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 $\rightarrow$ 6 Classes Total
Clothes Recognition

How to reduce the large configuration space?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points $\rightarrow$ 6 Classes Total

Training Database:
Clothes Recognition

How to reduce the large configuration space?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points $\rightarrow$ 6 Classes Total

Training Database:
Clothes Recognition

How to reduce the large configuration space?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points $\rightarrow$ 6 Classes Total

Training Database:
Clothes Recognition

How to reduce the large configuration space?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points $\rightarrow$ 6 Classes Total

Training Database:

Grasp lowest point $\rightarrow$ 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:
Clothes Recognition

How to reduce the large configuration space?

Grasp Lowest Hanging Point First:

6 non-symmetric lowest points $\rightarrow$ 6 Classes Total

Training Database:

Grasp lowest point $\rightarrow$ Rotate 360° $\rightarrow$ 28,800 Training Images
Clothes Recognition

Random Forest Training
Clothes Recognition

Random Forest Training

Cloth Training Samples
Clothes Recognition

Random Forest Training

Tree 1

\( h(V,C) \)

\( h(V,C) \)

\( \ldots \)

\( h(V,C) \)

\( h(V,C) \)

\( h(V,C) \)

\( h(V,C) \)

\( \bigcirc \)

\( \bigcirc \)

\( \bigcirc \)

\( \bigcirc \)

\( \bigcirc \)

\( \bigcirc \)

\( \bigcirc \)
Clothes Recognition

Random Forest  Training

Cloth Training Samples

Tree 1  Tree N

$f(V, C_i)$  Split Function
Clothes Recognition

Random Forest  

Training

Cloth Training Samples

Tree 1  Tree N

\( f(V, C_i) \)  Split Function

Depth Channel

\[
f(V, C_1) = d_u - d_v \\
f(V, C_1) = (d_u - d_w) - (d_w - d_v)
\]
Clothes Recognition

Random Forest  Training

Cloth Training Samples

Tree 1  Tree N

\( f(V, C_i) \)  Split Function

Depth Channel

\[
\begin{align*}
    f(V, C_1) &= d_u - d_v \\
    f(V, C_1) &= (d_u - d_w) - (d_w - d_v)
\end{align*}
\]

Curvature Channel

\[
f(V, C_2) = |c_u|
\]
Clothes Recognition

Random Forest Training

Cloth Training Samples

Tree 1

Tree N

\( f(V, C_i) \) Optimization
Clothes Recognition

Random Forest Training

Cloth Training Samples

Tree 1

Tree N

\( f(V, C_t) \) Optimization

Generate Random Set \( \{f, t\} \)
Clothes Recognition

**Random Forest**

**Training**

$f(V, C_i)$ Optimization

Generate Random Set \( \{f, t\} \)

Split Samples

\[ f(V, C_i) > t \quad f(V, C_i) < t \]

Tree 1

Tree N

\( f(V, C) \)
Clothes Recognition

Random Forest  

Training

Cloth Training Samples

Tree 1

Tree N

$f(V, C_i)$ Optimization

Generate Random Set  \( \{ f, t \} \)

Split Samples

\( f(V, C_i) > t \)  \( f(V, C_i) < t \)

Measure Entropy

\[
H = - (\left| L \right| p_L \log p_L + \left| R \right| p_R \log p_R)
\]
Clothes Recognition

Random Forest

Training

Cloth Training Samples

Tree 1

Tree N

\( f(V, C_i) \) Optimization

Generate Random Set \( \{f, t\} \)

Split Samples

\( f(V, C_i) > t \)

\( f(V, C_i) < t \)

Measure Entropy

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

Greedily Optimize

\[
\{f, t\} = \arg\min_{f, t} H
\]
Clothes Recognition

Random Forest  Training
Clothes Recognition

Random Forest

Testing

Tree 1

Test Garment

Tree N
Clothes Recognition

Random Forest  Testing
Clothes Recognition

Random Forest

Testing

Test Garment

Tree 1

Tree N

$P(c)$ $P(c)$ $P(c)$ $P(c)$ $P(c)$ $P(c)$ $P(c)$ $P(c)$
Clothes Recognition

Random Forest

Testing
Clothes Recognition

Random Forest

Testing

Tree 1

Tree N

Test Garment

\[ f(V,C) \]

\[ f(V,C) \]

\[ f(V,C) \]

\[ f(V,C) \]
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

\[ f(V, C) \]

\[ \mathcal{L} = \{ P_1, P_2, P_3, \text{INV} \} \]
Grasp Point Detection

Desired grasp Points:

Hough Forest

\[ f(V, C_i) \text{ Optimization} \]

\[ \mathcal{L} = \{ P_1, P_2, P_3, \text{INV} \} \]
Grasp Point Detection

Desired grasp Points:

Hough Forest

\( f(V, C_i) \) Optimization

Minimize:

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

or

\[ D = \sum_{\text{samples}} d(p_s, p_M) \] 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

Test Garment

Tree 1

Tree N

\( f(V,C) \)

\( f(V,C) \)

\( f(V,C) \)

\( f(V,C) \)

\( P(c) \)

\( P(c) \)

\( P(c) \)

\( P(c) \)

\( L \)

\( L \)

\( L \)

\( L \)

\( L \)
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?
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

<table>
<thead>
<tr>
<th>Shirt</th>
<th>Trousers</th>
<th>Shorts 1</th>
<th>Shorts 2</th>
<th>T-shirt 1</th>
<th>T-shirt 2</th>
</tr>
</thead>
</table>

Active Planning

Active Recognition → POMDP solution

States (S): 6 Classes

Actions (A): Rotate Cloth — Take Final Decision
Active Planning

Active Recognition — POMDP solution

States ($S$): 6 Classes

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

Observations ($O$): $P(e)$ (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: \( P(O | S, A) \rightarrow \) 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:** $P(O | S, A) \rightarrow$ Measured Experimentally

**Transition Probabilities:**

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

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

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 ($\mathcal{S}$): 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)
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 ($O$):  $P_{\text{grasppoint}}(i, j)$ (quantized) from Hough Image
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 ($O$): $P_{\text{grasppoint}}(i, j)$ (quantized) from Hough Image

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

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

→ Measured Experimentally
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 ($O$): $P_{\text{grasppoint}}(i, j)$ (quantized) from Hough Image

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

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

→ Measured Experimentally
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

\[A( b(s) ) \rightarrow \text{Optimal Action}\]
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
- \(x6\)  \(x6\)  \(x6\)  \(x6\)  \(x6\)

Testing Database
- \(x3\)  \(x3\)  \(x3\)  \(x3\)

28,800 Images

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

Depth Images Captured with Xtion
1,440 Images
Results
Results

positive examples

negative examples
### Results

<table>
<thead>
<tr>
<th></th>
<th>Experiments</th>
<th>Successful Unfoldings</th>
<th>Successful Recognition</th>
<th>Average rotations</th>
<th>Estimation Errors</th>
<th>Grasping Errors</th>
<th>Average Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shirts</td>
<td>30</td>
<td>27</td>
<td>30</td>
<td>0,8</td>
<td>2</td>
<td>4</td>
<td>150 sec.</td>
</tr>
<tr>
<td>Trousers</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>1,1</td>
<td>0</td>
<td>3</td>
<td>136 sec.</td>
</tr>
<tr>
<td>Shorts</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>2,7</td>
<td>0</td>
<td>2</td>
<td>127 sec.</td>
</tr>
<tr>
<td>T-Shirts</td>
<td>30</td>
<td>25</td>
<td>30</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>161 sec.</td>
</tr>
<tr>
<td>Total</td>
<td>120</td>
<td>112</td>
<td>120</td>
<td>2,4 (avg.)</td>
<td>7</td>
<td>15</td>
<td>143,5 sec. (avg)</td>
</tr>
<tr>
<td>%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>93,3%</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Experiments</th>
<th>Successful Unfoldings</th>
<th>Successful Recognition</th>
<th>Average rotations</th>
<th>Estimation Errors</th>
<th>Grasping Errors</th>
<th>Average Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shirts</strong></td>
<td>30</td>
<td>27</td>
<td>30</td>
<td>0,8</td>
<td>2</td>
<td>4</td>
<td>150 sec.</td>
</tr>
<tr>
<td><strong>Trousers</strong></td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>1,1</td>
<td>0</td>
<td>3</td>
<td>136 sec.</td>
</tr>
<tr>
<td><strong>Shorts</strong></td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>2,7</td>
<td>0</td>
<td>2</td>
<td>127 sec.</td>
</tr>
<tr>
<td><strong>T-Shirts</strong></td>
<td>30</td>
<td>25</td>
<td>30</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>161 sec.</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>120</td>
<td>112</td>
<td>120</td>
<td>2,4 (avg.)</td>
<td>7</td>
<td>15</td>
<td>143,5 sec.</td>
</tr>
<tr>
<td><strong>%</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>93,3%</strong></td>
<td></td>
<td><strong>100%</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 in the air (fast)
- Grasp lowest point once
- Unfolding using table (slow)
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
- Unfolding using table (slow)
- Unfolding in the air (fast)
- Baby clothes
- Regular-sized clothes
Supplementary Material

Supplementary Material can be found at:

http://clopema.iti.gr/icra_2014/
Thank you!
clo pena.eu certh.gr imperial.ac.uk