Karlsruhe Institute of Technology

Cv:hci
Institute for Anthropomathics

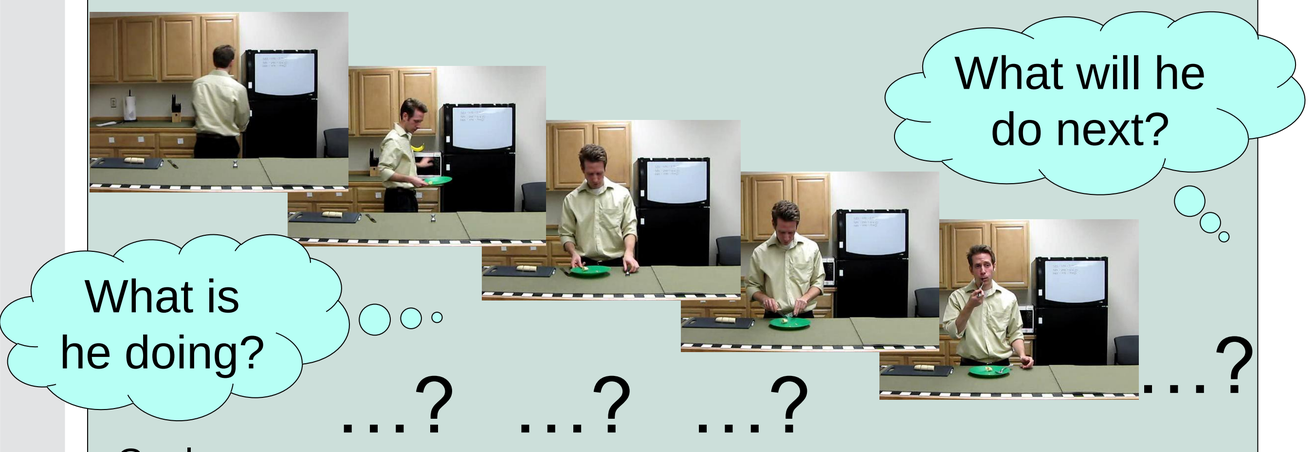
ANALYSING COMPLEX ACTIVITIES IN VIDEO SEQUENCES

Hildegard Kuehne (kuehne@kit.edu), Dirk Gehrig (dirk.gehrig@kit.edu)

Abstract

Video-based applications rely more and more on the fast and reliable recognition of human actions. Although current recognition algorithms compete for better recognition performance, their usability for real-world applications is still limited in terms of complexity, runtime and robustness. In this context, a system for the continuous online recognition of human actions from video-based motion information is proposed. It uses histograms of sparse feature point flow with a Hidden-Markov-Model (HMM) based decoder system inspired from speech recognition. To allow a recognition of ongoing tasks, complex sequences are split into action units and during the recognition assembled by a context free grammar. The evaluation of the presented system shows a good performance even compared to marker-based recognition approaches.

Understanding complex tasks



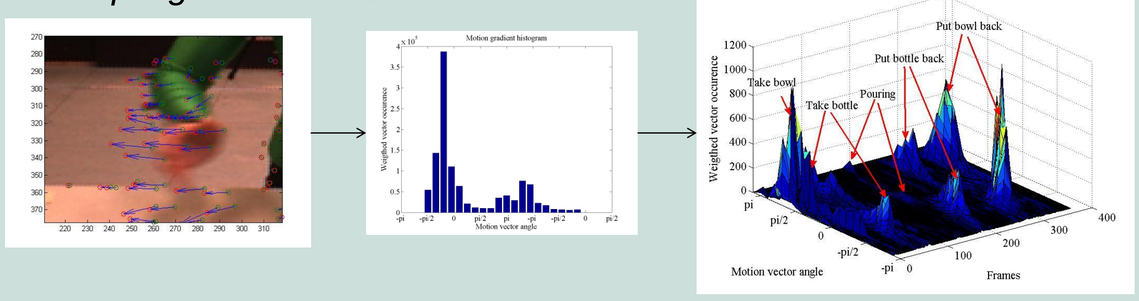
Goals:
1) Continuous analysis & recognition
- Recognition before tasks are finished
2) Anticipation
- Interpretation of what the user aims at

Recognition system

Action recognition:
- On the level of action units = HMM modeling
- Action unit = a 4-state left-to-right HMM
- On the level of action sequences = Combination of HMMs with a context free grammar
- Implicit automatic segmentation of the action sequences into action units during recognition

Video processing: Flow histograms

- Calculate motion vector of feature points of images *I(t)* and *I(t+1).*
→ motion direction θ and motion intensity γ
- Weighted histogram for frame t :
k-th bin = sum the intensity of all motion vectors with direction from *[1/n (k2Π), 1/n ((k+1)2Π)].*
- Sampling: over 20 - 89 bins



Evaluation

Datasets :
- Dataset I :10 tasks,48 action units, 20-30 samples/task. [1]
-'Activities of Daily Living Dataset' : 10 tasks, 71 action units. , 15 samples/task. [2]

Performance of optical flow and feature flow:

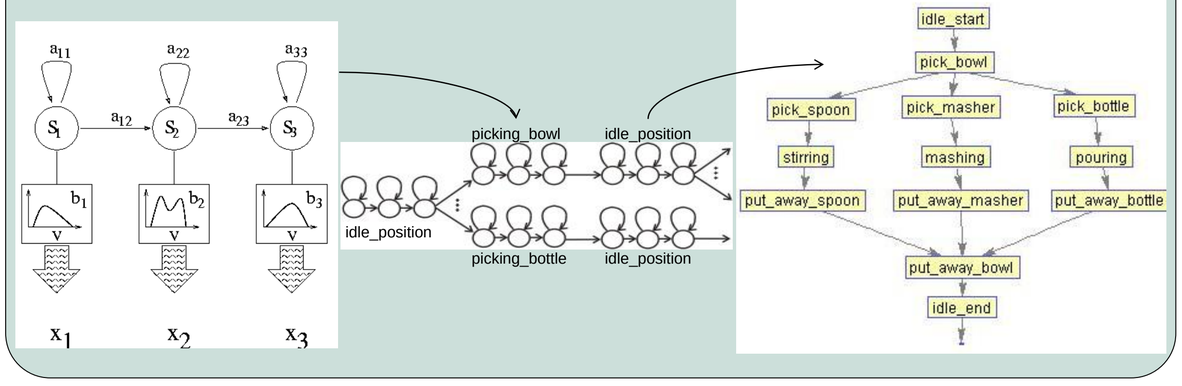
Dataset I	HoOF	HoFF
Sequence recog.	100.0 %	100.0 %
Unit recog.	97.7 %	96.6 %
ADL Dataset	**HoOF**	**HoFF**
Sequence recog.	82.0 %	71.3 %
Unit recog.	63.5 %	55.0 %

Performance of flow compared to 24 joint angles :
- over 5 tasks simultaneously recorded with Vicon

	Marker based	HoOF	HoFF
Input vector dimension	24	30	30
Gaussians per state	16	16	16
States per action unit	4	4	4
Unit recog.	98.3 %	96.9 %	97.5 %
Sequence recog.	100.0 %	100.0 %	100.0 %

Action units and grammar

- Action units = a motion that is performed without interruption → the smallest entity, whose order can be changed during the execution
-Action grammar = All meaningful tasks are executed in order → formulate a grammar



References

[1] Dirk Gehrig, Hildegard Kühne, Annika Wörner, Tanja Schultz, „HMM-based Human Motion Recognition with Optical Flow Data" Humanoids 2009, Paris, France
[2] C. Pal R. Messing and H. Kautz, "Activity recognition using the velocity histories of tracked keypoints," ICCV, 2009, IEEE Computer Society.