Ergonomic Gestures Recognition

Tom Forrer

DIVA

27.01.2011

Outline

Ergonomic gestures

Gestures for controlling an application that have:

- limited action space
- precision
- wrist or arm support
- comfort

Goals

- Recognize spatial and temporal aspects of ergonomic gestures
- Build an architecture capable of real-time gesture recognition
- Implement a demonstration program using this architecture

State of the art: tracking

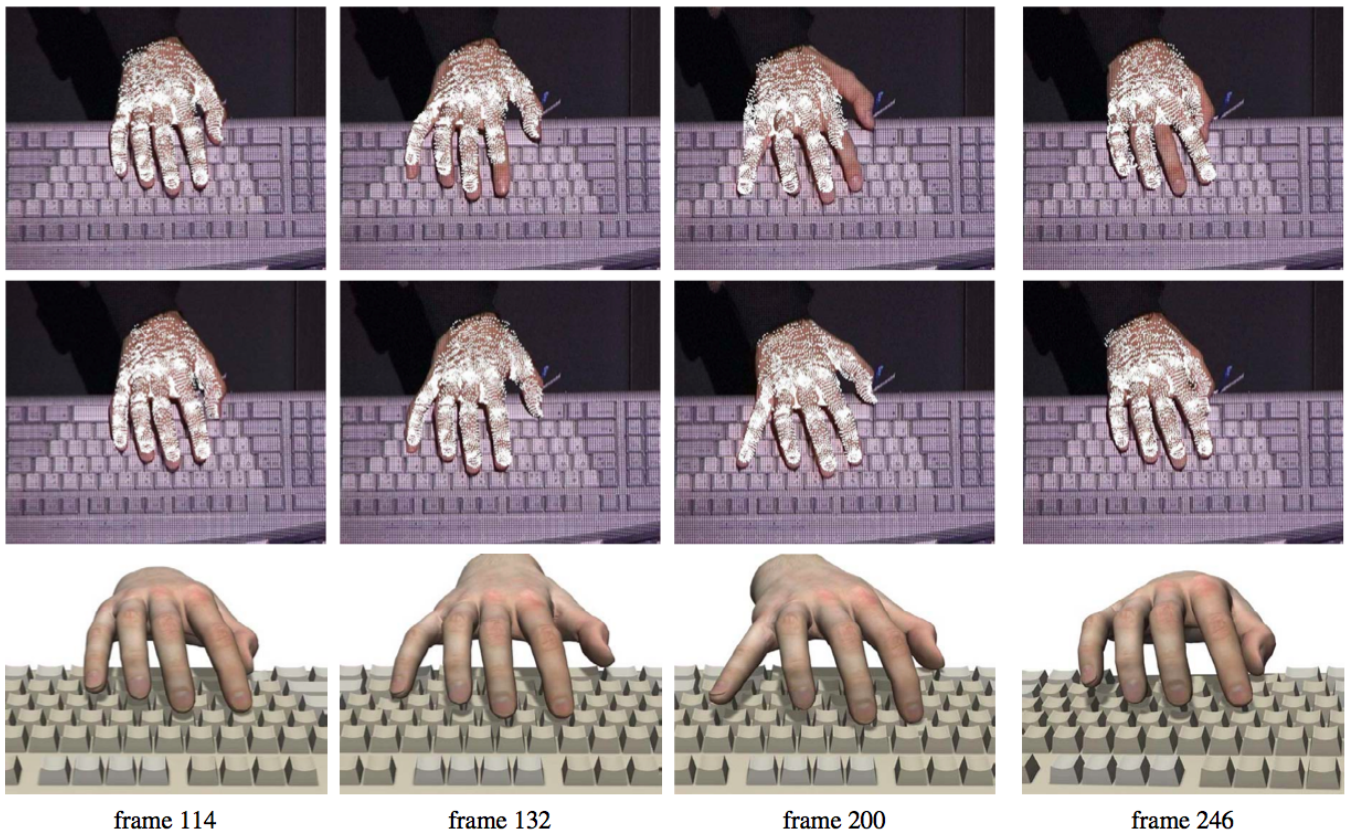


Figure: Bray, Koller-Meier, Van Gool: SPF particle filter, 3D hand-model (15 DOF)

State of the art: tracking

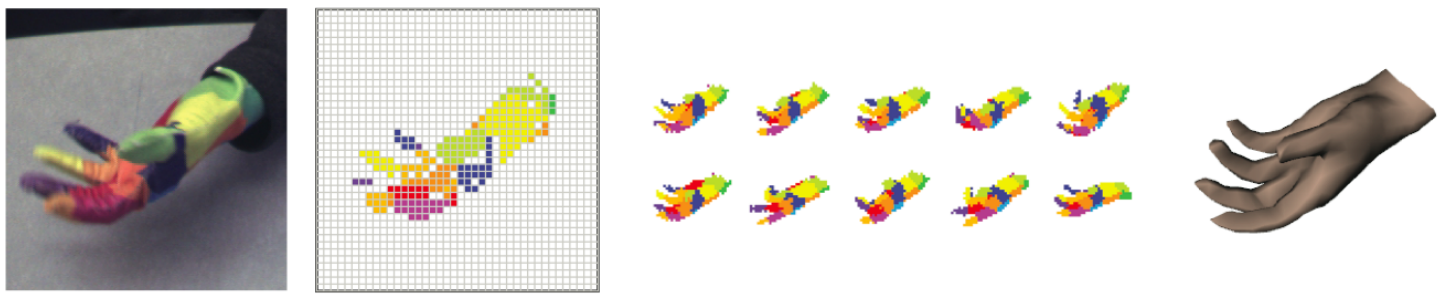


Figure: Wang, Popovic: Real-time hand tracking with color glove

State of the art: recognition

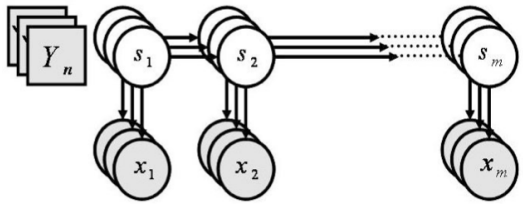


Figure: Hidden Markov Model

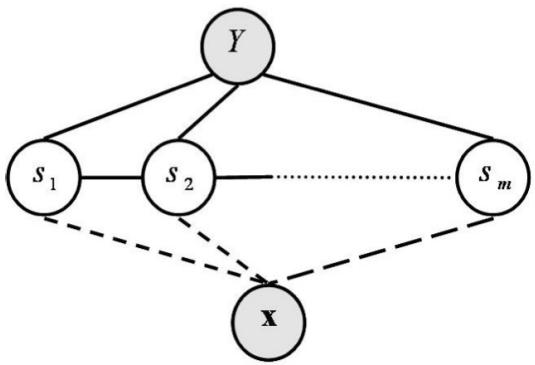


Figure: Sy Bor Wang et al: Hidden conditional random fields

Camera

Figure: Playstation 3 Eye Camera

- Resolution: 640×480
- Frame rate: 60 fps @ 640×480, 120 fps @ 320×240
- Color and image fidelity
- Price

Glove

Figure: Textile glove with color markers

- Color-marked finger positions (green)
- Color-marked wrist position (red)
- Easy to put on and off

Setup

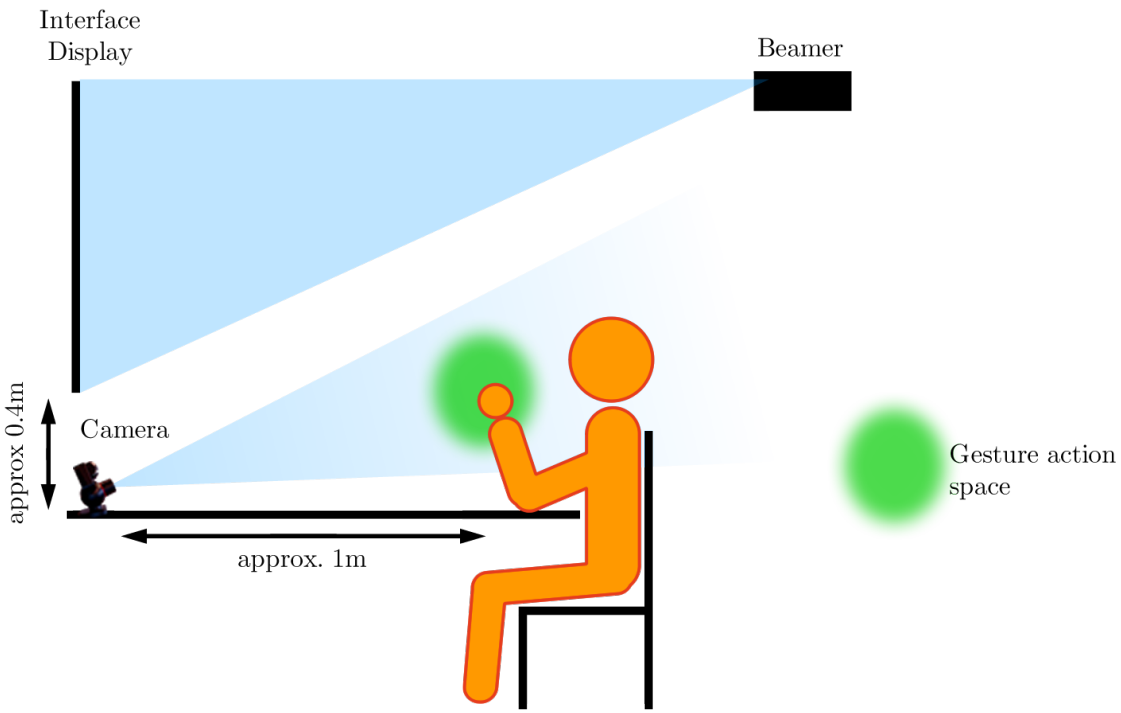


Figure: Environmental setup

Software

- Visual Studio 2008: C++
- OpenCV 2.0
- dlib C++ ML
- Qt 4.7
- libconfig

Chosen gestures

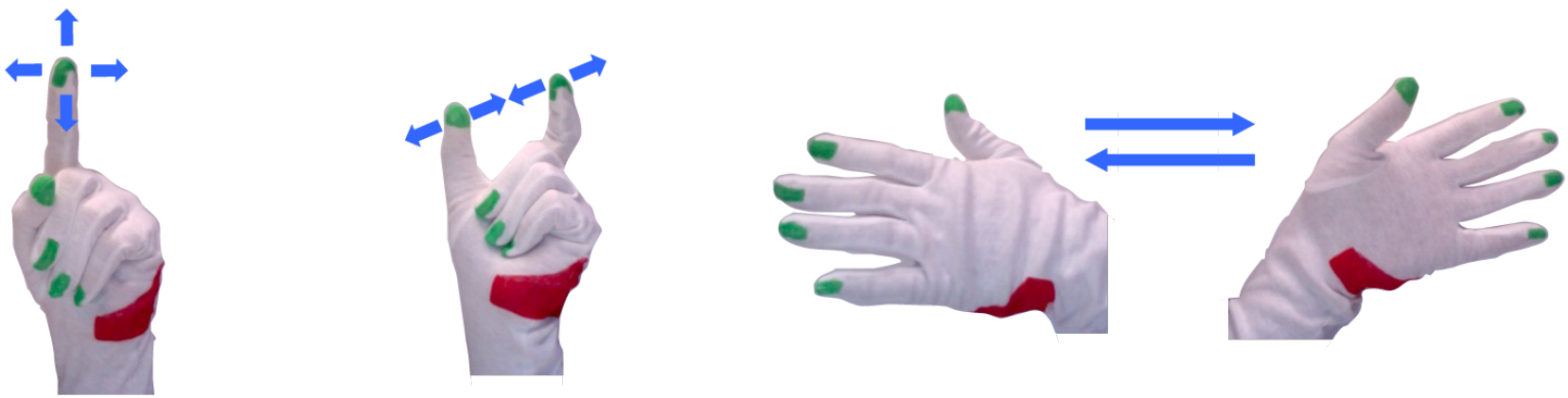
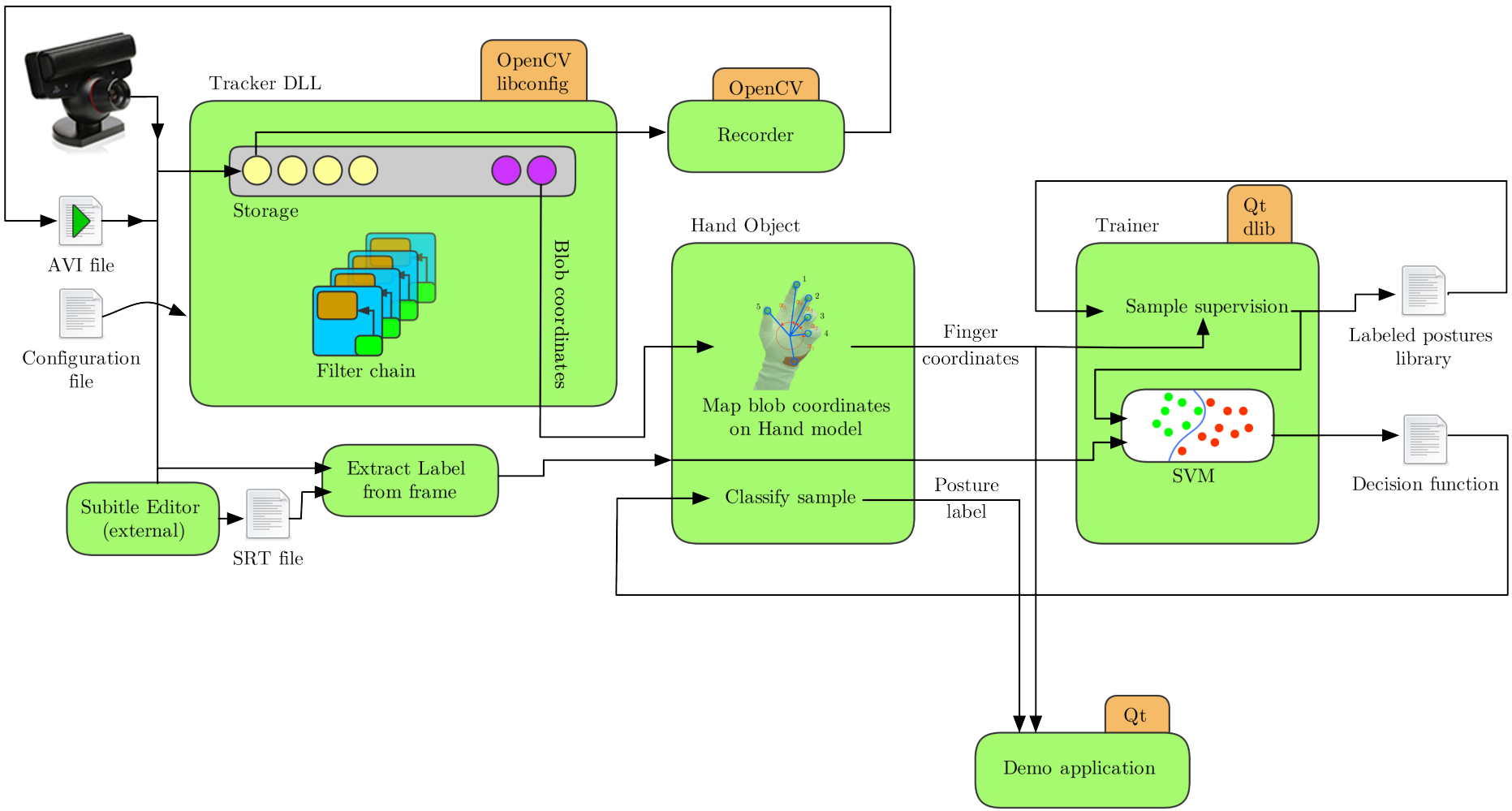


Figure: Pointing, zooming and horizontal swiping gesture

Architecture



Tracker configuration

- reconfigurable filter arrangement
- on-line parameter adjustement
- optimized for speed
- extendable image processing operations list: Lab Thresholding, RGB Thresholding, Binary Thresholding, Erosion, Dilation, Smooth, Blob detection, FAST corner detection, Add

Filter chain

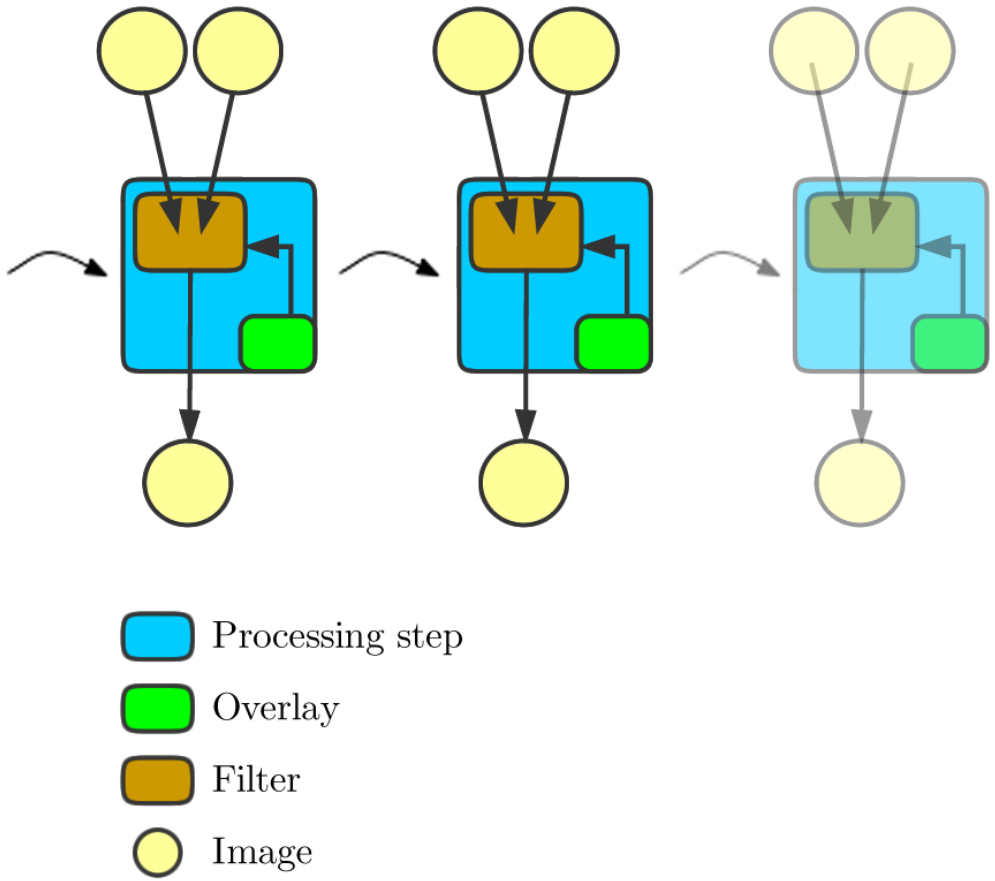


Figure: Chain of reconfigurable processing elements

Filter chain

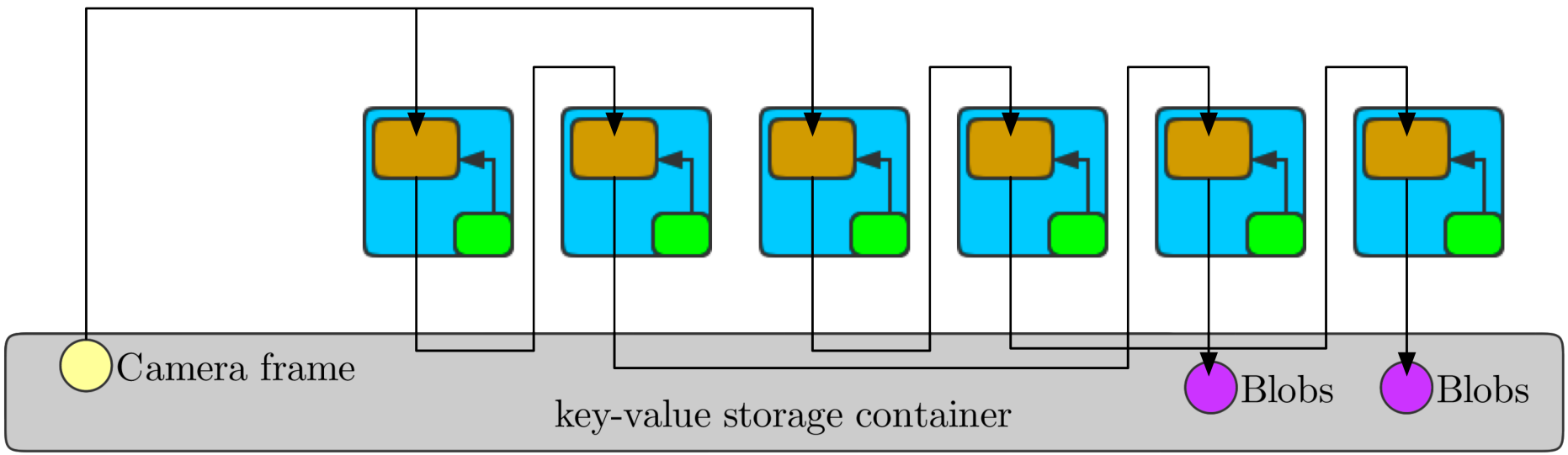


Figure: Chain of reconfigurable processing elements

Filter chain

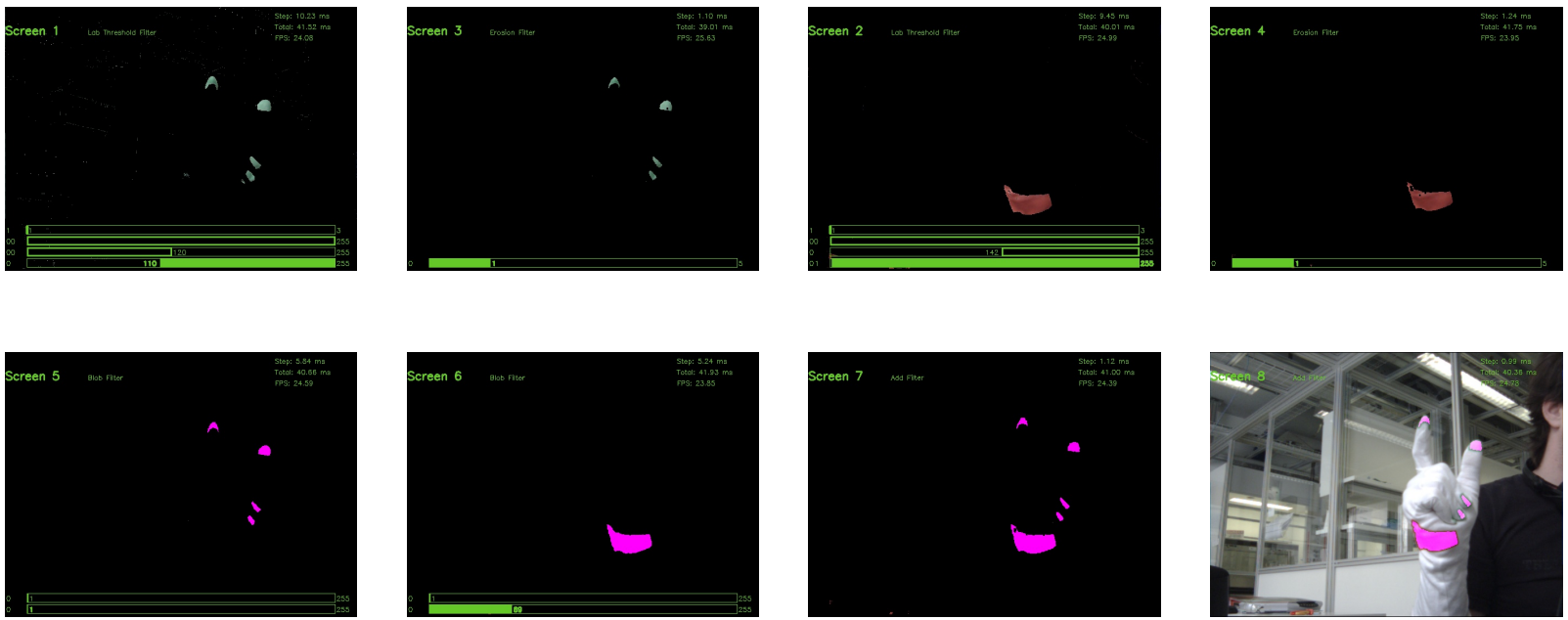


Figure: Resulting images from tracker with filter chain

Hand model

- Five fingers
- One wrist
- Fingers not crossed
- Angle
- Finger and wrist velocity

Mapping

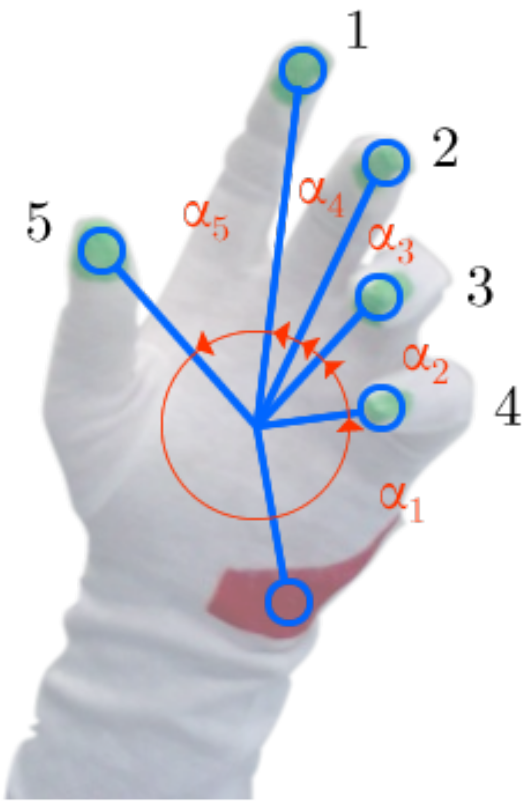


Figure: Hand-model mapping

Sample collection

- Record video of gestures with glove while respecting setup constraints
- Annotate postures with subtitle editor
- Collect labeled postures through tracker, hand model and subtitle file
- 55 seconds, 23 gestures, 23 FPS, 1251 postures

Sample supervision

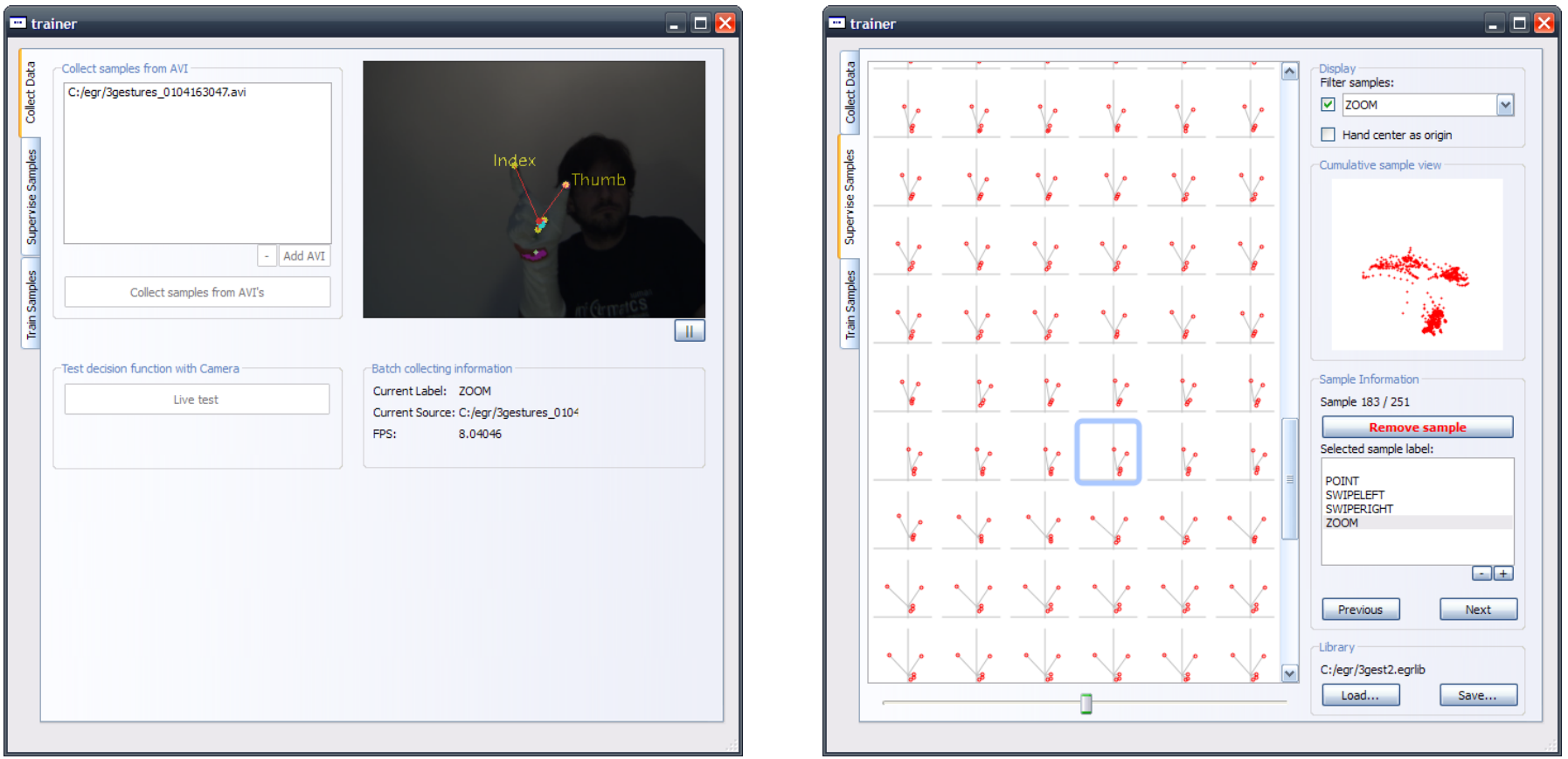


Figure: Sample collection and supervision

Sample supervision

Figure: Sample collection and supervision

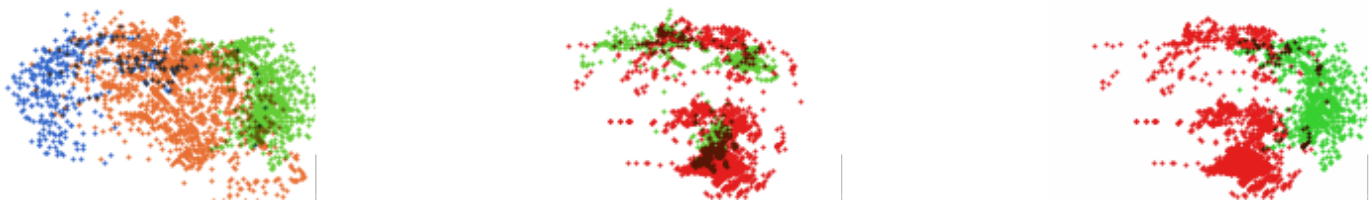


Figure: Sample collection and supervision

Posture training

Train a one-vs-one multiclass ν-Support Vector Machine with polynomial kernels

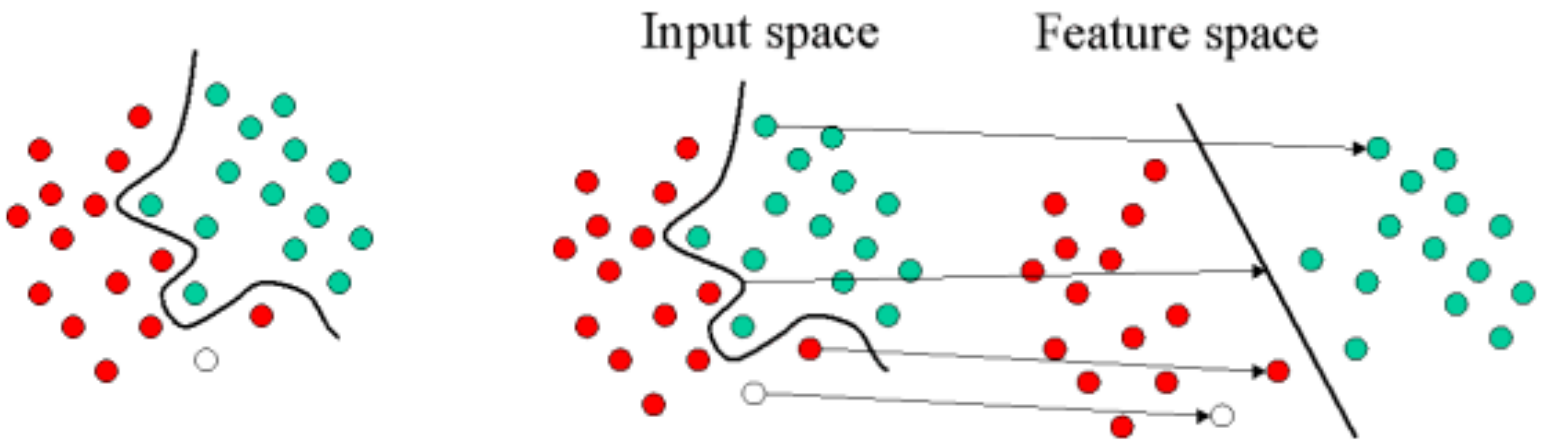


Figure: Principle of support vector machines

Gesture classification

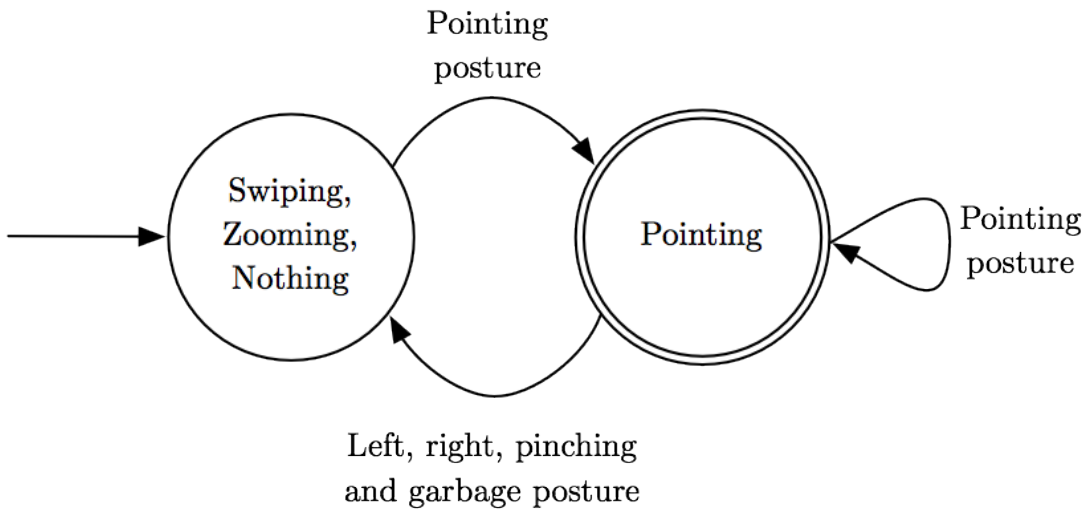


Figure: States of pointing gesture

Gesture classification

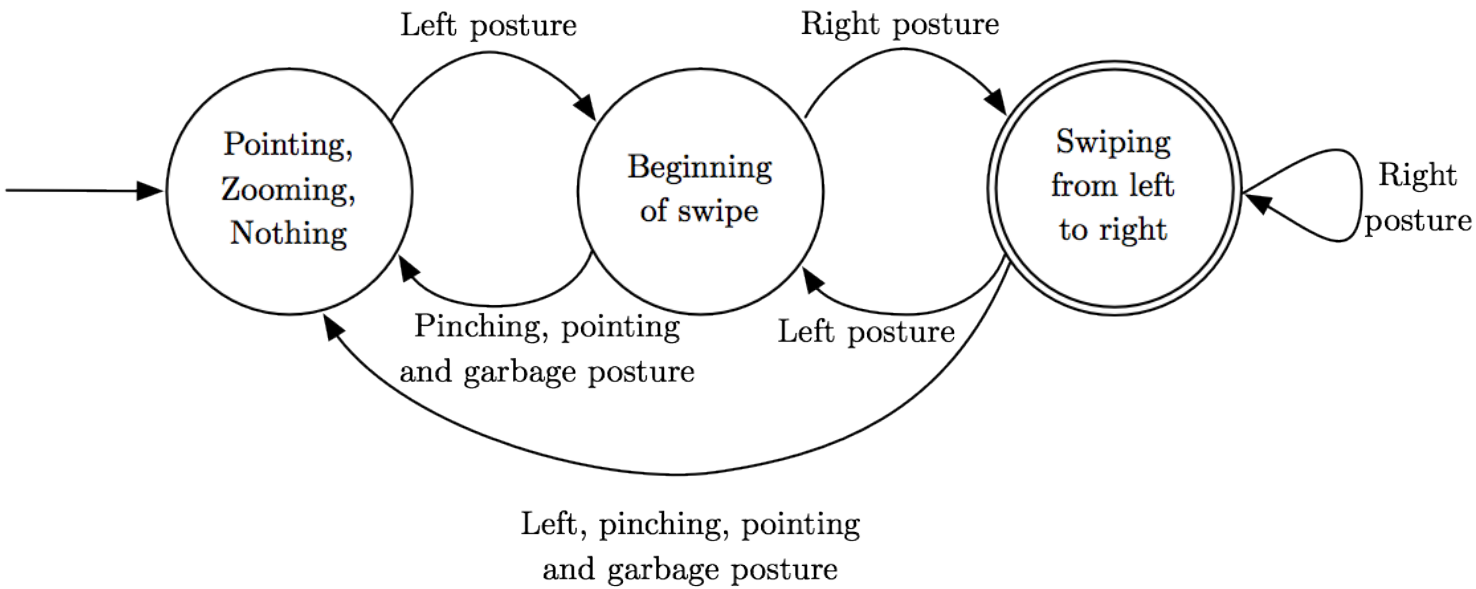


Figure: States of swiping gesture

Evaluation

Table: Postures and gestures of training video

gesture name	**number of gestures**	**number of postures**
Nothing	24	317
Pointing	5	397
Zooming	6	253
Swiping	12	78 (left)
		174 (right)
Total	**47**	**1219**

Evaluation

Table: Confusion matrix of the trained decision function

	Garbage	Pointing	Left	Right	Zoom
Garbage	283	11	8	12	3
Pointing	14	383	0	0	0
Left	6	0	71	1	0
Right	8	0	0	166	0
Zoom	4	1	0	0	247

Demo

Conclusion

- Gesture recognition system has fast reaction time ($<$0.4s)
- Gestures are easy to learn by other users
- Limitations: color markers have always to be seen
- Gestural interfaces should rely on ergonomic features

Future extensions

- train more postures, investigate parametrization of SVM
- use HMM to classify gestures with more complex spatial and temporal dependencies
- replace tracking module with some other real-time capable hand-tracking system without gloves

Thank you!