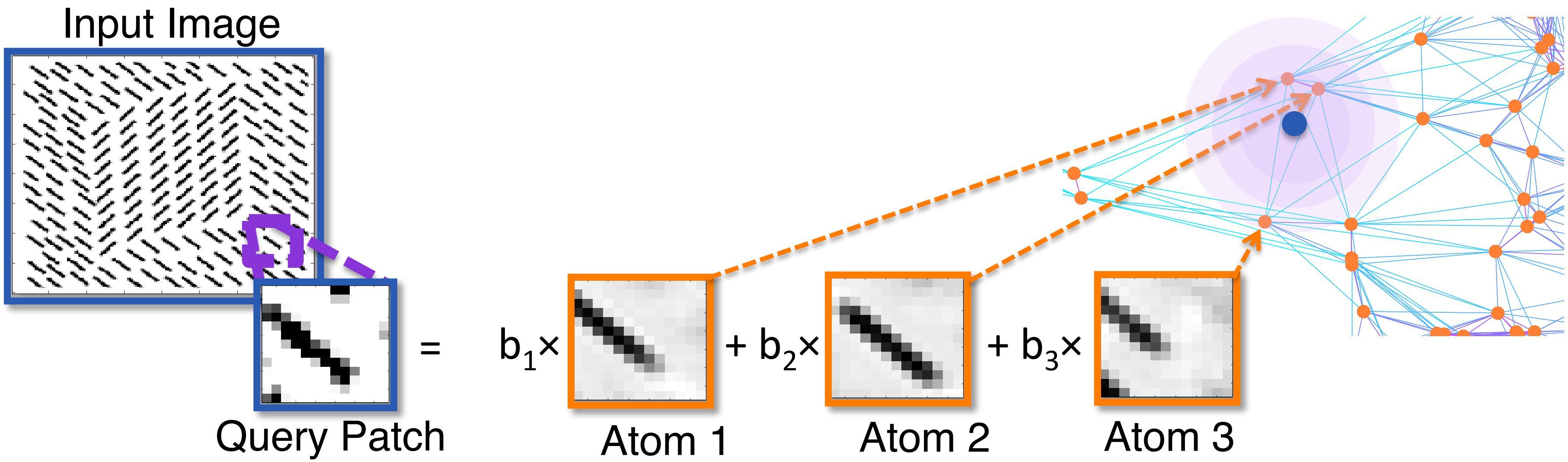
Laplacian Graph

$$\mathcal{L} = \mathbf{D} - \mathbf{W}$$

Eigen decomposition

$$\mathcal{L} = \mathbf{V}^\top \Lambda\, \mathbf{V}$$

Sparse and Local Coding for Spectral Representation



Locality constrained linear coding

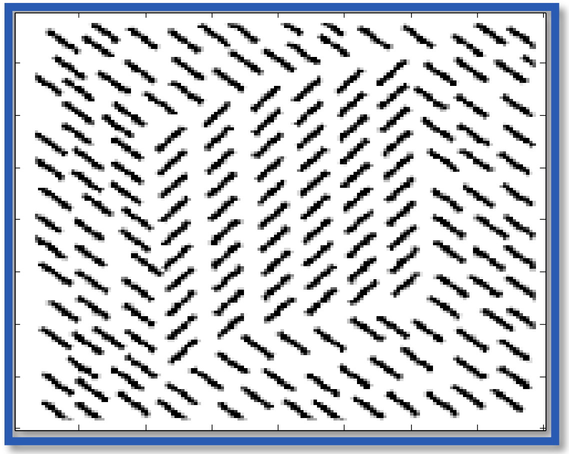
$$\min_{\mathbf{X}} \sum_{n=1}^{N} \|\mathbf{y}_n - \mathbf{C}\tilde{\mathbf{x}}_n\|^2 + \lambda \|\mathbf{b}_n \odot \tilde{\mathbf{x}}_n\|^2 \text{ s.t. } \forall n \,, \mathbf{1}^\top \tilde{\mathbf{x}}_n = 1$$

$$b_{(n,m)} = \exp\left(\|\mathbf{y}_n - \mathbf{c}_m\|^2 / \sigma\right)$$

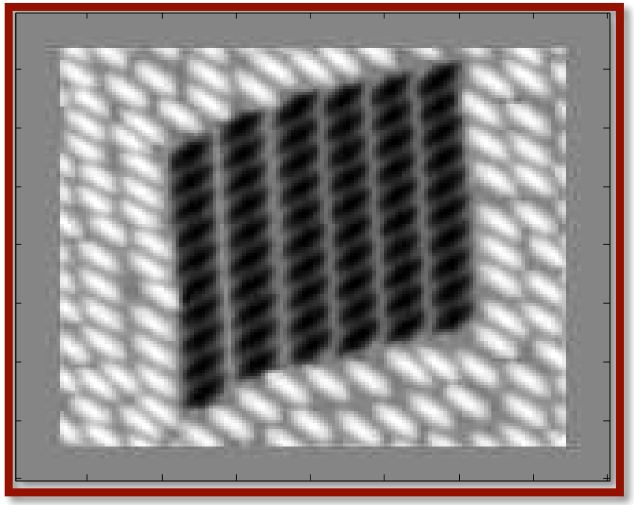
$$\mathbf{b}_n = [b_{(n,1)}, \ldots, b_{(n,M)}]$$

1. K. Yu et al. : "Nonlinear learning using local coordinate coding" NIPS 2009
2. J. Wang et al. : "Locality-constrained linear coding for image classification" CVPR 2010

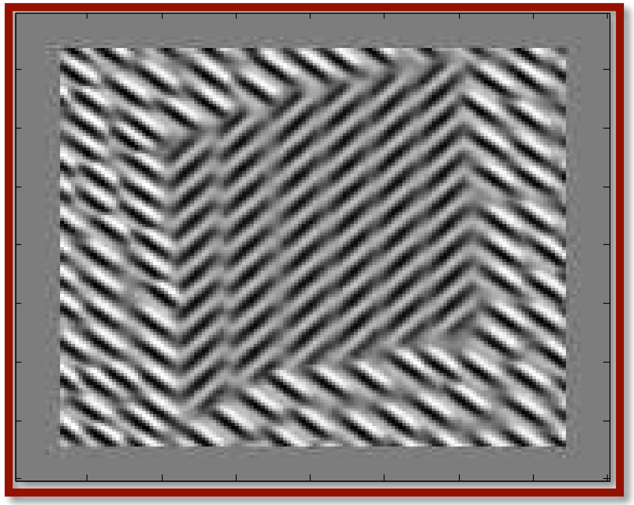
Dictionary Based Spectral Representation



Input Image

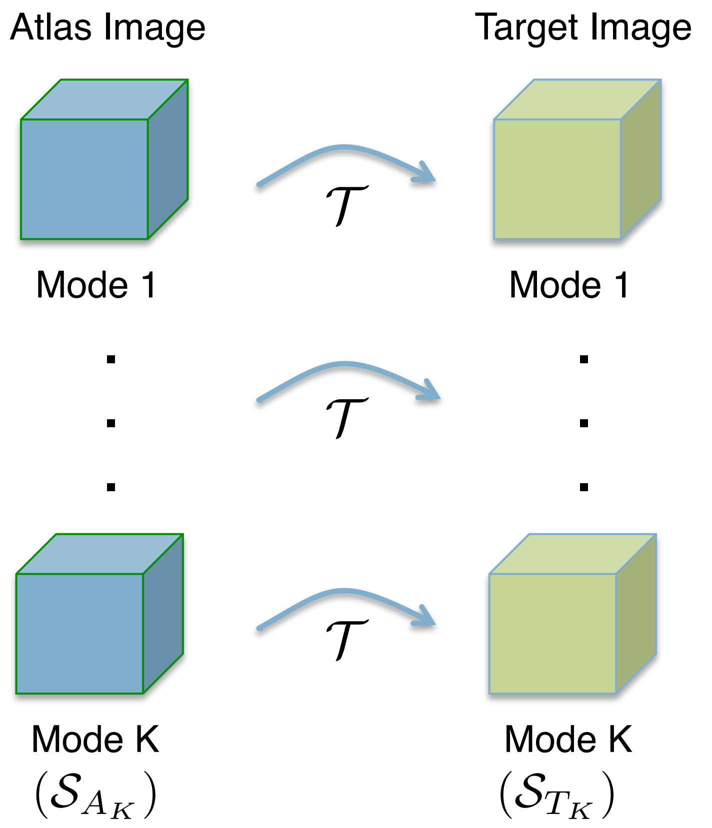


First Spectral Mode



Second Spectral Mode

Registration Strategy



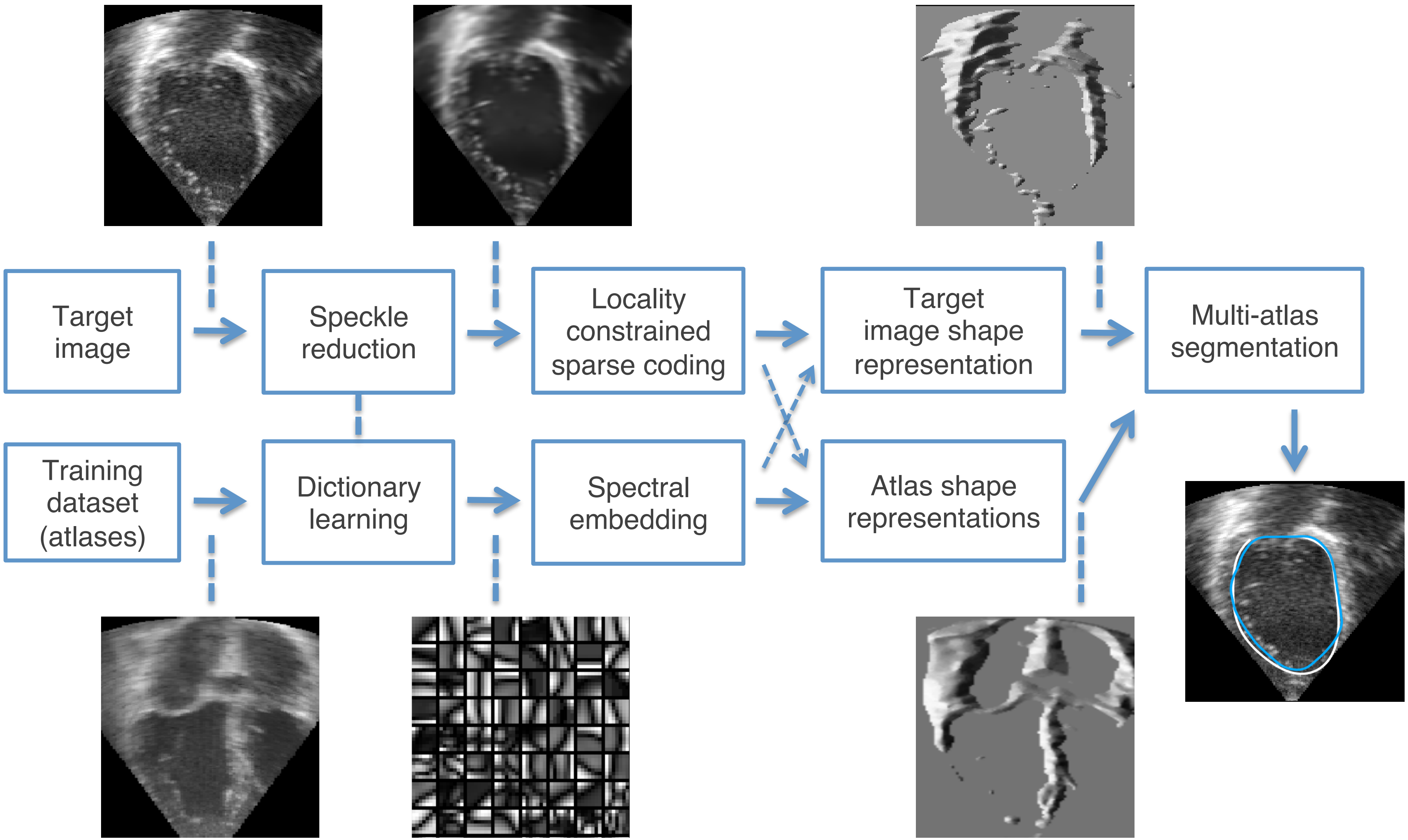
Single deformation field $(\mathcal{T})$
for all 3D-3D spectral image pairs

$$\sum_{k=1}^{K} \|\mathcal{S}_{A_k}(\,\mathcal{T}(\mathbf{p})\,) - \mathcal{S}_{T_k}(\mathbf{p})\|^2 + \beta\,\mathcal{R}(\mathbf{p})$$

$$\mathbf{p} \in \mathbb{R}^3 \;,\; \mathcal{T} : \mathbb{R}^3 \mapsto \mathbb{R}^3$$

- B-spline free-form deformations. [Rueckert et al. TMI 99]
- Sum of squared differences similarity measure.
- Number of modes K = 4.

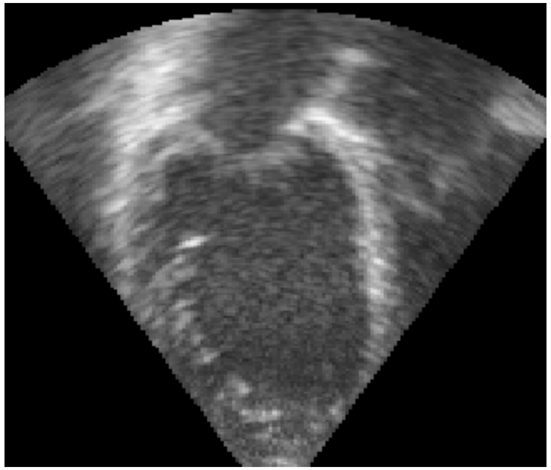
The Proposed Segmentation Framework



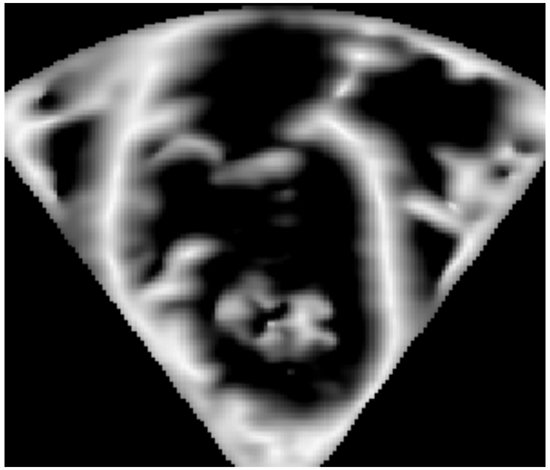
Validation Dataset

1. Training Dataset (Atlases) (15 Patients)
   - 3D+T echo scans.
   - Cross-validation is performed on the training dataset.
   - Ground-truth segmentations are available for ED and ES frames.
2. Testing Dataset (15 Patients)
   - 3D+T echo scans, obtained from different view angles.
   - Only the ED and ES frames are segmented

Other Image Representations



Original echo image



Local phase image [2]

Boundary Image [1]

Spectral Representation

Intensity and phase features

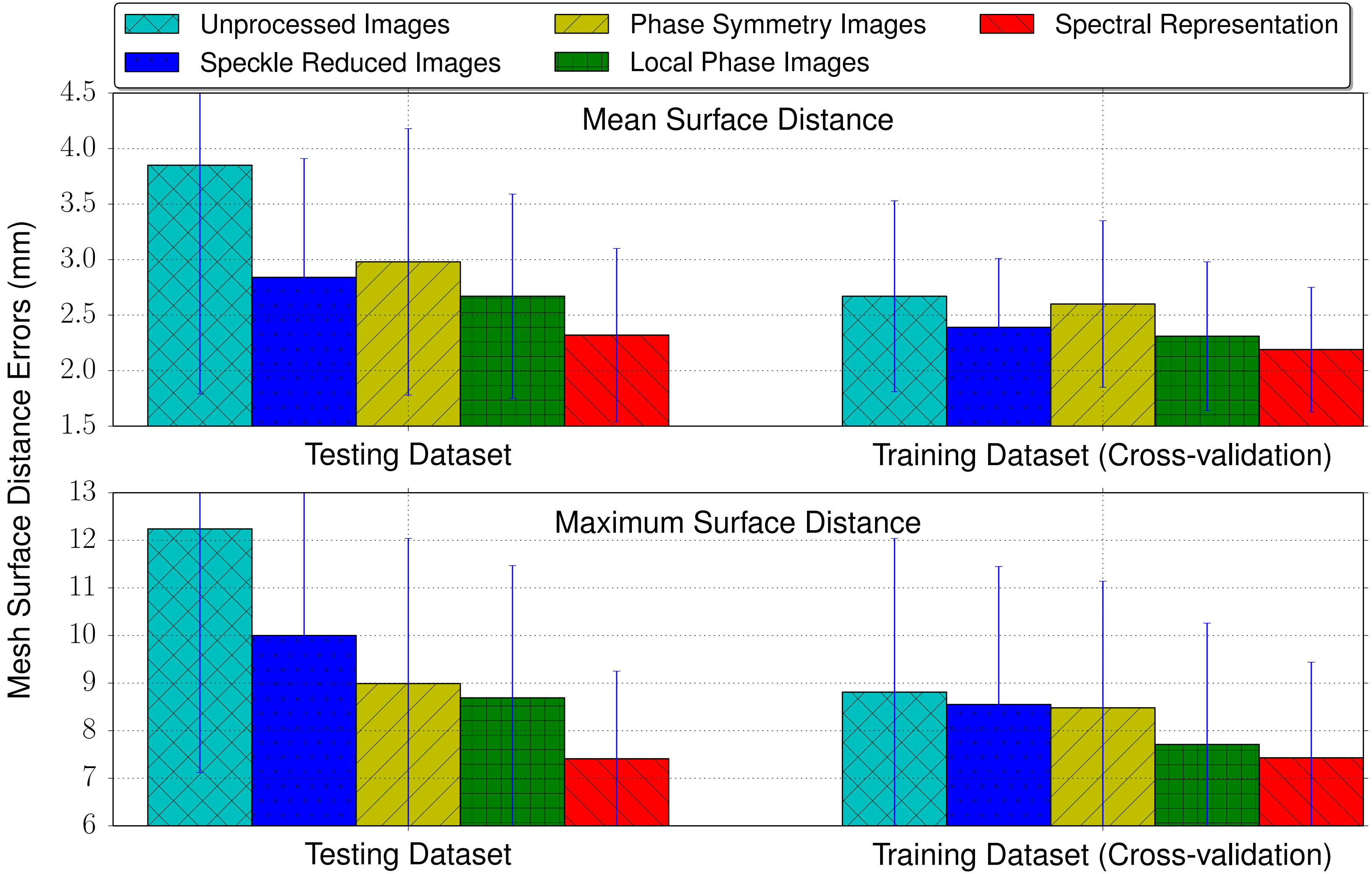
- Encodes only the tissue boundary information.
- It is not sufficient for image analysis applications

Spectral representation

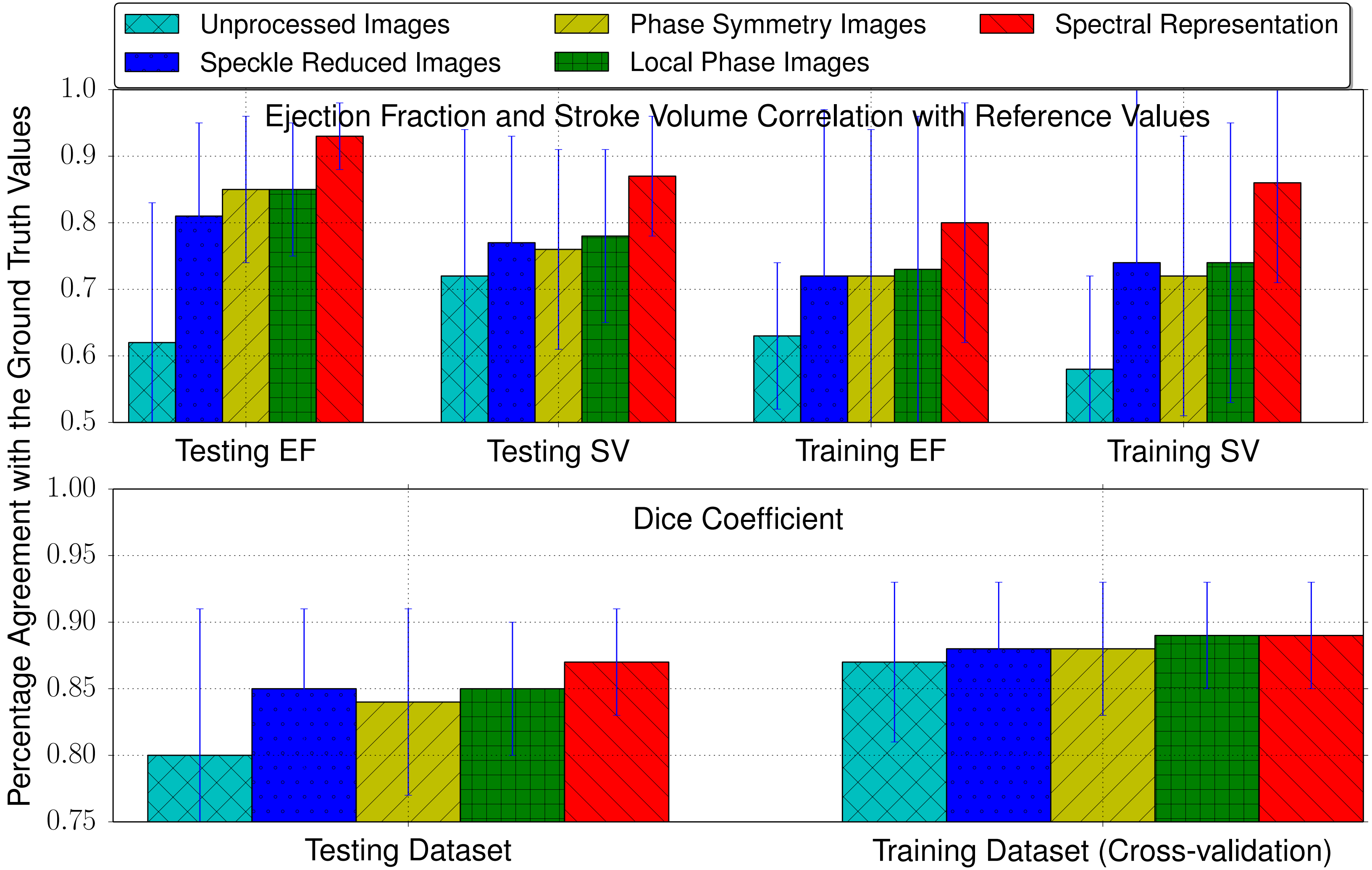
- Encodes the contextual information

1. Rajpoot, K. et al.: ISBI 2009
2. Zhuang, X. et al.: ISBI 2010

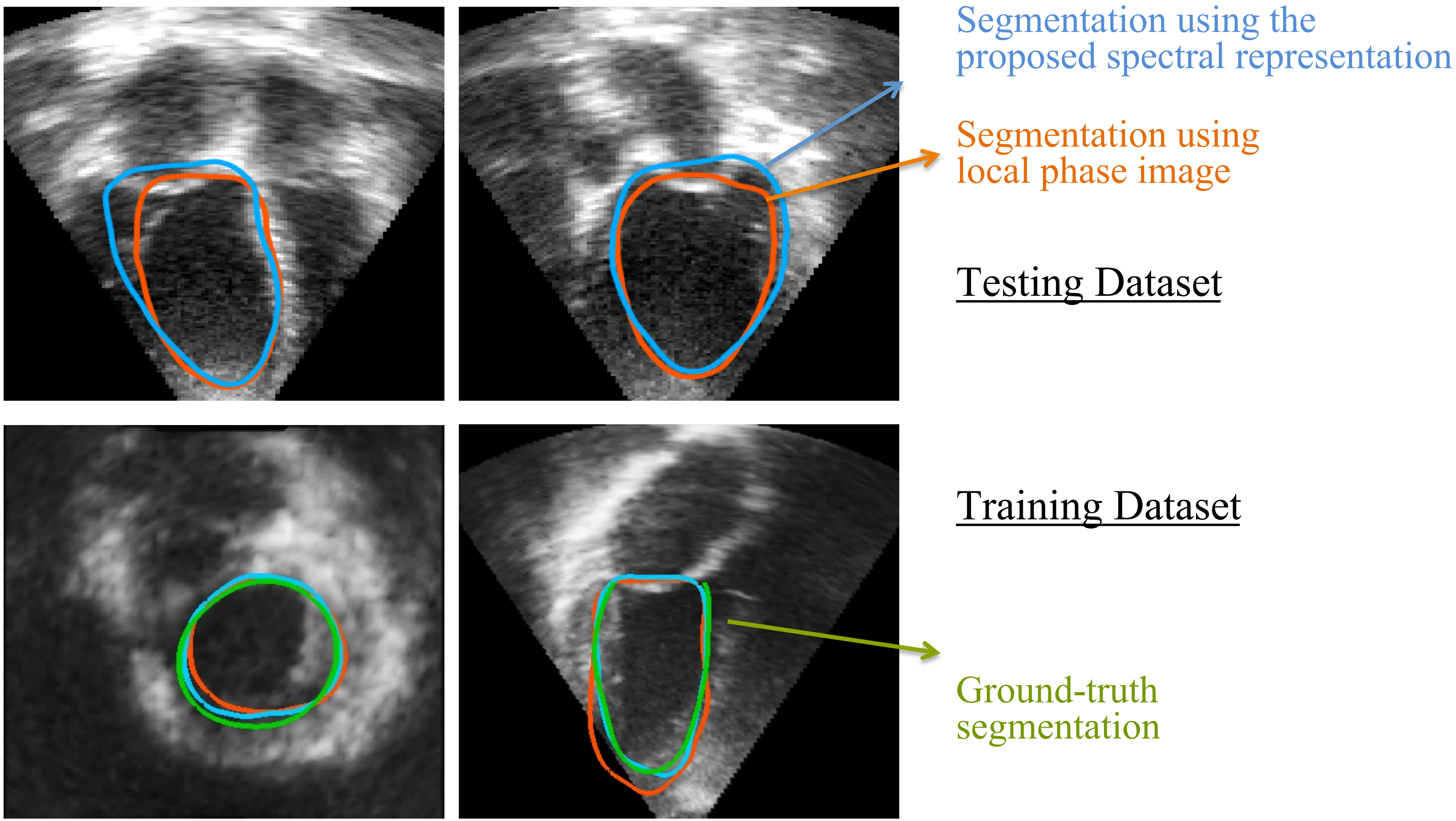
Surface to Surface Distance Errors



Estimation of Clinical Indices



Qualitative Results



Conclusion & Acknowledgements

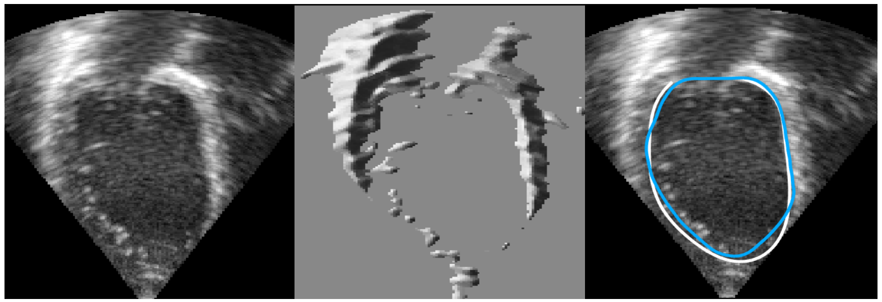
- Conclusion
  - We propose a novel structural representation for echocardiography
  - This representation enables accurate multi-atlas segmentation
- Future work
  - The linear approximation of non-linear manifold can be improved
  - Application of the proposed feature in echocardiography strain imaging
  - Explore application to multi modal image registration (echo / CT)

Sparsity Based Spectral Embedding: Application to Multi-Atlas Echocardiography Segmentation

Ozan Oktay*, Wenzhe Shi, Jose Caballero,

Kevin Keraudren, and Daniel Rueckert

*E: o.oktay13@imperial.ac.uk



- Acknowledgements:

Imperial College London

Comparison against the state-of-the-art echocardiographic image segmentation methods

Table 1: Comparison of the proposed multi-atlas approach (A) against the state-of-the-art echocardiogaphy segmentation: active surfaces [1] and active shape model [2]. Estimated ejection fraction (EF) and end-diastolic volume (EDV) are compared against their reference values. The correlation accuracy is reported in terms of Pearson's coefficient (R) and Bland-Altman's limit of agreement (BA).

	Mean (mm)	R_{EF}	BA_{EF} $(\mu \pm 2\sigma)$	R_{EDV}	BA_{EDV} $(\mu \pm 2\sigma)$	# of Patients
(A)	**2.32±0.78**	**0.923**	**-0.74±6.26**	**0.926**	**12.88±35.71**	**15**
[1]	-	0.907	-2.4±23	0.971	-24.60±21.80	24
[2]	1.84±1.86	-	0±19	-	3.06±46.86	10

1. Barbosa, D., et al.: Fast and fully automatic 3-D echocardiographic segmentation using B-spline explicit active surfaces. Ultrasound in medicine and biology (2013)

2. Butakoff, C., et al: Order statistic based cardiac boundary detection in 3D+T echocardiograms. FIMH. Springer (2011)

Accuracy of the Derived Clinical Indices

Table 2: This table shows the accuracy of the derived clinical indices for the training (Patient 1 to 15) and testing datasets (Patient 16 to 30). Pearson's correlation coefficient (PCC) and Bland-Altman's limit of agreement ($\mu \pm 1.96\sigma$) values are given for the following indices: ejection fraction, stroke volume, end-systolic volume, and end-diastolic volume.

Testing dataset	PCC	LOA ($\mu \pm 1.96\sigma$)
ED volume (ml)	0.926	12.81±33.77
ES volume (ml)	0.936	-7.77±28.27
Ejection fraction (%)	0.923	0.74±7.58
Stroke volume (ml)	0.832	-5.05±12.49

Training dataset	PCC	LOA ($\mu \pm 1.96\sigma$)
ED volume (ml)	0.983	9.80±45.66
ES volume (ml)	0.961	11.21±56.91
Ejection fraction (%)	0.787	-1.07±18.52
Stroke volume (ml)	0.856	-1.38±27.08

Method runtime

- Results on 3 GHz machine
- Speckle reduction: 3 min
- Shape representation: 2.5 min
- Deformable registration (5 atlases): 30 min