ICRA 2014

Autonomous Active Recognition & Unfolding of Clothes using Random Decision Forests & Probabilistic Planning

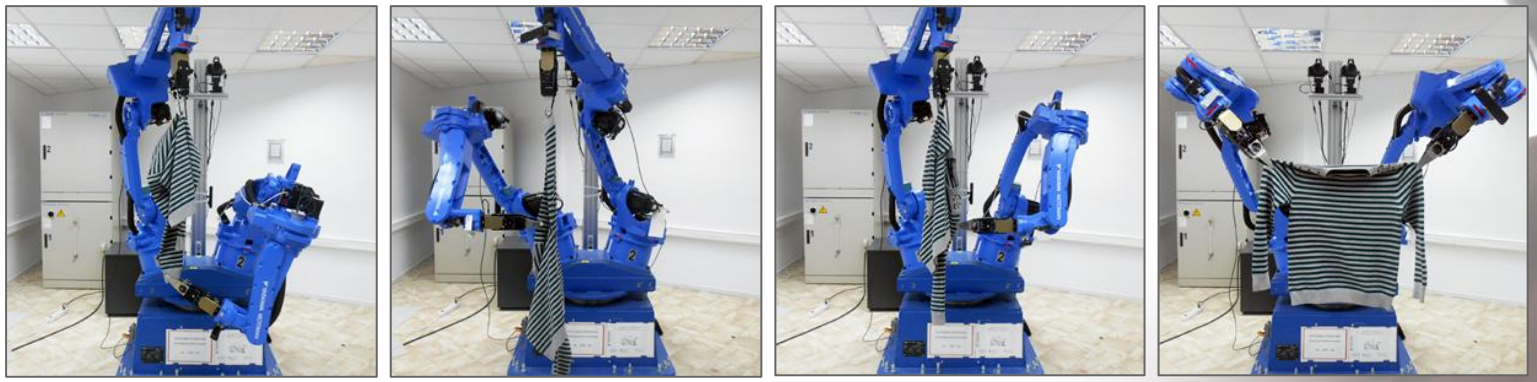
A. Doumanoglou, A. Kargakos, T-K Kim, S. Malassiotis

Robot Challenge

Autonomously Unfold clothes (in order to fold them)

Robot Challenge

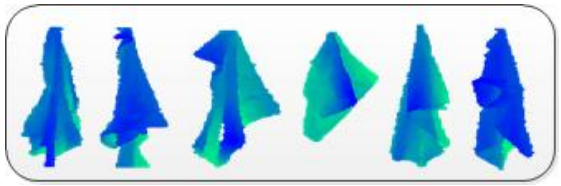
Autonomously Unfold clothes (in order to fold them)



Task Overview

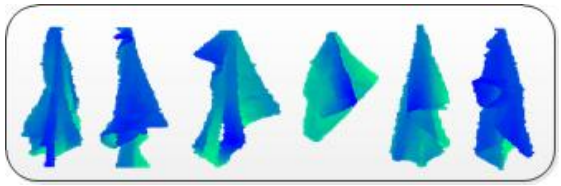
Task Overview

Clothes Recognition

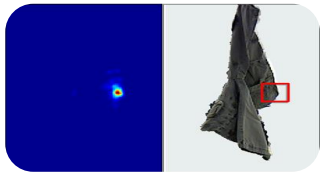


Task Overview

Clothes Recognition

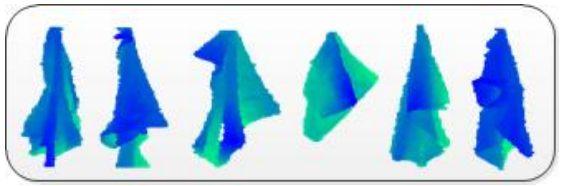


Grasp Point Detection

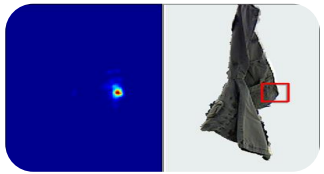


Task Overview

Clothes Recognition



Grasp Point Detection



Active Planning

Clothes Recognition

Clothes Recognition

How to reduce the large configuration space ?

Clothes Recognition

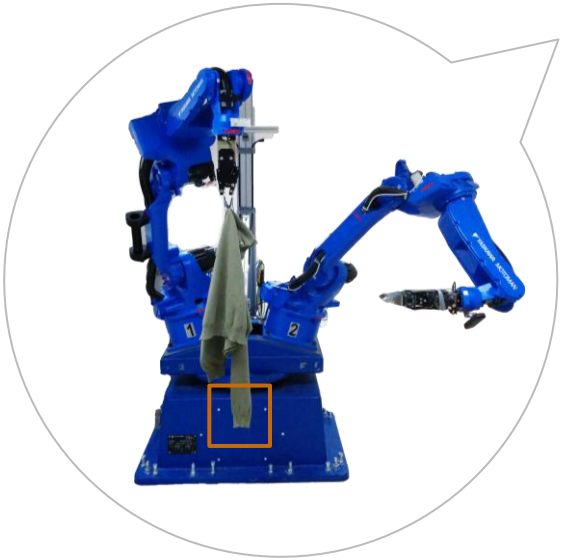
How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:

Clothes Recognition

How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:



Clothes Recognition

How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points → **6 Classes** Total

Clothes Recognition

How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points → **6 Classes** Total

Training Database:

Clothes Recognition

How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points → **6 Classes** Total

Training Database:

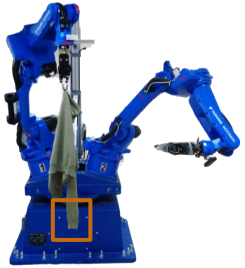
Clothes Recognition

How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points → **6 Classes** Total

Training Database:



Grasp lowest point →

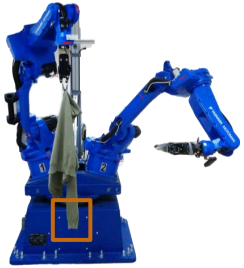
Clothes Recognition

How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points → **6 Classes** Total

Training Database:



Grasp lowest point →

Rotate 360° →

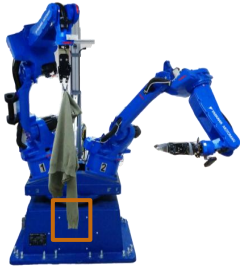
Clothes Recognition

How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:

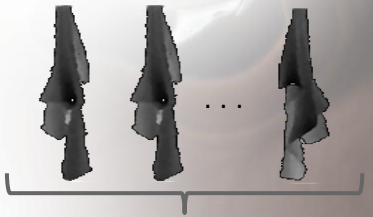
6 non-symmetric lowest points → **6 Classes** Total

Training Database:



Grasp lowest point →

Rotate 360° →



40 depth images

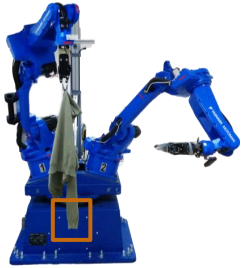
Clothes Recognition

How to reduce the large configuration space ?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points → **6 Classes** Total

Training Database:



Grasp lowest point →

Rotate 360° →

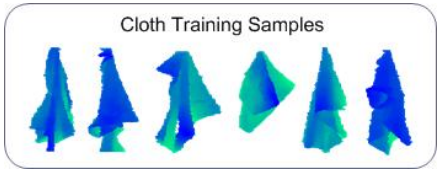
28,800
Training Images

Clothes Recognition

Random Forest **Training**

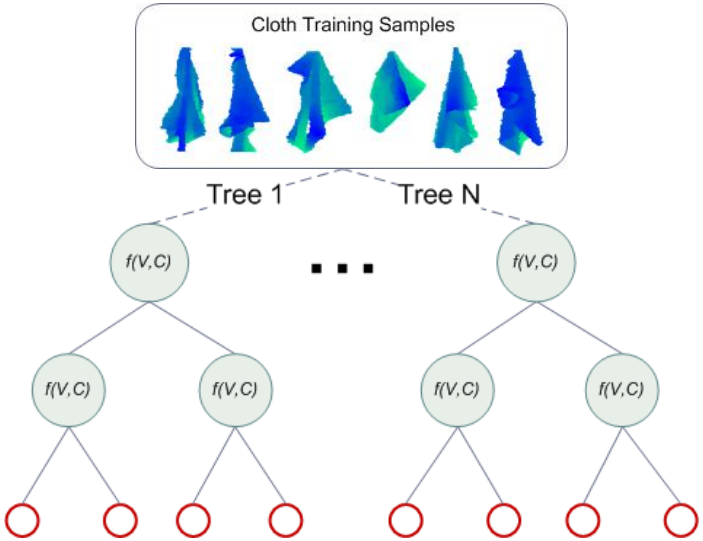
Clothes Recognition

Random Forest **Training**



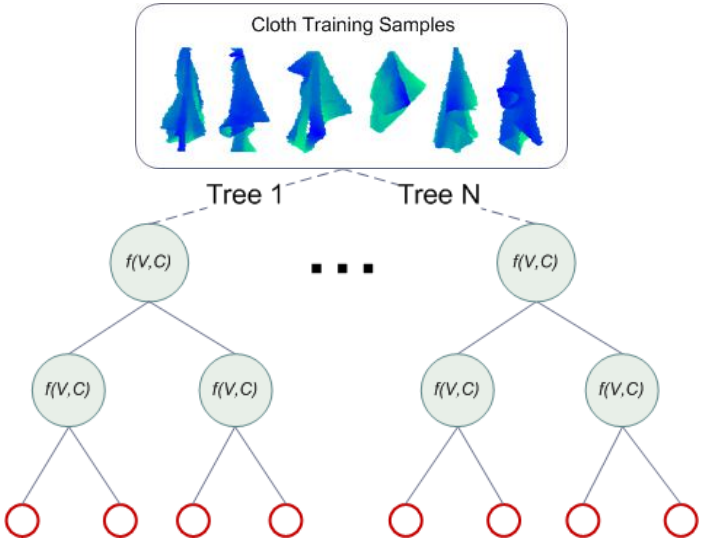
Clothes Recognition

Random Forest **Training**

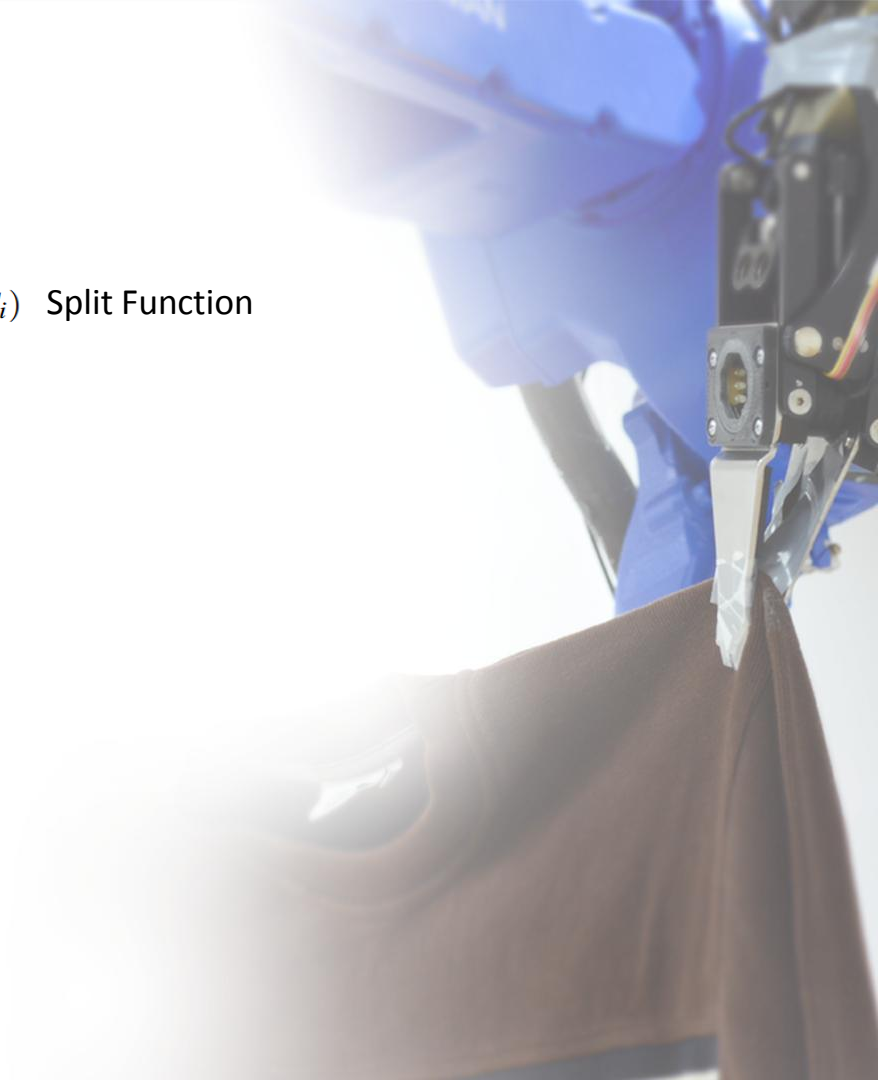


Clothes Recognition

Random Forest **Training**

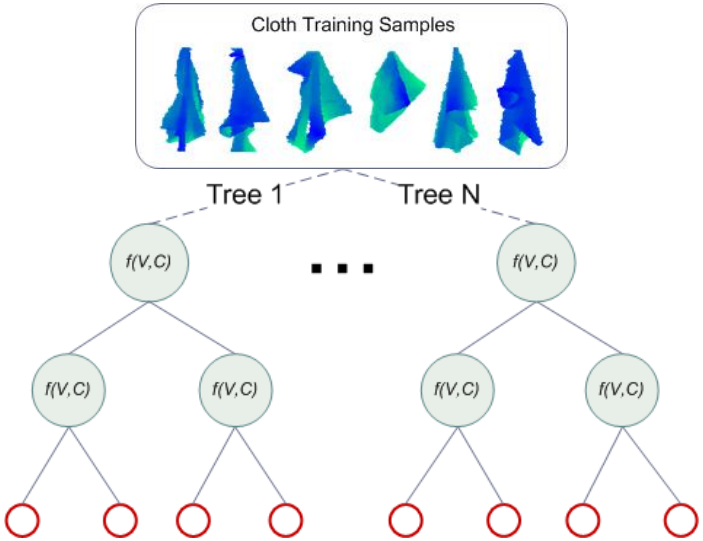


$f(\mathbf{V}, C_i)$ Split Function

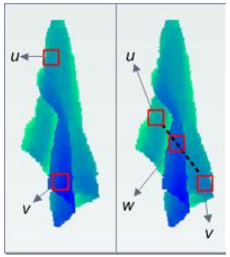


Clothes Recognition

Random Forest **Training**



$f(\mathbf{V}, C_i)$ Split Function

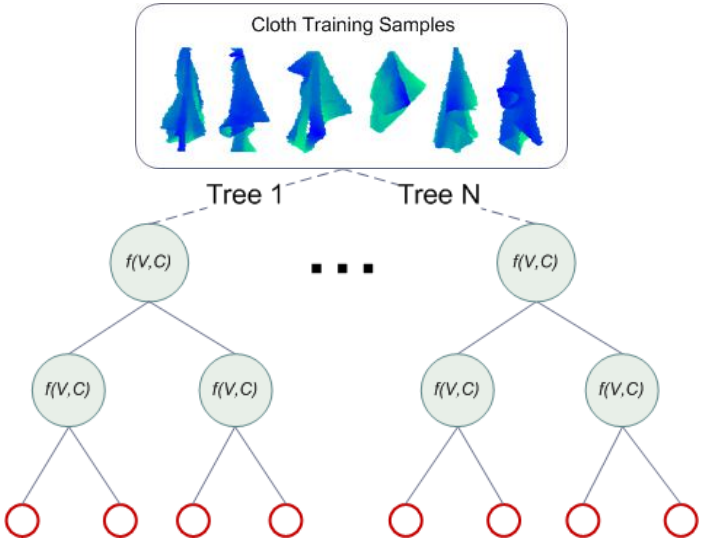


Depth Channel

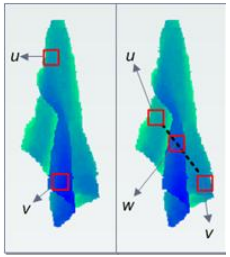
$$f(\mathbf{V}, C_1) = d_u - d_v$$
$$f(\mathbf{V}, C_1) = (d_u - d_w) - (d_w - d_v)$$

Clothes Recognition

Random Forest **Training**

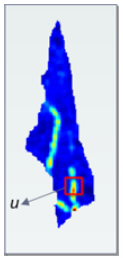


$f(\mathbf{V}, C_i)$ Split Function



Depth Channel

$$f(\mathbf{V}, C_1) = d_u - d_v$$
$$f(\mathbf{V}, C_1) = (d_u - d_w) - (d_w - d_v)$$

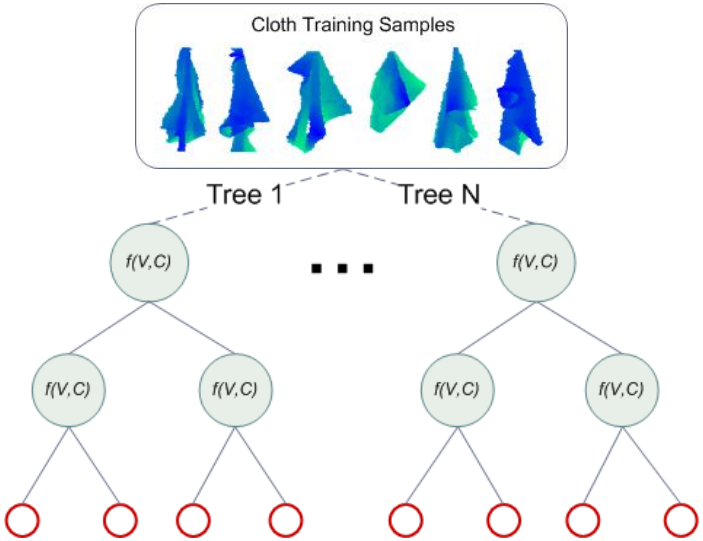


Curvature Channel

$$f(\mathbf{V}, C_2) = |c_u|$$

Clothes Recognition

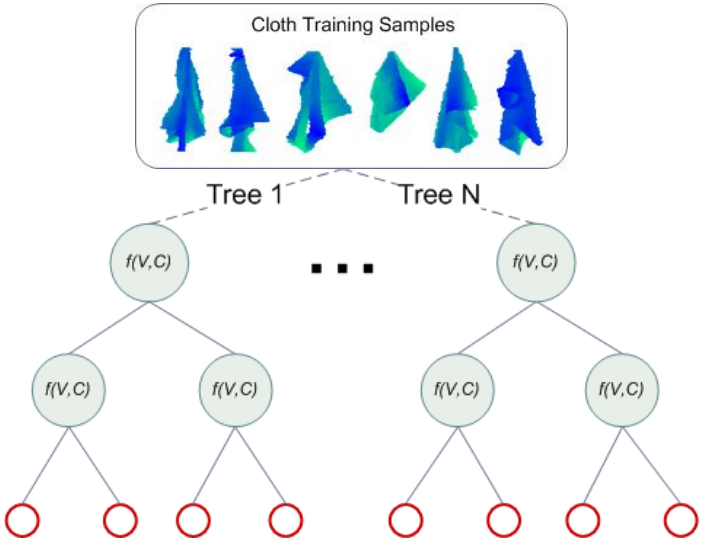
Random Forest **Training**



$f(\mathbf{V}, C_i)$ Optimization

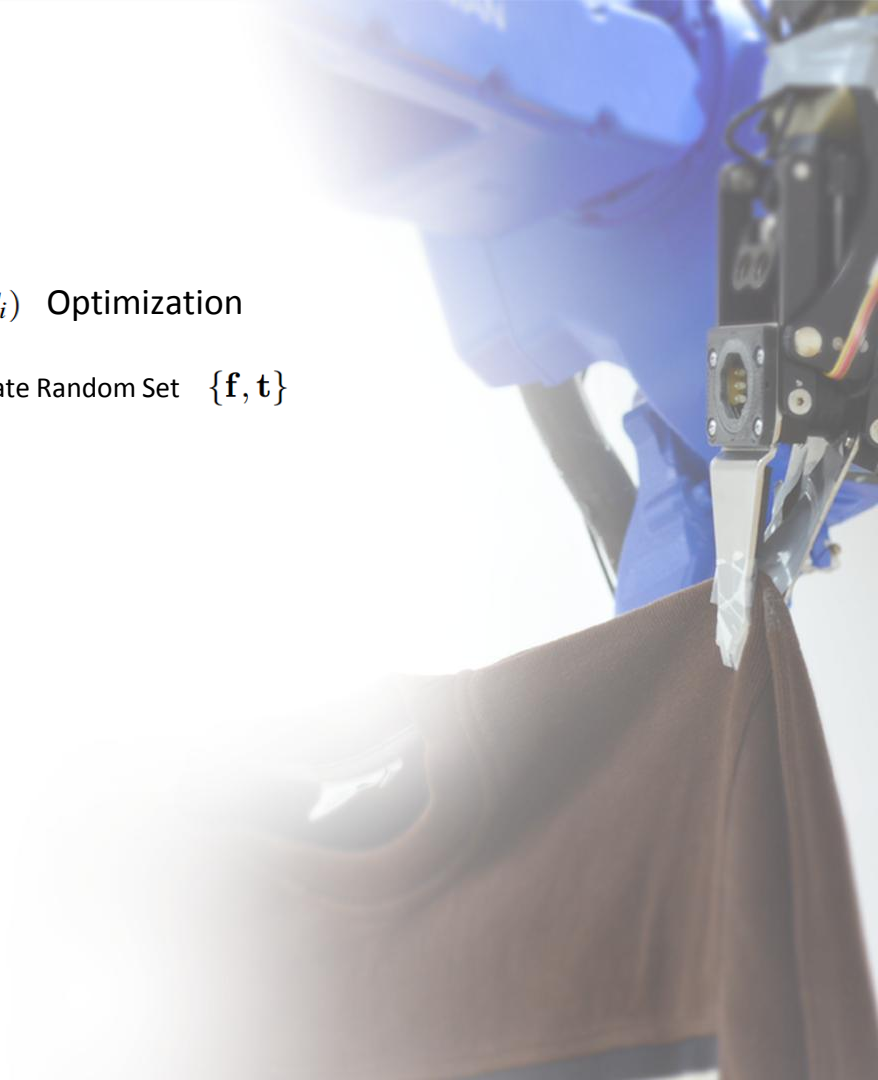
Clothes Recognition

Random Forest **Training**



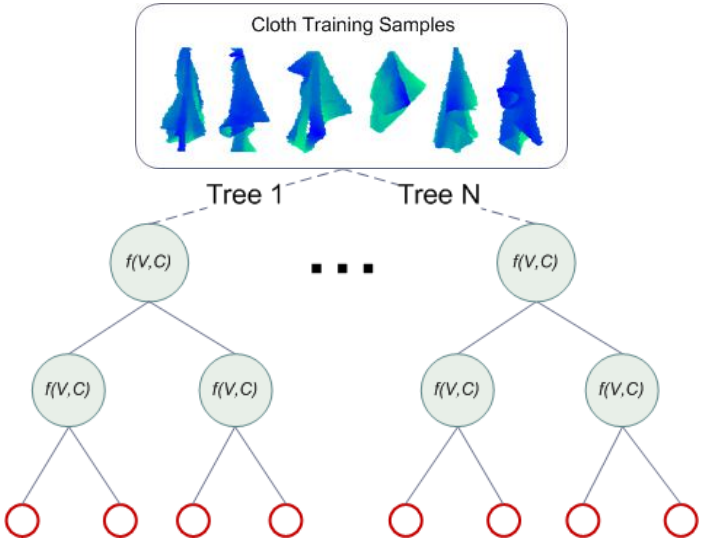
$f(\mathbf{V}, C_i)$ Optimization

Generate Random Set $\{\mathbf{f}, \mathbf{t}\}$



Clothes Recognition

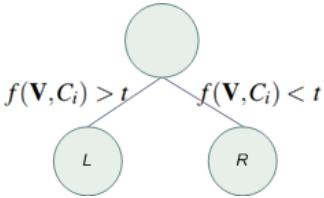
Random Forest **Training**



$f(\mathbf{V}, C_i)$ Optimization

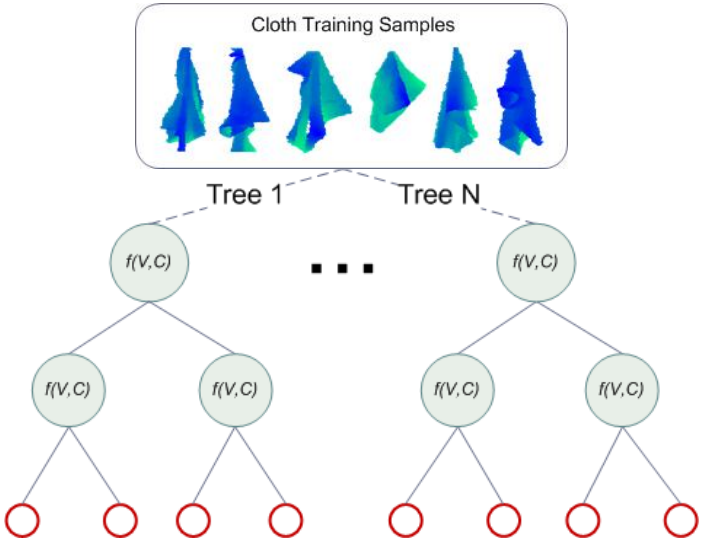
Generate Random Set $\{\mathbf{f}, \mathbf{t}\}$

Split Samples



Clothes Recognition

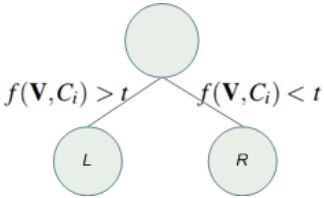
Random Forest **Training**



$f(\mathbf{V}, C_i)$ Optimization

Generate Random Set $\{\mathbf{f}, \mathbf{t}\}$

Split Samples

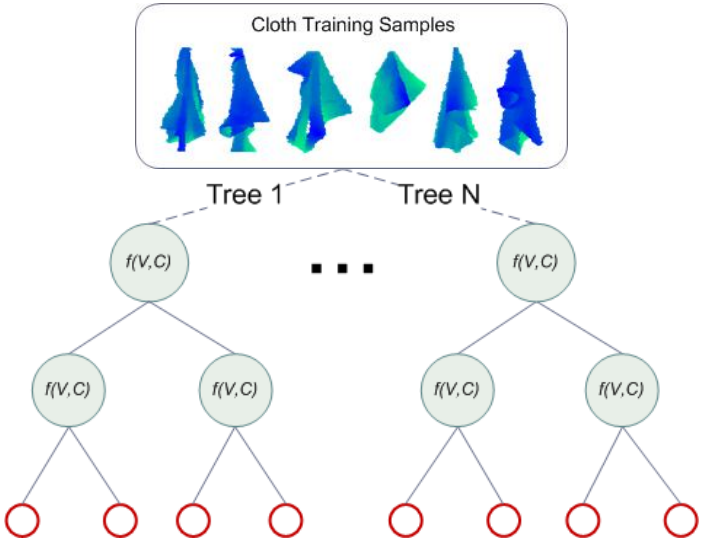


Measure Entropy

$$H = -\left(|L| p_L \log p_L + |R| p_R \log p_R\right)$$

Clothes Recognition

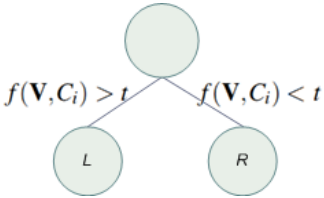
Random Forest **Training**



$f(\mathbf{V}, C_i)$ Optimization

Generate Random Set $\{\mathbf{f}, \mathbf{t}\}$

Split Samples



Measure Entropy

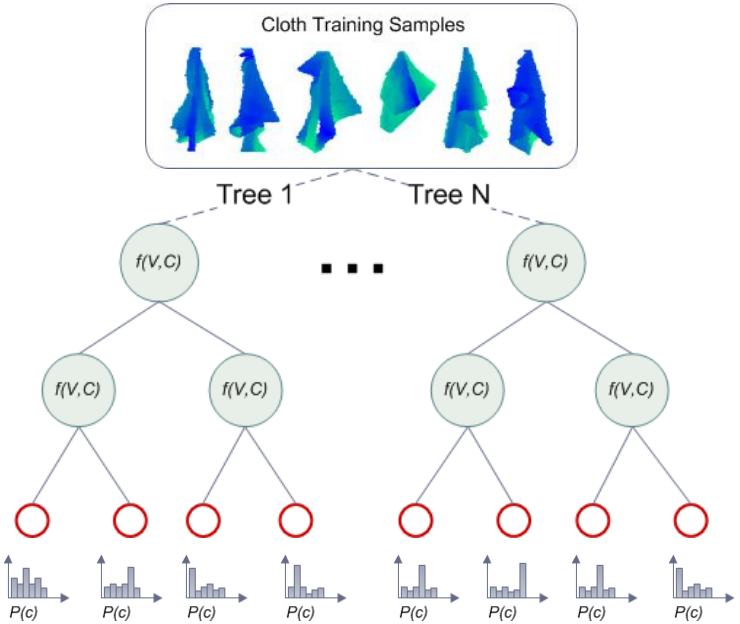
$$H = -\left(|L| p_L \log p_L + |R| p_R \log p_R\right)$$

Greedily Optimize

$$\{f, t\} = argmin_{f,t} H$$

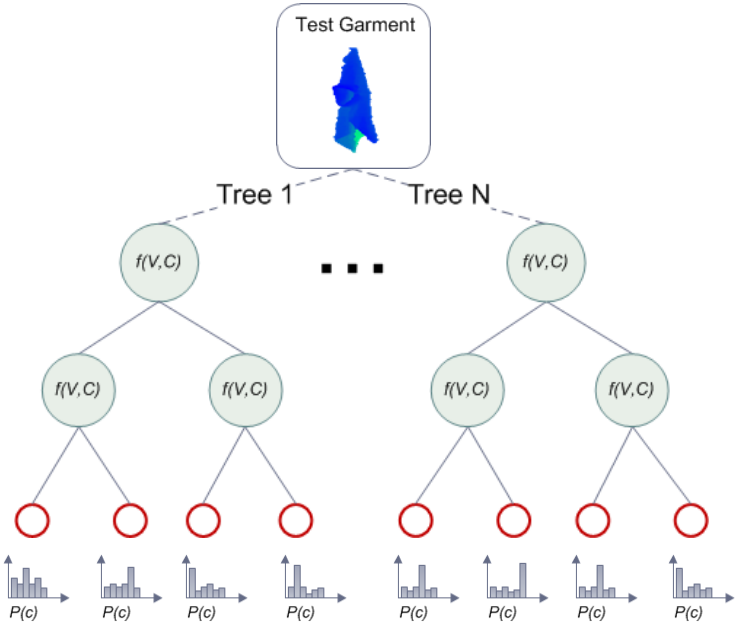
Clothes Recognition

Random Forest **Training**



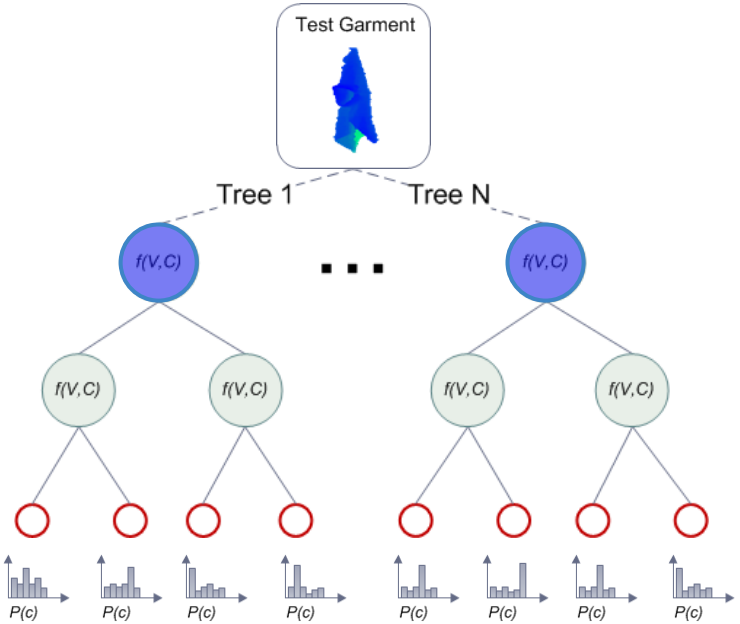
Clothes Recognition

Random Forest **Testing**



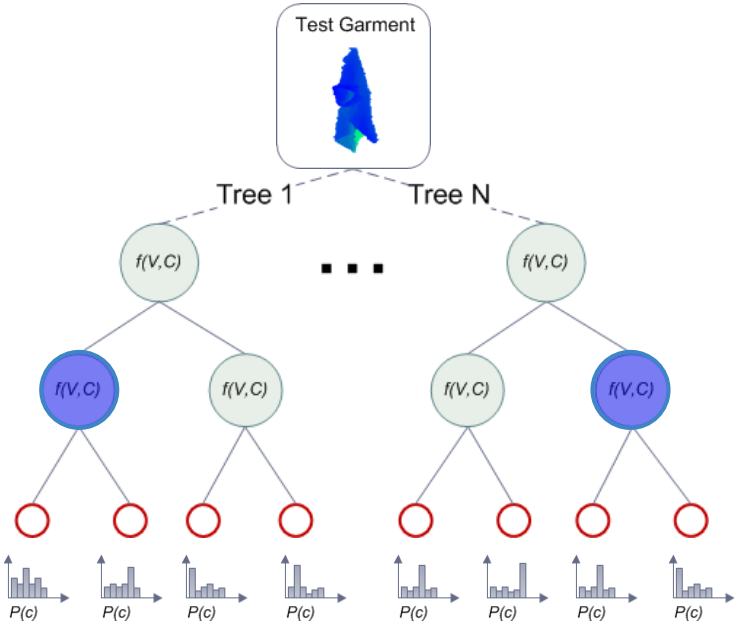
Clothes Recognition

Random Forest **Testing**



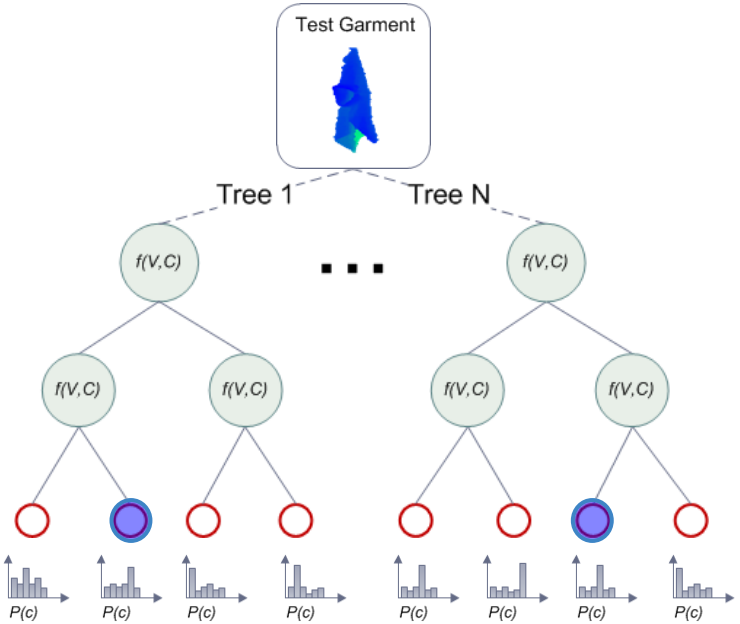
Clothes Recognition

Random Forest **Testing**



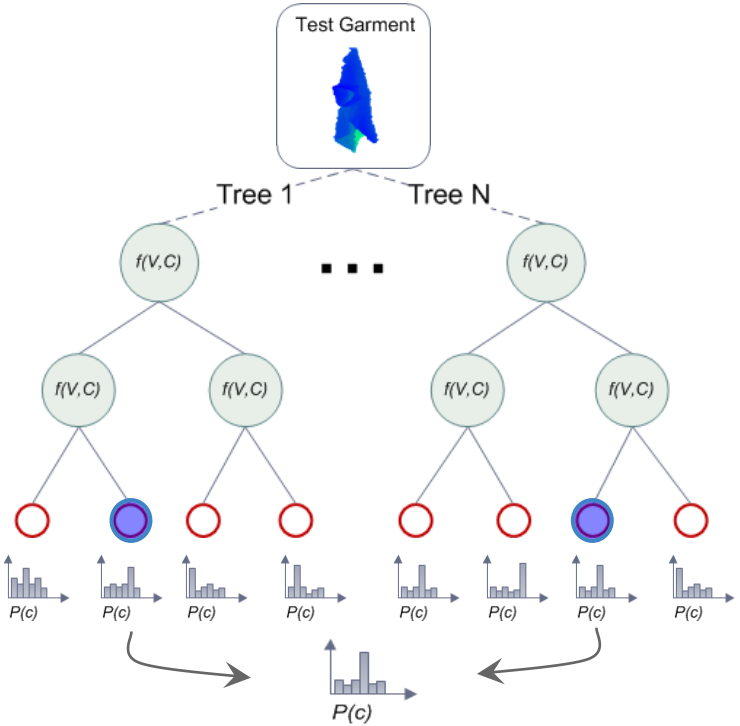
Clothes Recognition

Random Forest **Testing**



Clothes Recognition

Random Forest **Testing**



Grasp Point Detection

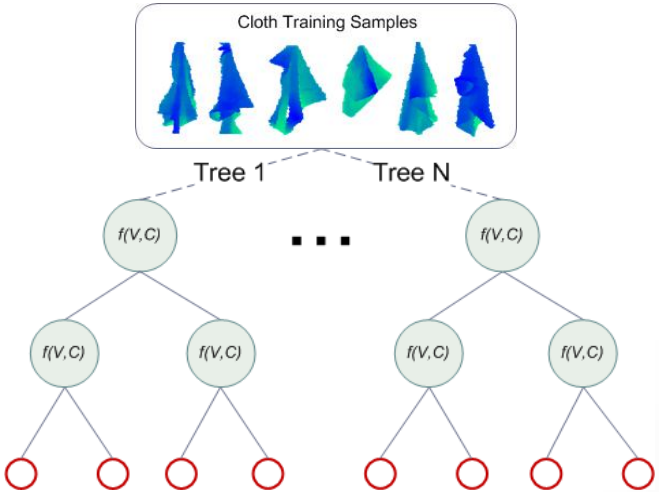
Grasp Point Detection

Desired grasp Points:

Grasp Point Detection

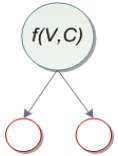
Desired grasp Points:

Hough Forest



Grasp Point Detection

Desired grasp Points:

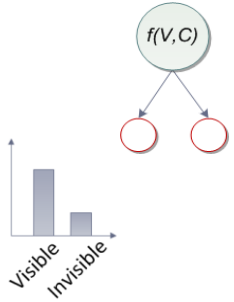


Hough Forest

Grasp Point Detection

Desired grasp Points:

Hough Forest

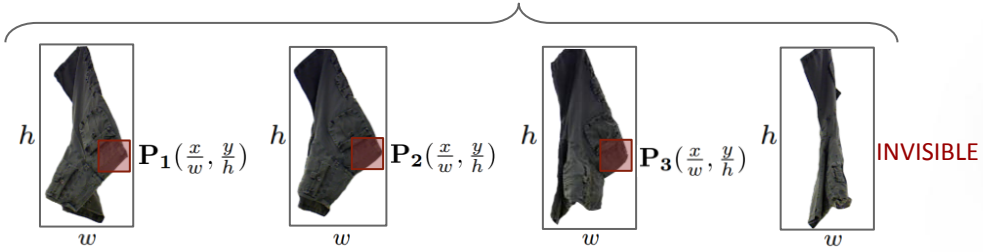


Grasp Point Detection

Desired grasp Points:

Hough Forest

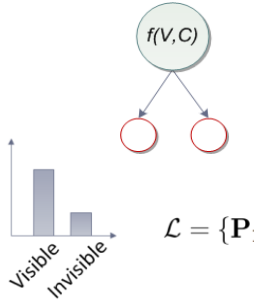
$$\mathcal{L} = \{\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}, \mathbf{INV}\}$$



Grasp Point Detection

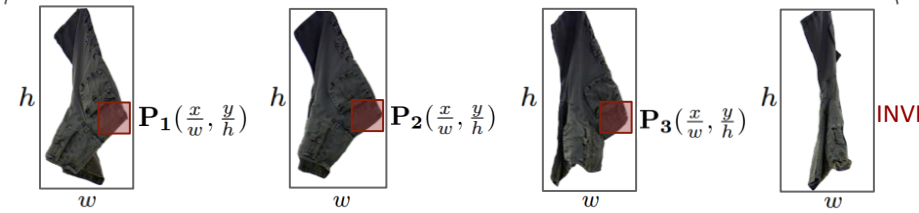
Desired grasp Points:

Hough Forest



$\mathcal{L} = \{\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}, \mathbf{INV}\}$

$f(\mathbf{V}, C_i)$ Optimization



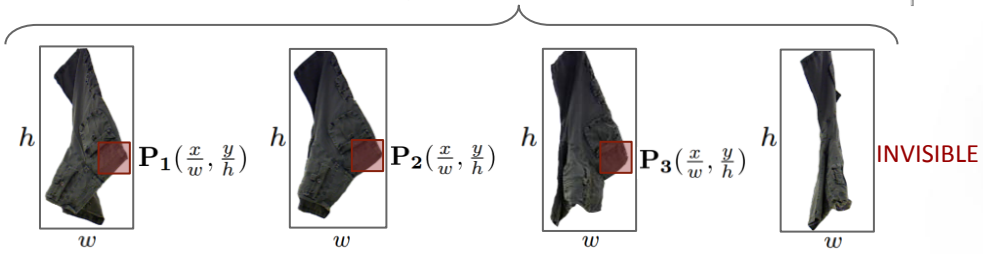
INVISIBLE

Grasp Point Detection

Desired grasp Points:

Hough Forest

$\mathcal{L} = \{\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}, \mathbf{INV}\}$



$f(\mathbf{V}, C_i)$ Optimization

Minimize:

$H = -(|L| p_L \log p_L + |R| p_R \log p_R)$

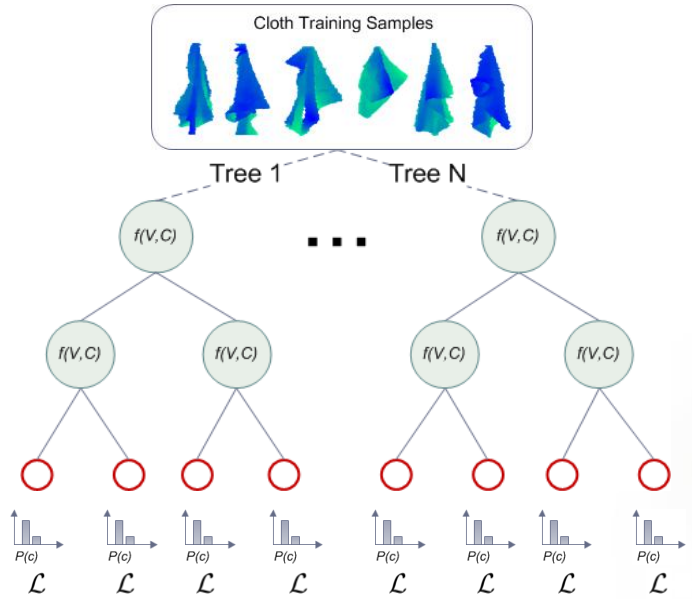
or

$D = \sum_{samples} d(\mathbf{p_s}, \mathbf{p_M})$ Variance in each child

Grasp Point Detection

Desired grasp Points:

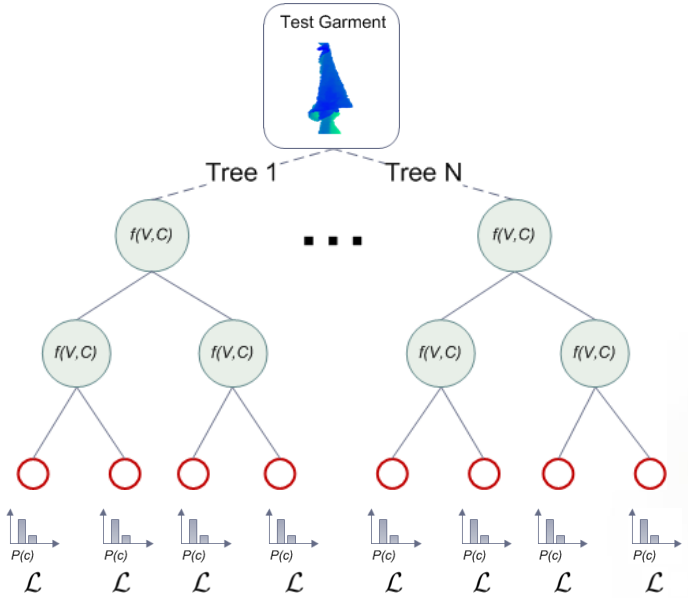
Hough Forest



Grasp Point Detection

Desired grasp Points:

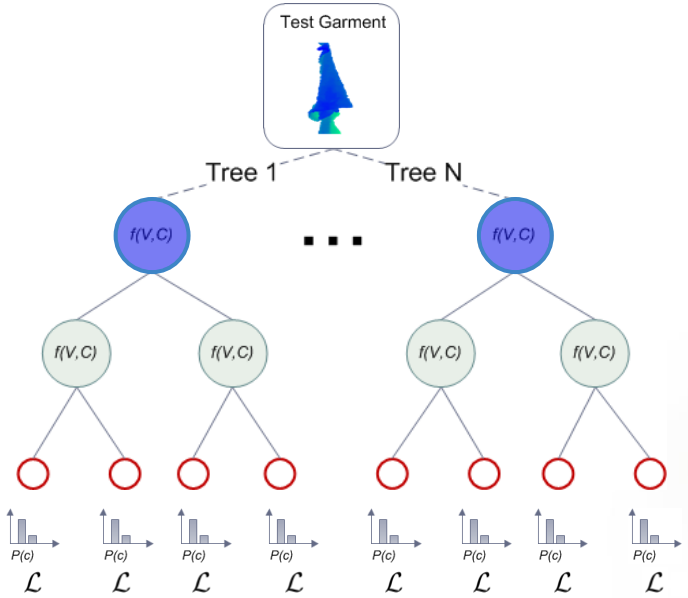
Hough Forest



Grasp Point Detection

Desired grasp Points:

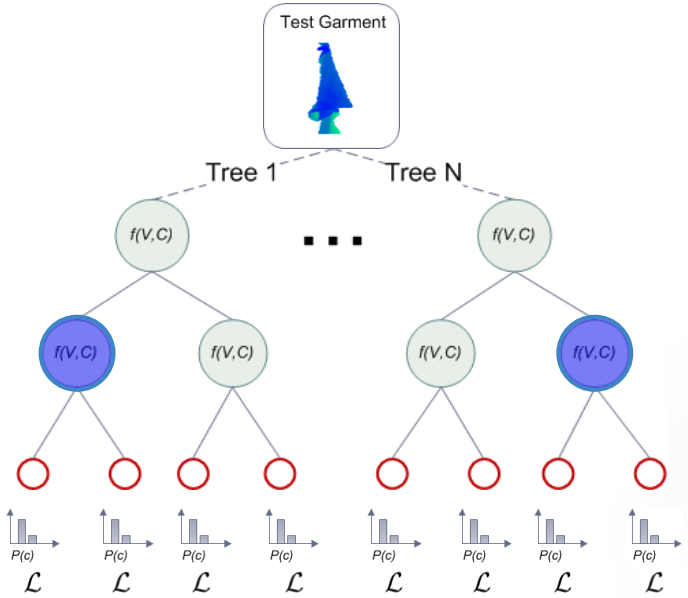
Hough Forest



Grasp Point Detection

Desired grasp Points:

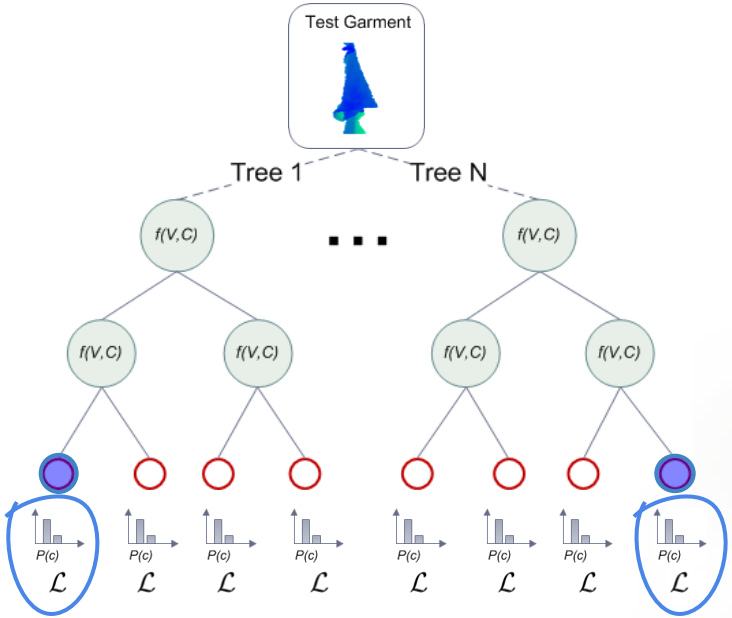
Hough Forest



Grasp Point Detection

Desired grasp Points:

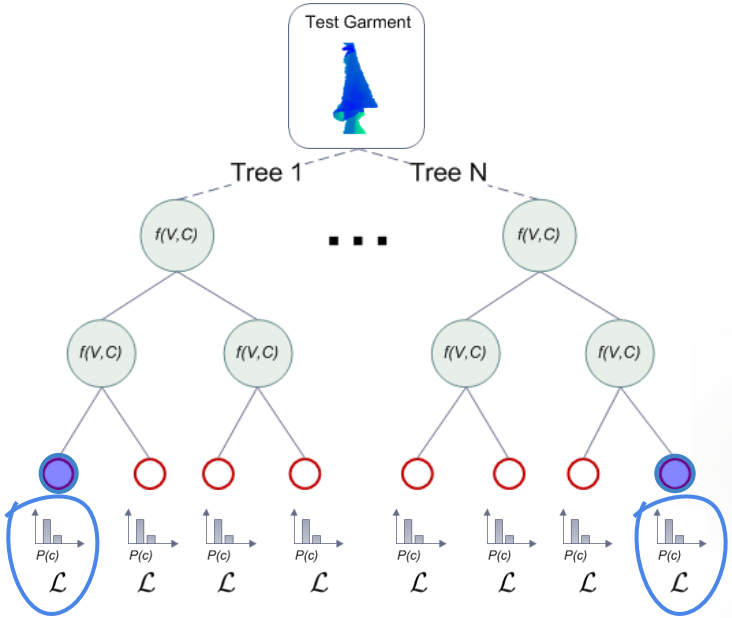
Hough Forest



Grasp Point Detection

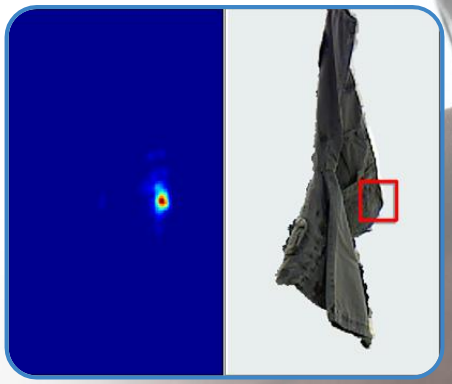
Desired grasp Points:

Hough Forest



Hough Voting Image

Point Prediction



Active Planning

Single view

success ~ 90%

Crucial Decisions

Active Planning

Single view

success ~ 90%

Crucial Decisions

How can other views help ?

Active Planning

Single view

success ~ 90%

Crucial Decisions

How can other views help ?

Approach

Active Planning

Single view

success ~ 90%

Crucial Decisions

How can other views help ?

Approach

Keep looking sequential views

Active Planning

Single view

success ~ 90%

Crucial Decisions

How can other views help ?

Approach

Keep looking sequential views

Until we reach a certain **degree of confidence**

Active Planning

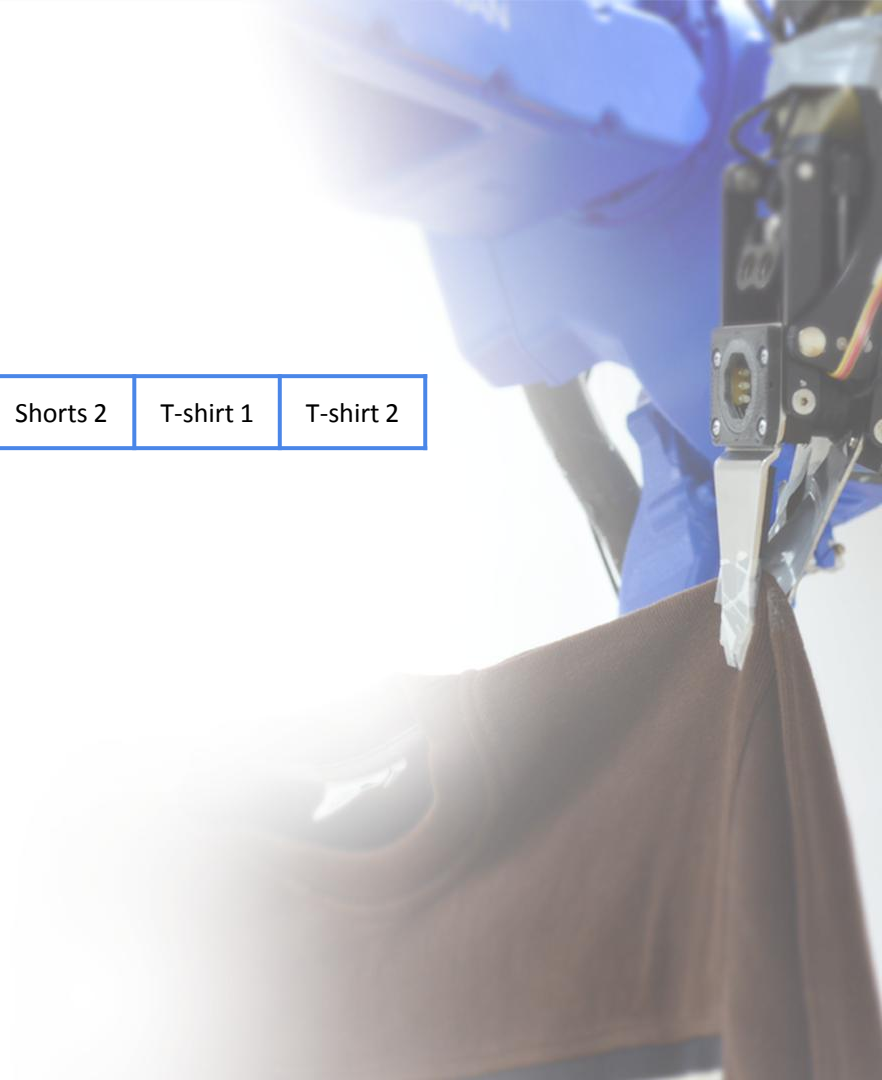
Active Recognition — **POMDP solution**

Active Planning

Active Recognition — **POMDP solution**

States (**S**): 6 Classes

Shirt	Trousers	Shorts 1	Shorts 2	T-shirt 1	T-shirt 2



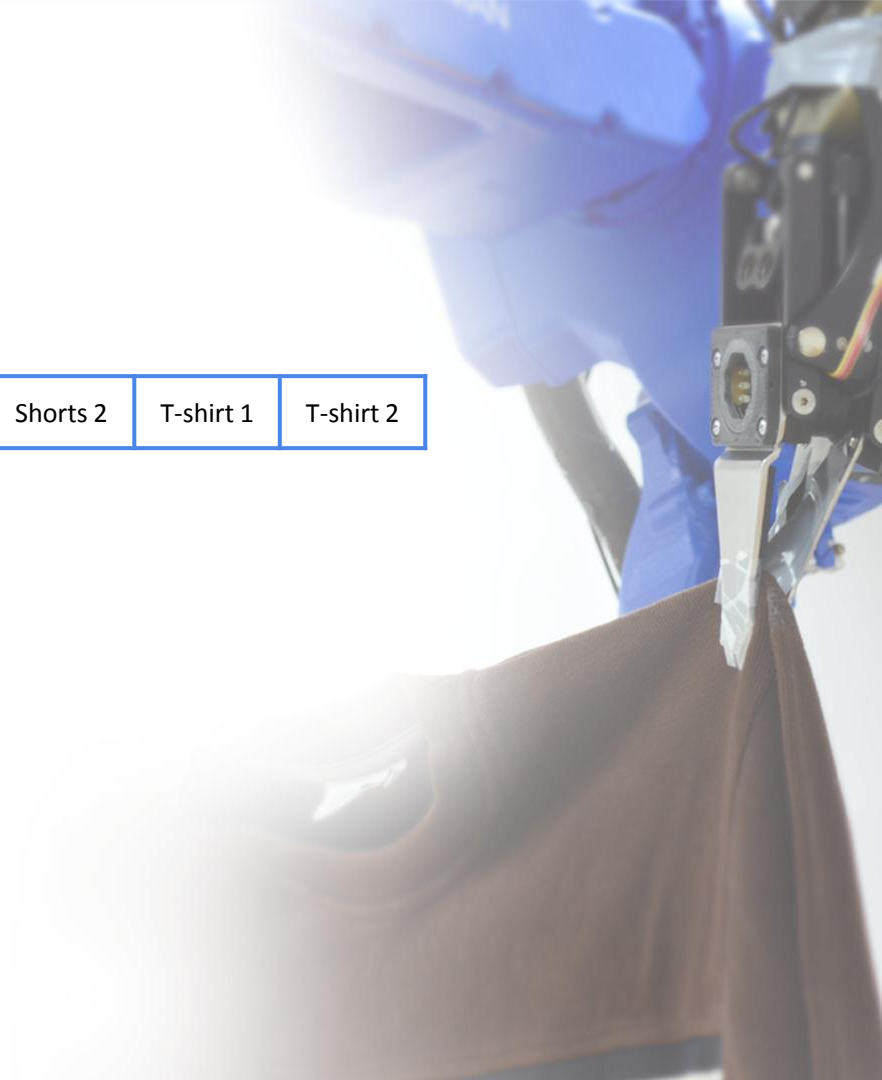
Active Planning

Active Recognition — **POMDP solution**

States (**S**): 6 Classes

Shirt	Trousers	Shorts 1	Shorts 2	T-shirt 1	T-shirt 2

Actions (**A**): Rotate Cloth — Take Final Decision



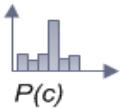
Active Planning

Active Recognition — **POMDP solution**

States (**S**): 6 Classes

Shirt	Trousers	Shorts 1	Shorts 2	T-shirt 1	T-shirt 2

Actions (**A**): Rotate Cloth — Take Final Decision

Observations (**O**): 

(quantized)

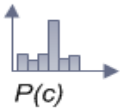
Active Planning

Active Recognition — **POMDP solution**

States (**S**): 6 Classes

Shirt	Trousers	Shorts 1	Shorts 2	T-shirt 1	T-shirt 2

Actions (**A**): Rotate Cloth — Take Final Decision

Observations (**O**): 

(quantized)

Observation Probabilities: **P (O | S, A)** → Measured Experimentally

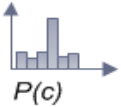
Active Planning

Active Recognition — **POMDP solution**

States (**S**): 6 Classes

Shirt	Trousers	Shorts 1	Shorts 2	T-shirt 1	T-shirt 2

Actions (**A**): Rotate Cloth — Take Final Decision

Observations (**O**): 

(quantized)

Observation Probabilities: **P (O | S, A)** → Measured Experimentally

Transition Probabilities: $$T(S_i|A_{rotate}, S_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

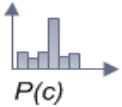
Active Planning

Active Recognition — **POMDP solution**

States (**S**): 6 Classes

Shirt	Trousers	Shorts 1	Shorts 2	T-shirt 1	T-shirt 2

Actions (**A**): Rotate Cloth — Take Final Decision

Observations (**O**): 

(quantized)

Observation Probabilities: **P (O | S, A)** → Measured Experimentally

Transition Probabilities: $$T(S_i|A_{rotate}, S_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

Rewards:
- R > 0, Correct Decision
- R << 0, Wrong Decision
- R < 0, Rotation

Active Planning

Active Point Estimation — **POMDP solution**

Active Planning

Active Point Estimation — **POMDP solution**

States (**S**): 65 — 8x8 grid quantization, or (INV)

(i, j)

Active Planning

Active Point Estimation — **POMDP solution**

States (**S**): 65 — 8x8 grid quantization, or (INV)

Actions (**A**): Rotate Cloth — Grasp Garment at (i, j)

(i, j)

Active Planning

Active Point Estimation — **POMDP solution**

States (**S**): 65 — 8x8 grid quantization, or (INV)

Actions (**A**): Rotate Cloth — Grasp Garment at (i, j)

Observations (**O**): $P_{grasppoint}(i, j)$ (quantized) from Hough Image

(i, j)

Active Planning

Active Point Estimation — **POMDP solution**

States (**S**): 65 — 8x8 grid quantization, or (INV)

(i, j)

Actions (**A**): Rotate Cloth — Grasp Garment at (i, j)

Observations (**O**): $P_{grasppoint}(i, j)$ (quantized) from Hough Image

Observation Probabilities: **P (O | S, A)**

→ Measured Experimentally

Transition Probabilities: **T (S_i | A, S_j)**

Active Planning

Active Point Estimation — **POMDP solution**

States (**S**): 65 — 8x8 grid quantization, or (INV)

(i, j)

Actions (**A**): Rotate Cloth — Grasp Garment at (i, j)

Observations (**O**): $P_{grasppoint}(i, j)$ (quantized) from Hough Image

Observation Probabilities: **P (O | S, A)**

→ Measured Experimentally

Transition Probabilities: **T (S_i | A, S_j)**

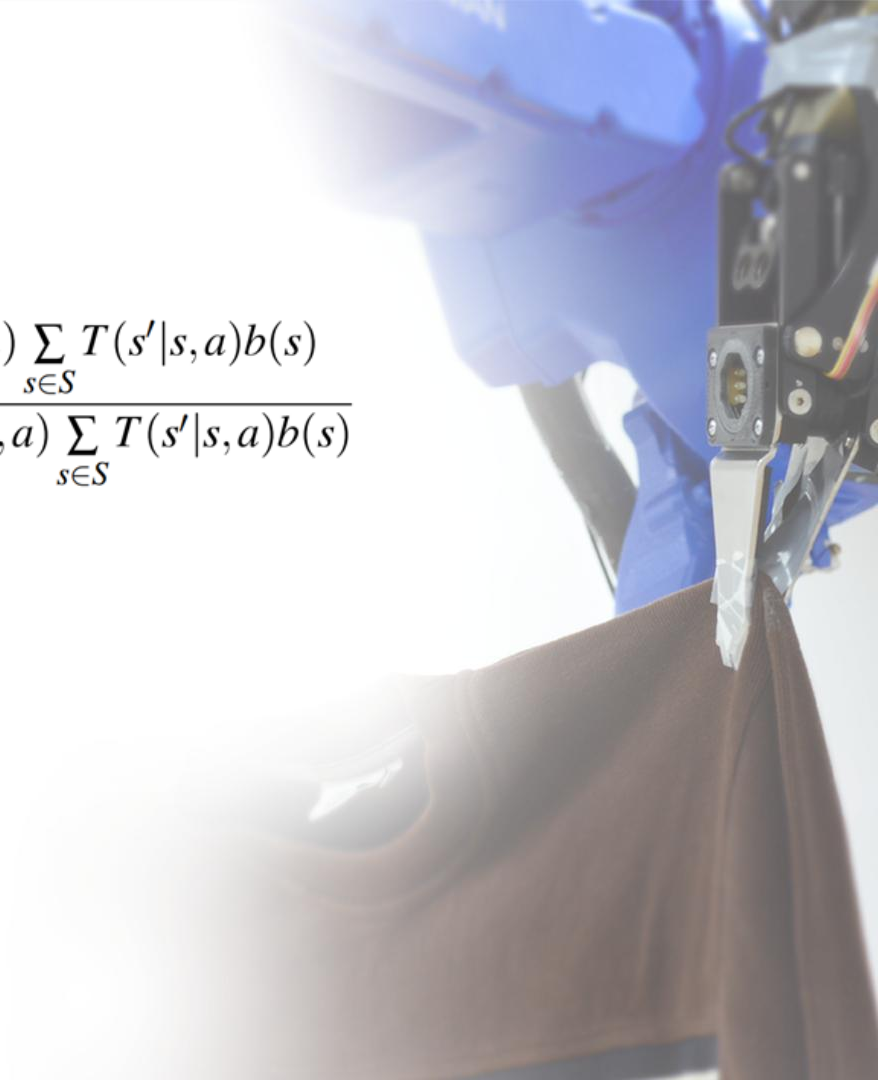
Rewards:

- R >> 0, Correct Estimation
- R << 0, Wrong Estimation
- R < 0, Rotation if Point is visible
- R > 0, Rotation if Point is invisible

Active Planning

POMDP belief update

$$b'(s') = \frac{P(o|s',a) \sum_{s \in S} T(s'|s,a) b(s)}{\sum_{s' \in S} P(o|s',a) \sum_{s \in S} T(s'|s,a) b(s)}$$



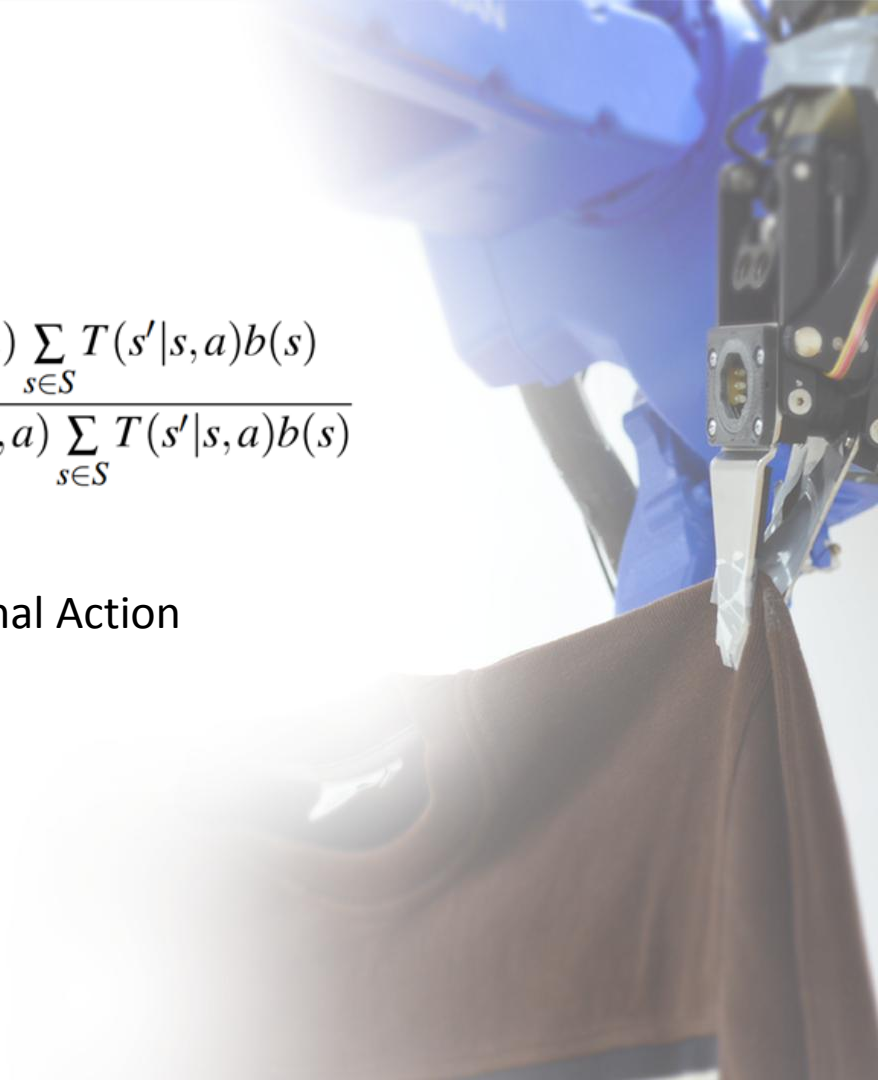
Active Planning

POMDP belief update

$$b'(s') = \frac{P(o|s',a) \sum_{s \in S} T(s'|s,a) b(s)}{\sum_{s' \in S} P(o|s',a) \sum_{s \in S} T(s'|s,a) b(s)}$$

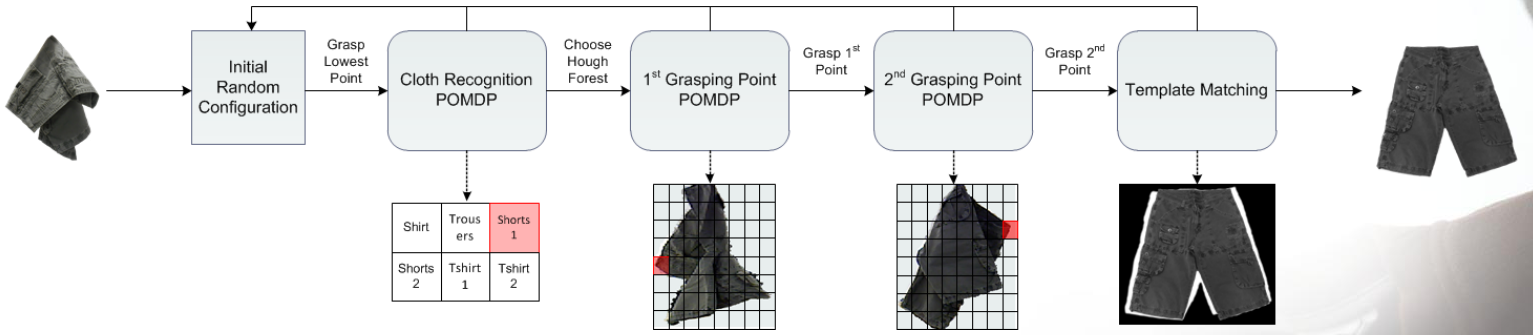
POMDP solution policy

A(b(s)) → Optimal Action



Block Diagram

Unfolding Process



Results

Training Database

Results

Training Database

x6 x6 x6 x6

Results

Training Database

x6 x6 x6 x6 **28,800 Images**

Results

Training Database

x6 x6 x6 x6 **28,800 Images**

Testing Database

Results

Training Database

x6	x6	x6	x6	**28,800 Images**

Testing Database

x3	x3	x3	x3	not in training

Results

Training Database

x6	x6	x6	x6	**28,800 Images**

Testing Database

x3	x3	x3	x3	not in training
				1,440 Images

Results

Training Database

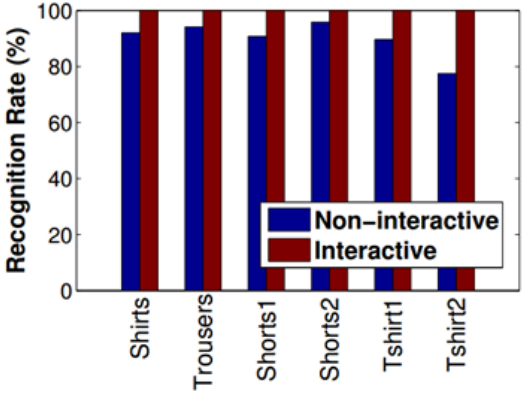
x6	x6	x6	x6	**28,800 Images**

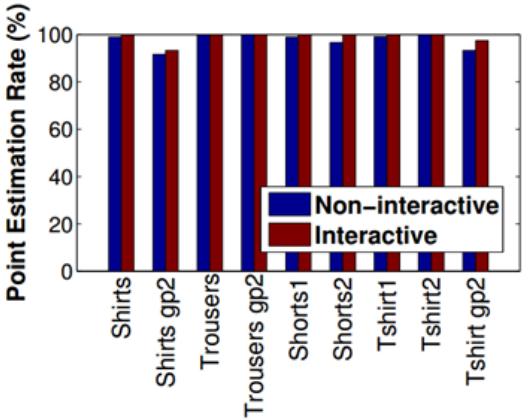
Testing Database

x3	x3	x3	x3	not in training
				1,440 Images

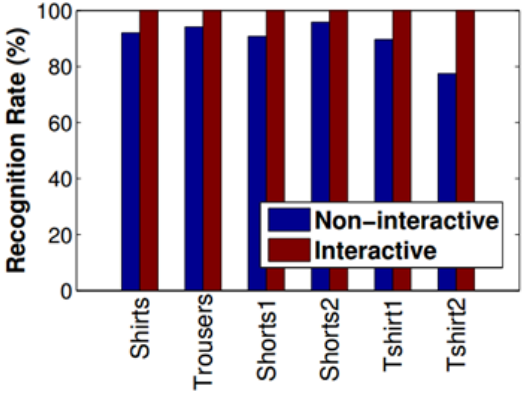
Depth Images Captured with Xtion

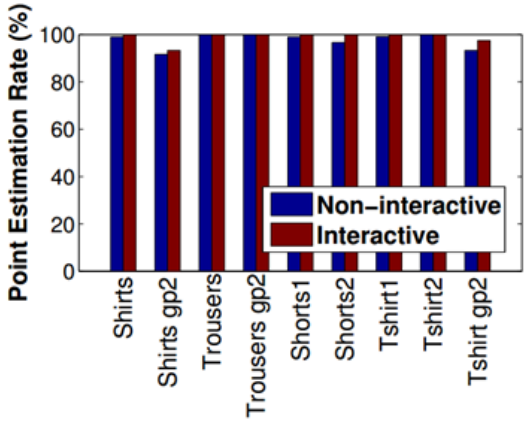
Results


Recognition Rate (%)
100
80
60
40
20
0
Non−interactive
Interactive
Shirts
Trousers
Shorts1
Shorts2
Tshirt1
Tshirt2


Point Estimation Rate (%)
100
80
60
40
20
0
Non−interactive
Interactive
Shirts
Shirts gp2
Trousers
Trousers gp2
Shorts1
Shorts2
Tshirt1
Tshirt2
Tshirt gp2

Results



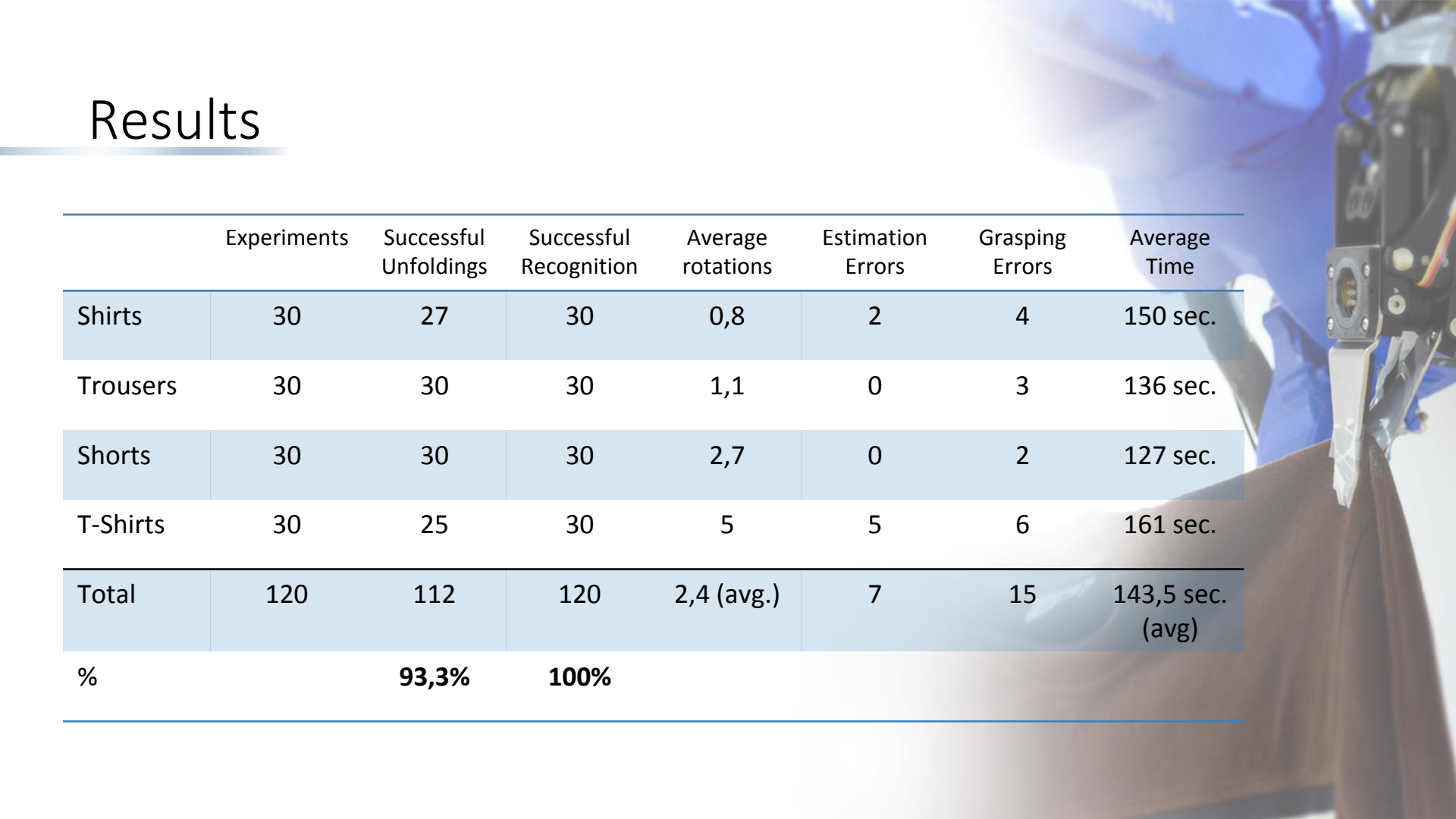


positive examples

negative examples

Results

	Experiments	Successful Unfoldings	Successful Recognition	Average rotations	Estimation Errors	Grasping Errors	Average Time
Shirts	30	27	30	0,8	2	4	150 sec.
Trousers	30	30	30	1,1	0	3	136 sec.
Shorts	30	30	30	2,7	0	2	127 sec.
T-Shirts	30	25	30	5	5	6	161 sec.
Total	120	112	120	2,4 (avg.)	7	15	143,5 sec. (avg)
%		**93,3%**	**100%**				

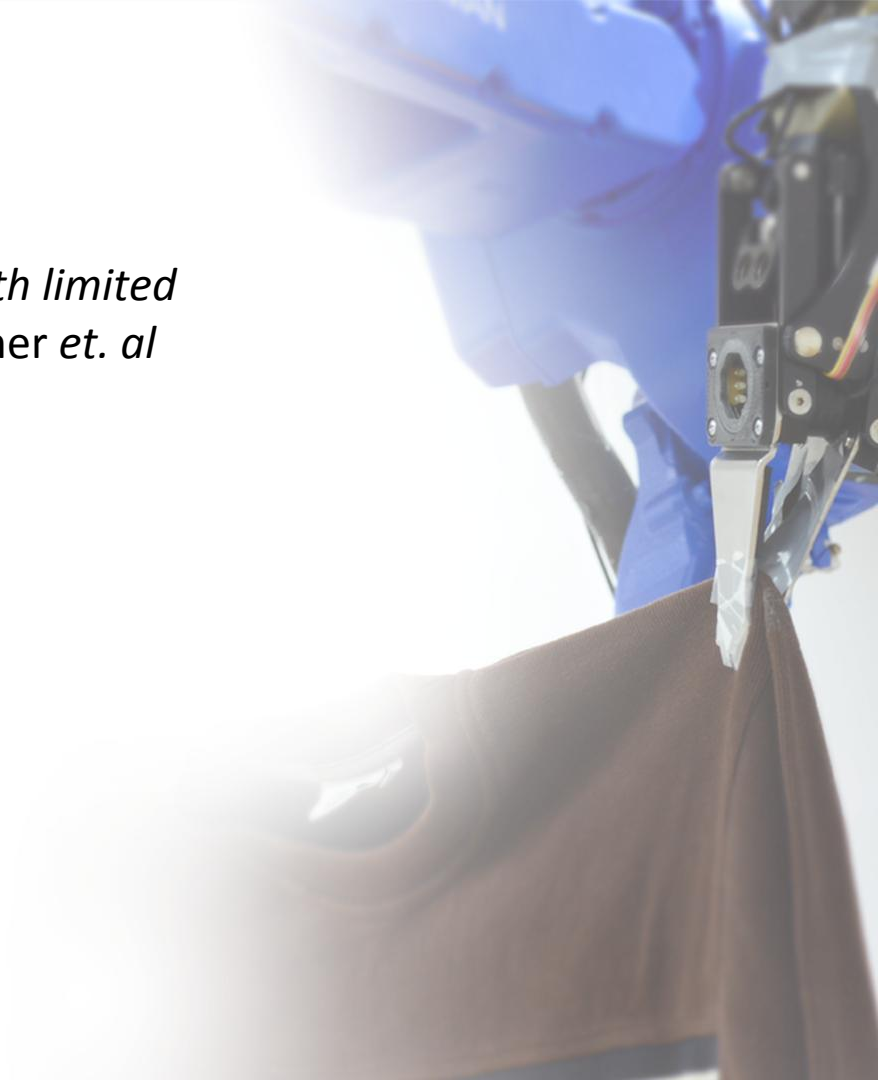


Results

	Experiments	Successful Unfoldings	Successful Recognition	Average rotations	Estimation Errors	Grasping Errors	Average Time
Shirts	30	27	30	0,8	2	4	150 sec.
Trousers	30	30	30	1,1	0	3	136 sec.
Shorts	30	30	30	2,7	0	2	127 sec.
T-Shirts	30	25	30	5	5	6	161 sec.
Total	120	112	120	2,4 (avg.)	7	15	143,5 sec. (avg)
%		**93,3%**	**100%**				

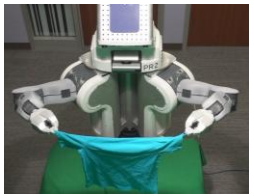
State of the Art

Bringing clothing into desired configurations with limited perception, ICRA 2011 — M. Cusumano-Towner *et. al*



State of the Art

Bringing clothing into desired configurations with limited perception, ICRA 2011 — M. Cusumano-Towner *et. al*

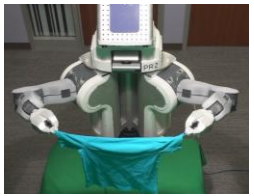


grasp
lowest point
twice

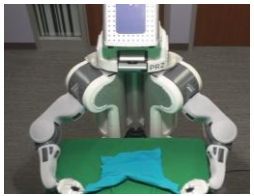
grasp
lowest point
once

State of the Art

Bringing clothing into desired configurations with limited perception, ICRA 2011 — M. Cusumano-Towner *et. al*



grasp
lowest point
twice



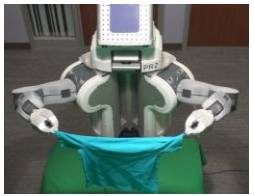
unfolding
using table
(slow)

grasp
lowest point
once

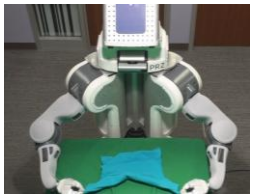
unfolding
in the air
(fast)

State of the Art

Bringing clothing into desired configurations with limited perception, ICRA 2011 — M. Cusumano-Towner *et. al*



grasp
lowest point
twice



unfolding
using table
(slow)

baby clothes

grasp
lowest point
once

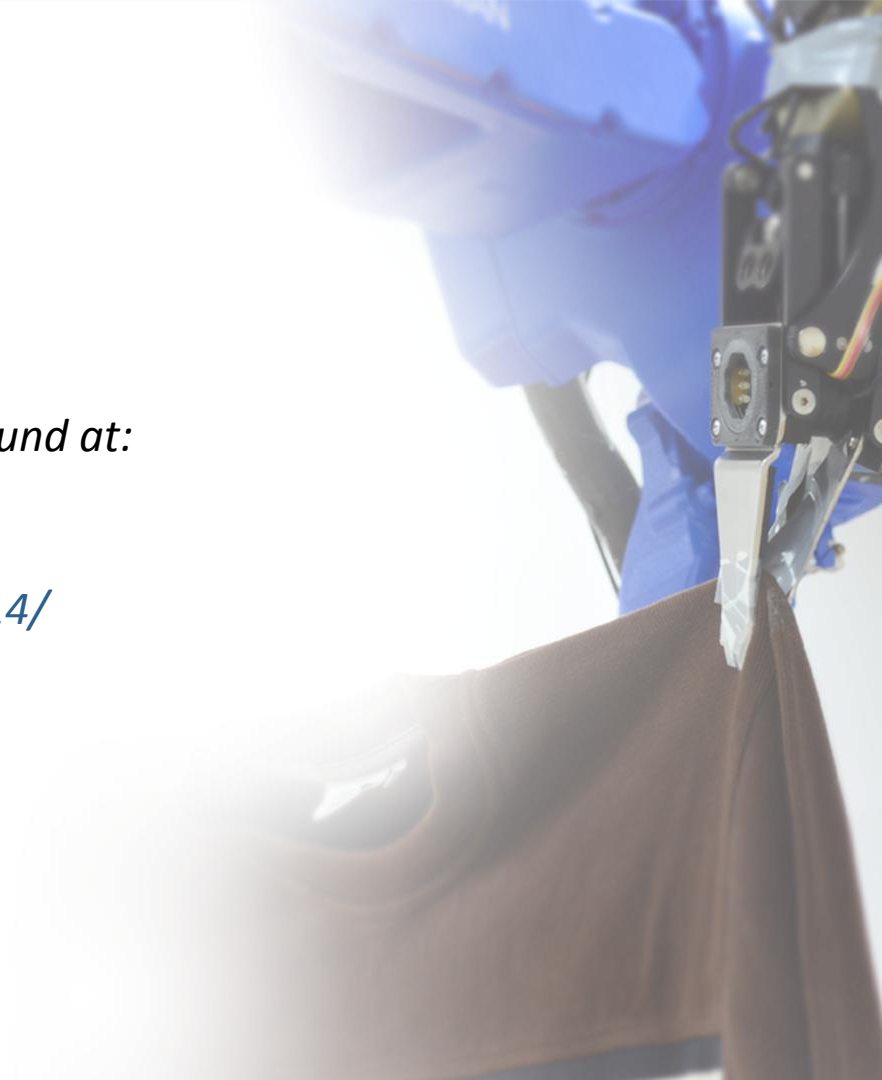
unfolding
in the air
(fast)

regular-sized
clothes

Supplementary Material

Supplementary Material can be found at:

http://clopema.iti.gr/icra_2014/



Thank you!

clopema.eu

certh.gr

imperial.ac.uk