Recognition of Deformable Object Category and Pose

Columbia University

Yinxiao Li
Chih-Fan Chen
Peter Allen

What is a deformable object (DO)?

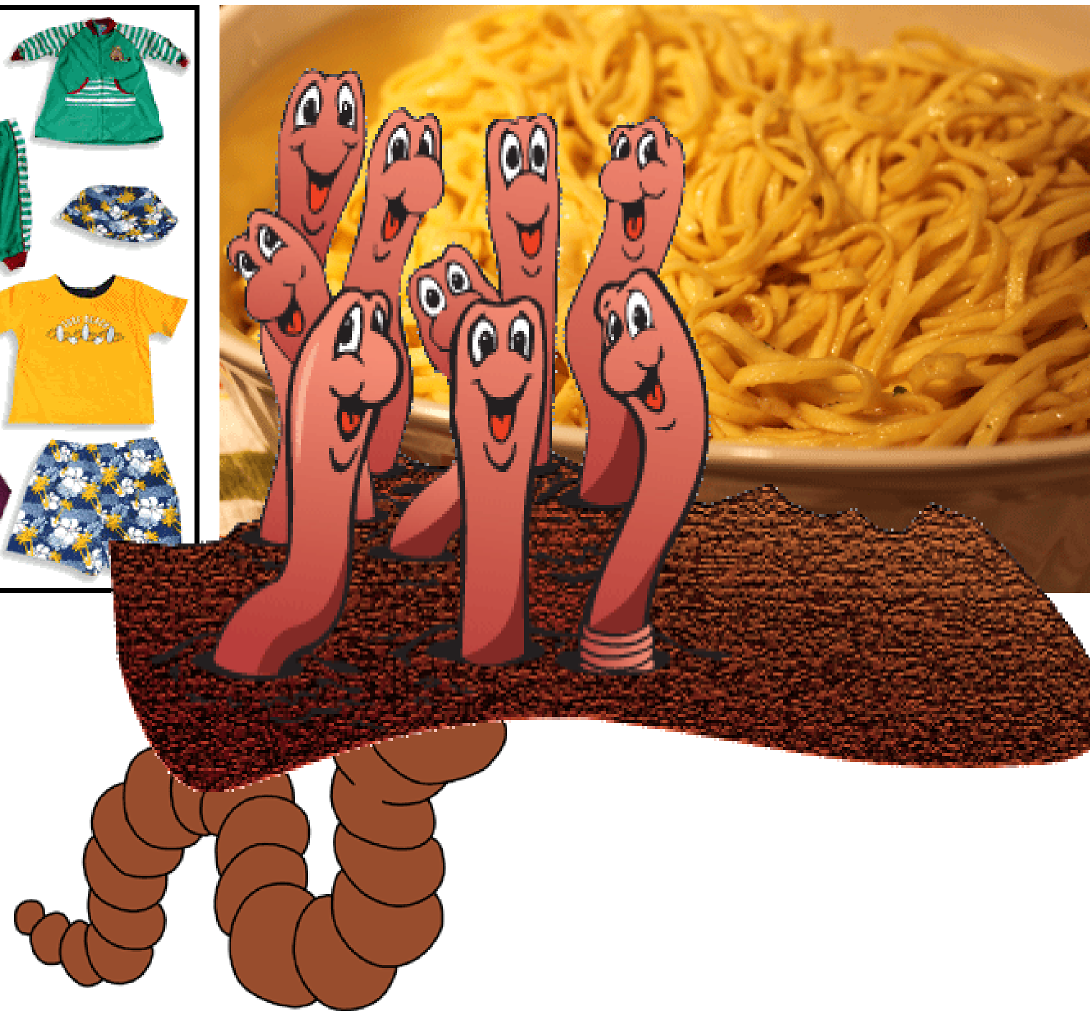
What is a deformable object (DO)?

http://new-delhi.all.biz/ready-made-garments-for-children-g132102#show0

What is a deformable object (DO)?

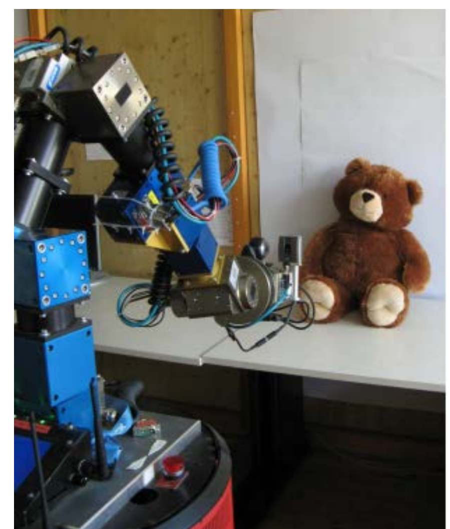
http://www.melovinglife.com/food/main-course/crustaceans-inspired-roasted-crab-garlic-noodles/

What is a deformable object (DO)?



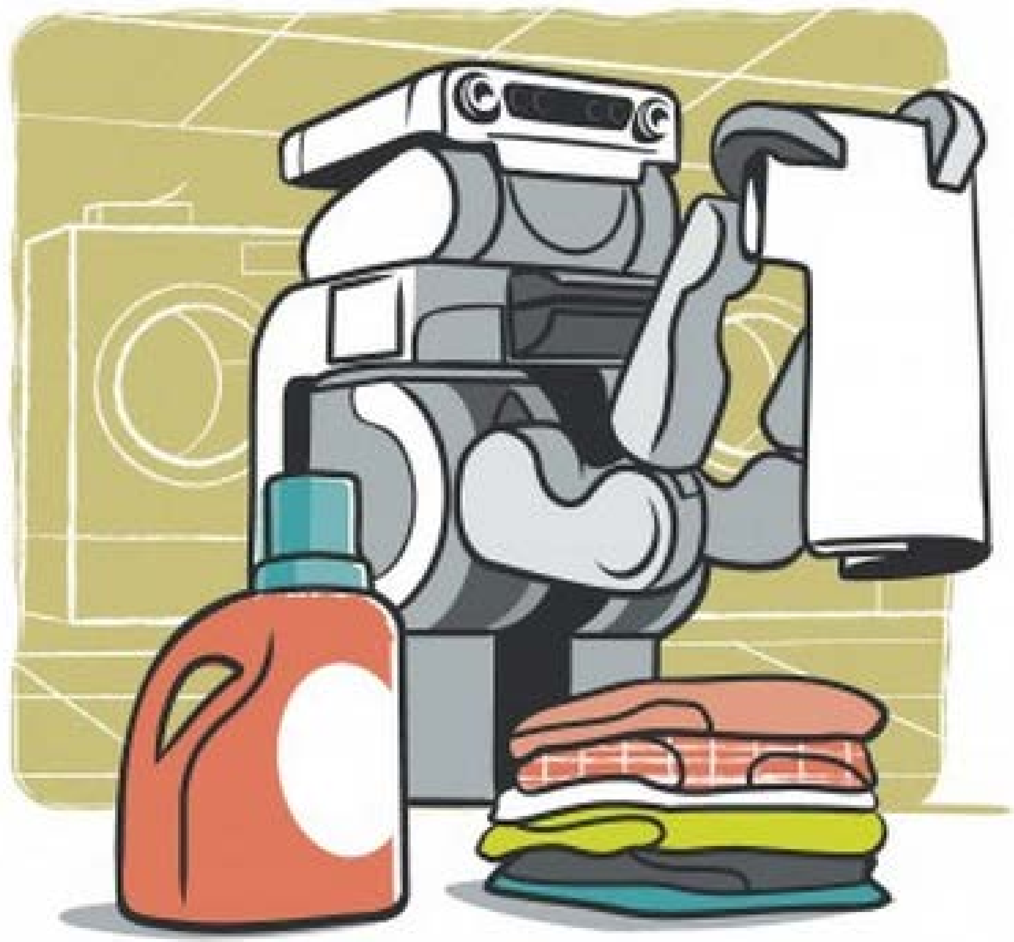
http://www.kidnesia.com/Kidnesia/Potret-Negeriku/Flona/Fauna/Cacing-Tanah

Why do robots need to recognize DO?

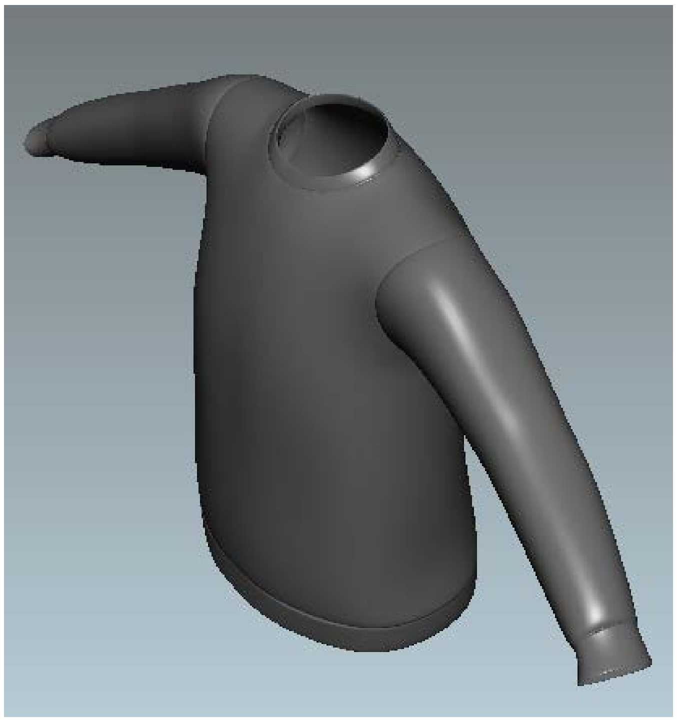


http://mynorthwest.com/920/2312665/Why-robots-could-soon-replace-fast-food-workers-demanding-higher-minimum-wage
http://www.cs.berkeley.edu/~pabbeel/
http://mynorthwest.com/920/2312665/Why-robots-could-soon-replace-fast-food-workers-demanding-higher-minimum-wage
http://www2.informatik.uni-freiburg.de/~stachnis/pdf/frank10rssws.pdf

Our goal



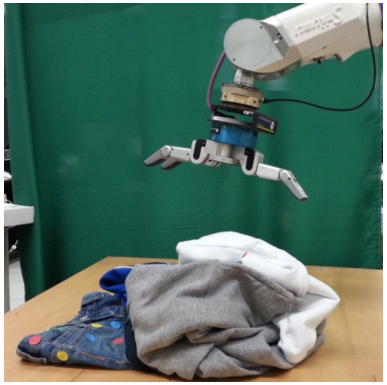
Challenge – Find Category & Pose



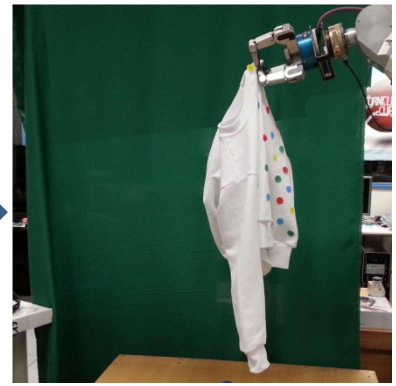
Outline

- Entire Pipeline of Manipulation of DO
- Training and Testing Flow
  – Defining Deformable Poses
  – Generating Training Exemplars
  – Estimating Poses of DO
  – Two-Layer Classifier
- Experimental Results
- Conclusion

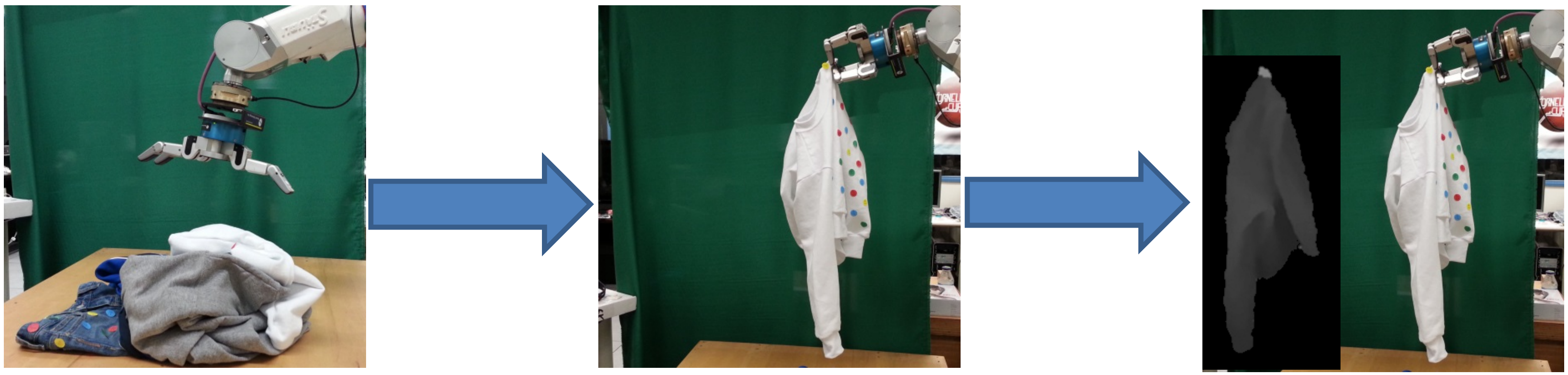
Pipeline for Manipulation of DO



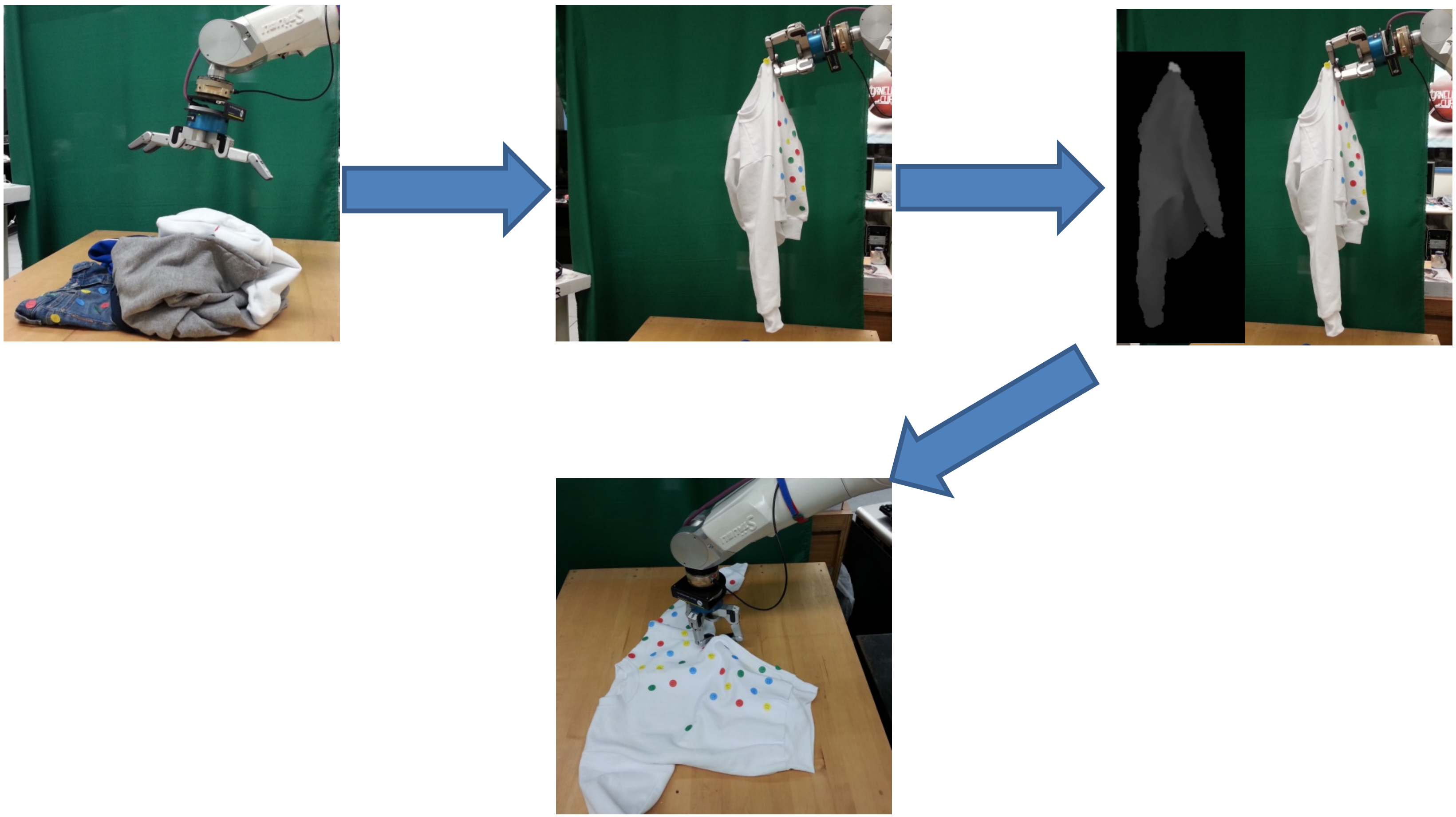
Pipeline for Manipulation of DO



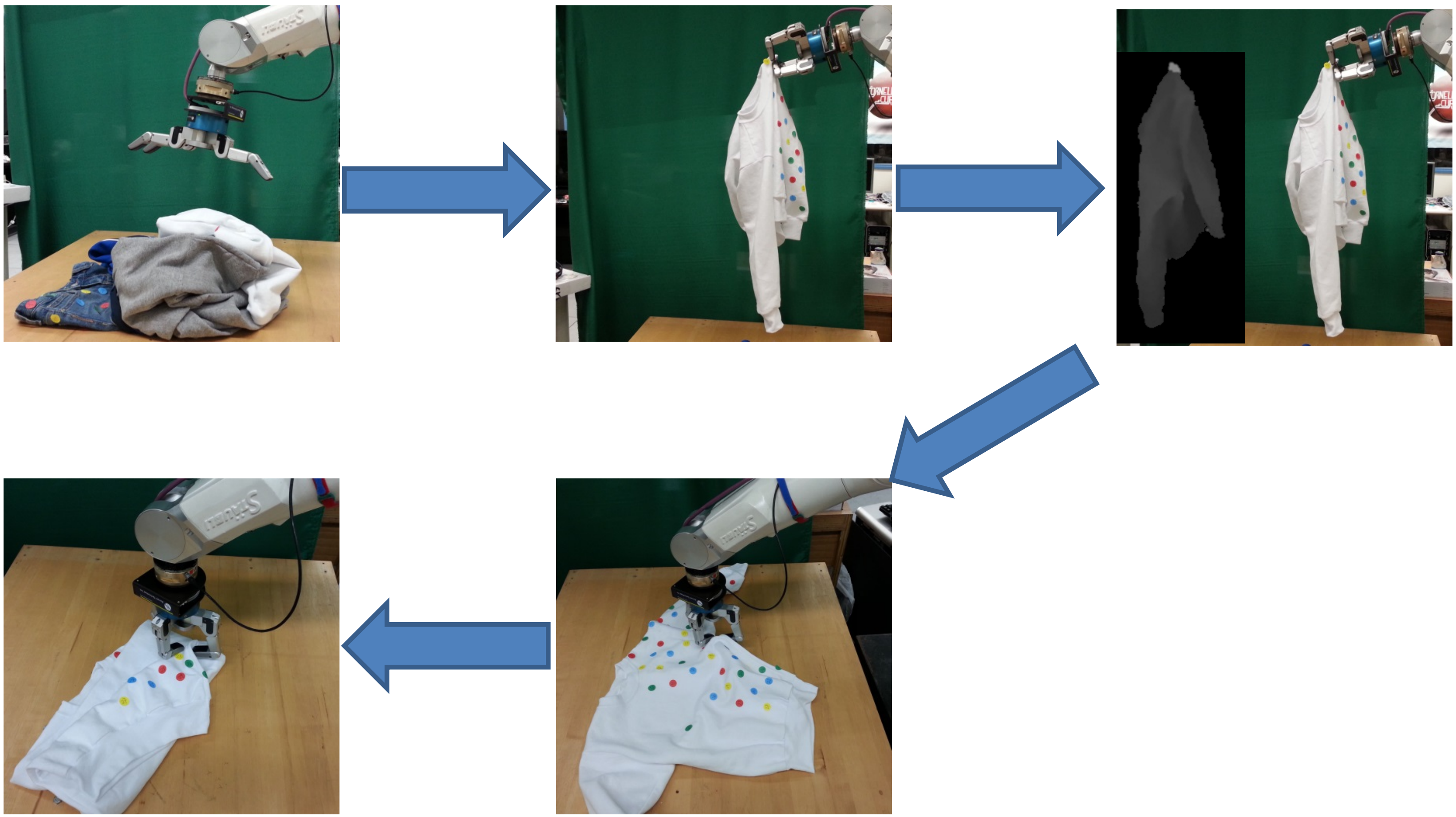
Pipeline for Manipulation of DO



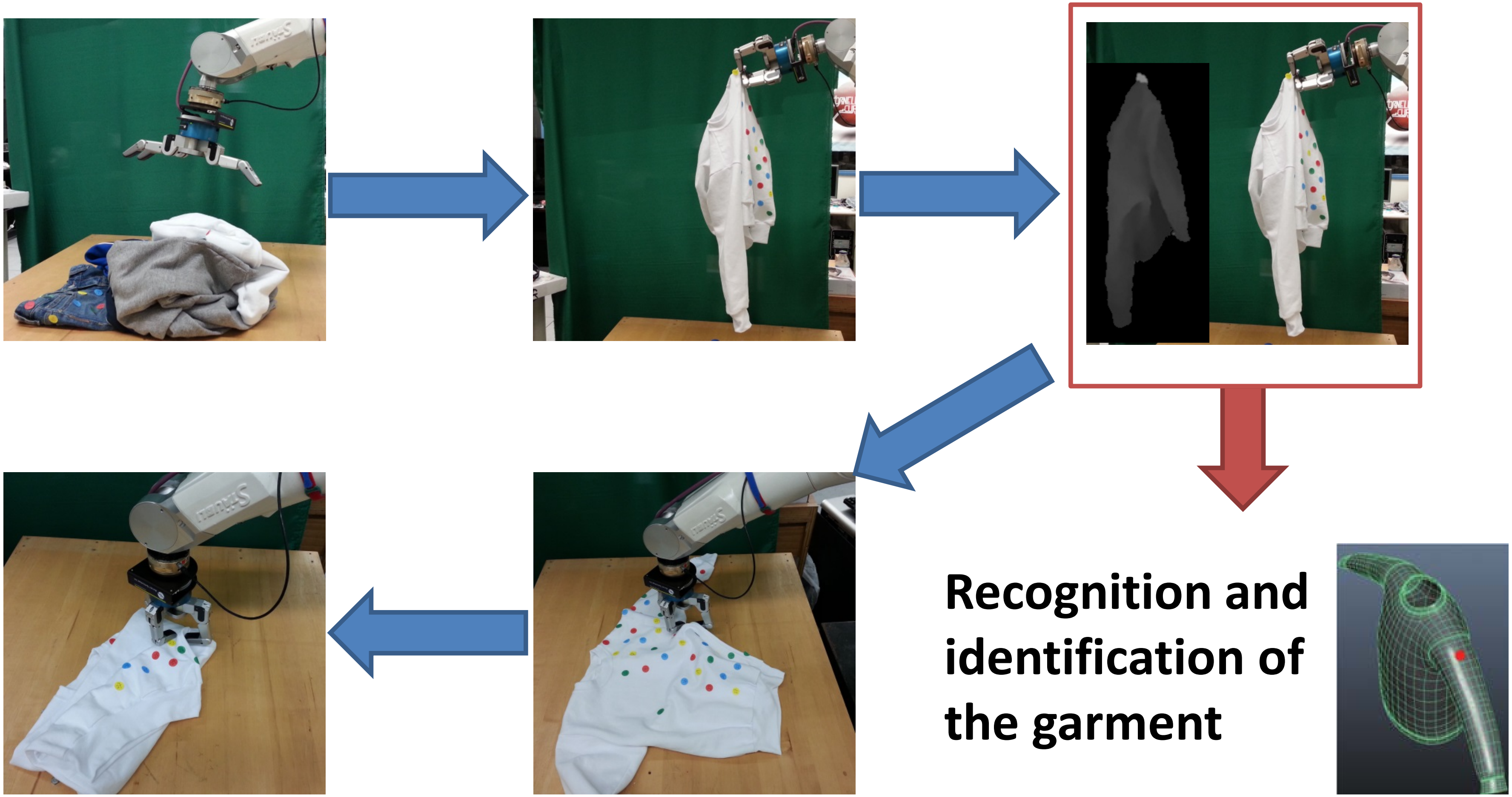
Pipeline for Manipulation of DO



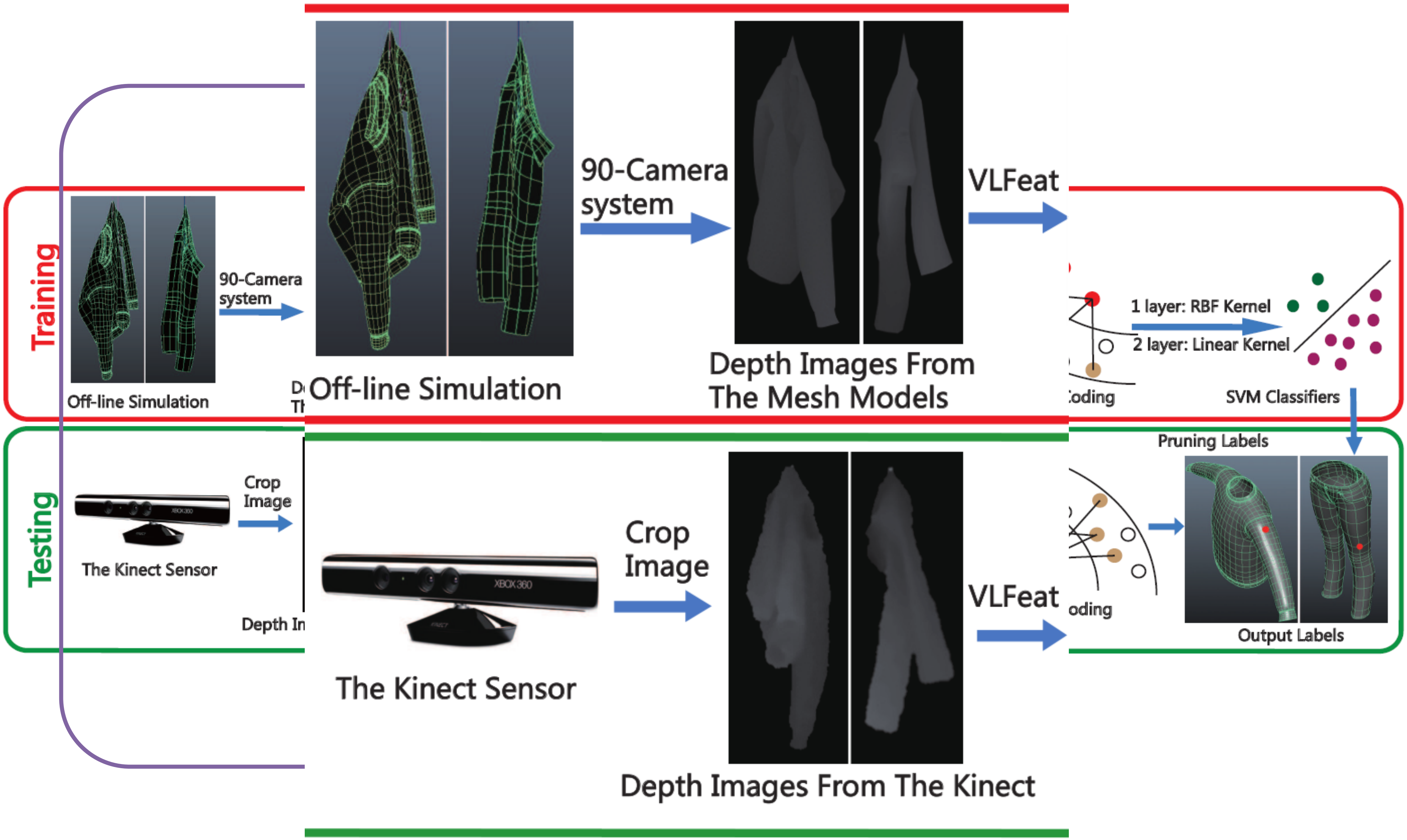
Pipeline for Manipulation of DO



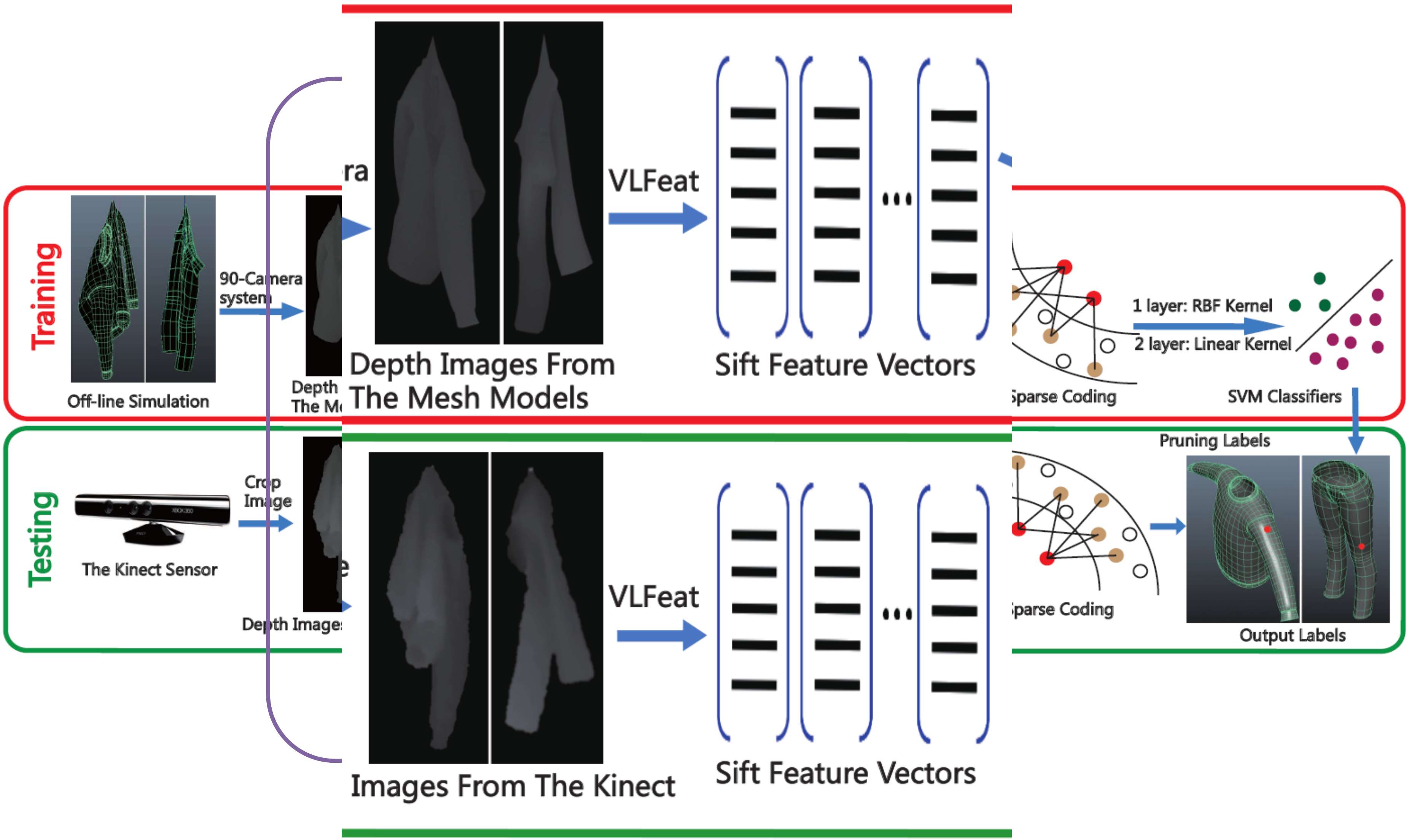
Pipeline for Manipulation of DO



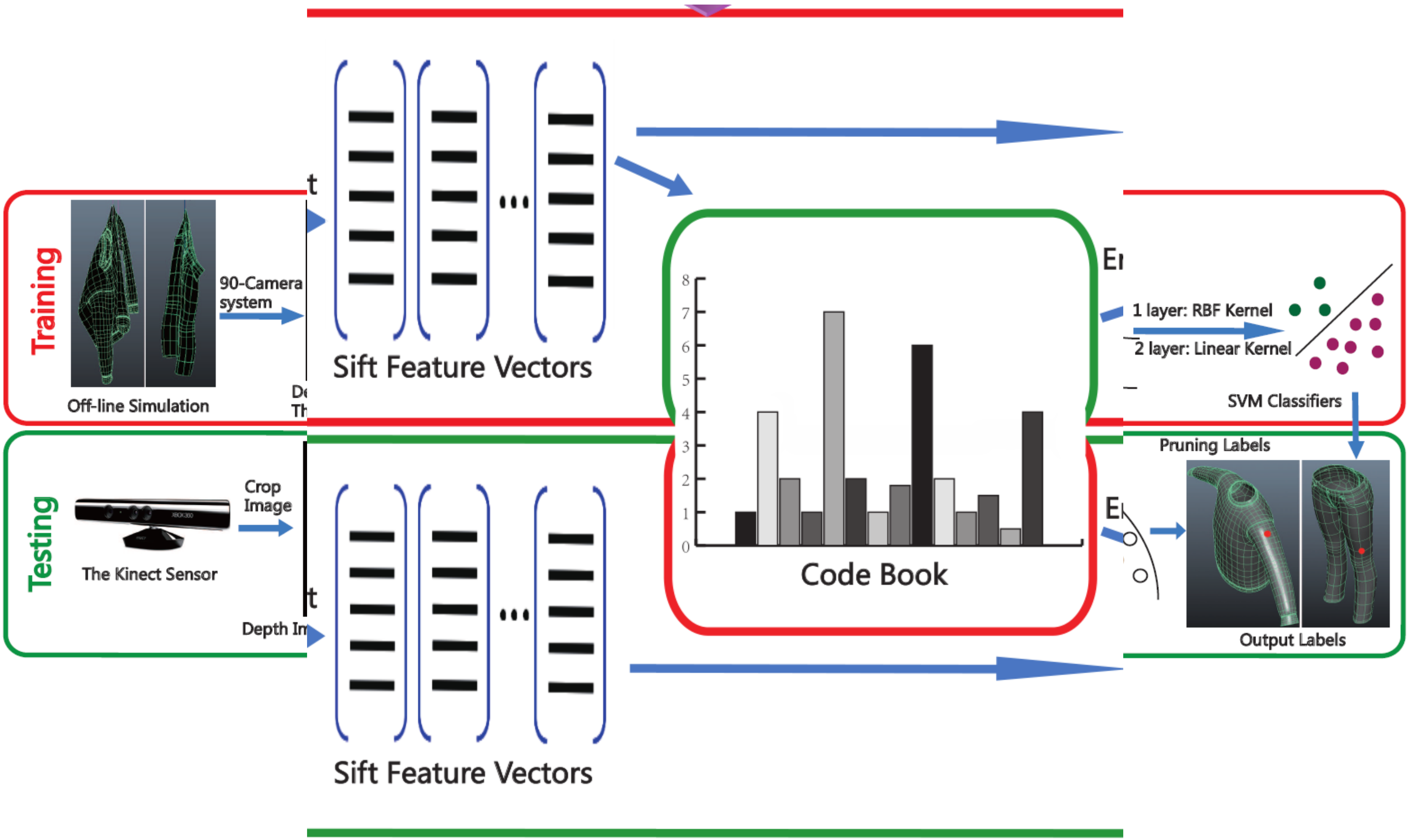
Entire Training and Testing Flow



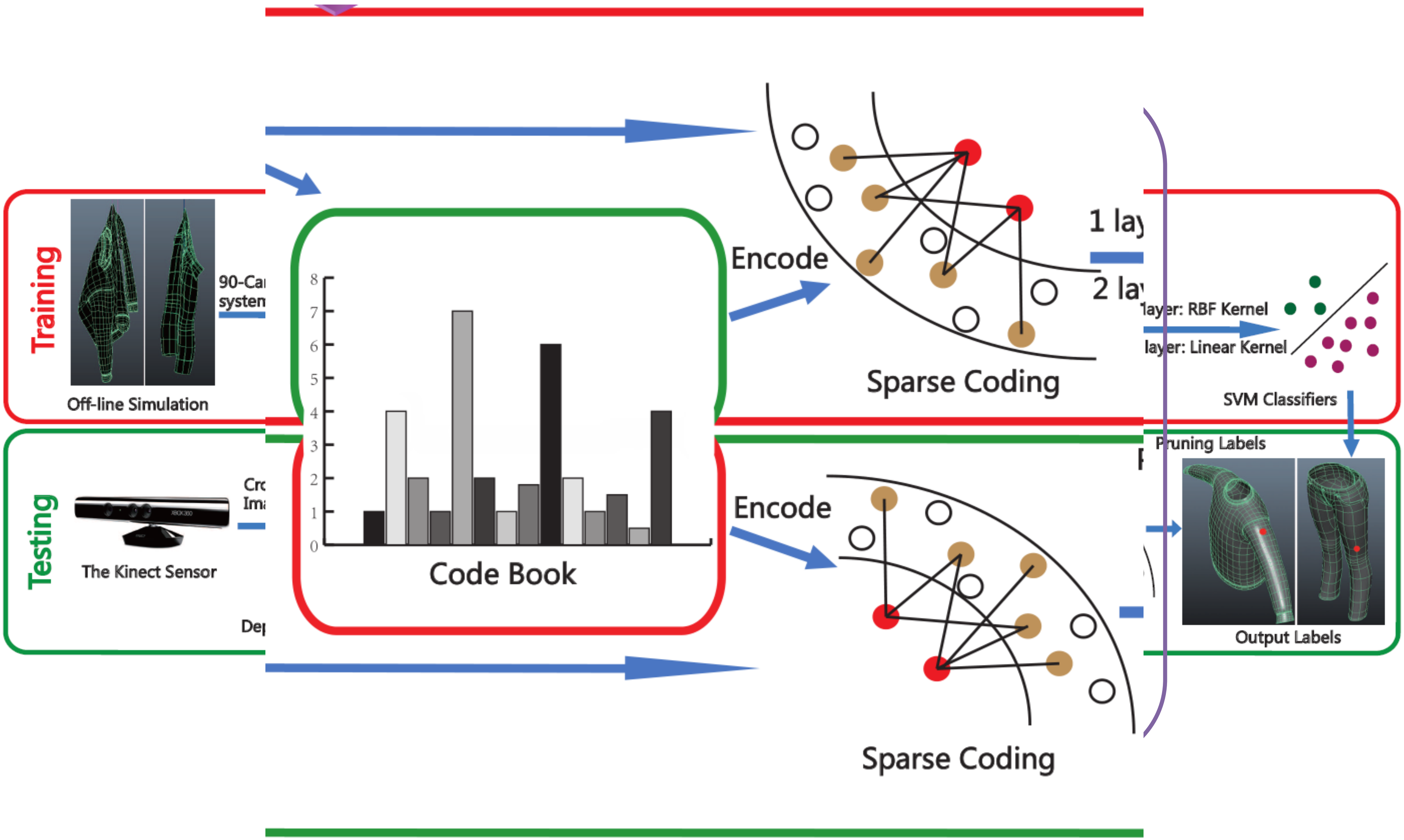
Entire Training and Testing Flow



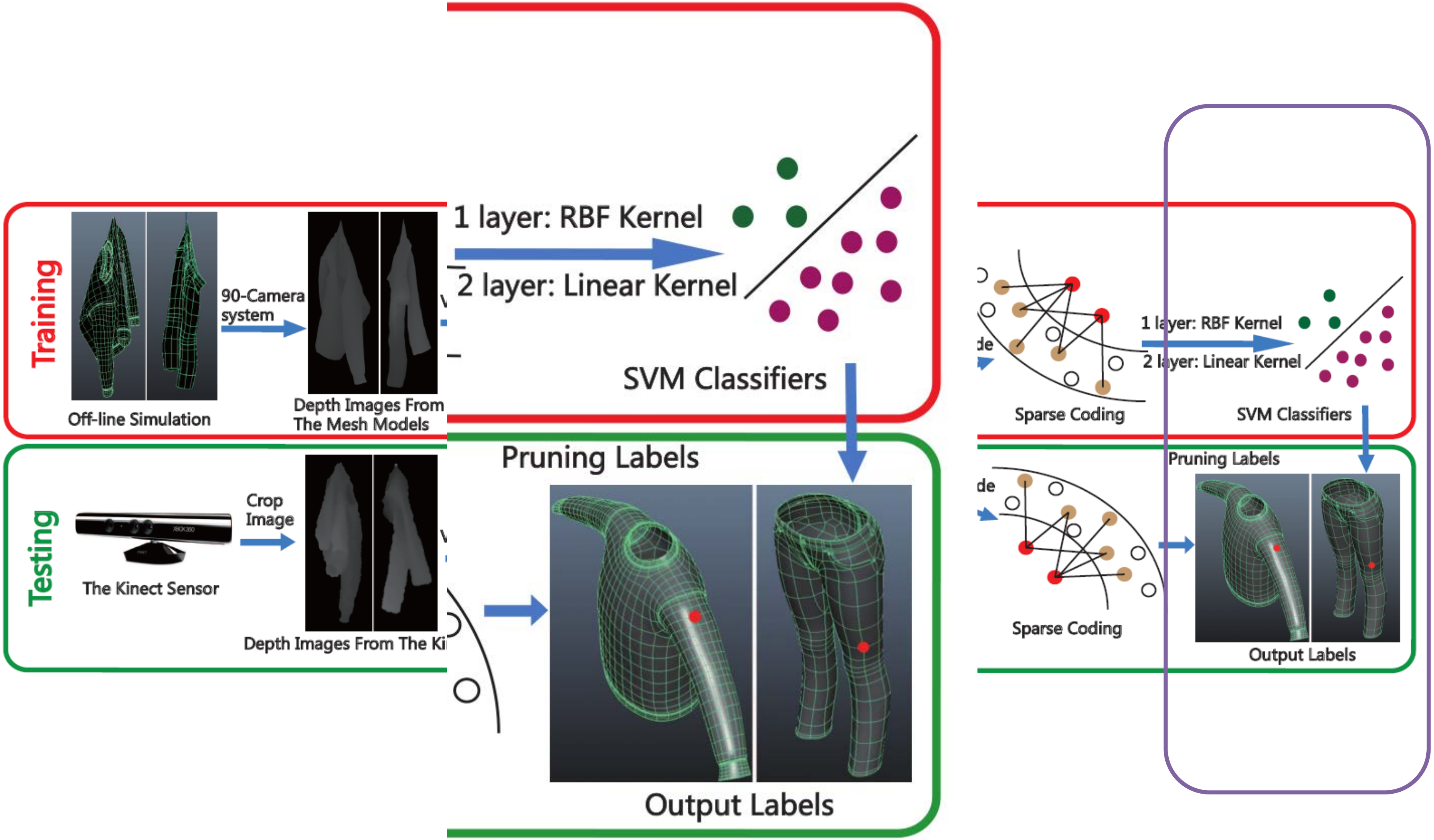
Entire Training and Testing Flow



Entire Training and Testing Flow

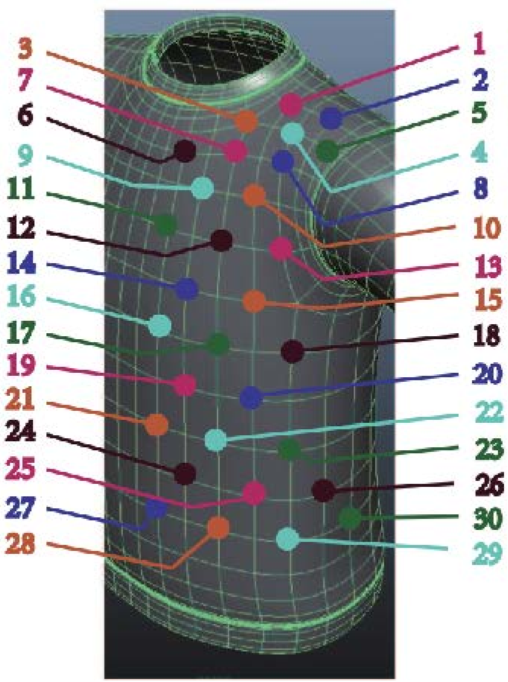


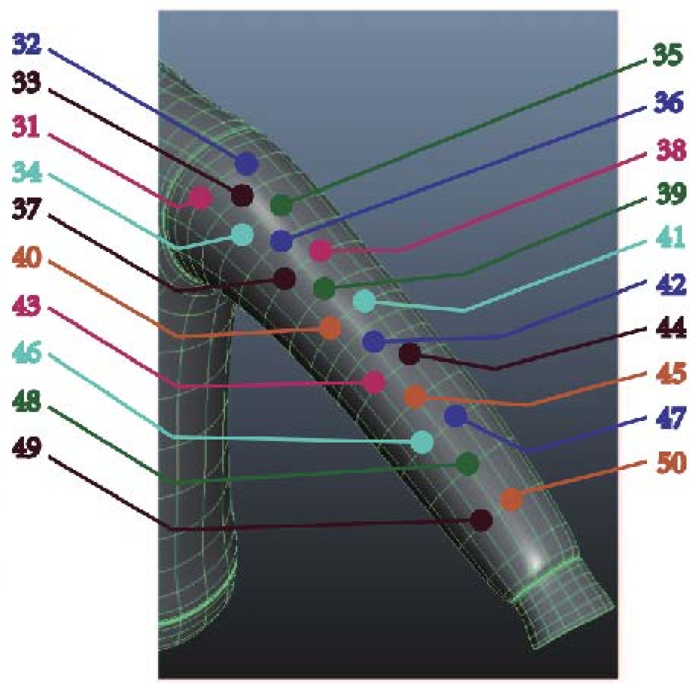
Entire Training and Testing Flow



Defining Deformable Poses

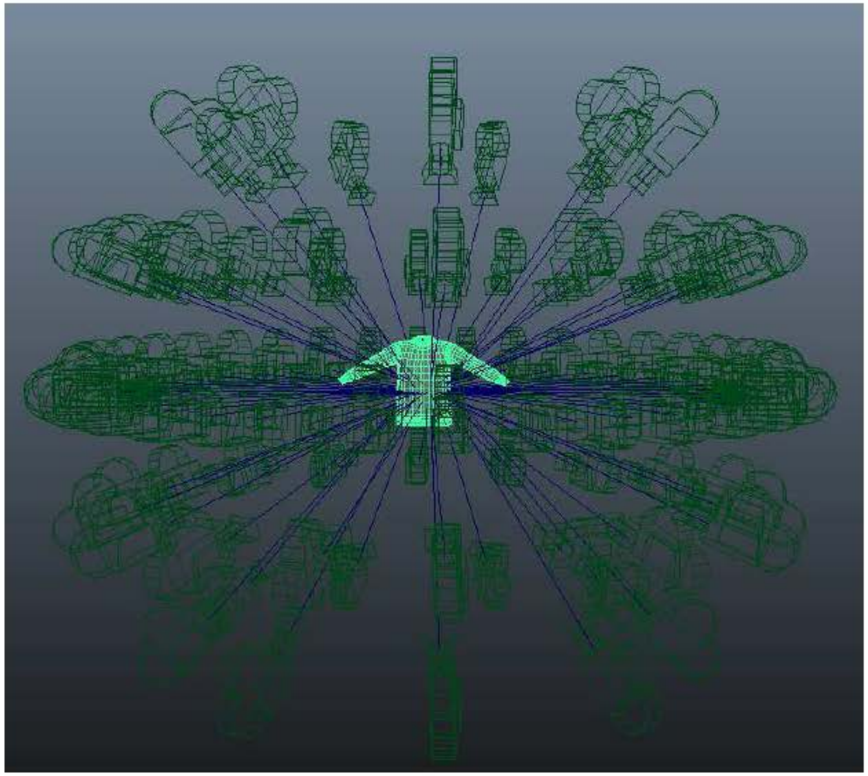
- 20-50 points on each garment
- Color coding label pads on real garment

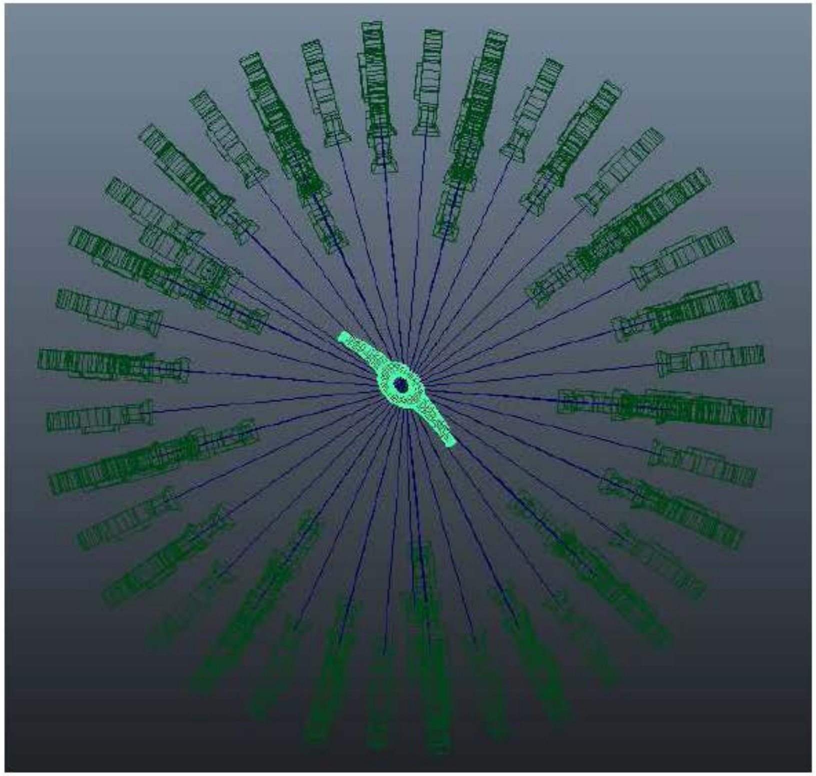




Generating Training Exemplars

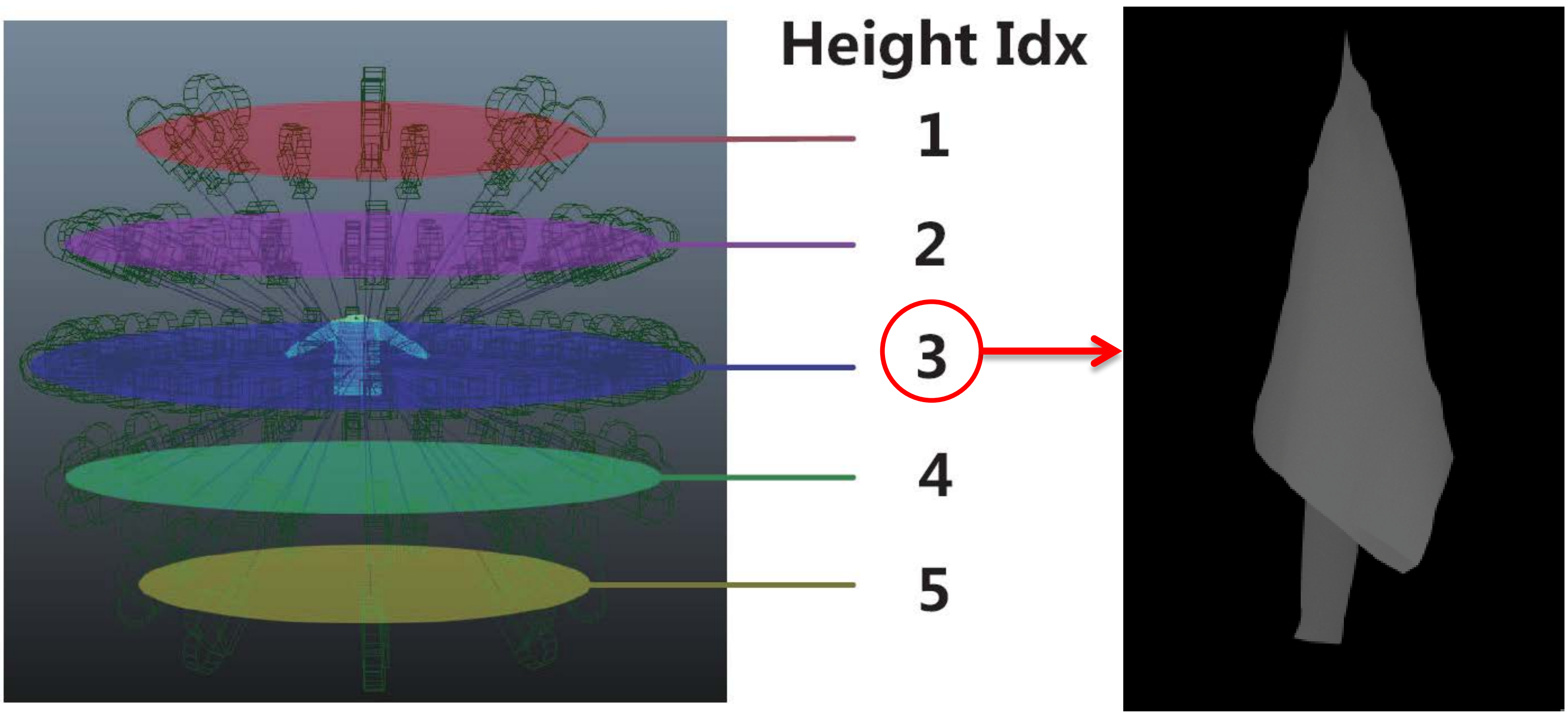
- Virtual cameras in Maya (cloth simulator)
- 90-Camera System on a geodesic dome





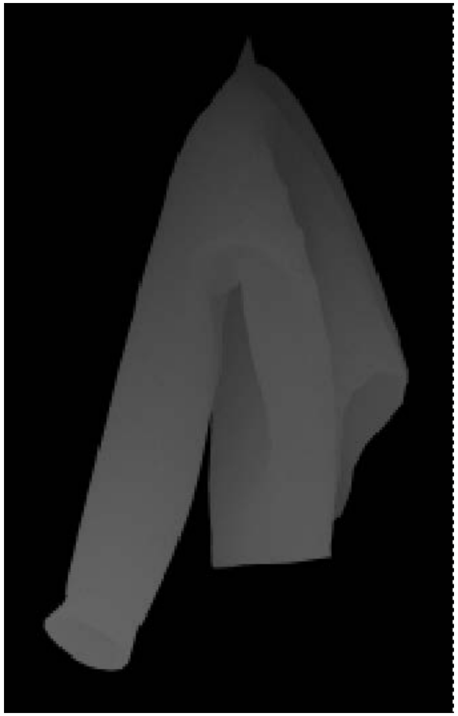
Generating Training Exemplars

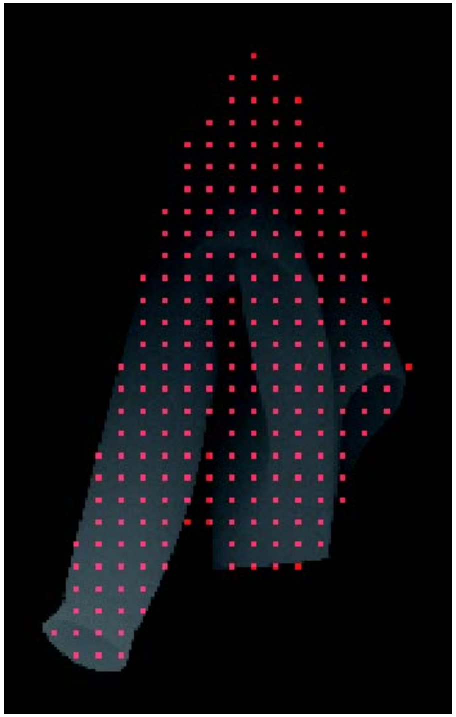
- Virtual cameras in Maya (cloth simulator)
- 90-Camera System on a geodesic dome



Generating Training Exemplars

- Crop image in terms of the object
- Dense SIFT feature over the entire image





Learning feature signature

- Sparse coding

Feature Vectors

Code Book

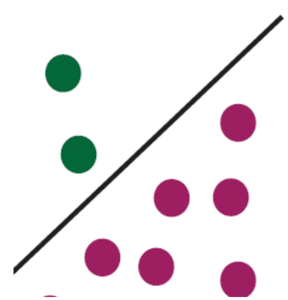
$$\min_{\mathbf{W},\mathbf{V}} \sum_{1}^{N} \left\| \mathbf{X}_i - \mathbf{w}_i \mathbf{V} \right\|^2 + \lambda \left| \mathbf{w}_i \right|$$

Regularization Term

- Max pooling

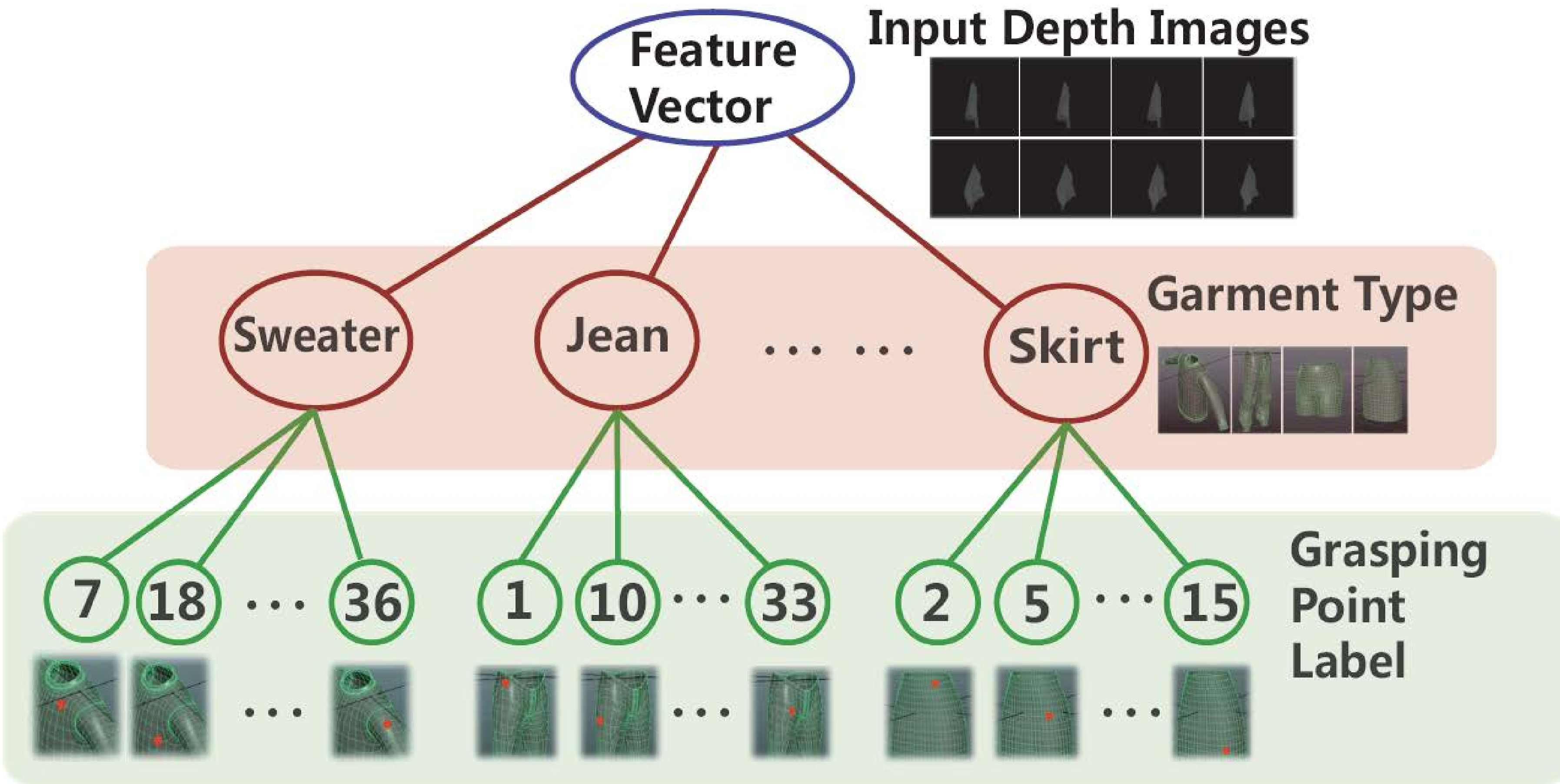
$$r_j = \max\{ \left| w'_{1j} \right|, \left| w'_{2j} \right| \ldots \left| w'_{Nj} \right| \}$$

- SVM

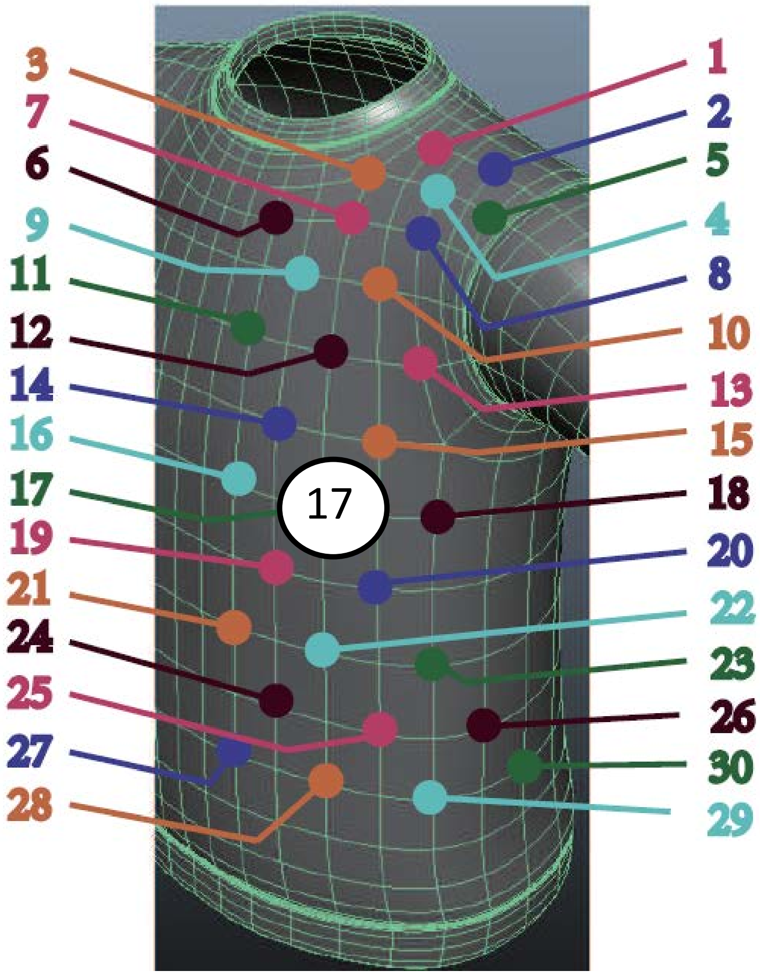


Two-Layer Classifier

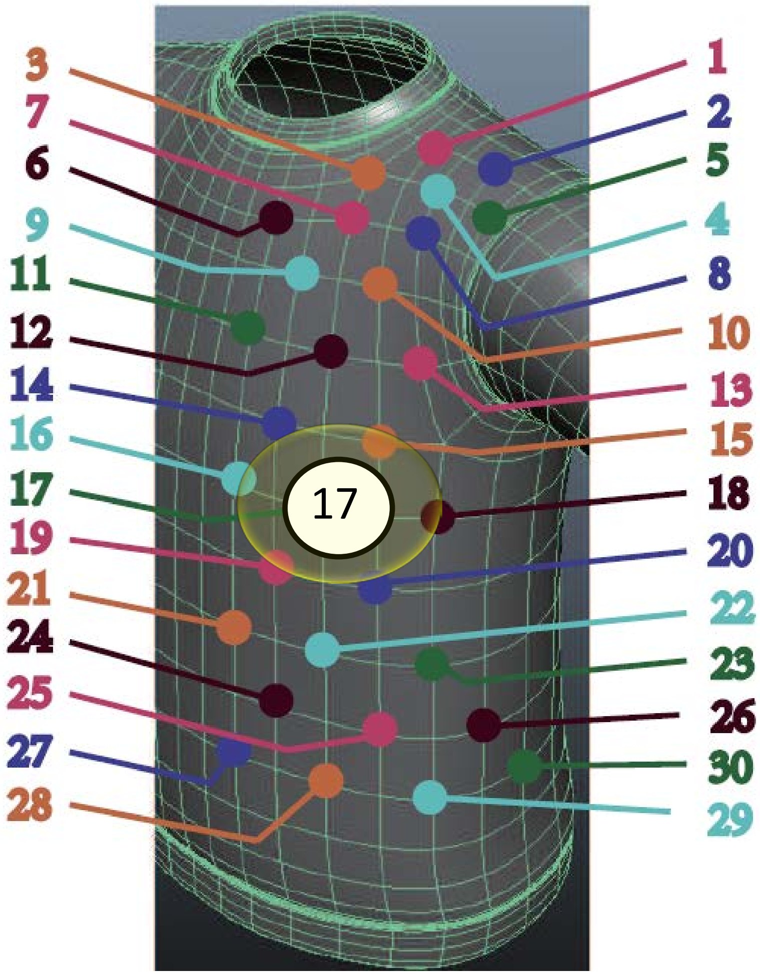
- First Layer: Category classification (SVM_1)
- Second Layer: Pose identification (SVM_2)



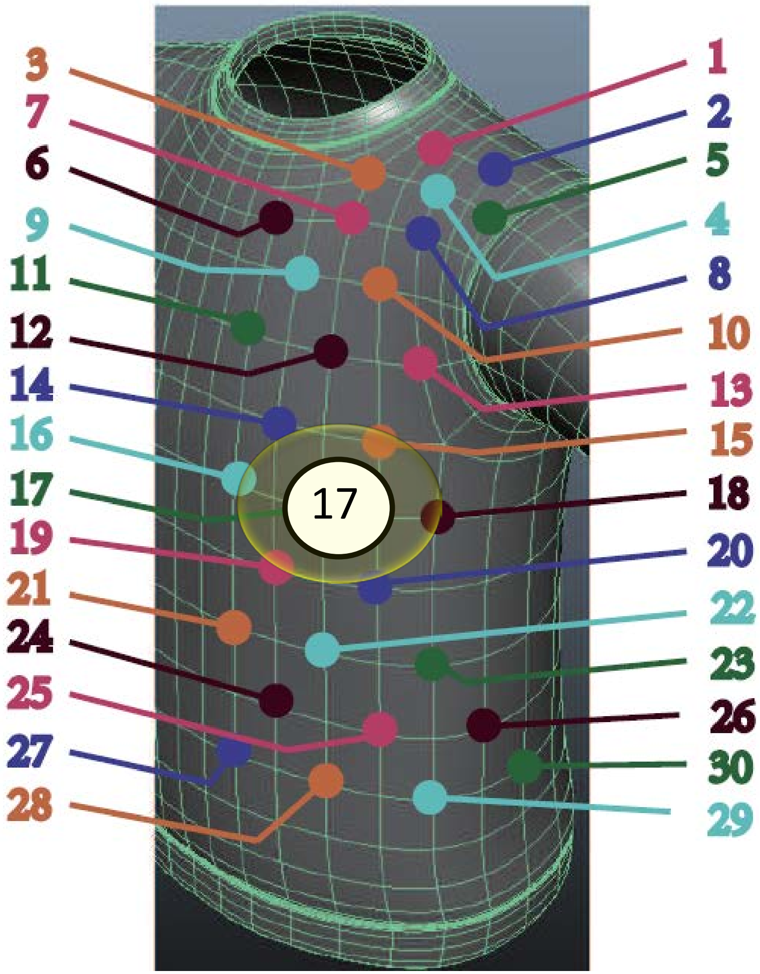
Test using simulation data

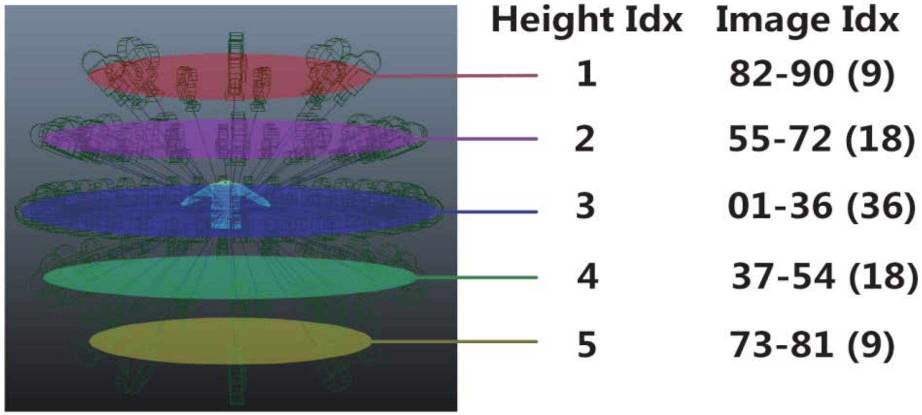


Test using simulation data



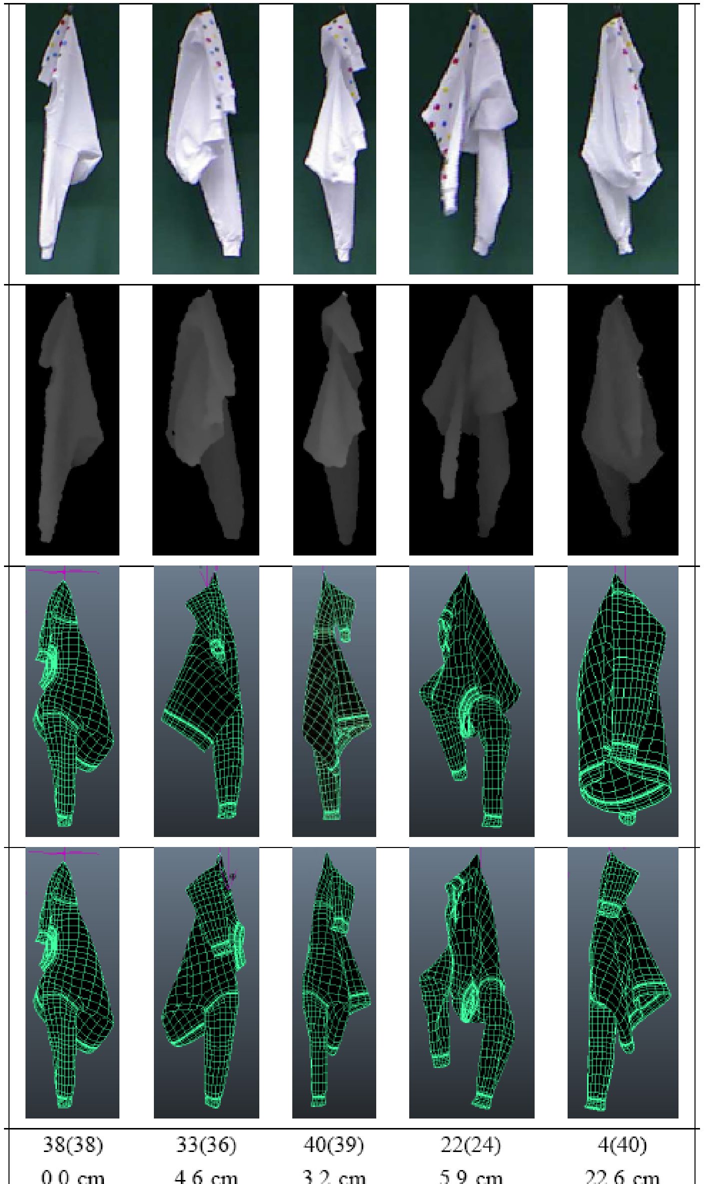
Test using simulation data





Height Group	# of Cam	Sweater Rank 1	Jean Rank 1	Short Pant Rank 1
Group 1	9	6.06	6.12	7.19
Group 2	18	4.93	5.12	5.67
Group 3	36	**4.52**	**4.44**	**5.48**
Group 4	18	4.84	5.06	6.29
Group 5	9	5.38	6.05	7.29
Group 2, 3, 4	72	5.02	5.09	5.81
All	90	5.79	5.73	6.76

Test using depth image from the Kinect



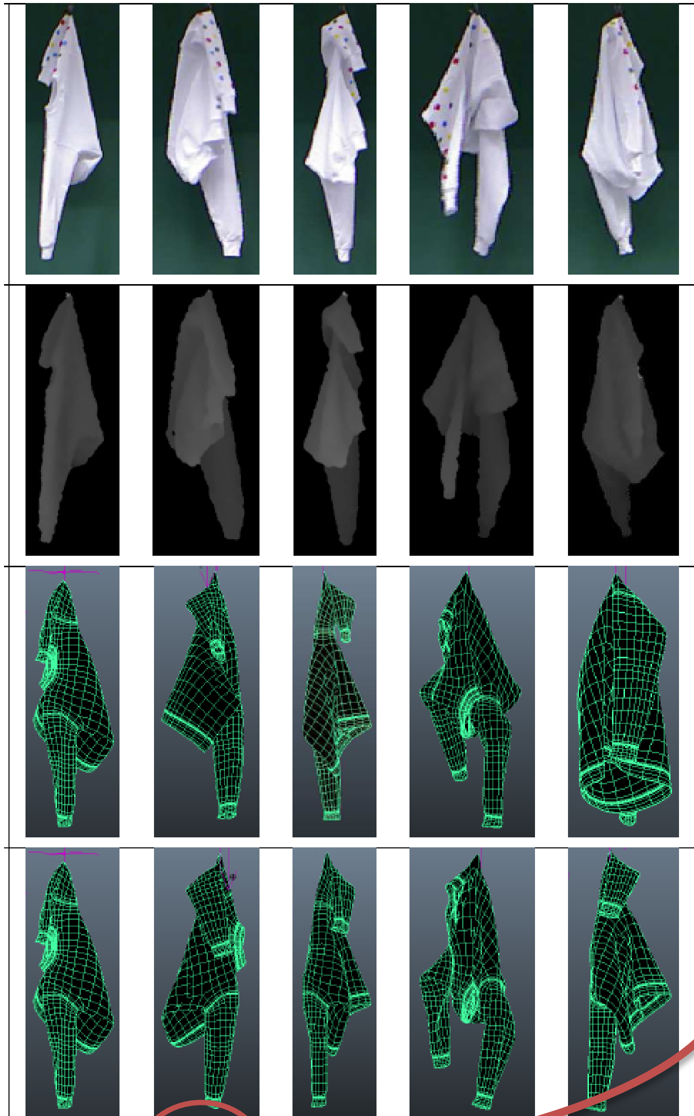
Color Image

Depth image

Ground truth from database

Predicted result from database

Test using depth image from the Kinect



38(38)	33(36)	40(39)	22(24)	4(40)
0.0 cm	4.6 cm	3.2 cm	5.9 cm	22.6 cm

Color Image

Depth image

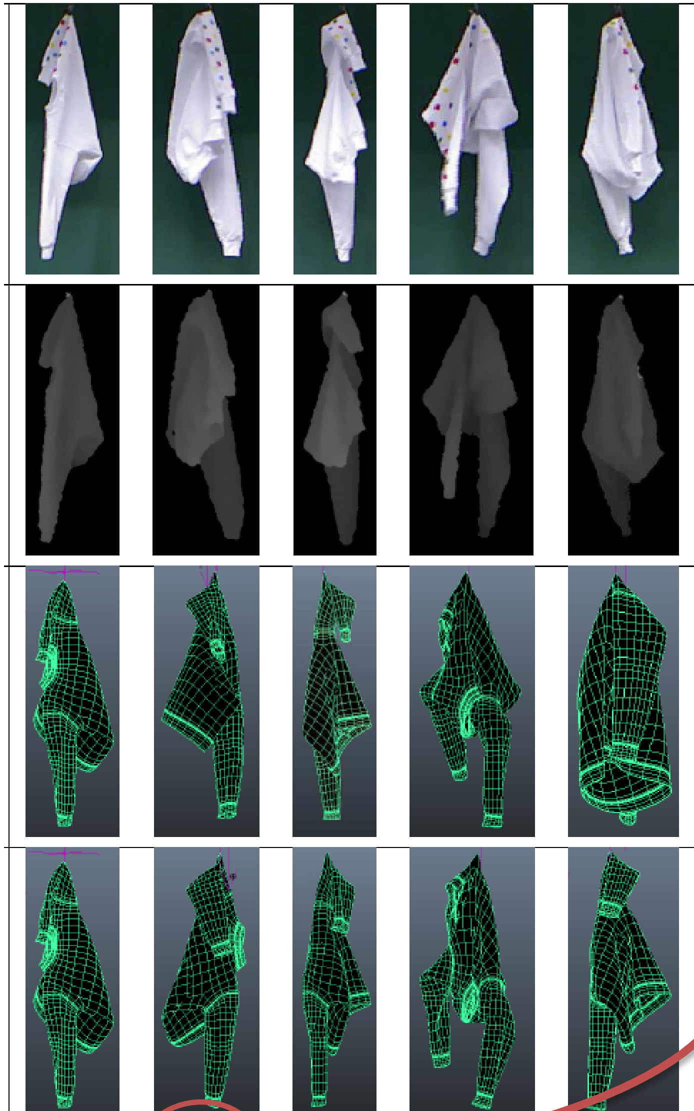
Ground truth from database

Predicted result from database

33(36)

4.6 cm

Test using depth image from the Kinect



38(38)	33(36)	40(39)	22(24)	4(40)
0.0 cm	4.6 cm	3.2 cm	5.9 cm	22.6 cm

Color Image

Depth image

Ground truth from database

Predicted result from database

Ground truth grasping point

Predicted grasping point

33(36)

4.6 cm

Distance between the two points

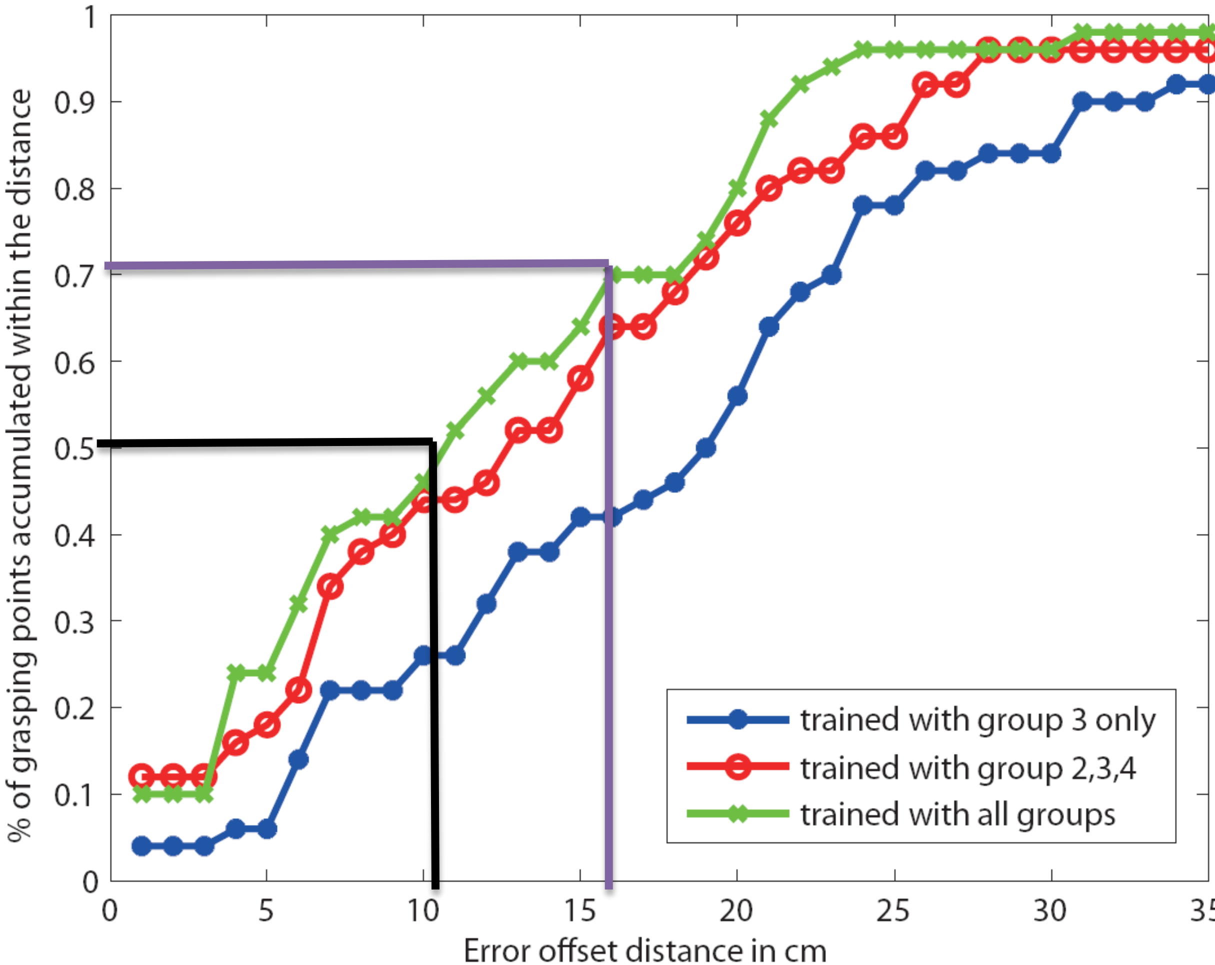
Experimental Results

- Test using depth image from the Kinect



	Garment Categories		
Height Group	Sweater	Jean	Short Pant
Group 3	75.79%	63.33%	84.71 %
Group 2, 3, 4	**79.61%**	62.83%	90.76 %
All	73.77%	**72.73%**	**91.40%**

Accuracy plot for sweater

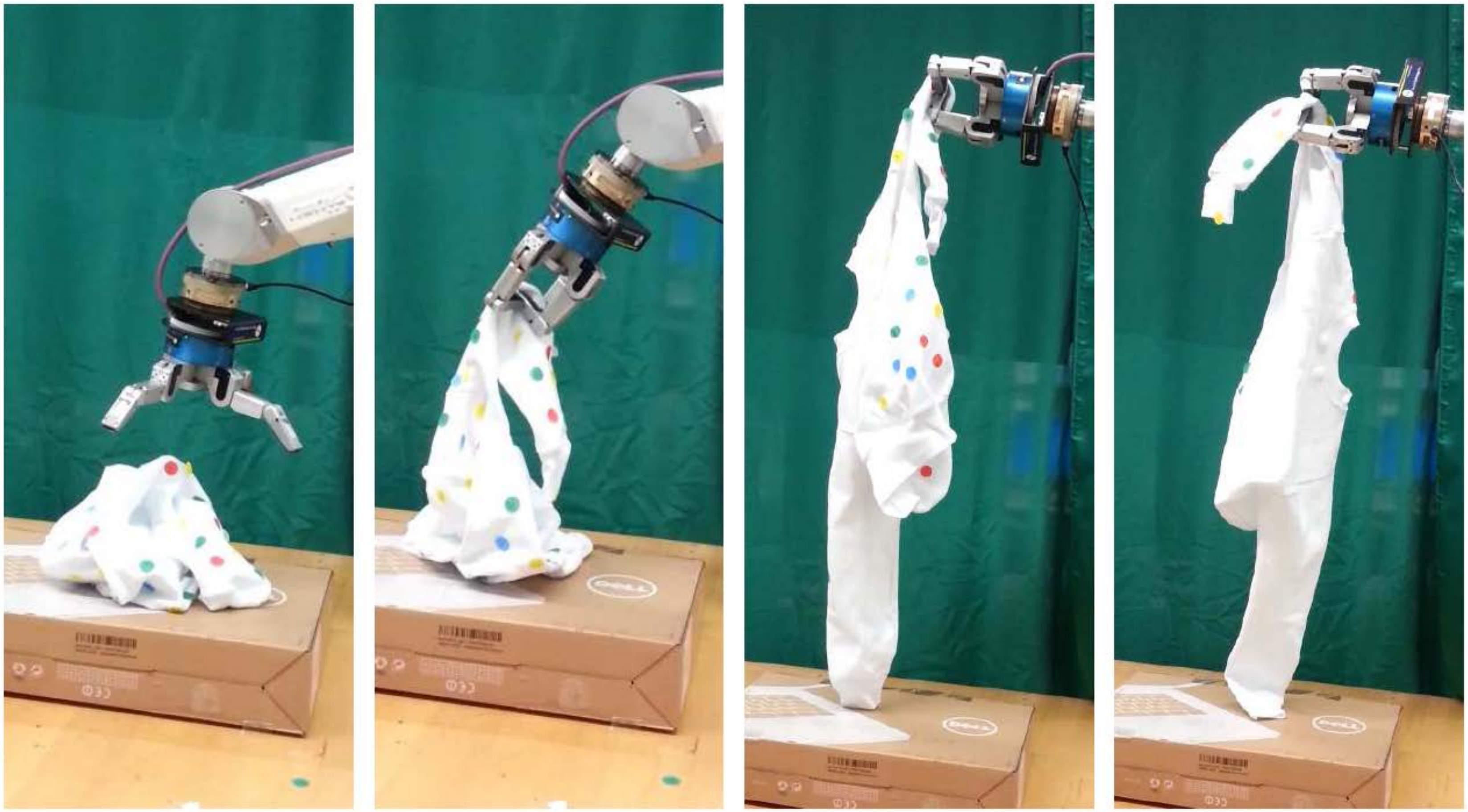


Max distance between two points is 65 cm

15cm offset is good enough for a robot to manipulate a garment!

Implementation on a robot -- Video

Implementation on a robot



1. Grasping
2. Picking up
3. Rotation and Recognition

Sweater: avg distance = <u>7.8</u> cm
Pants: avg distance = <u>10.2</u> cm

Conclusion

- A framework of recognizing deformable clothing object **category and poses**
- **Data-Driven** approach: off-line simulation, on-line recognition
- Experiments on both simulation data and real depth images– **high recognition rate / accuracy**

Future Work

- **Kinect fusion** algorithm to reconstruct
- Employing **color and texture** features
- Publish a **database** of deformable object with simulation
- Further tasks of manipulation – **regrasping**, placing flat, folding clothing

Acknowledgement

- **Columbia Computer Vision and Graphics Center (CVGC)**
  - Eitan Grinspun, Austin Reiter, Jon Weisz, Akash Grag, Thomas Berg
- **Anonymous reviewers**

Recognition of Deformable Object Category and Pose

Q&A

Columbia University

Yinxiao Li, Chih-Fan Chen, and Peter Allen