



Long horizon segmentation

Andrea Pierré

November 3, 2025

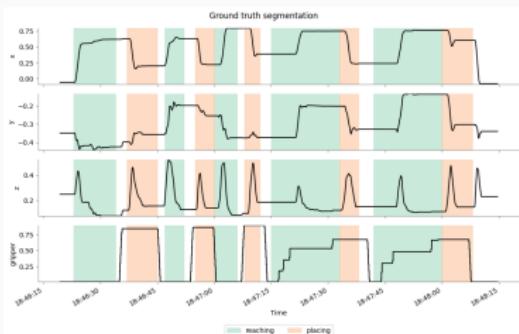
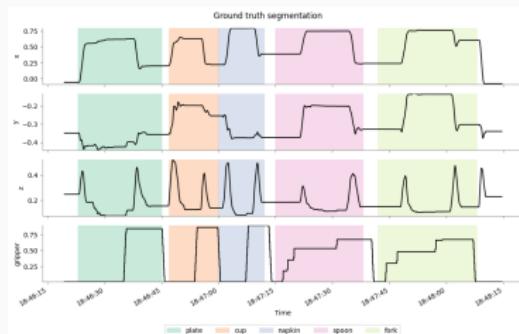
Outline

1. Context
2. Unsupervised learning – State Detection
3. Unsupervised learning – Change Point Detection
4. Supervised learning
5. Conclusion

Outline

1. Context
2. Unsupervised learning – State Detection
3. Unsupervised learning – Change Point Detection
4. Supervised learning
5. Conclusion

Robot learning of long horizon demonstrations

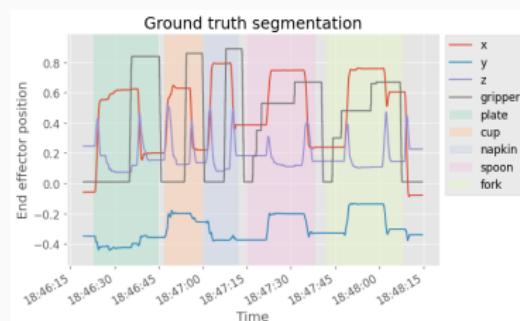
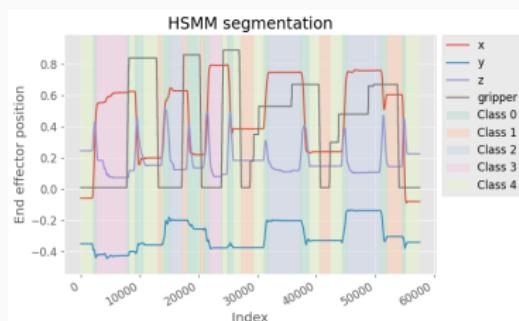
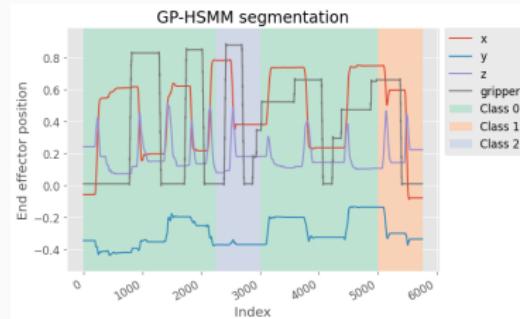
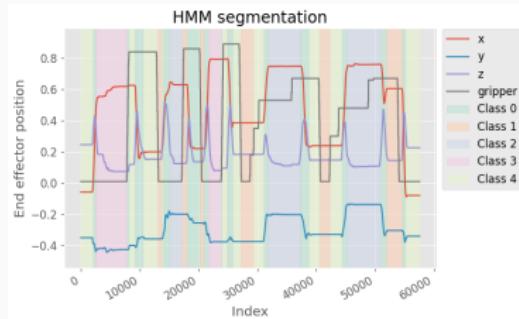


→ Extremely **small** data i.e., few demonstrations

Outline

1. Context
2. Unsupervised learning – State Detection
3. Unsupervised learning – Change Point Detection
4. Supervised learning
5. Conclusion

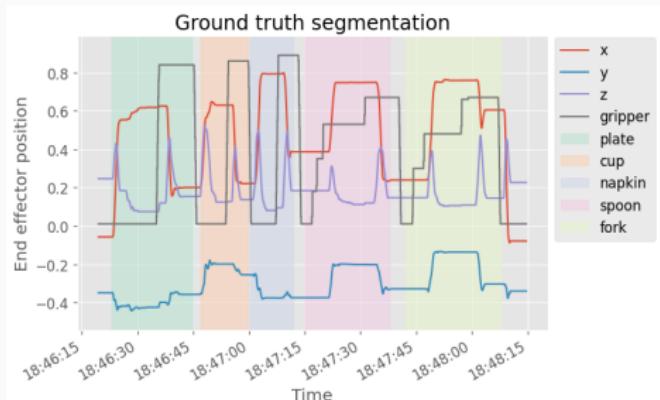
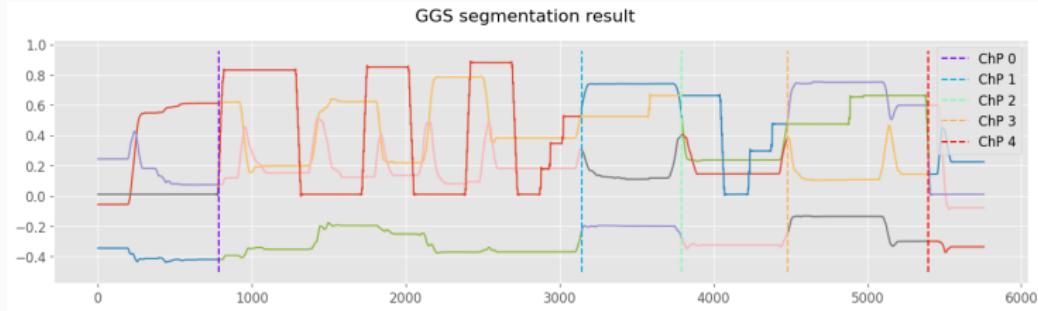
H(S)MM – Higher level movements ✕



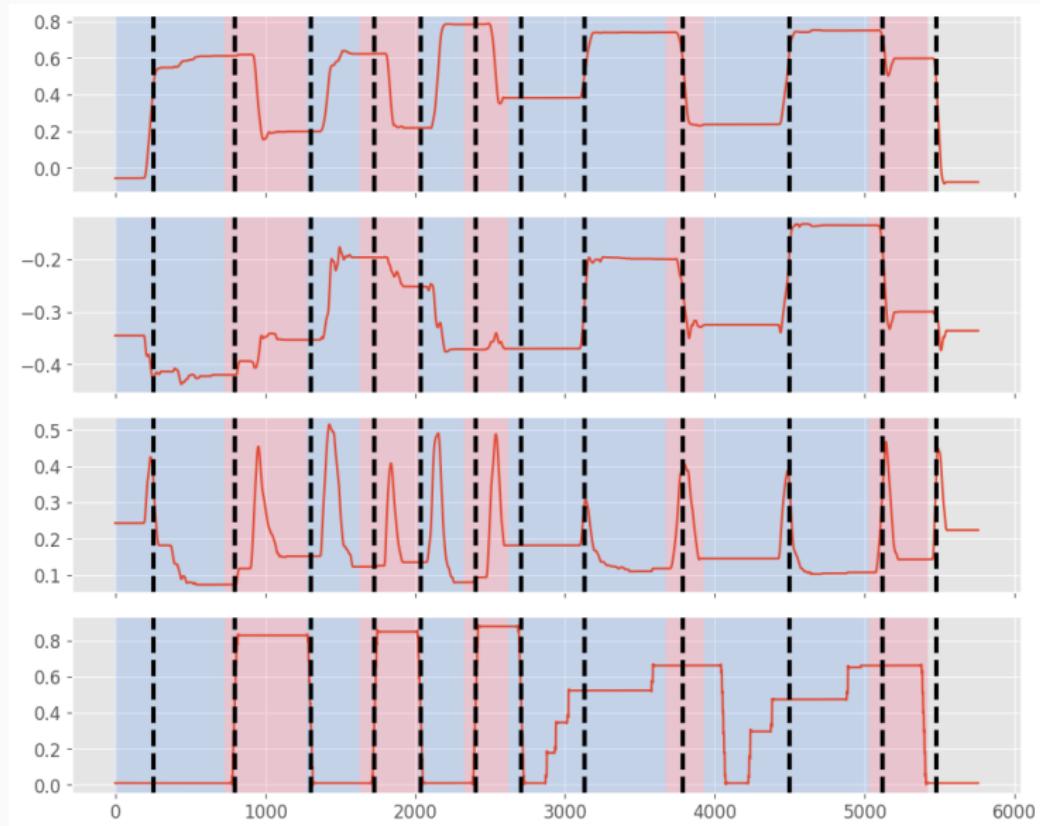
Outline

1. Context
2. Unsupervised learning – State Detection
3. Unsupervised learning – Change Point Detection
4. Supervised learning
5. Conclusion

Greedy Gaussian Segmentation – Higher level movements ✗



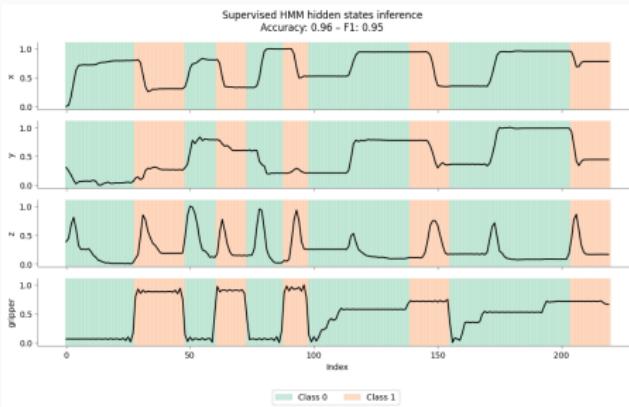
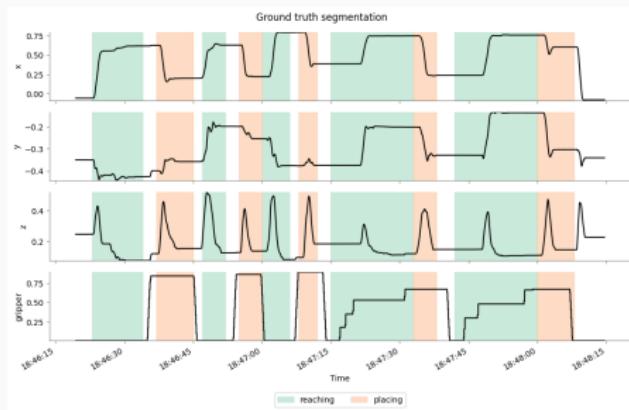
ruptures – Lower level movements



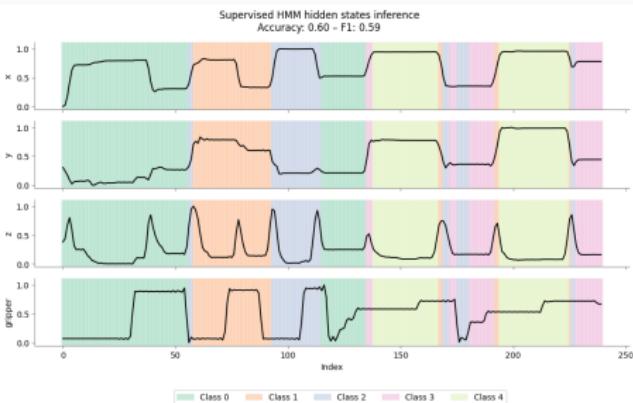
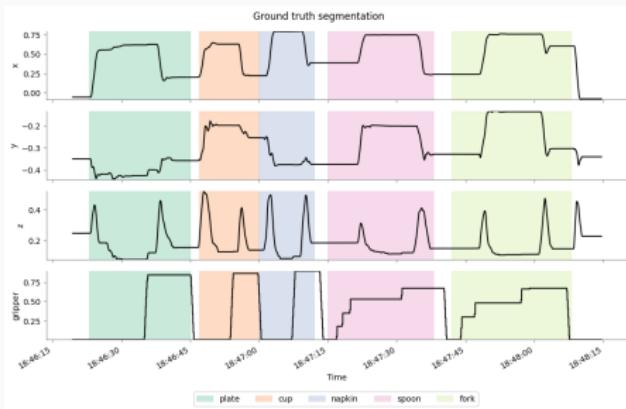
Outline

1. Context
2. Unsupervised learning – State Detection
3. Unsupervised learning – Change Point Detection
4. Supervised learning
5. Conclusion

Supervised HMM – Lower level movements



Supervised HMM – Higher level movements ✕



XGBoost – Higher level movements ✕

Split	F1	Accuracy
1	0.2722	0.3741
2	0.1639	0.1872
3	0.1684	0.3381
4	0.1469	0.3501
5	0.1713	0.3458

Outline

1. Context
2. Unsupervised learning – State Detection
3. Unsupervised learning – Change Point Detection
4. Supervised learning
5. Conclusion

Conclusion

- Can model *lower* level movements with good performance
- Modeling *higher* level movements remains a challenge

Questions

- Any other method I may have missed that could help solve this problem?
 - Variational inference?
- How to integrate context, *e.g.*, video?
 - Multimodal transformer but probably too data hungry?

Questions ?