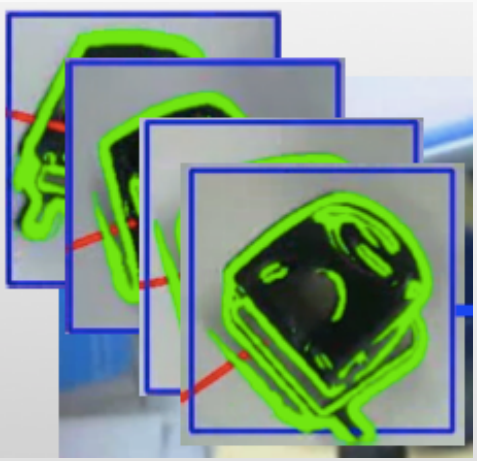
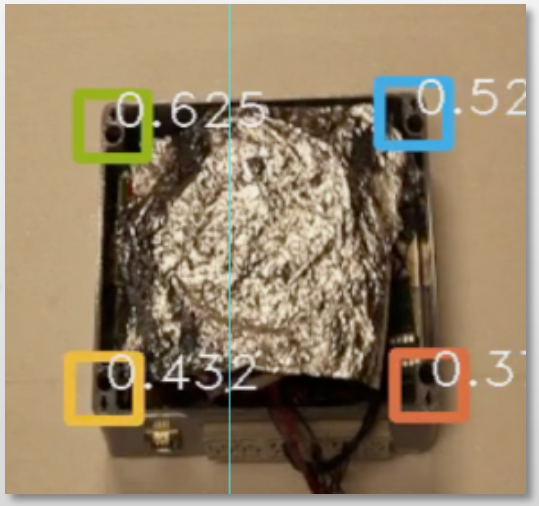
Novel Object and Part Representations for 3D Pose Estimation

Vincent Lepetit
TU Graz

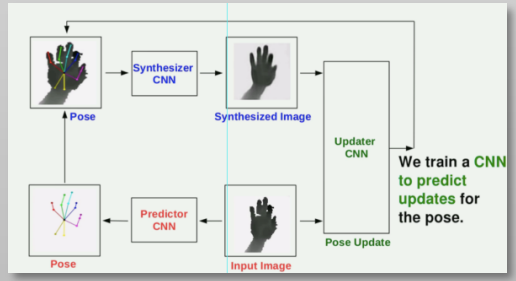


template-based approach

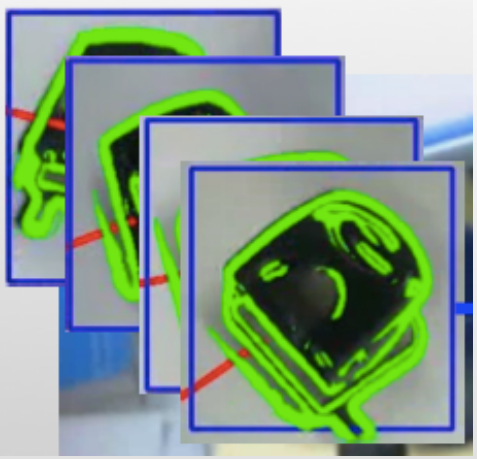


part-based approach

3D Pose Estimation of Objects



new optimization scheme

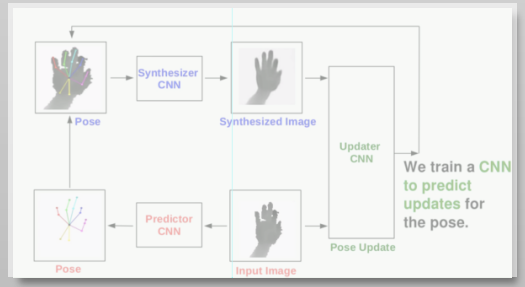


template-based approach

Paul Wohlhart and Vincent Lepetit.
Learning Descriptors for Object Recognition and 3D Pose Estimation.
CVPR'15.

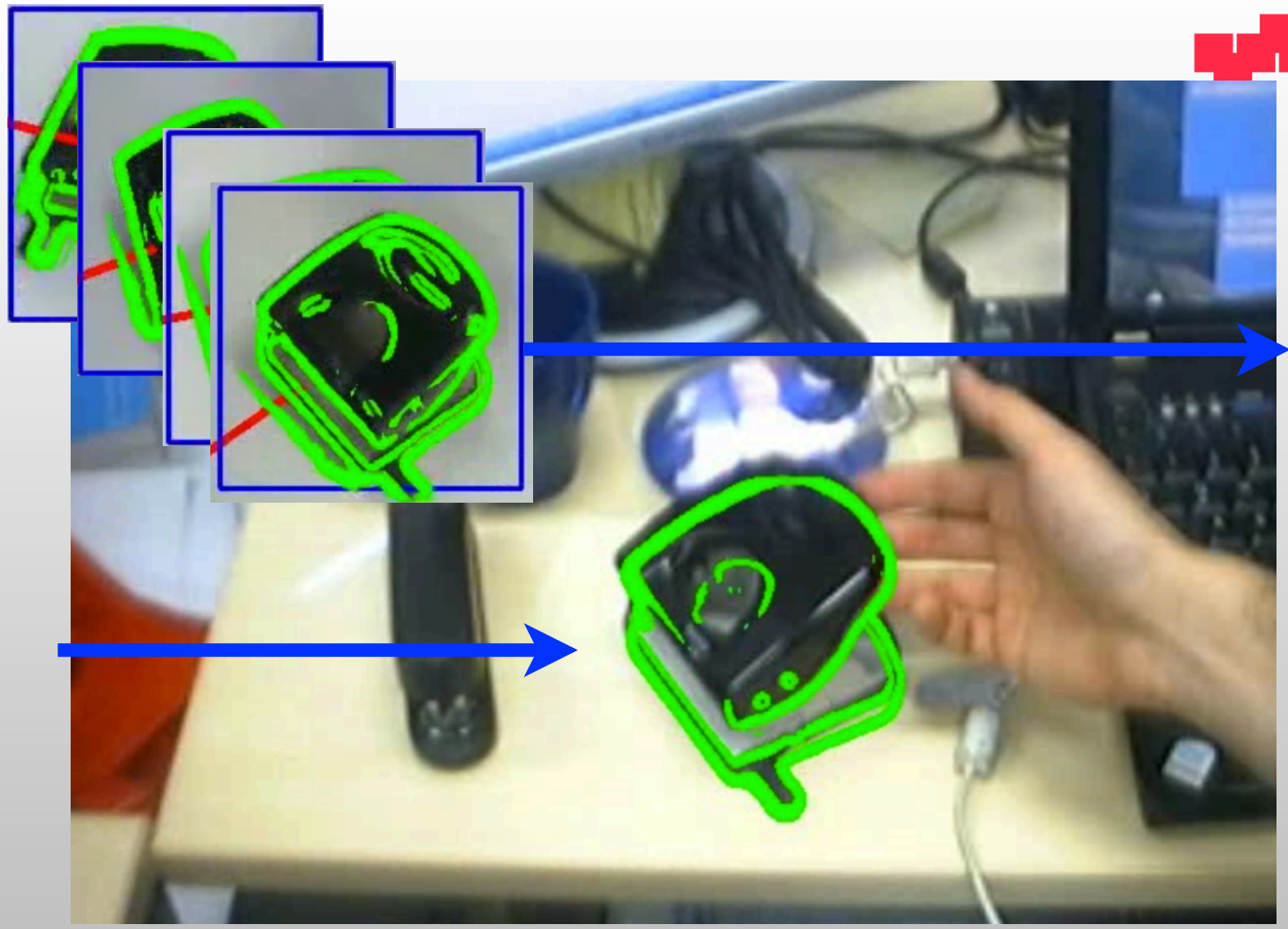
3D Pose Estimation of Objects

part-based approach

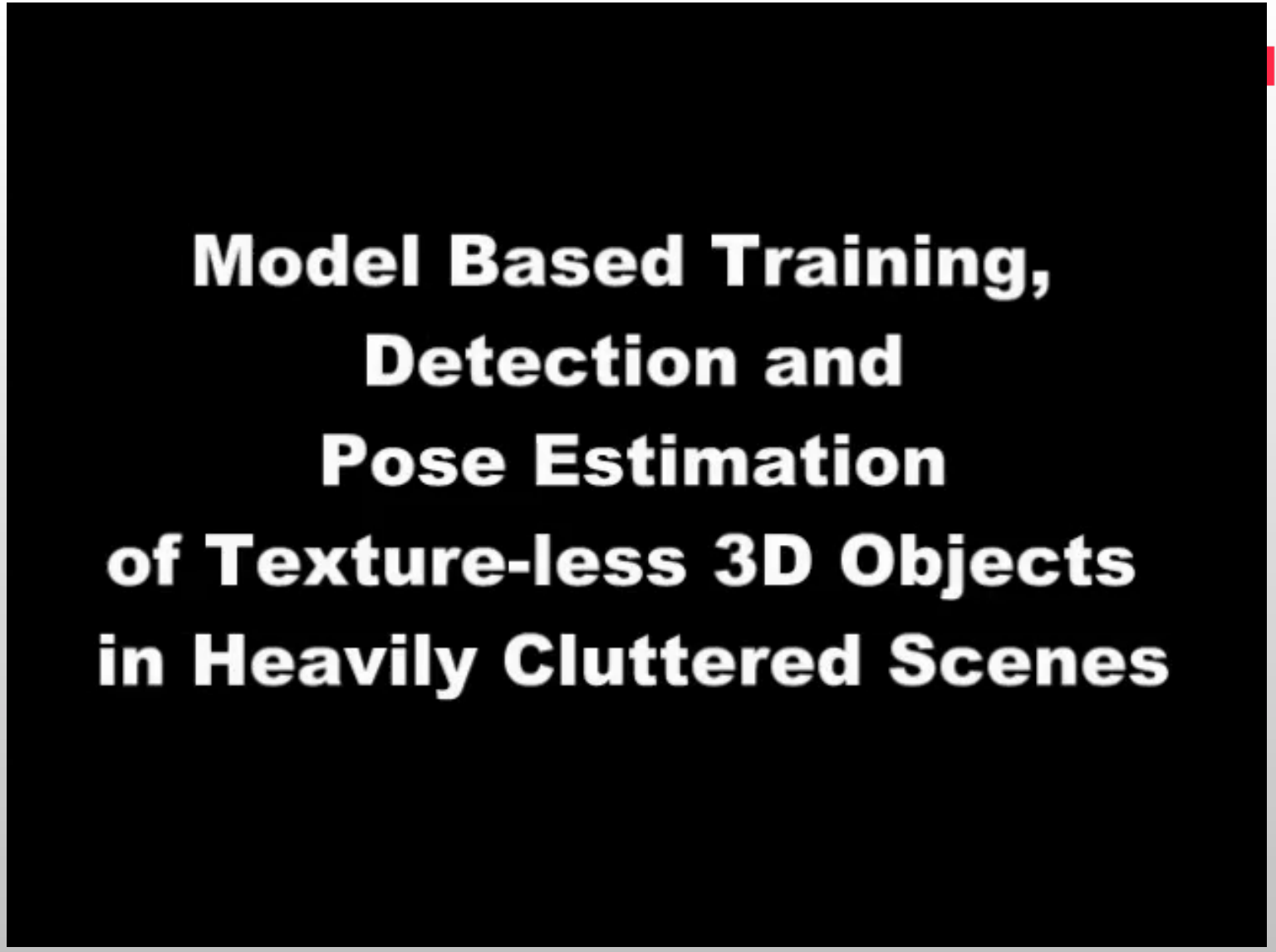


new optimization scheme

Many daily objects do not exhibit many feature points:



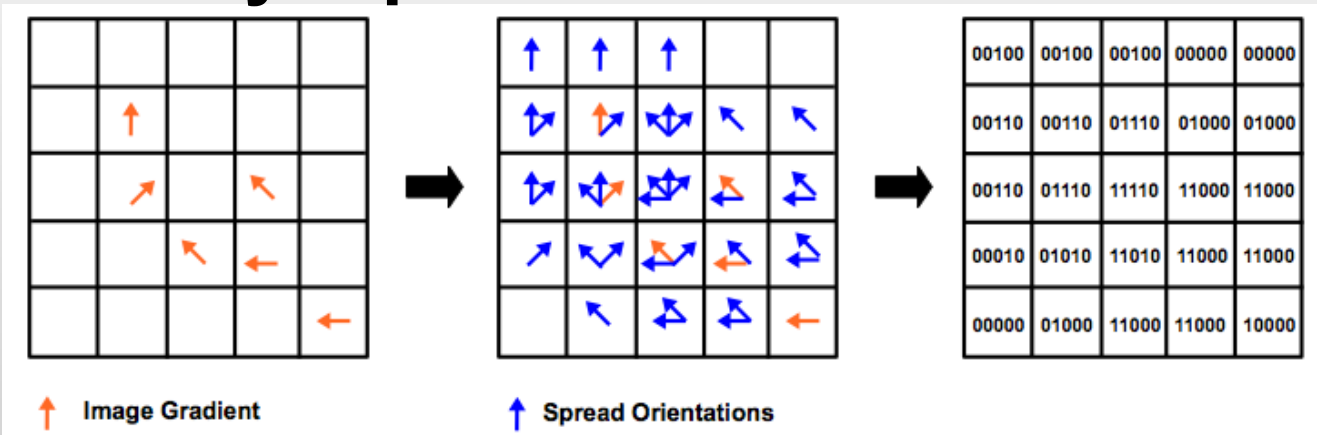
Template matching with an efficient representation of the images and the templates

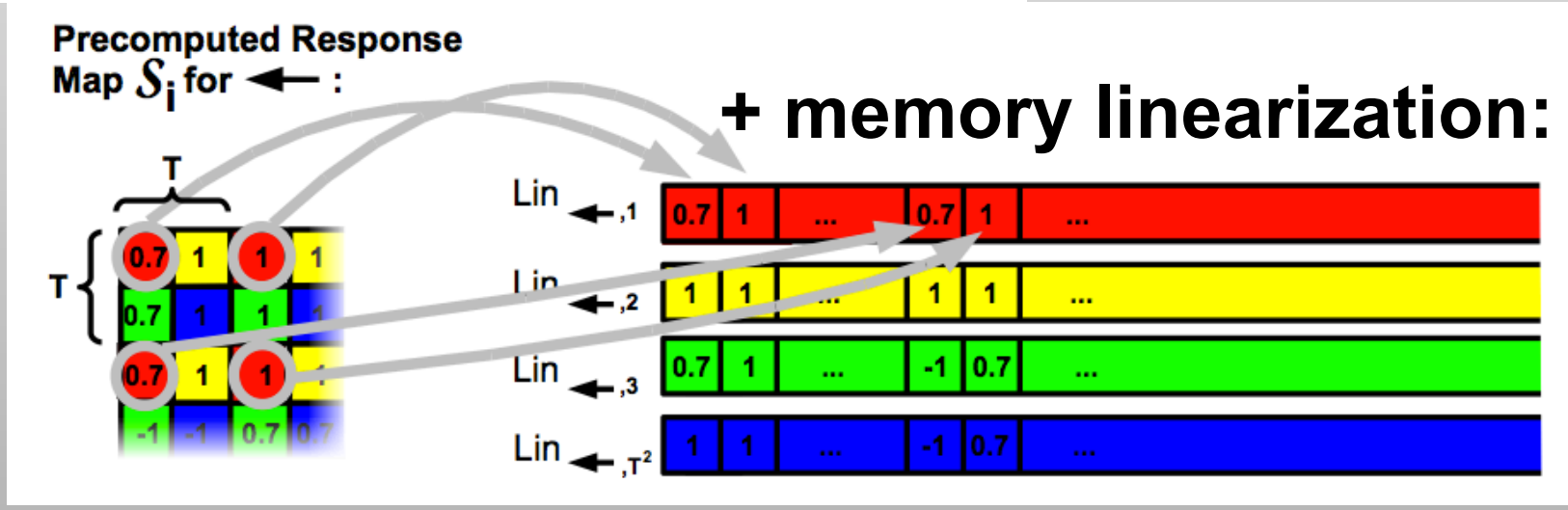


S. Hinterstoisser, C. Cagniart, S. Ilic, P. Sturm, N. Navab, P. Fua and V. Lepetit. "Gradient Response Maps for Real-Time Detection of Texture-Less Objects," PAMI 2011.

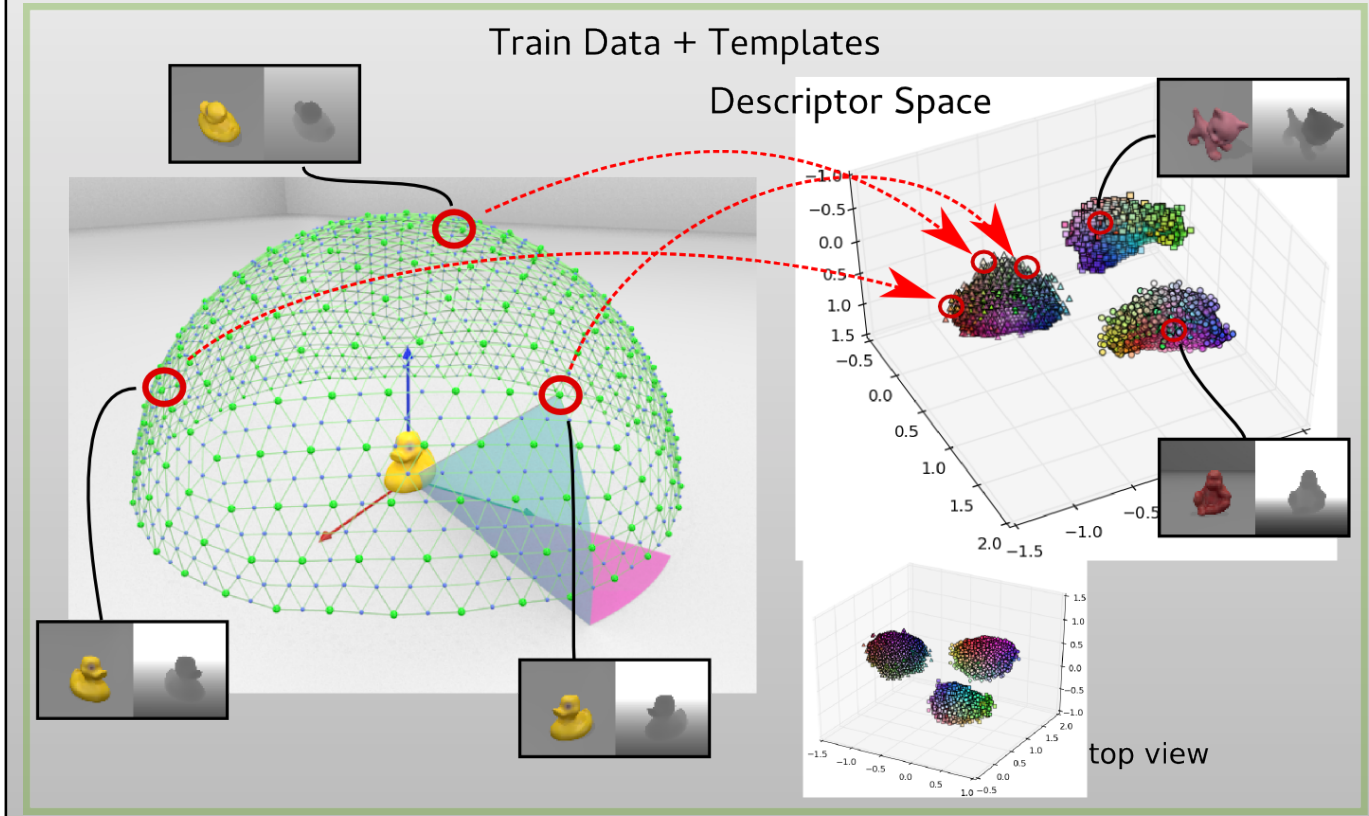
LineMOD: Handcrafted Representation

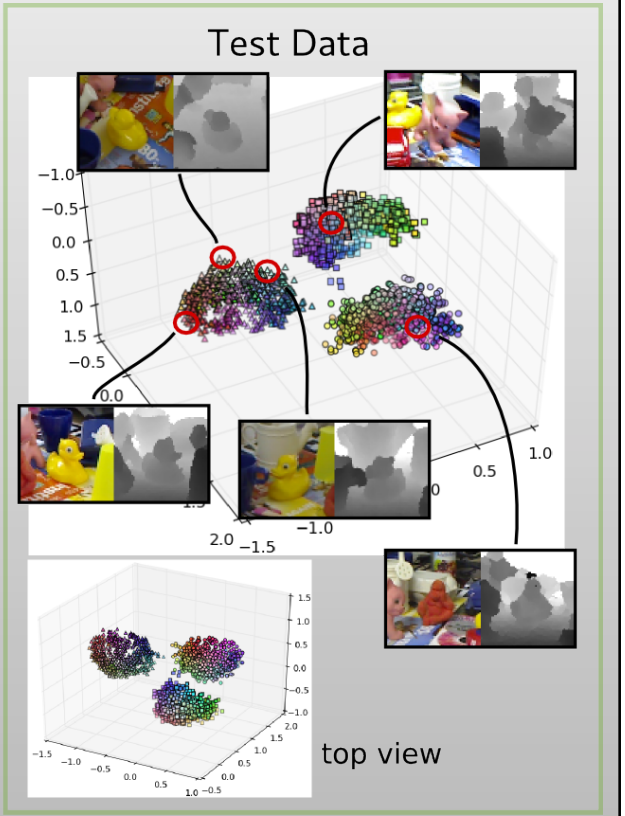
binary representation:



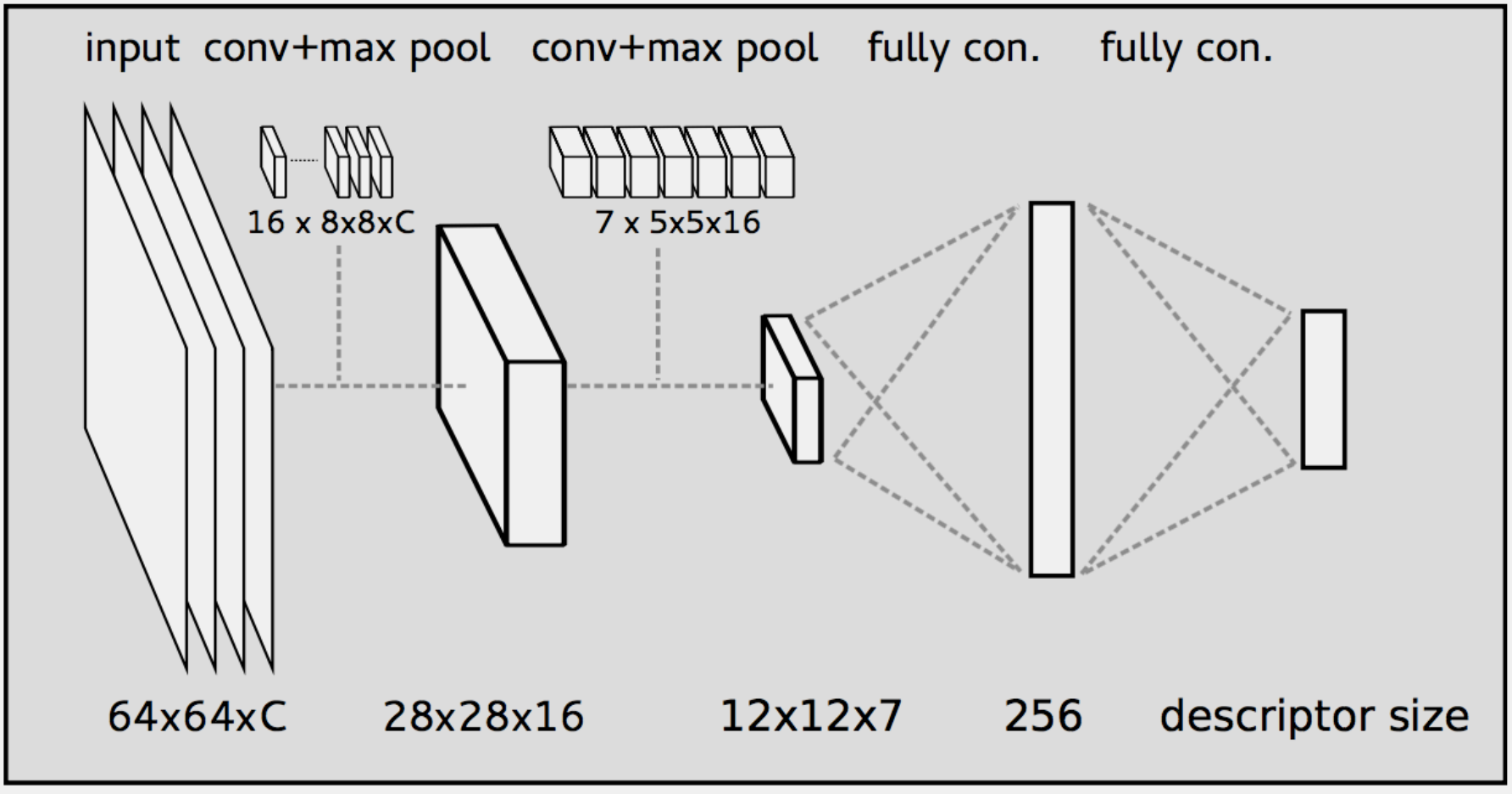


Learning Descriptors for Object Recognition and 3D Pose Estimation

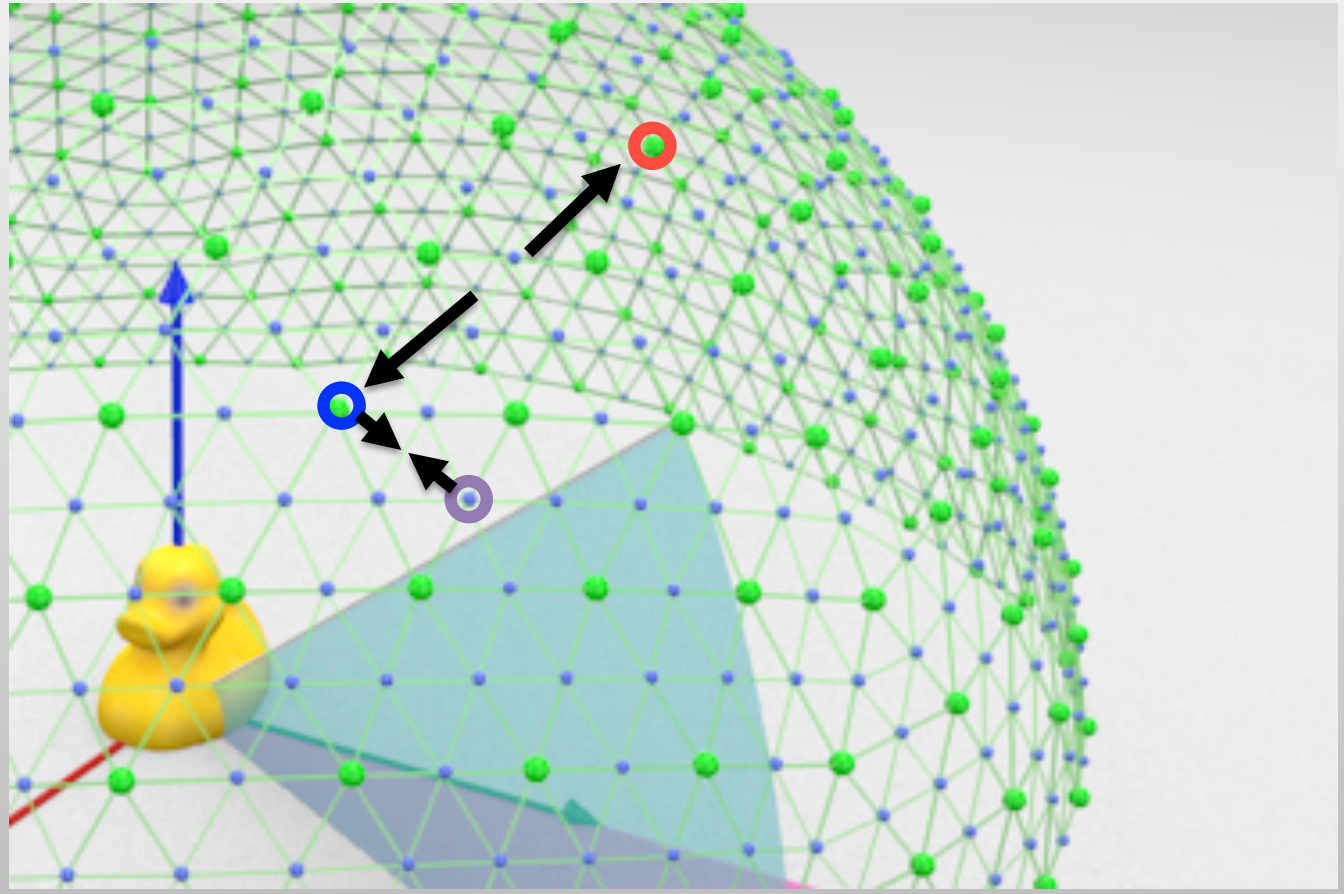




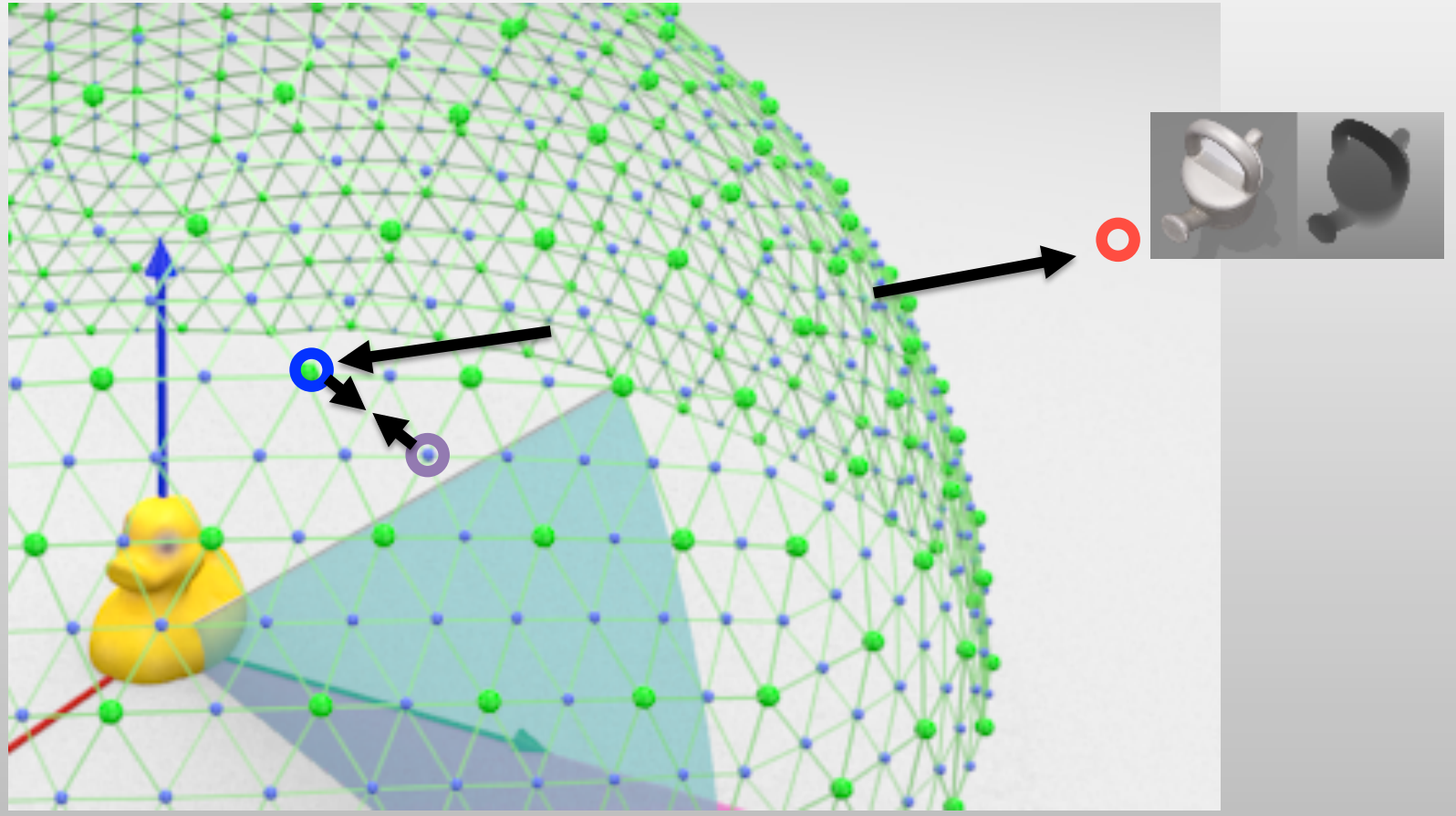
CNN Architecture



Triplet Constraints

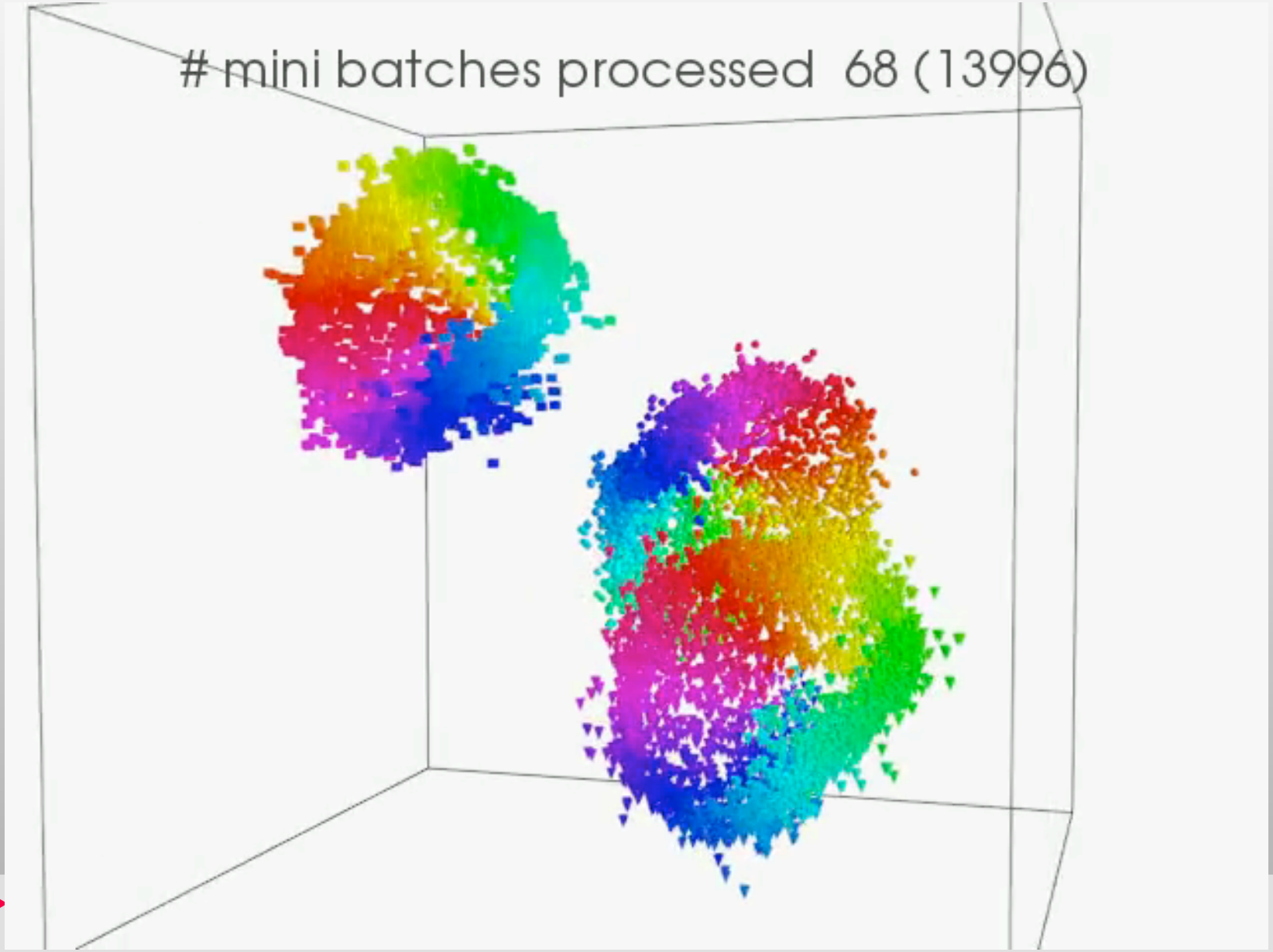


Triplet Constraints

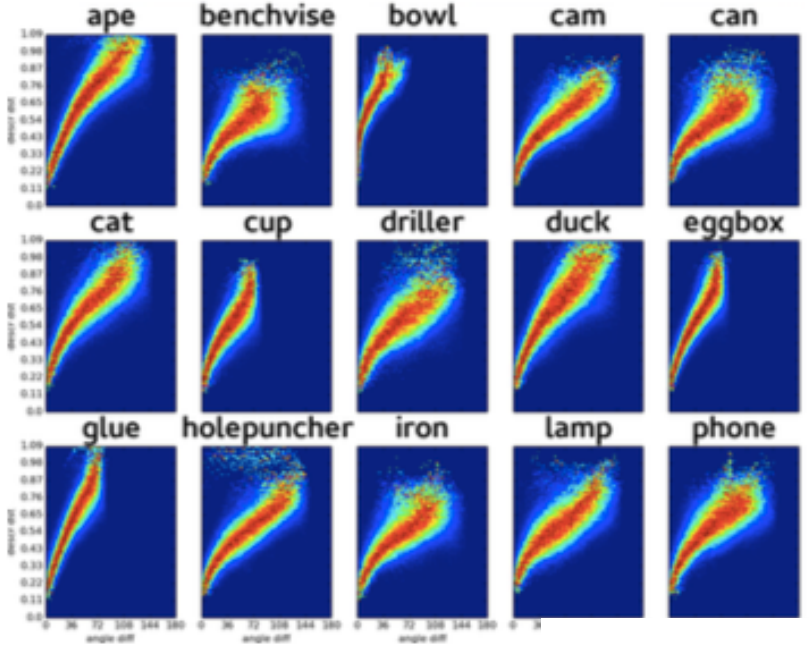


$$c(s_i, s_j, s_k) = \max\left(0, 1 - \frac{||f_w(x_i) - f_w(x_k)||_2}{||f_w(x_i) - f_w(x_j)||_2 + m}\right)$$

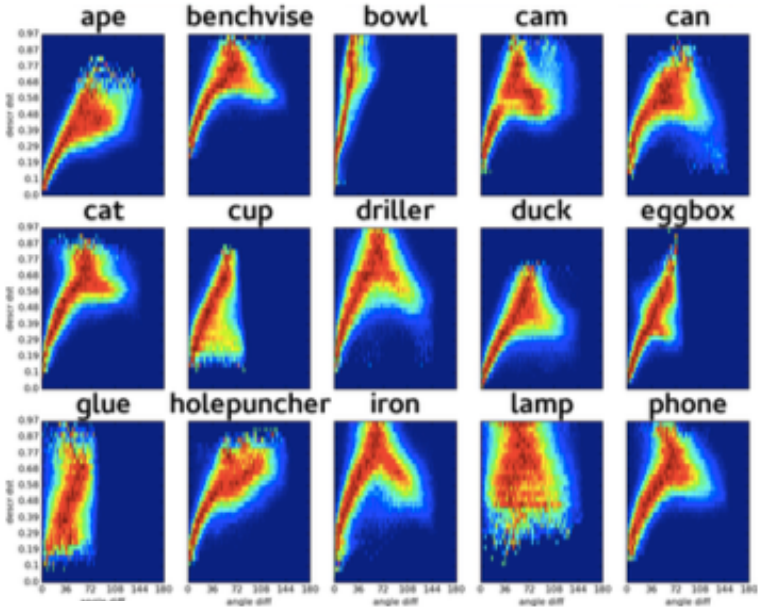
Toy data results: 3-descriptors for 3 objects



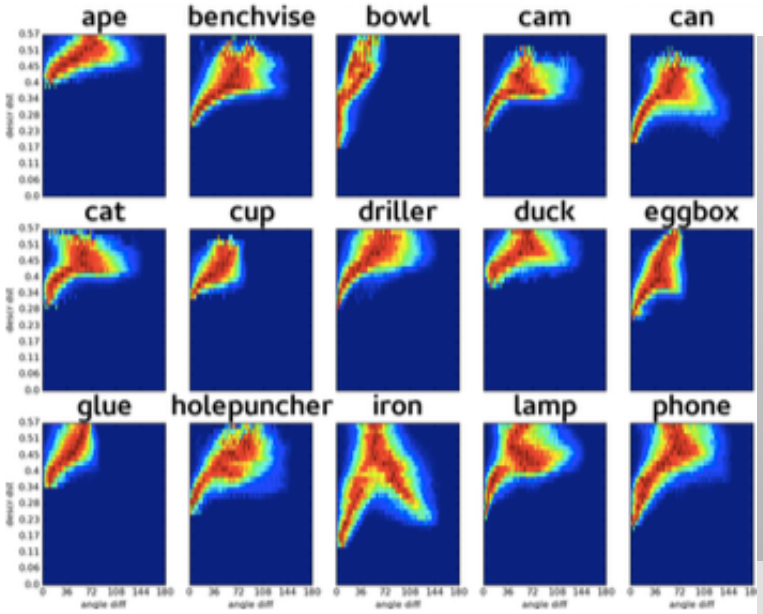
pose difference *vs* descriptor distance



(a) ours



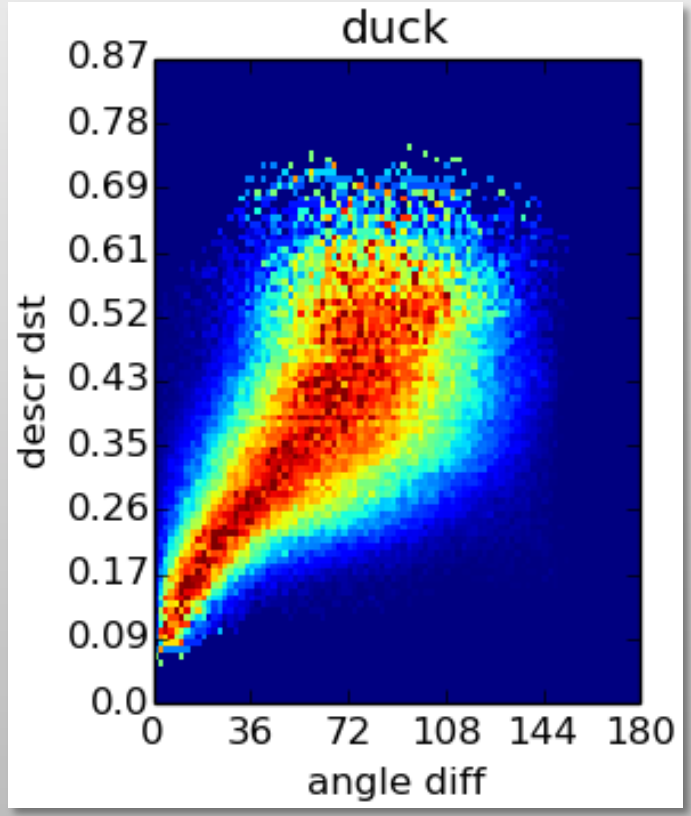
(b) LineMOD



(c) HOG

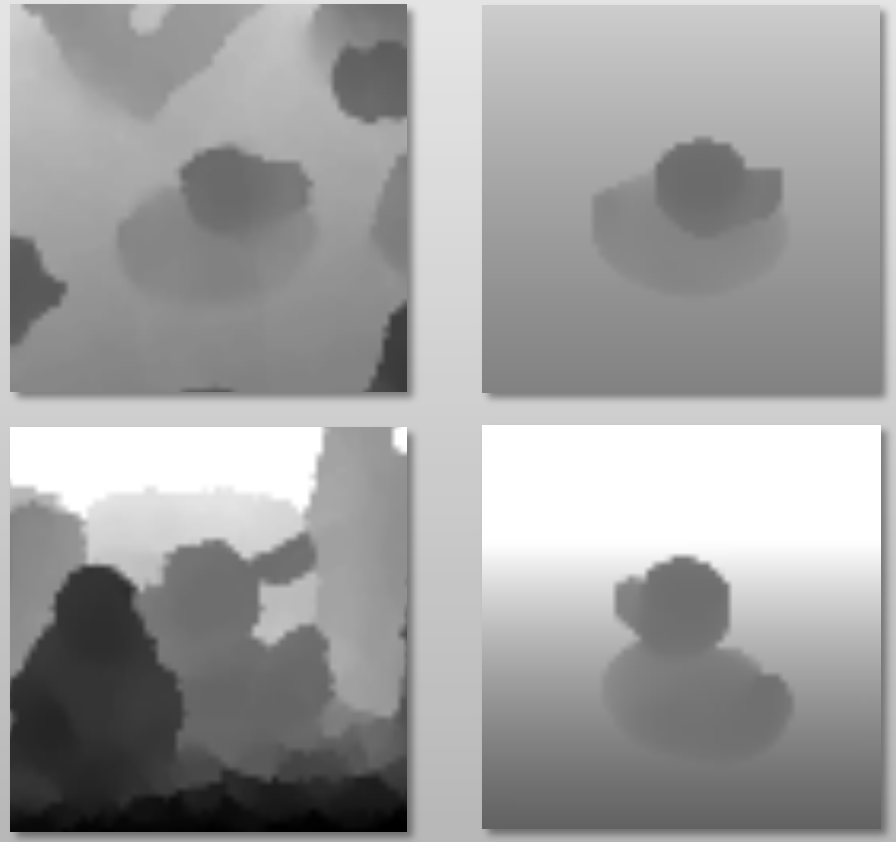
Generalization

Testing on “duck” not seen during training

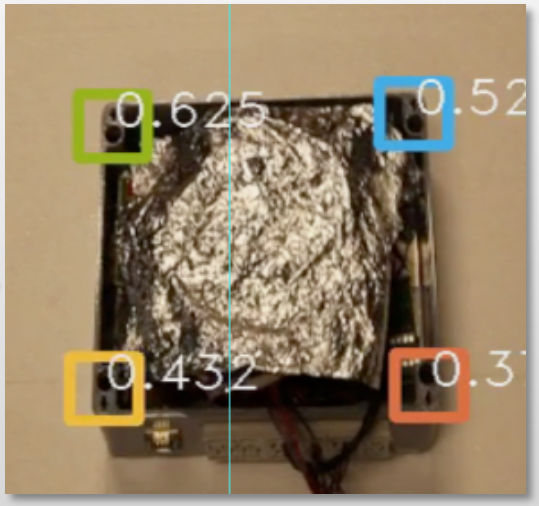


input image

closest template among those from 16 objects



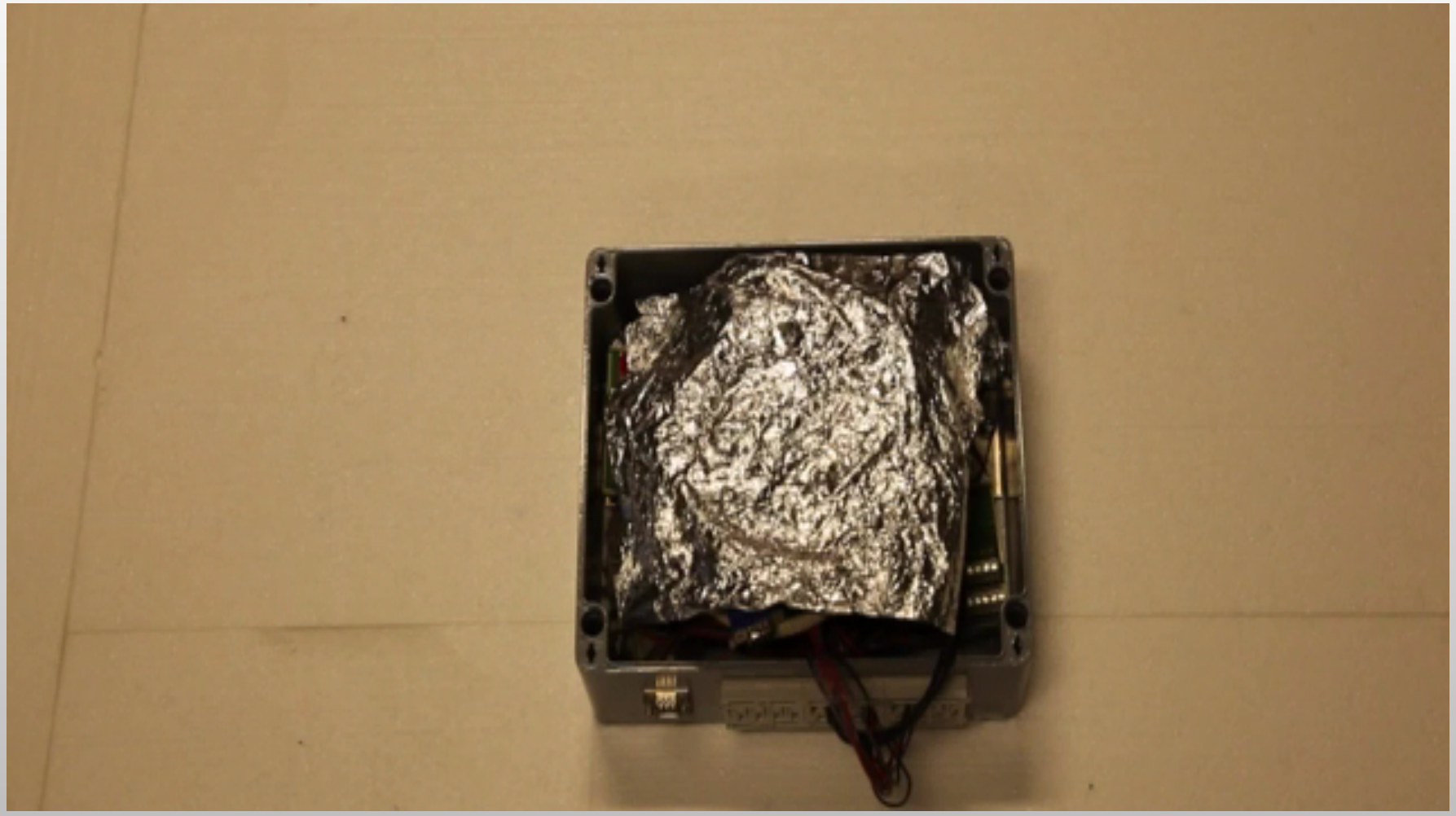
template-based approach



3D Pose Estimation
of Objects

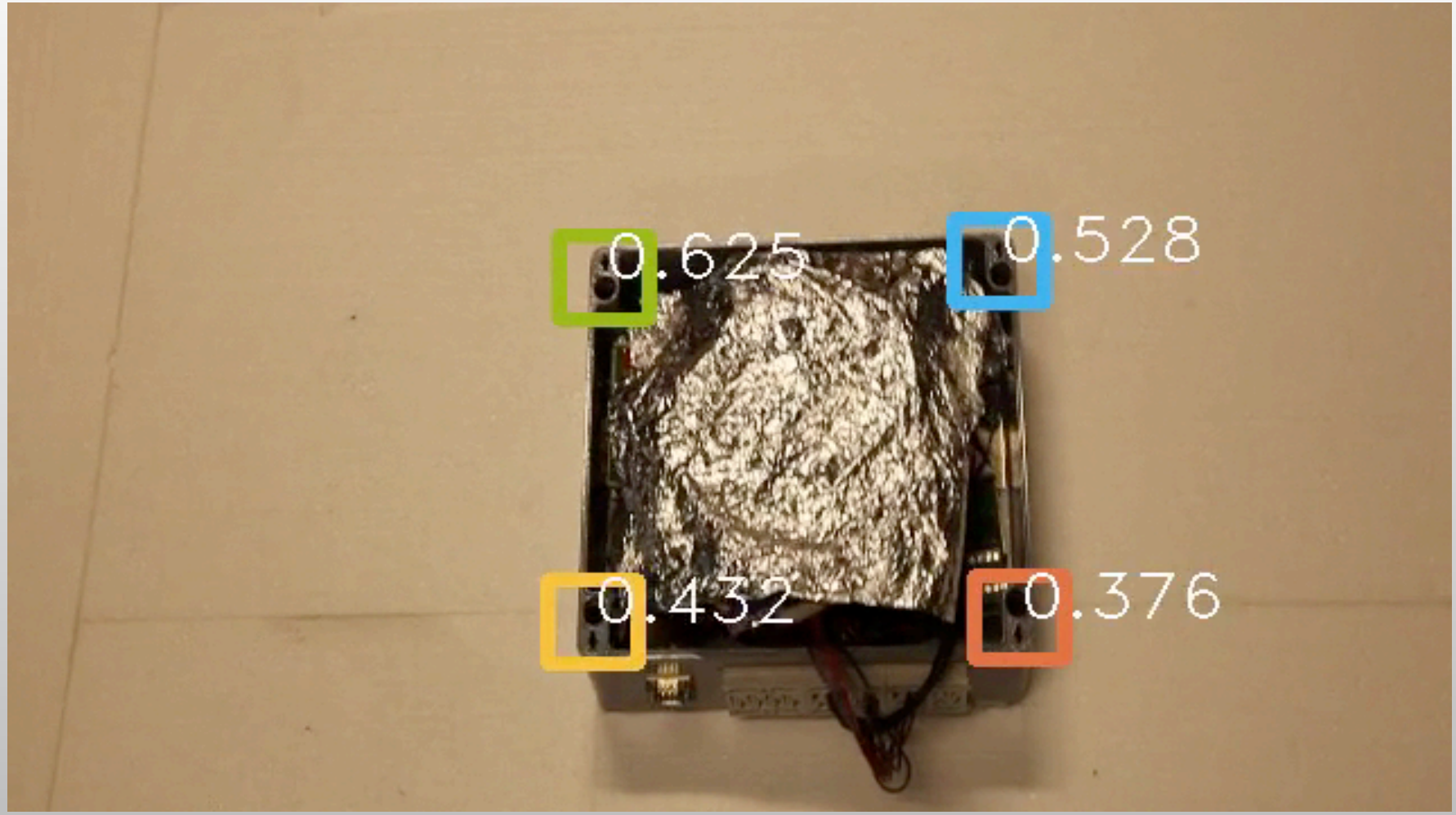
part-based approach

Alberto Crivellaro, Mahdi Rad, Yannick Verdie, Kwang Moo Yi, Pascal Fua, and Vincent Lepetit.
A Novel Representation of Parts for Accurate 3D Object Detection and Tracking in Monocular Images.
ICCV'15.

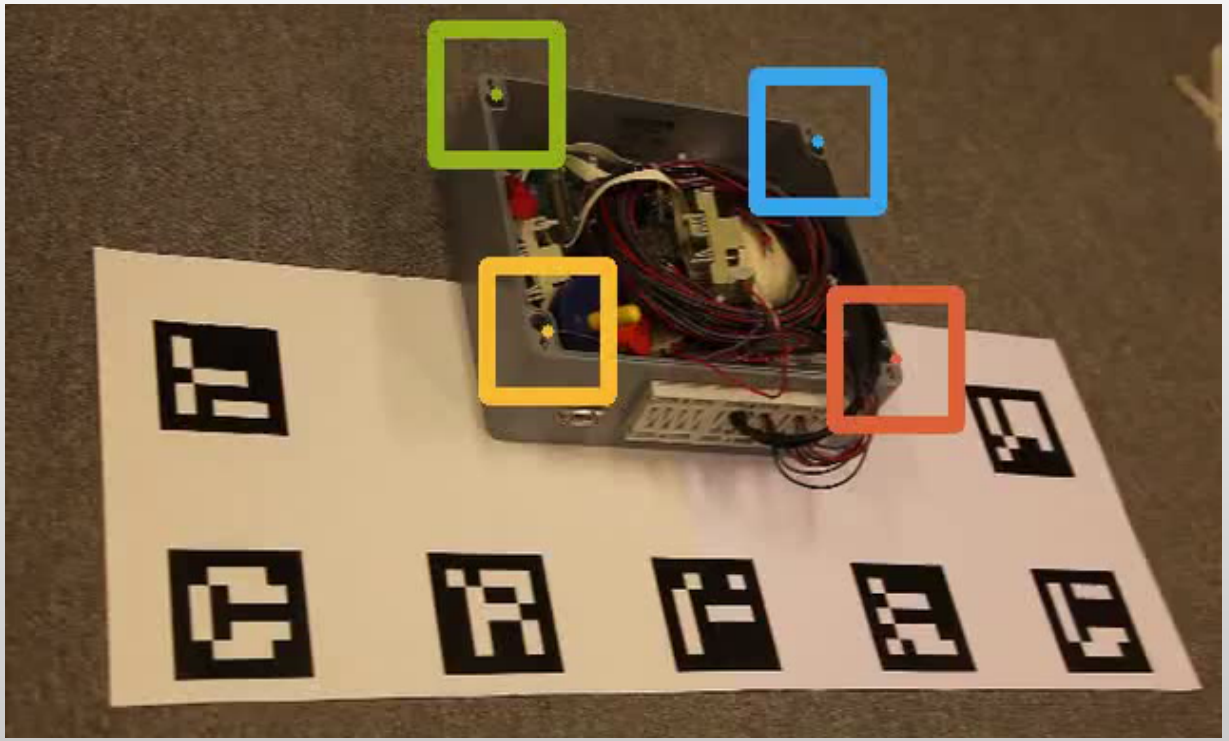


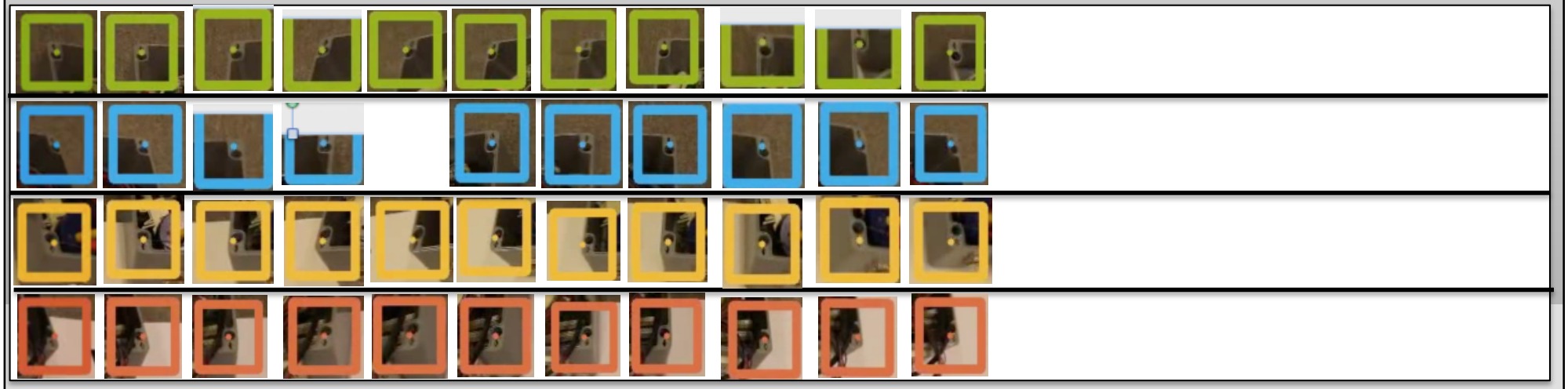
without texture, without a depth camera, under occlusions, etc.

Learning to Detect Parts of the Object

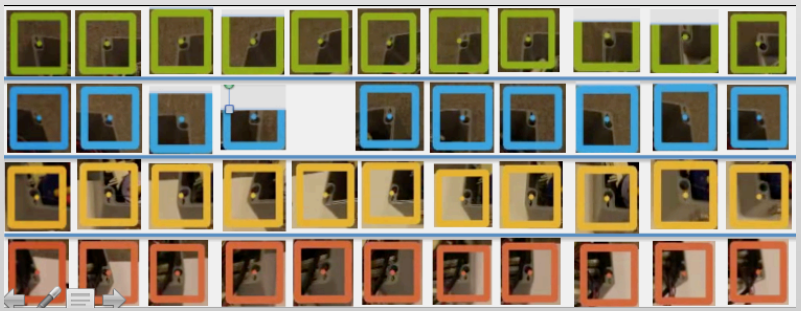


Learning to Detect Parts of the Object
Collecting Training Data





Learning to Detect Parts of the Object
Training

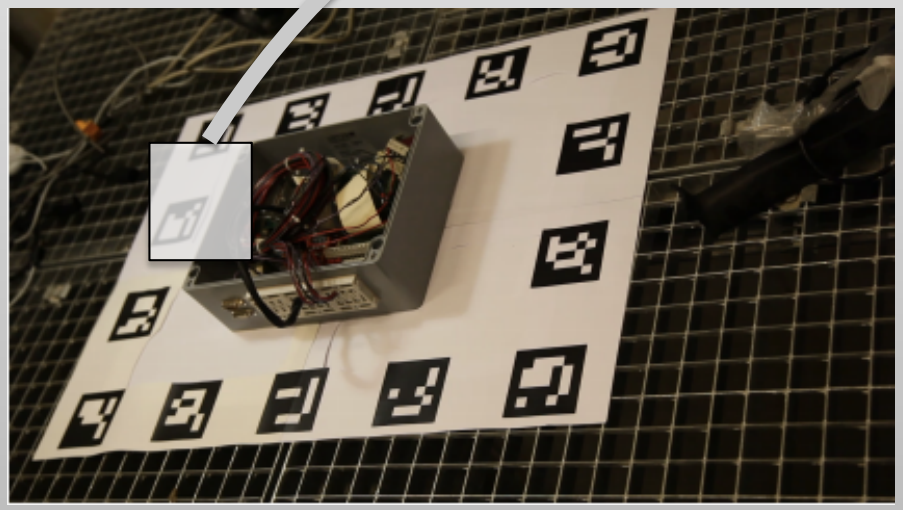


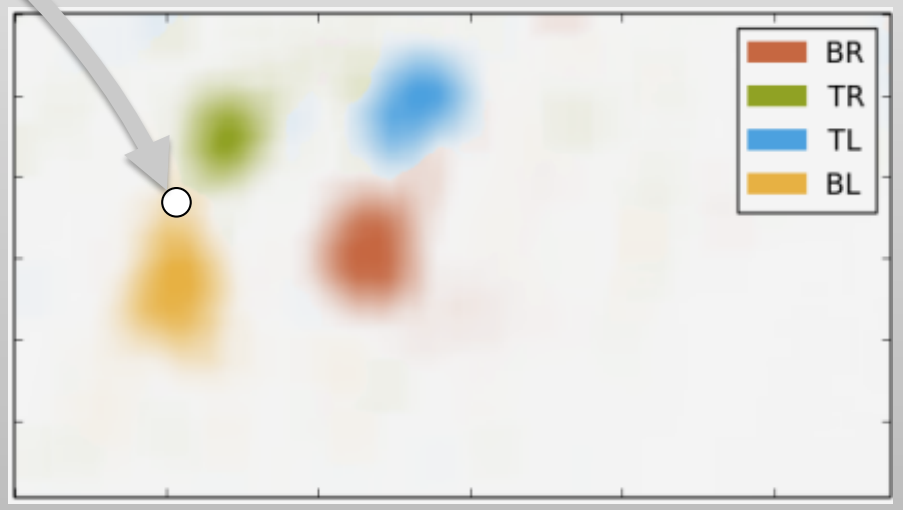
training

Convolutional Neural Network

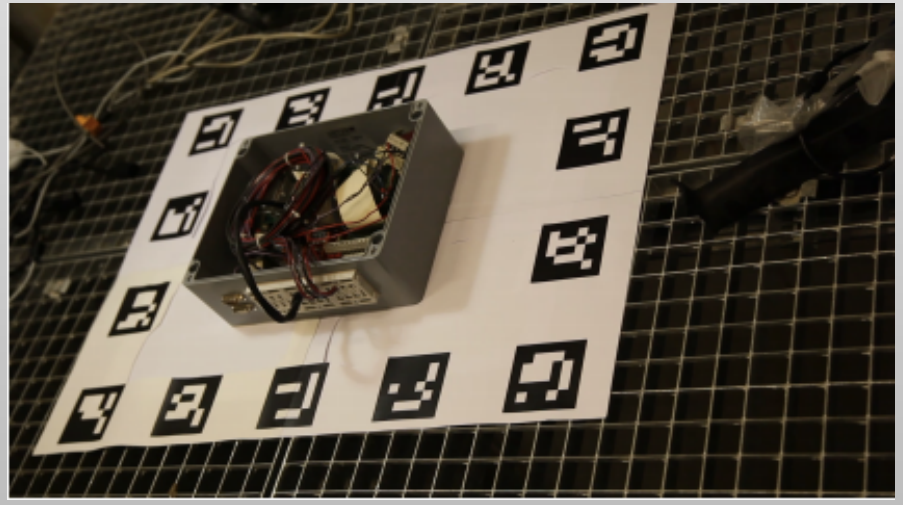
Learning to Detect Parts of the Object Testing

Convolutional Neural Network

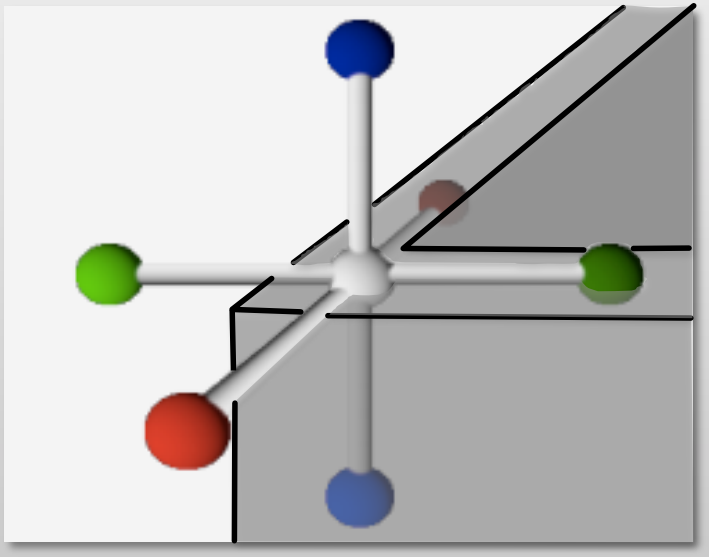




Learning to Detect Parts of the Object Testing

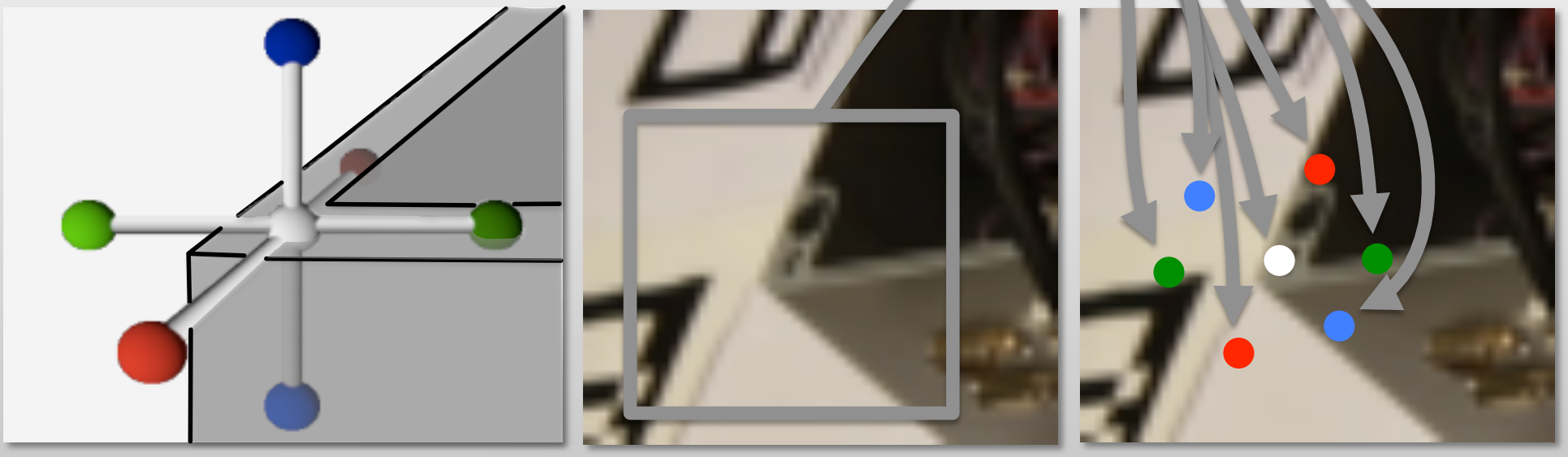


Can We Get More from a Part?



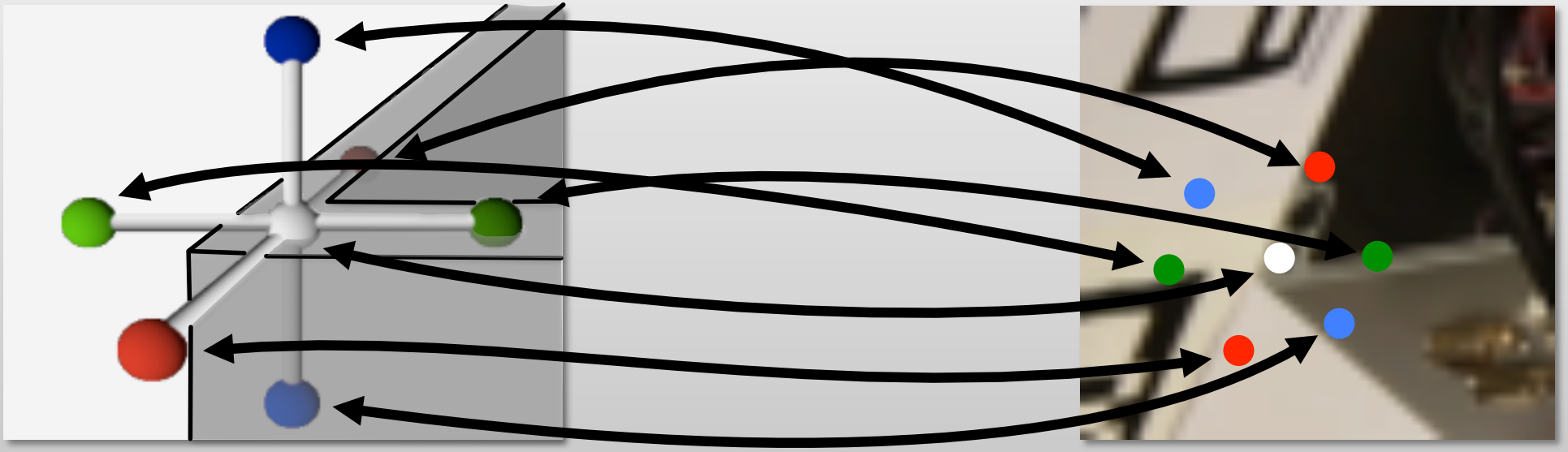
idea: representing the 3D pose of the part with (virtual) 3D control points

Convolutional Neural Network 2



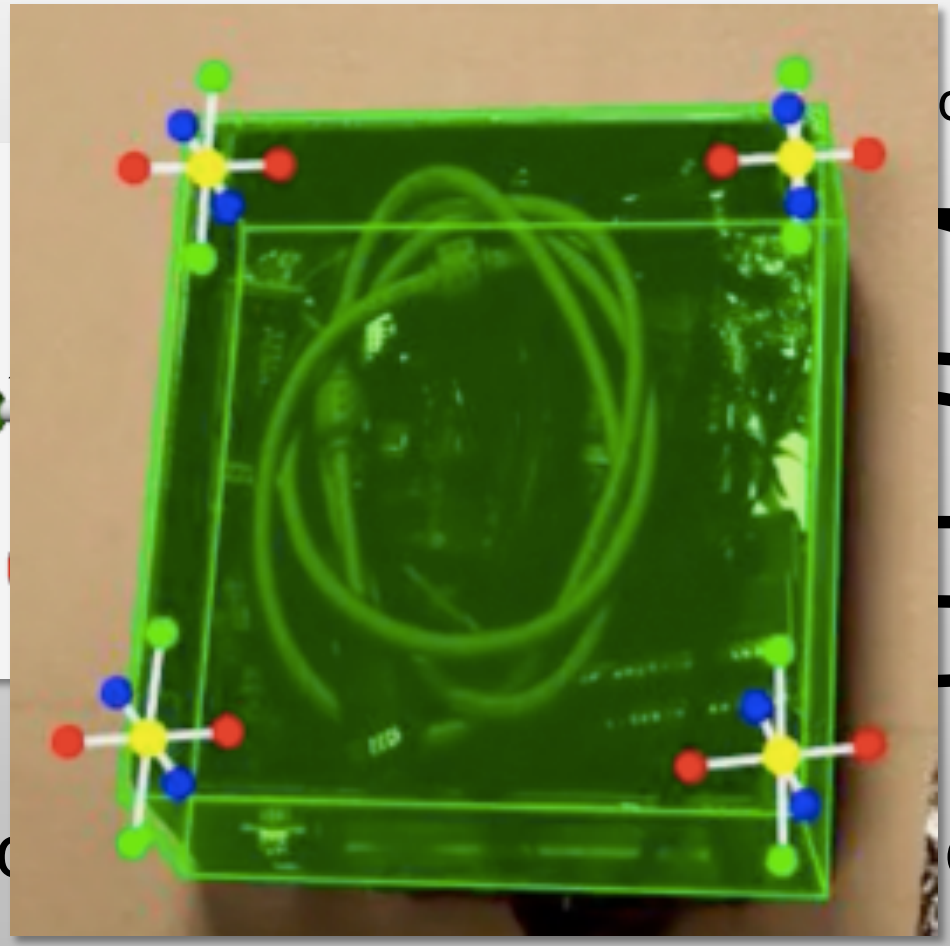
another CNN is trained to predict the *2D reprojections* of the control points

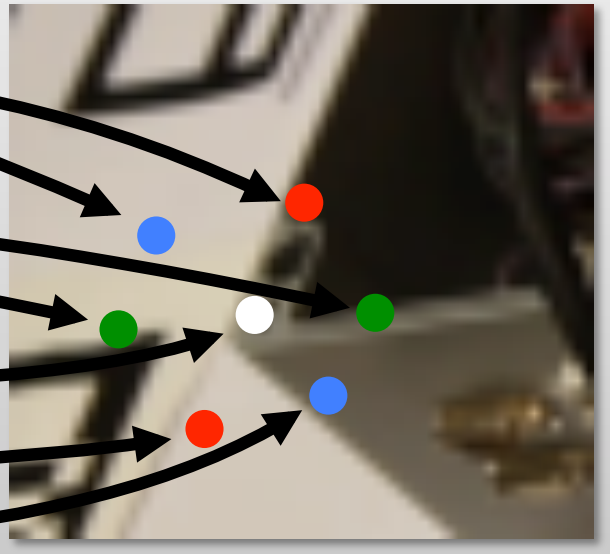
3D-2D correspondences



we can use a PnP algorithm to compute the pose!

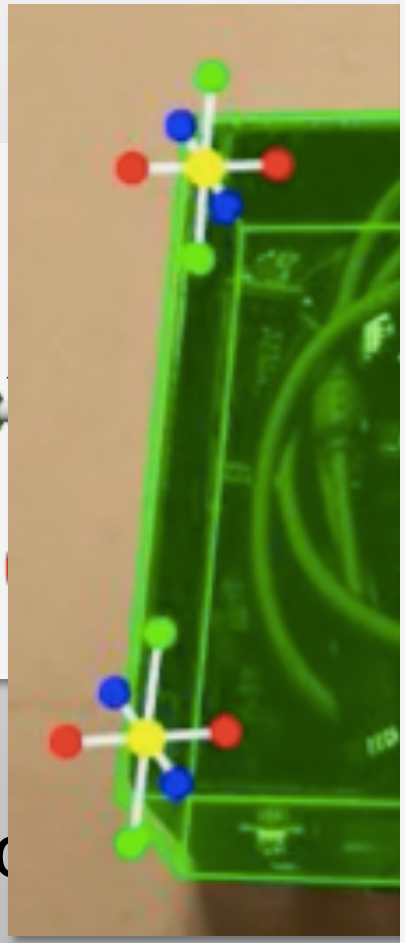
dences

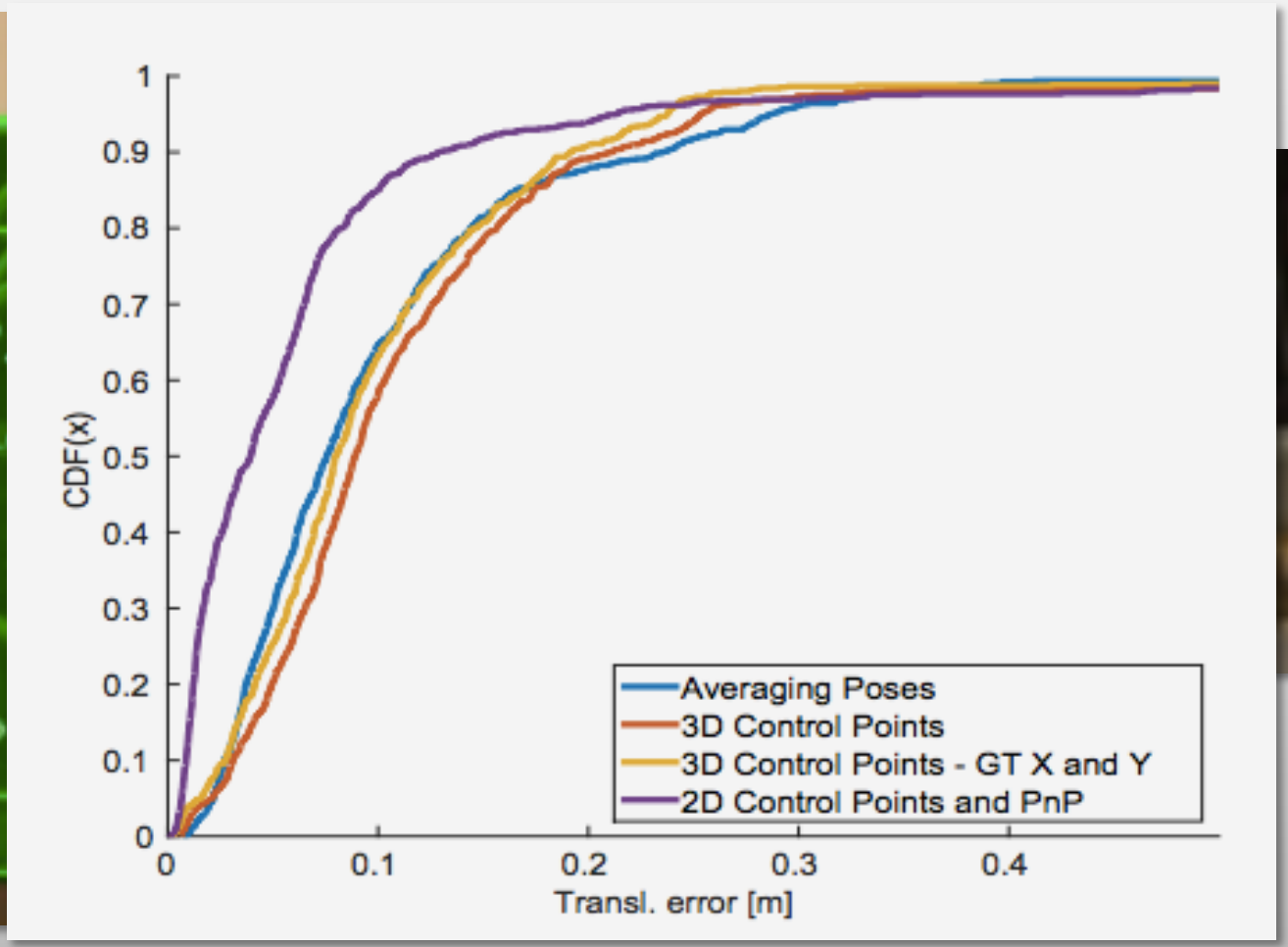


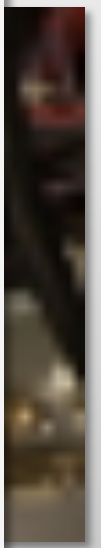


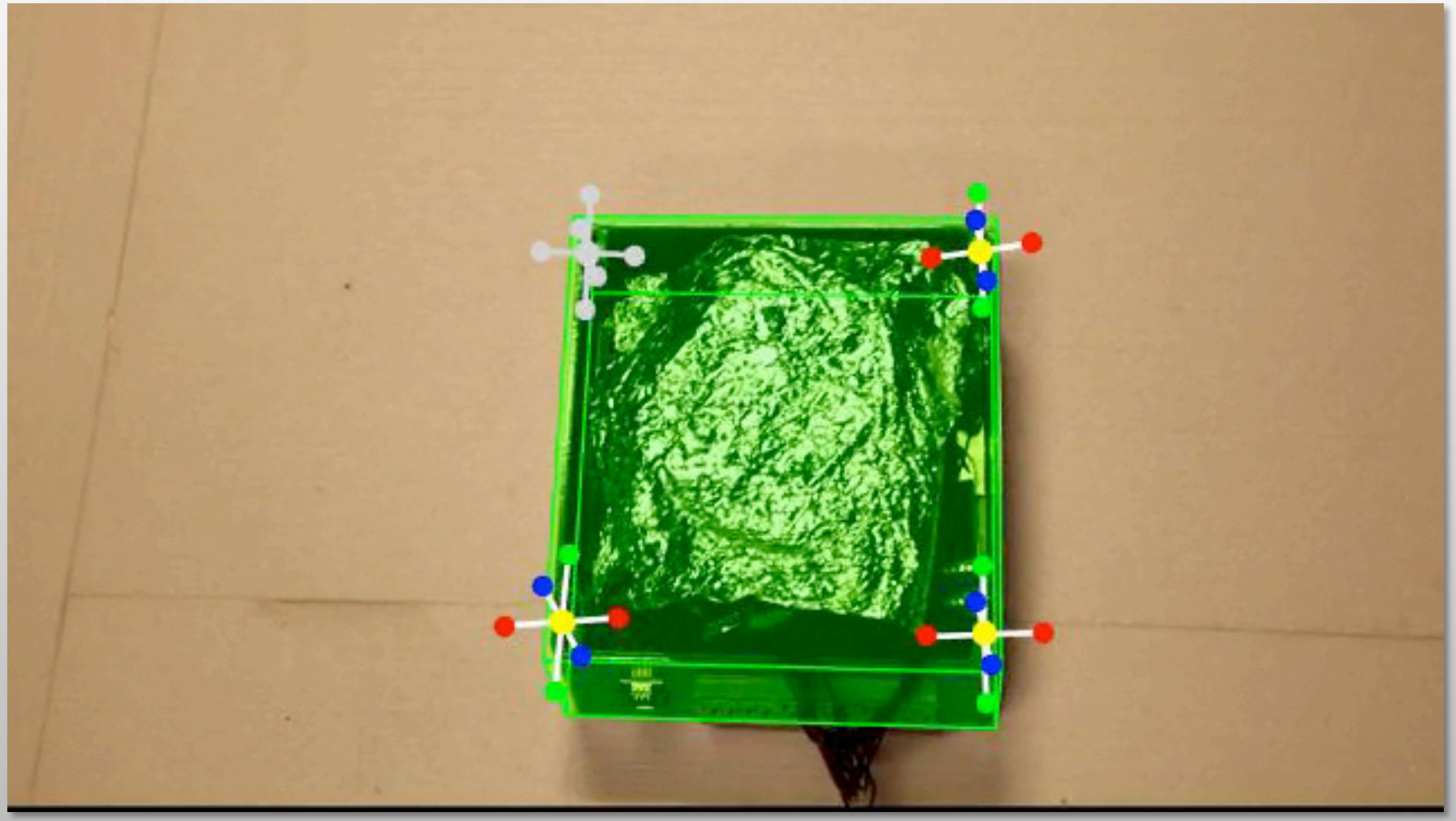
we c compute the pose!

we c




CDF(x)
Transl. error [m]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.1
0.2
0.3
0.4
Averaging Poses
3D Control Points
3D Control Points - GT X and Y
2D Control Points and PnP

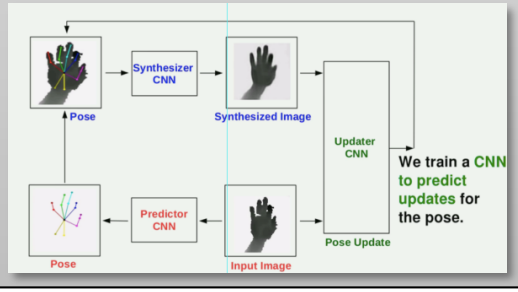




template-based approach

3D Pose Estimation
of Objects

part-based approach

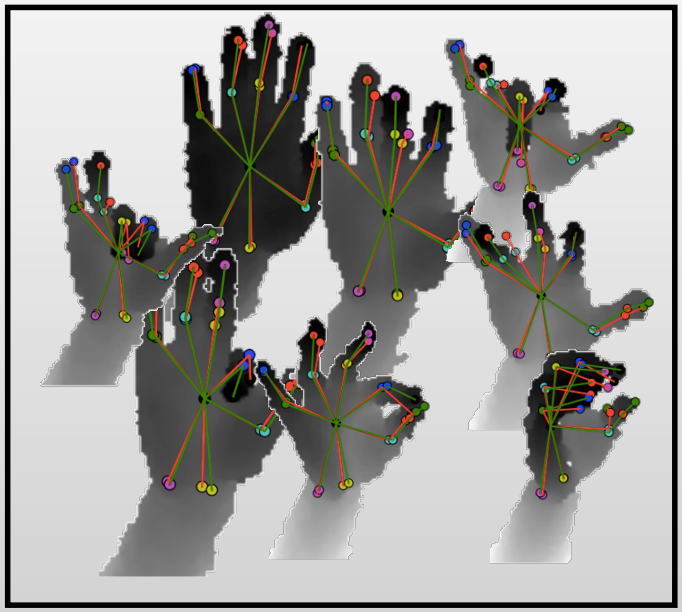


new optimization scheme

Markus Oberweger, Paul Wohlhart, and Vincent Lepetit.
Training a Feedback Loop for Hand Pose Estimation.
ICCV'15.

A First Network

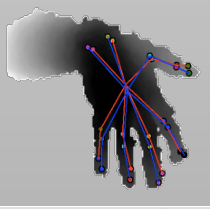
Training Set



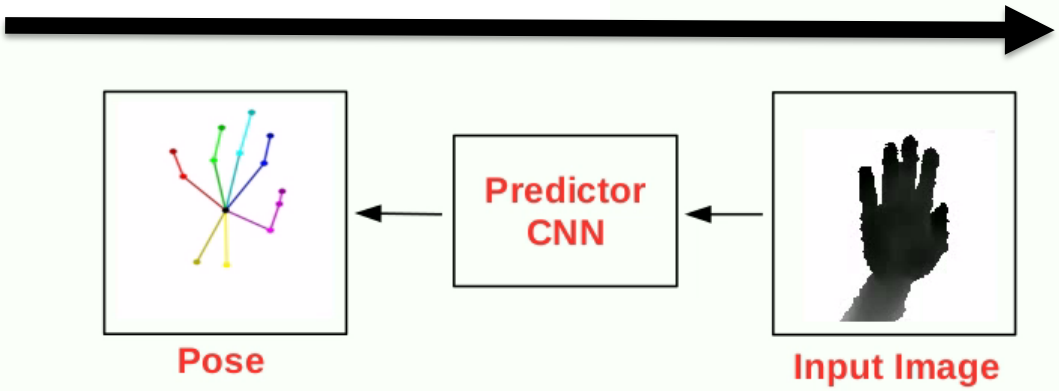
Input: A depth image

Convolutional Neural Network

Output: A **vector** made of the 3D coordinates of the joints

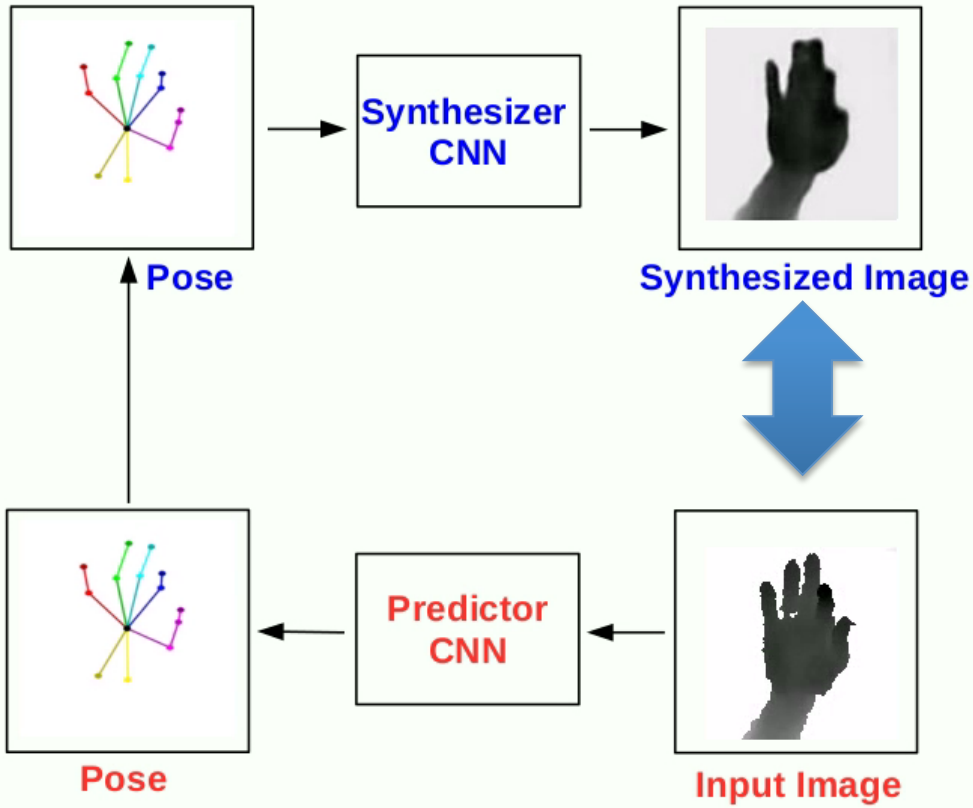


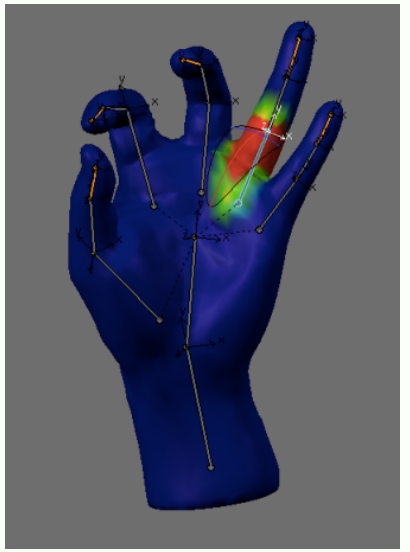
many neurons in the visual cortex provide *feedback*



Feed-forward network

Starting a feedback loop: We train a CNN to synthesize a depth image from the pose.

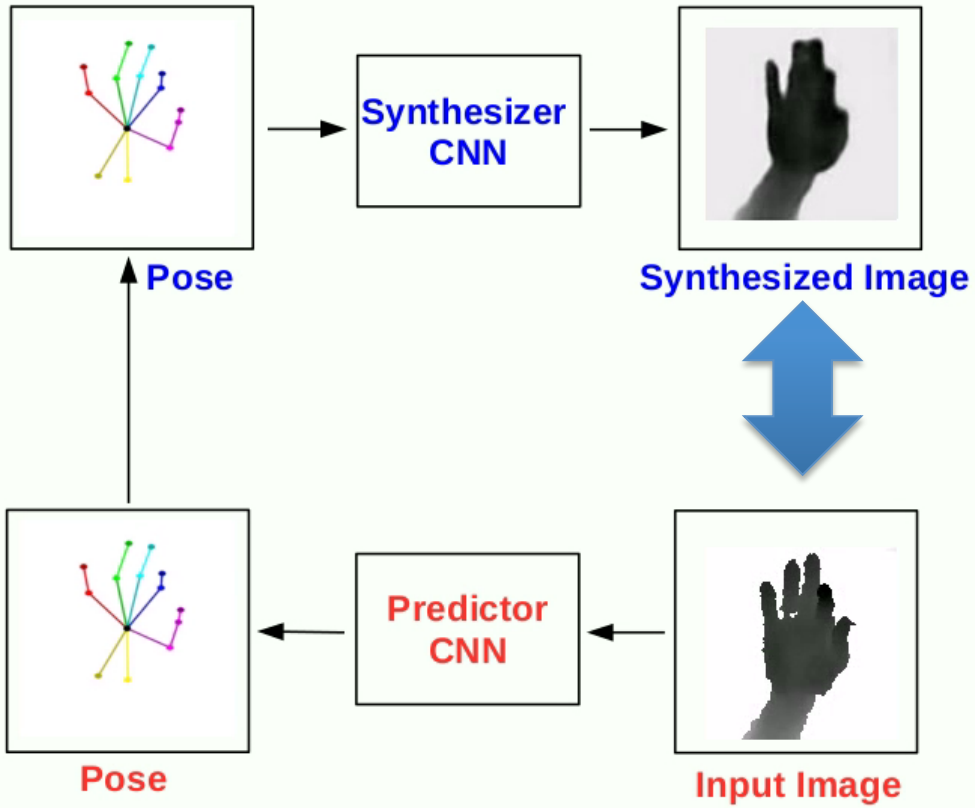




[de la Gorce et al.]

No need for a 3D model

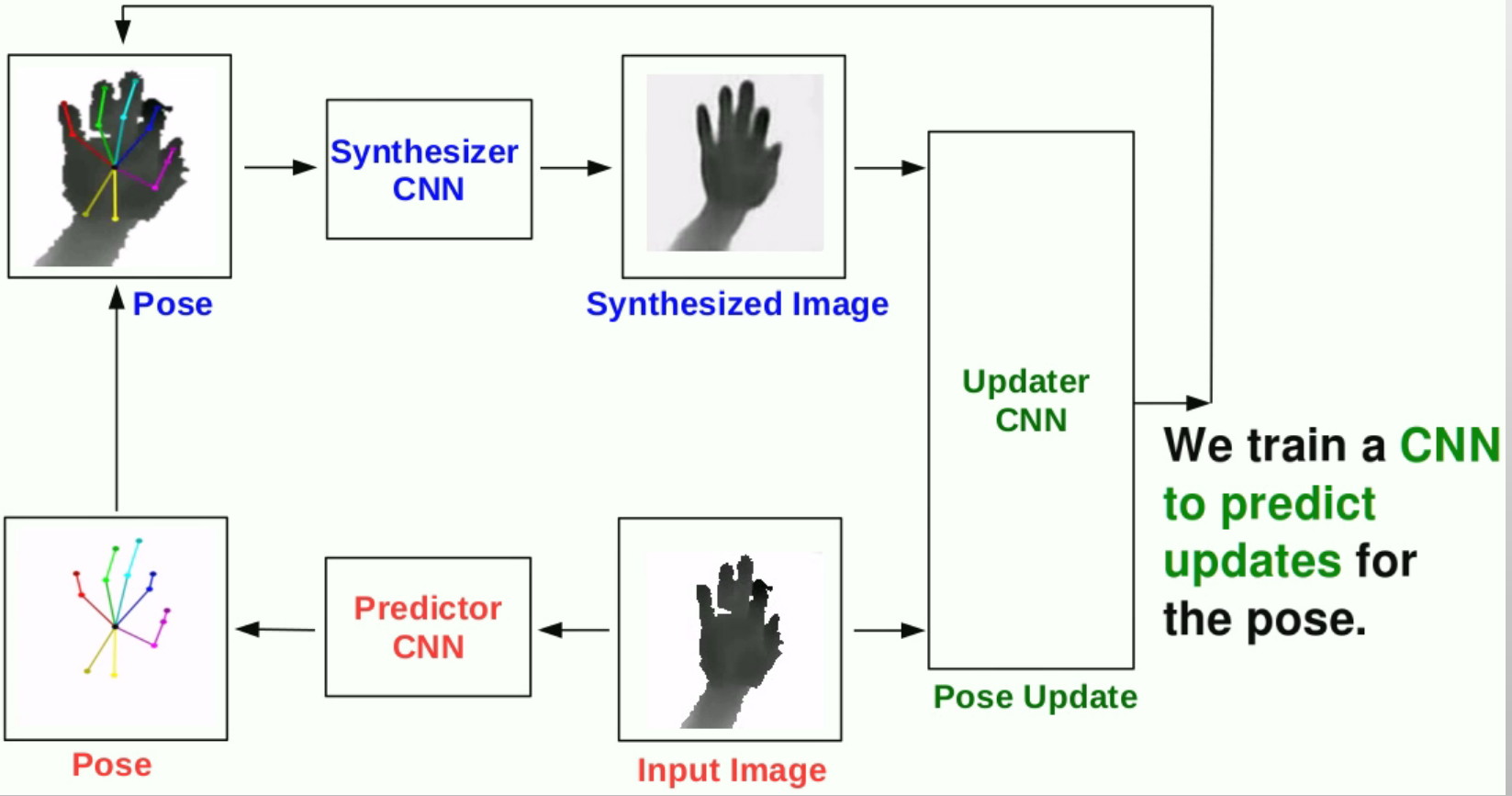
Starting a feedback loop: We train a CNN to synthesize a depth image from the pose.



Minimizing the difference between these two images to estimate the pose?

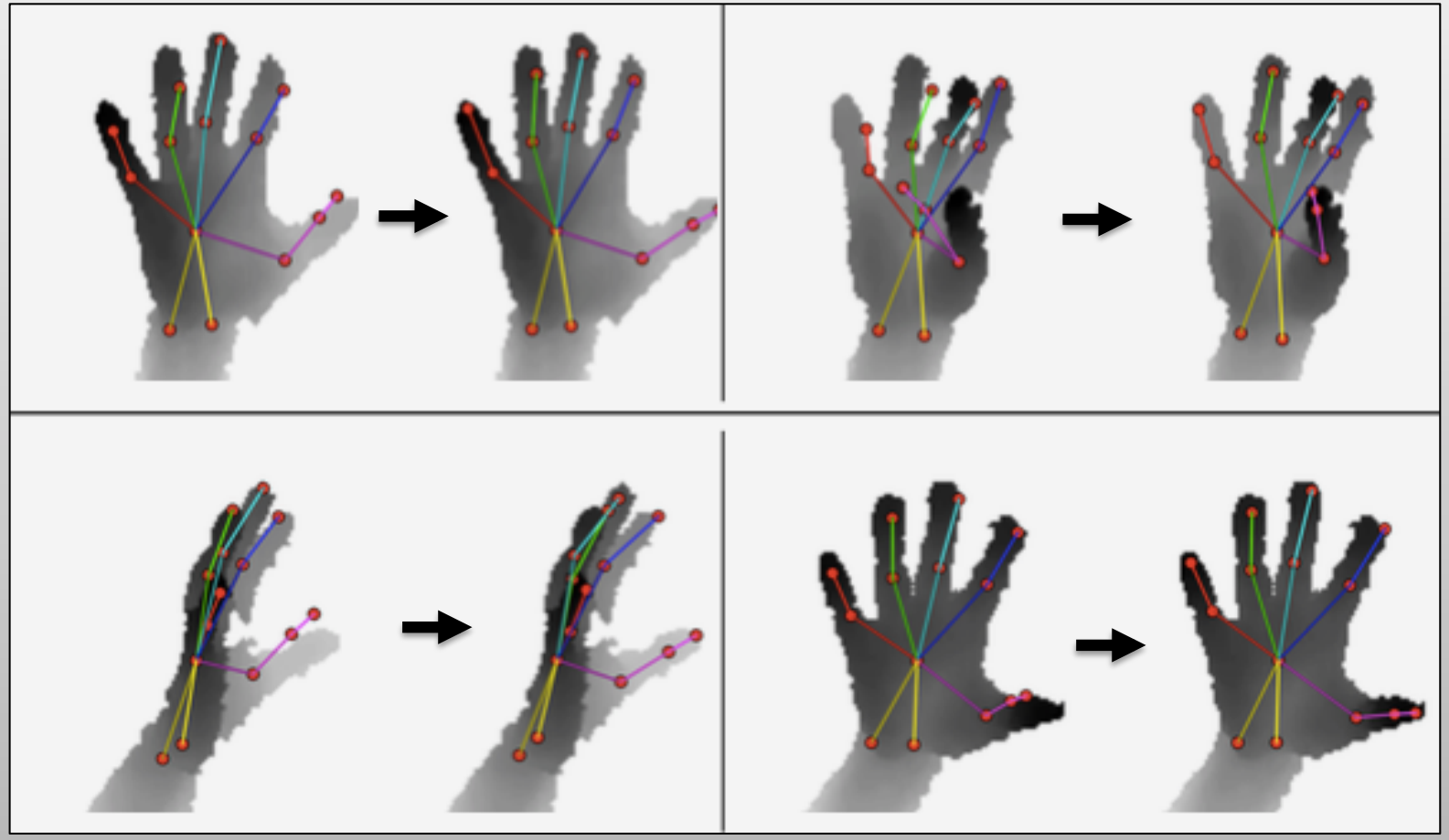
does not work well in practice

Final results when iterating the feedback loop.

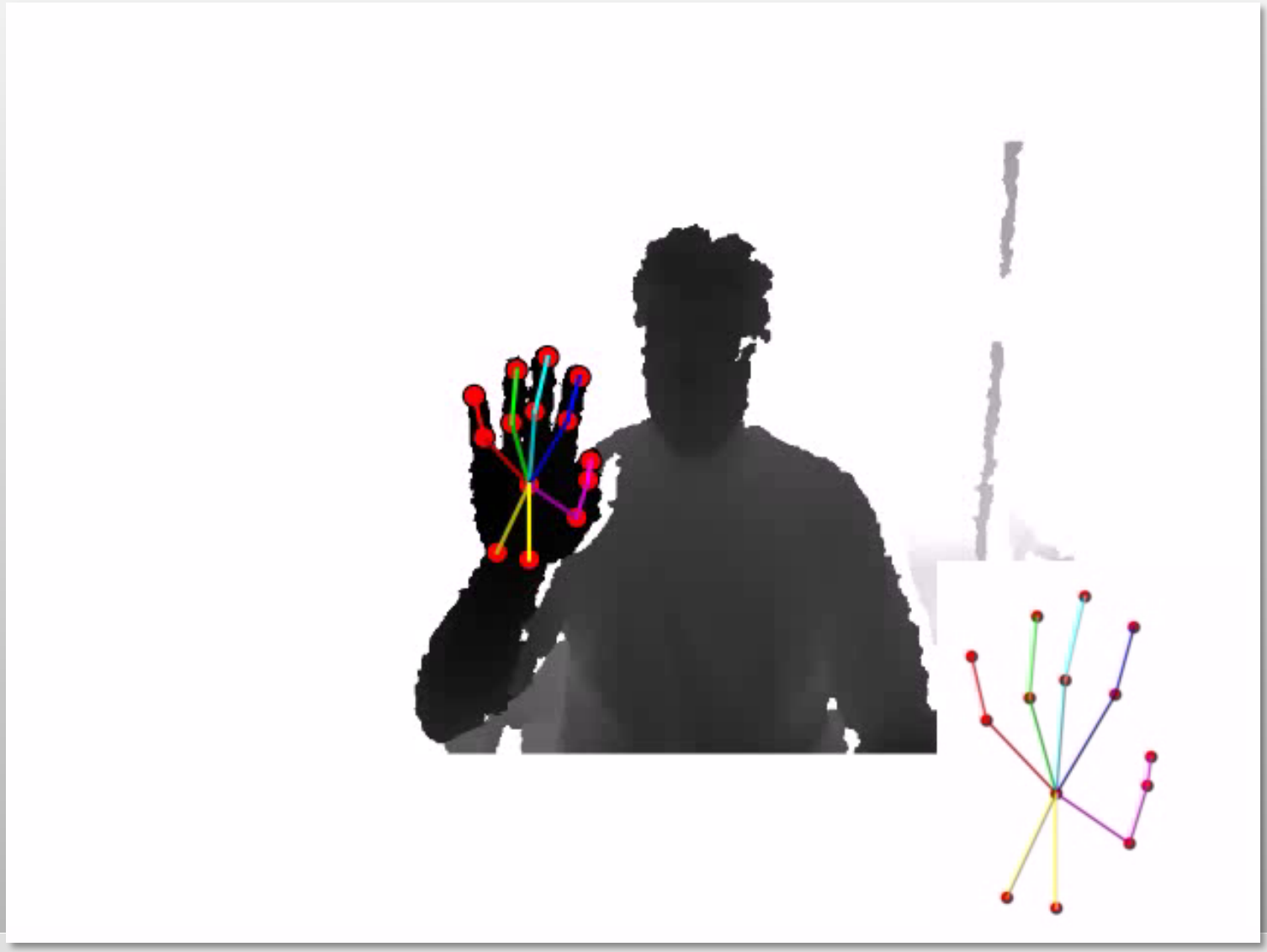


details
results

Some Results



Final Results

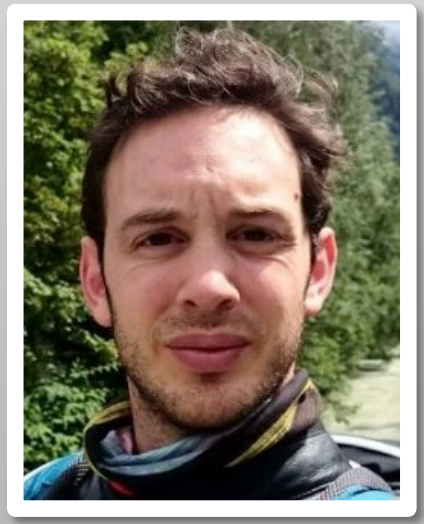


Dr. Paul Wohlhart

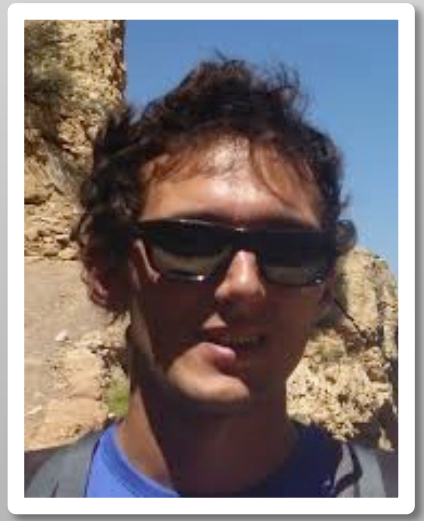
Markus Oberweger

Mahdi Rad

Dr. Kwang Moo Yi



Dr. Yannick Verdie



Alberto Crivellaro

Thanks!

Questions?