Rewards-Driven Control of Robot Arm by Decoding EEG Signals

Ajay K. Tanwani[*], José del R. Millán[**], Aude Billard[*]

{ajay.tanwani, jose.millan, aude.billard}@epfl.ch

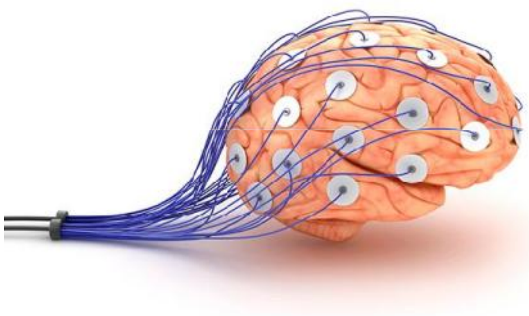
[*] Learning Algorithms and Systems Laboratory (LASA), EPFL, Switzerland

[**] Chair in Non-Invasive Brain-Machine Interface (CNBI), EPFL, Switzerland

Problem Scope

High-dimensionality

Reward function



Brain Controlled Prosthesis

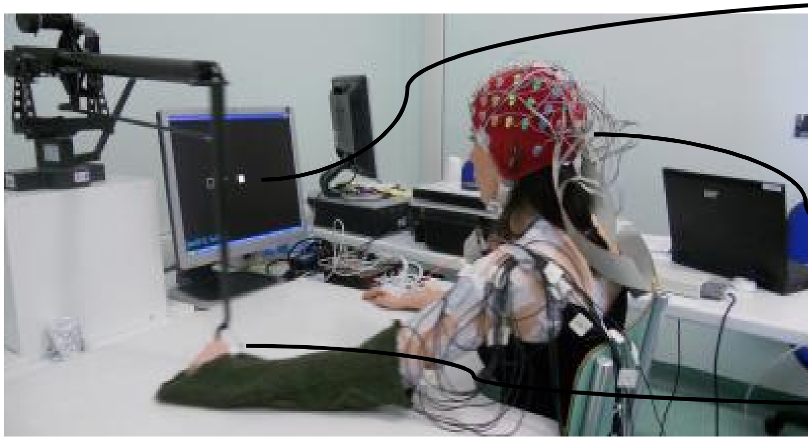
Non-stationary

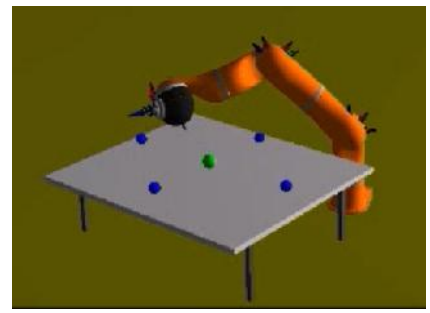
Large variability across users

Experimental Set-Up

- Self-paced planar center-out reaching movement task

- One healthy subject performed 3 runs for a total of 230 trials in the interval [-2 1] seconds





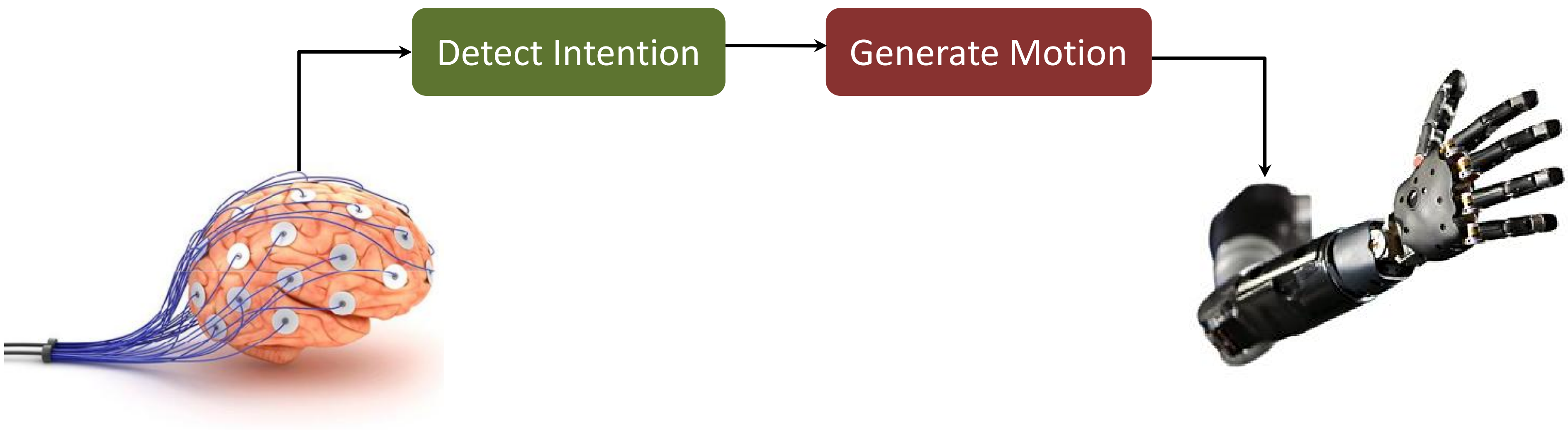
64 Channel EEG Data (2 kHz)

Kinematic Data (100 Hz)

$\{\boldsymbol{x}, \dot{\boldsymbol{x}} \in \mathbb{R}^2\}$ Cartesian planar position and velocity of robot arm

$\boldsymbol{u} \in \mathbb{R}^{64}$ 64 Channel EEG data

Framework

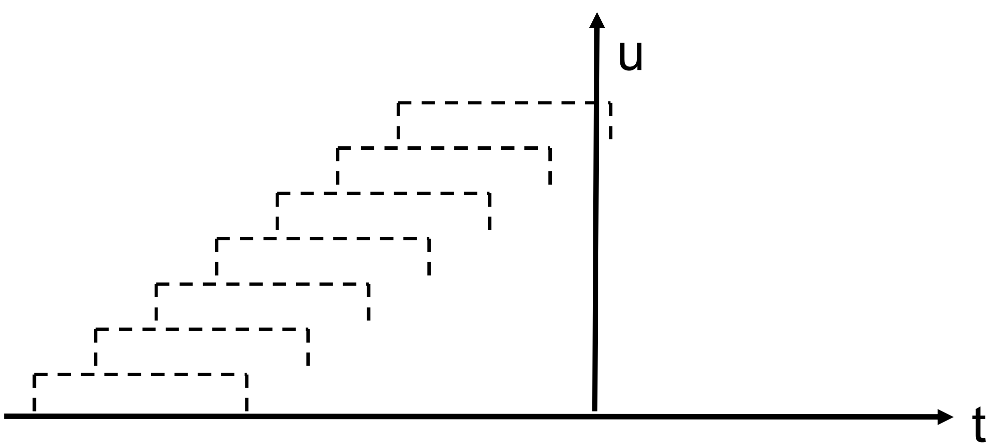


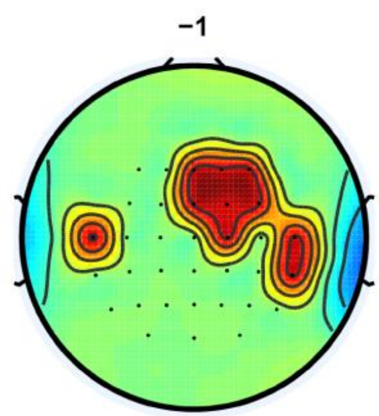
Intention Decoder

- EEG data is sampled with moving window of 250 ms overlapping every 62.5 ms

- 10 features with the highest discriminating power for every window using Canonical Variant Analysis (CVA)

- Brain activity is dominant in frontal-parietal regions





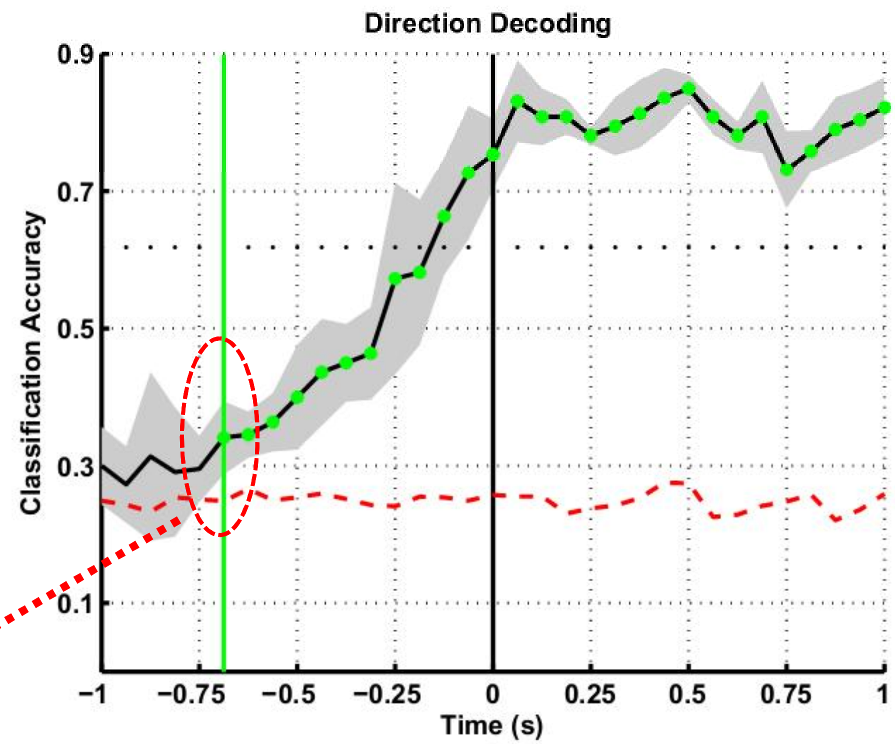
Intention Decoder

EEG Signals → Feature Selection → Classification

- LDA classifier is used to predict goal estimate x_{g_t} given the EEG feature vector in every time window

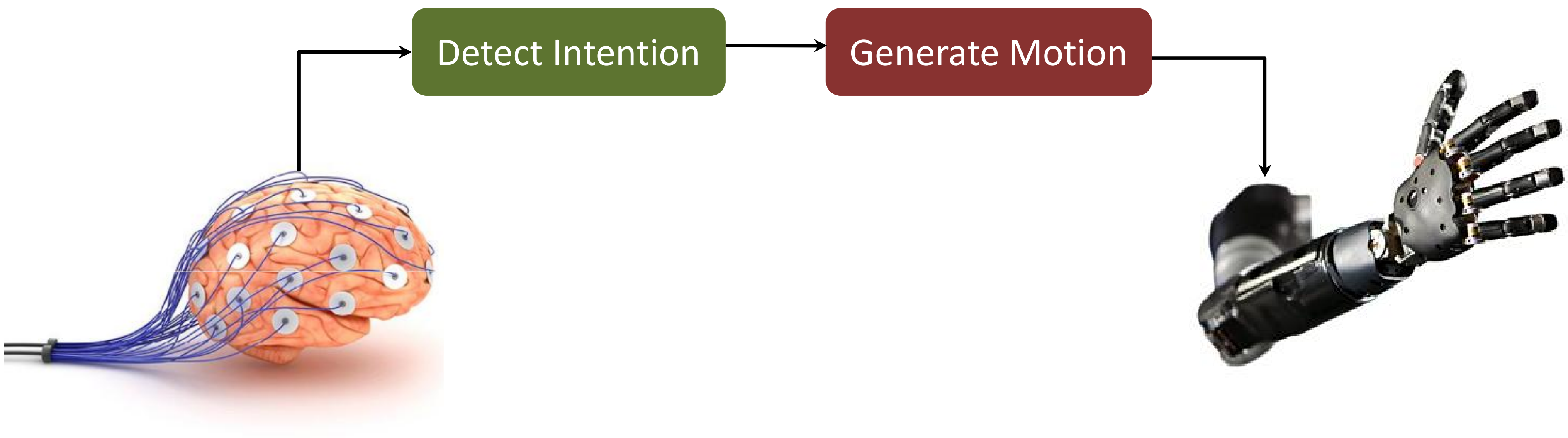
$$x_{g_t} = f(u_l) = \arg\max_{i=1...4} P(C = x_g^{(i)} | u_l)$$

- Goal is estimated starting from 687.5 ms before the movement onset



Classification accuracy exceeds chance level

Framework



Trajectory Decoder

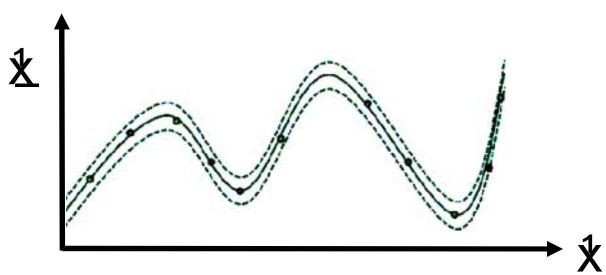
Kinematic data → Initial Policy → Optimized Policy

$$\dot{x} = f_i(\bar{x})$$

- Collect data $\{\boldsymbol{x}, \dot{\boldsymbol{x}}\}$ of 230 trials re-sampled at 5 Hz

- Learn the initial policy $\dot{\boldsymbol{x}} = f_i(\bar{\boldsymbol{x}})$ using Support Vector Regression (SVR), $\bar{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x}_g$

- Hyper-parameters of SVR found through grid search, $C = 1; \gamma = 0.5; \epsilon = 0.5$



Trajectory Decoder

$$\dot{\bar{x}} = f_i(\bar{x}) \qquad \dot{\bar{x}} = f_o(\bar{x})$$

- Support vectors of initial policy act as basis functions for optimized policy

- Weights of the support vector are optimized by stochastic gradient ascent on value function $J(\bar{x})$

$$J(\bar{x}) = \frac{1}{T} \sum_{t=0}^{T} r(\bar{x})$$

- Reward function for reaching the goal

$$r(\bar{x}) = -w_1 \bar{x}_f^T \bar{x}_f - w_2 \dot{\bar{x}}_f^T \dot{\bar{x}}_f - w_3 \ddot{\bar{x}}_t^T \ddot{\bar{x}}_t$$

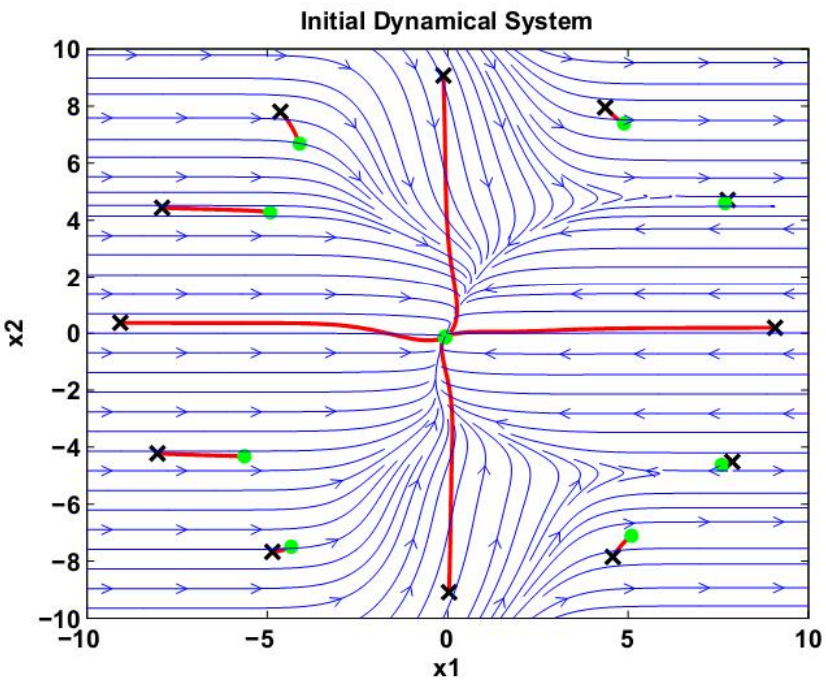
reach goal position at t=T

zero final velocity at t=T

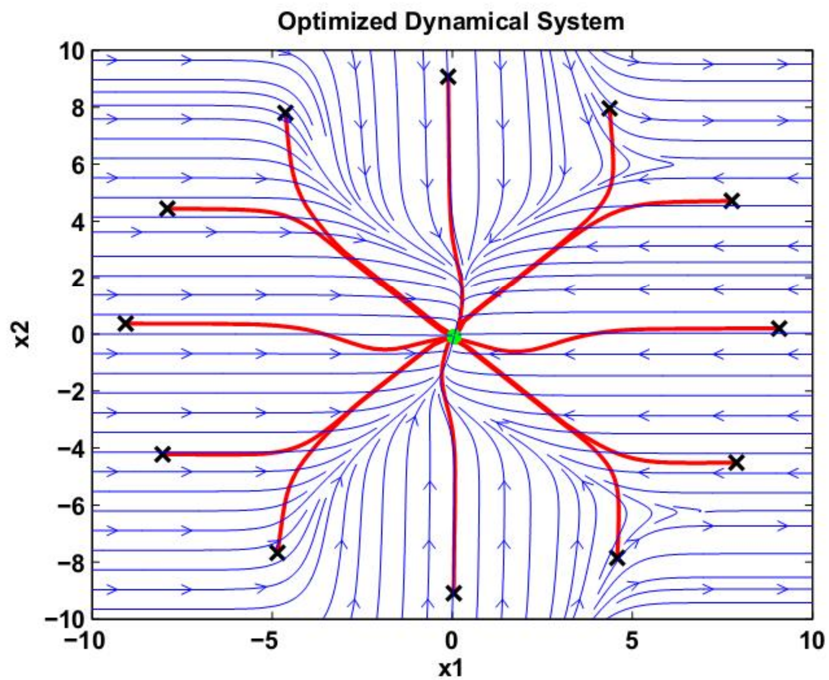
minimize squared acceleration norm during movement

Trajectory Decoder

$$\dot{\bar{x}} = f_i(\bar{x})$$

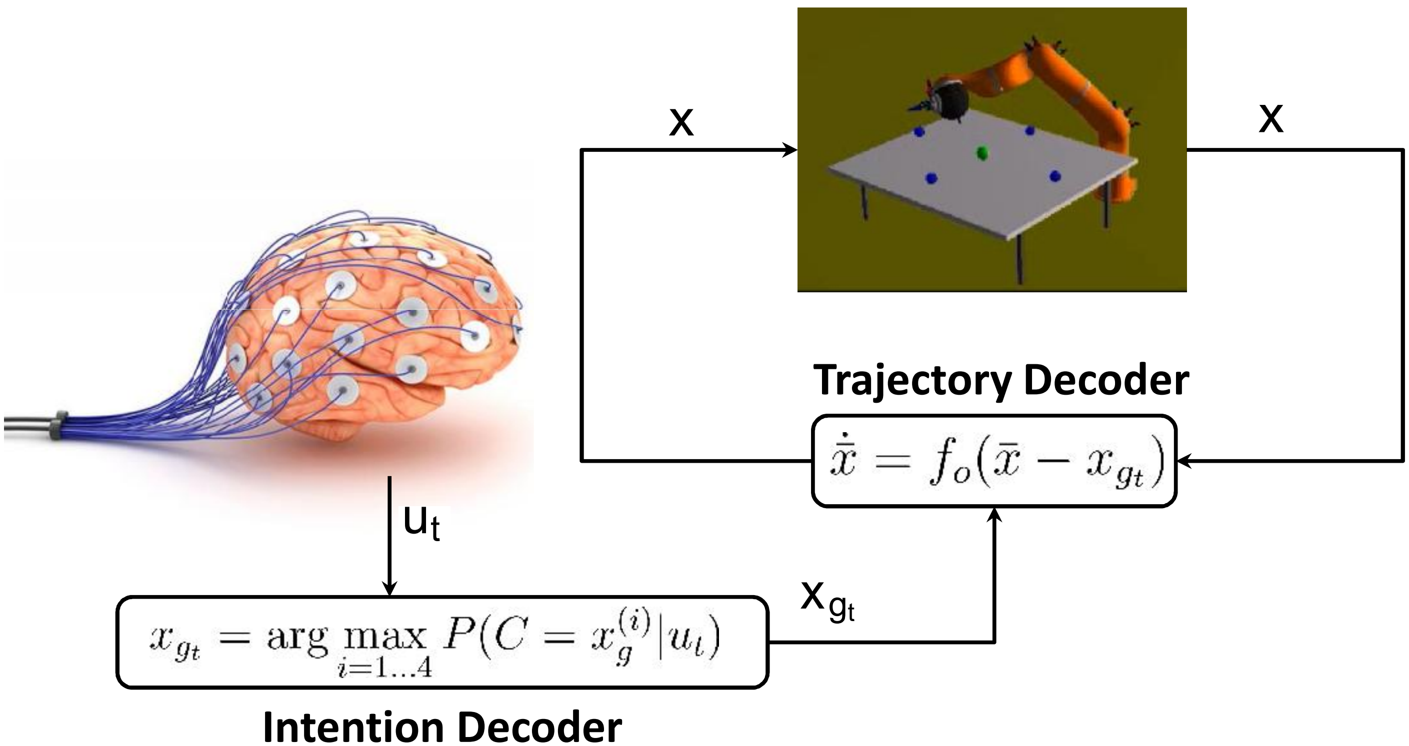


$$\dot{\bar{x}} = f_o(\bar{x})$$

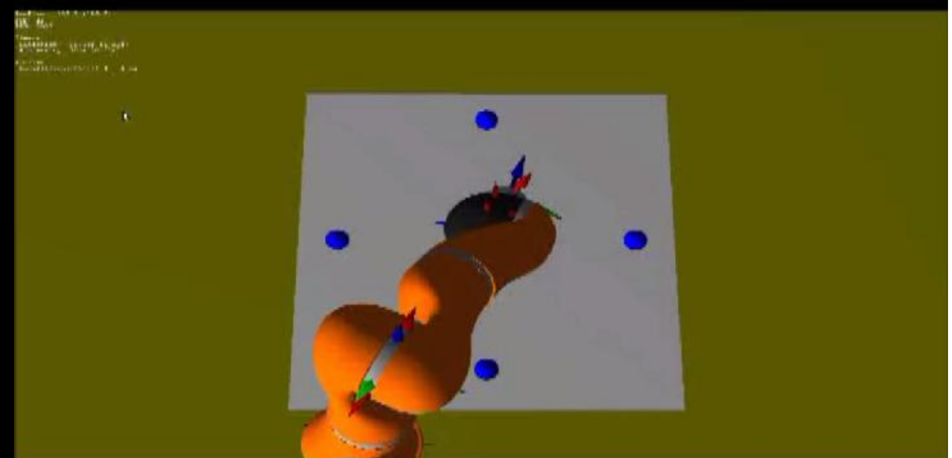


- Optimal motion plans are generated away from the training data

Combined Framework



Experimental set-up for collecting data in reaching task



Conclusions

We have proposed a brain signal decoder that:

generates optimal motion plans for robot control away from training data

starts to detect intention before the movement onset