

EPFL Smart Kitchen: Platform to analyze natural human motion

Haozhe Qi

PhD student EPFL (Mathis Group)





Natural human motions (like running) can be diverse under different situations



Human Motion Analysis

Sport Science

Clinical assessments

Pharmacology

Sociology

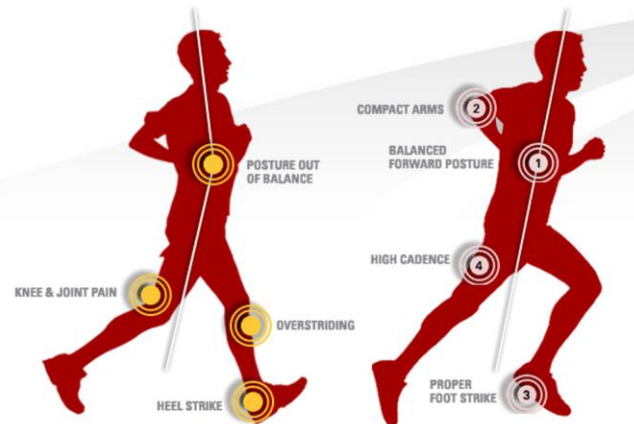
Therapy

...

Running techniques

FAMILIAR RUNNING FORM

GOOD RUNNING FORM



Human Motion Analysis

Sport Science

Clinical assessments

Pharmacology

Sociology

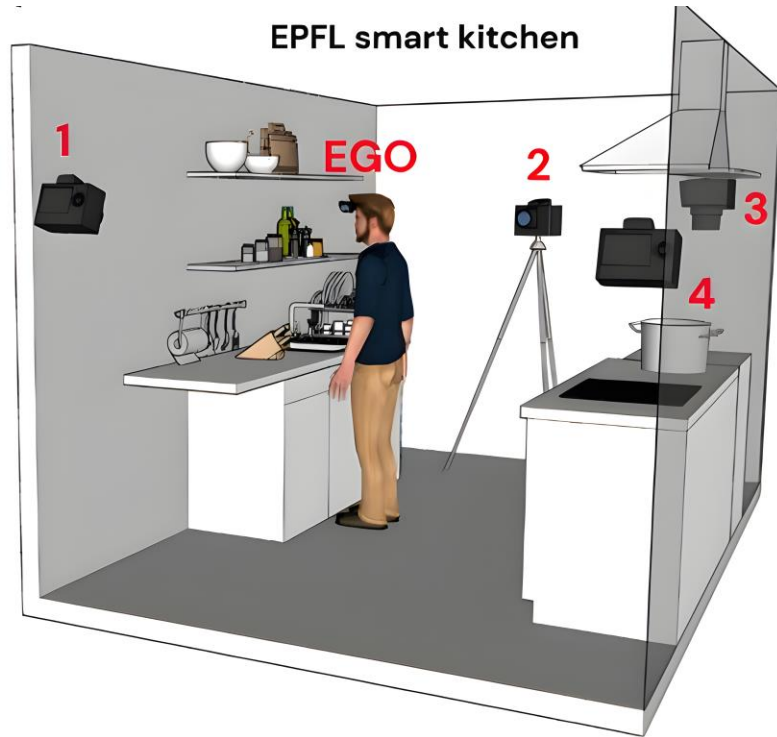
Therapy

...

Fugl-Meyer Assessment (standard for stroke)



Natural Human Motion Analysis



Kitchen environment:

- Diverse actions
- Many movements especially for hands

EPFL Smart Kitchen for unconstrained human motion analysis



Healthy subjects



Skill learning



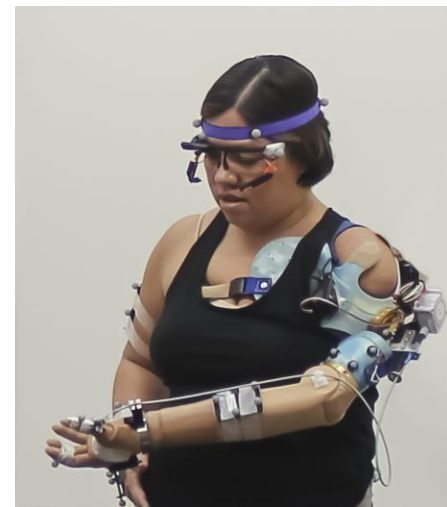
Stroke patients



Tailored therapeutics



Amputee patients



Novel control algorithms

EPFL Smart Kitchen for unconstrained human motion analysis



Healthy subjects



Skill learning



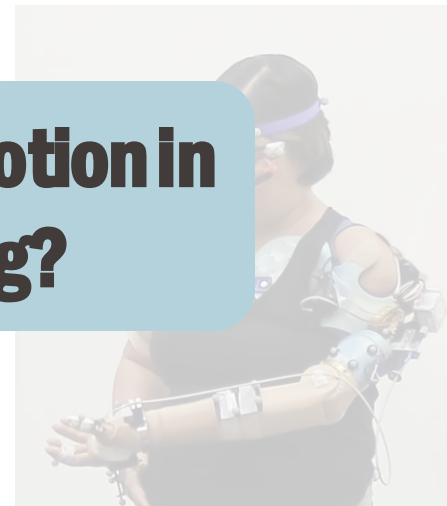
Stroke patients



Tailored therapeutics



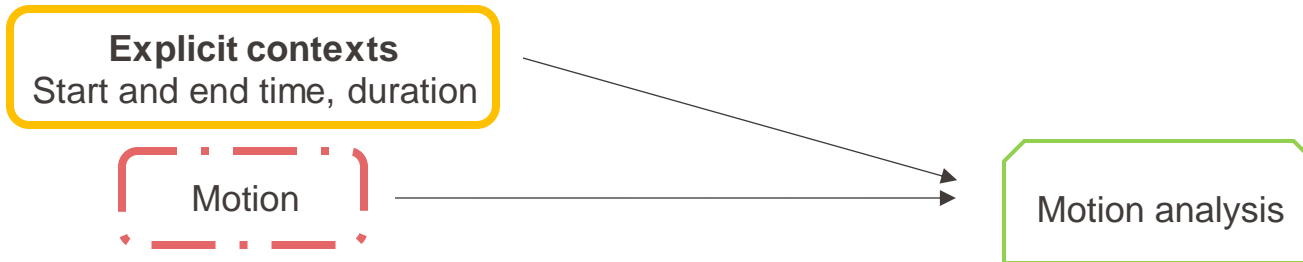
Amputee patients



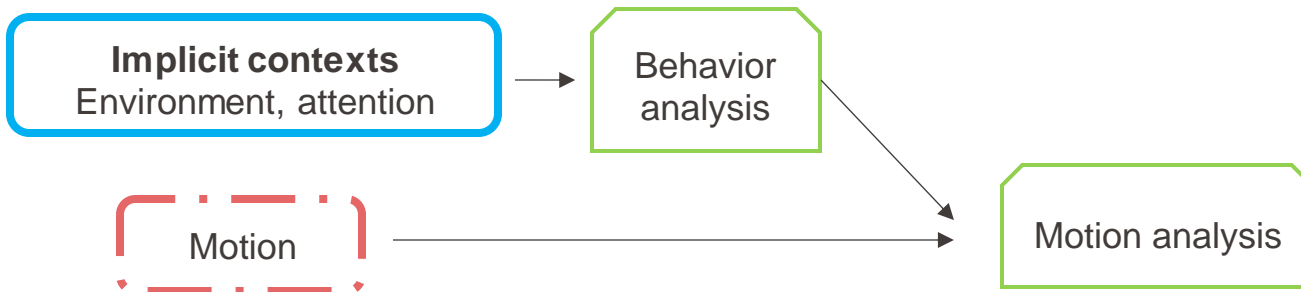
Novel control algorithms

Q: How do we study human motion in an unconstrained setting?

Constrained human motion analysis



Unconstrained human motion analysis



Movement information



Body movement



Hand movement

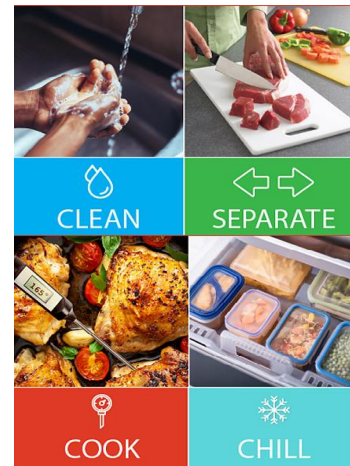
Context information



Eye gaze



Interacted objects



Actions

Q.1: How can we collect all of these in one dataset?

EPFL The Smart Kitchen platform

All Synchronized and Calibrated



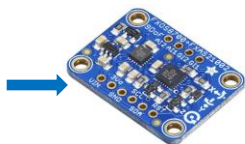
9 Kinect Azure Cameras



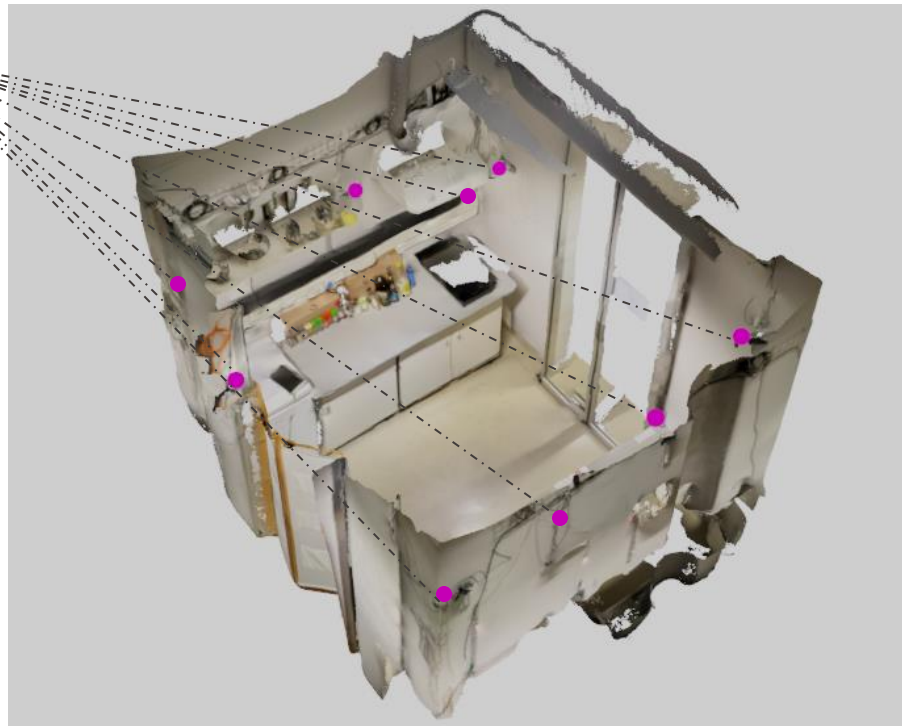
Ego HoloLens Cameras



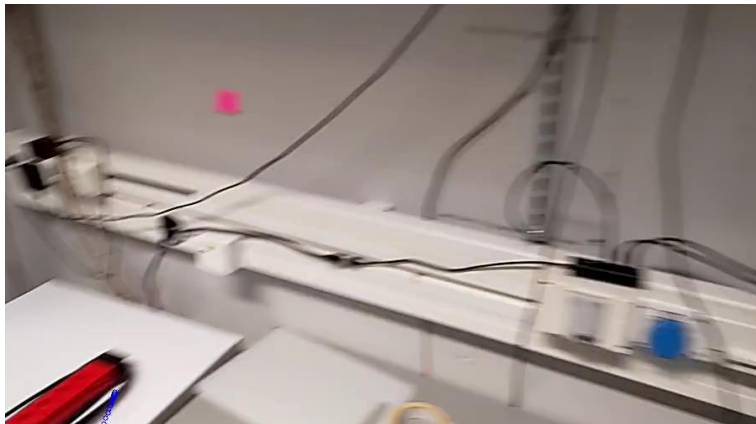
Arduino



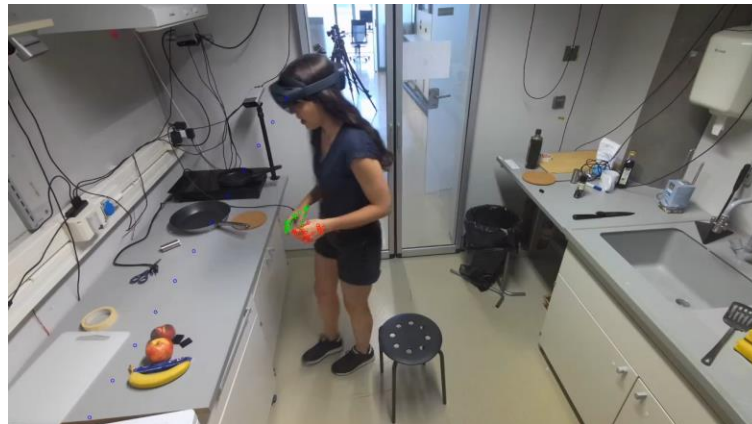
IMU sensors



HoloLens view

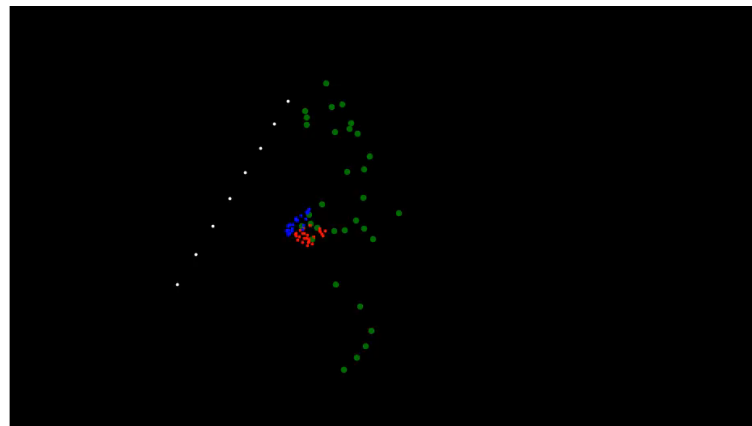
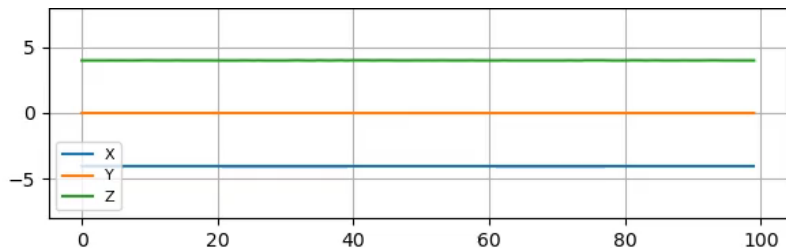


One Kinect view

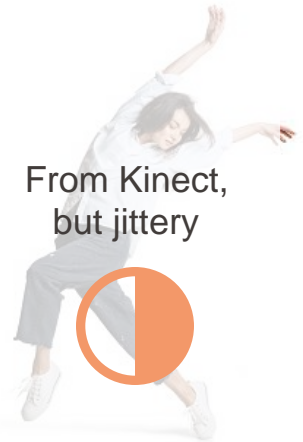


Hands+
Gaze

IMU data from the knife



Hands+
Gaze+
Body



From Kinect,
but jittery

Body movement



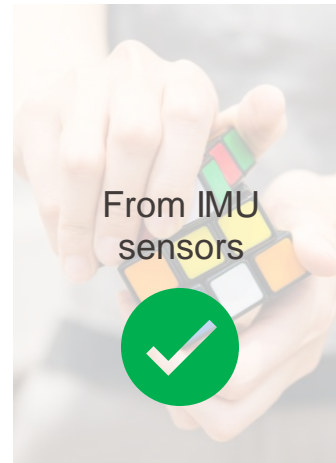
From Kinect,
but jittery

Hand movement



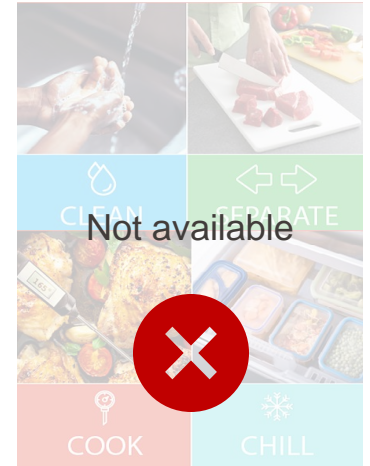
From HoloLens

Eye gaze



From IMU
sensors

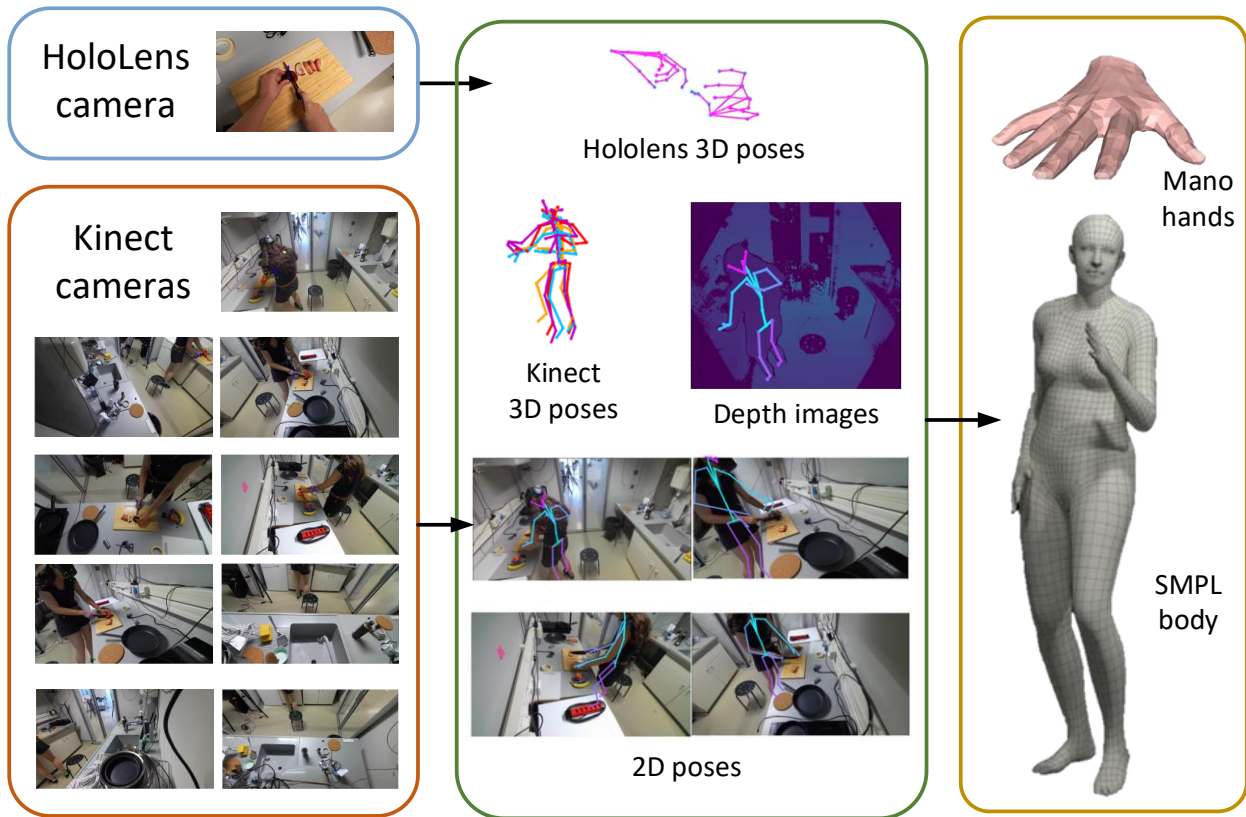
Interacted objects



Not available

Actions

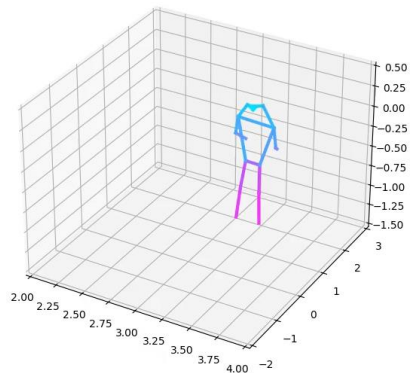
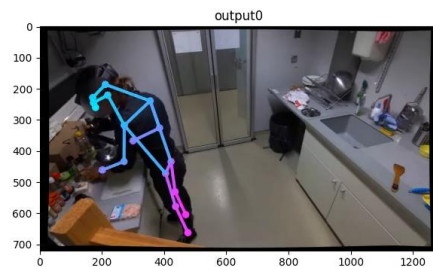
EPFL 3D pose improvements in the Smart Kitchen



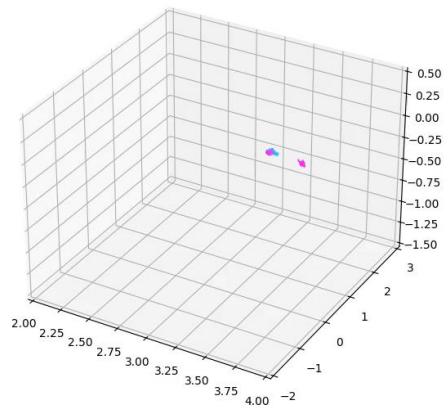
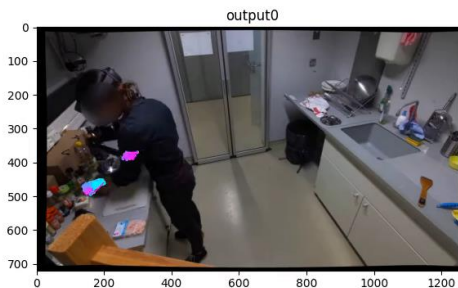
Methods inspired by our paper newly accepted by CVPR2024!

Qi H, Zhao C, Salzmann M, et al. HOISDF: Constraining 3D Hand-Object Pose Estimation with Global Signed Distance Fields. CVPR (in press), 2024.

Body Mesh fitting



Hand Mesh fitting



■ Rendered 3D meshes

Projected 3D poses

3D poses

Fine-grained Actions

Manually defined and annotated

33 Verbs

Open Put Carry
Read Tap Slide
Grab Clean Peel
Close Taste Stir
Add Grate Hold
Wash Move Cut
Switch Take off
Split Touch Pour
Press Sauté Wait
Throw Squat Dry
Put on Shake
Other action

78 Nouns

Food
Mushrooms Noodles
Radish Bean sprouts
Peanuts Salad Surimi
Frying Oil Stock cube
Avocado Sauce Pasta
Broth Tomatoes Rice
Fresh thyme Lemon
Processed ingredient
Tamarind paste Tofu
Zucchini Carrots Eggs
Onions Water Butter
Shrimps Cheese Salt
Seasoning Cucumber
Shallots Bell pepper
Egplant Other food

Objects

Pan Knife Grater Sponge
Salt Box Recipe Pot Soap
Bottle Tissue Towel Whip
Fork Bowl Trash Button
Doser Glass Brush Glove
Cutting board Colander
Salad bowl Spoon Trivet
Spatula Package Drawer
Other object

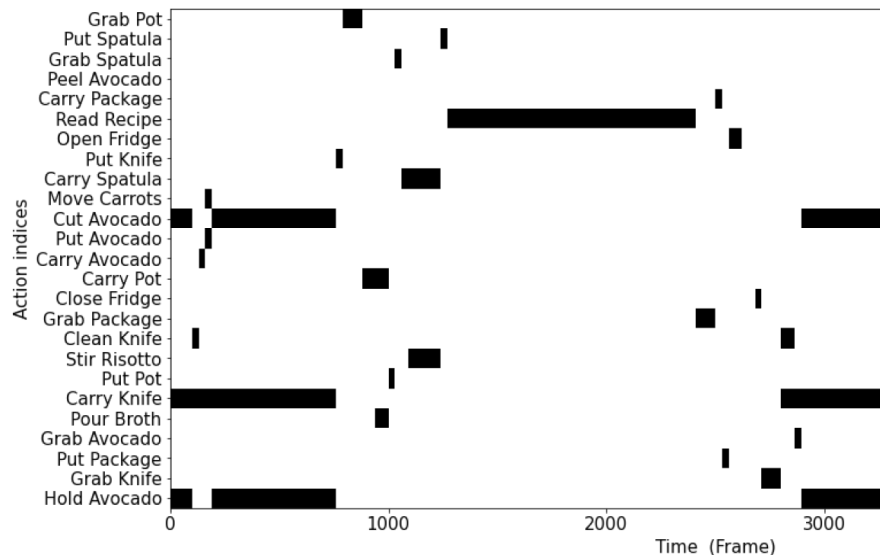
Appliances

Stoves
Fridge Sink
Cupboard
Ventilation
Trash bin

Recipes

Tomato salad
Green salad
Ratatouille
Pad Thai
Omelet Risotto

Ethogram example



We will release the whole dataset including action annotations to the CV community!



Omelette and tomato salad recipe

- Tomato salad**
- Dice 2 tomatoes
 - Mix one spoon of oil with salt, pepper and balsamic vinegar
 - Pour the dressing on the tomato salad
 - Stir the salad for at least 2 minutes to allow the tomatoes to soak up the dressing
- Omelette**
- Beat 3 eggs and season them with salt and pepper
 - Heat the oil in a pan over a medium-low heat
 - Pour the eggs into the pan, lift the pan ever so slightly from one side to another to allow the eggs to twist and cover the surface of the pan completely
 - Let the mixture cook for about 20 seconds then scrape a line through the middle with a spatula
 - Flip the pan again to allow it to fill back up with the runny egg
 - Repeat once or twice more until the egg has just set
 - Fold gently in half with the spatula
- Enjoy your meal!

1. Omelette



Ratatouille recipe

- 1 eggplant
- 1 zucchini
- 1 yellow pepper
- 2 ripe tomatoes
- 80g button mushrooms
- 1/2 a bunch of fresh basil (15g)
- olive oil
- a few sprigs of fresh thyme
- 1 tablespoon balsamic vinegar
- 1/2 a lemon
- Salt and pepper

1. Put a casserole of salted water to boil.
2. Pick the basil leaves and finely slice the basil stalks.
3. Heat 2 tablespoons of oil in a pan over a medium heat, add the basil stalks and thyme leaves. Cook on a medium heat for 2-3 minutes. During this time, chop the zucchini, eggplant and mushrooms. The zucchini and eggplant should be cut into quarters (lengthways twice) and then diced.
4. When the water boils, add the pasta for 7-10mins depending on the type. Then drain them, drizzle them with olive oil and keep them aside.

2. Ratatouille



Recette de risotto et salade verte

- 1 bouillon cube
- 200 g de riz pour risotto
- 10 cl de vin blanc
- 2 cuillères de beurre
- 3 cuillères de parmesan râpé
- 1 portion de radis
- 1 avocat
- 1/2 concombre
- 1/2 citron
- 1 portion de sésame
- huile de sésame

Portez à ébullition bouillon 1.5L. Étape dans une casserole et verser y un cube de bouillon de légumes, remuez. Dès l'ébullition, baissez sur feu doux et laissez mijoter le bouillon afin de le conserver chaud.

Faites fondre le beurre dans une seconde casserole sur feu moyen. Ajoutez 200 g de riz pour risotto, puis laissez cuire pendant 3 min environ, en remuant bien, jusqu'à ce qu'il devienne translucide. Versez ensuite 10 cl de vin blanc, puis laissez cuire, en remuant fréquemment, jusqu'à ce qu'il soit complètement absorbé par le riz.

3. Risotto



Ratatouille recipe

- 1 eggplant
- 1 zucchini
- 1 yellow pepper
- 2 ripe tomatoes
- 80g button mushrooms
- 1/2 a bunch of fresh basil (15g)
- olive oil
- a few sprigs of fresh thyme
- 1 tablespoon balsamic vinegar
- 1/2 a lemon
- Salt and pepper

1. Put a casserole of salted water to boil.
2. Pick the basil leaves and finely slice the basil stalks.
3. Heat 2 tablespoons of oil in a pan over a medium heat, add the basil stalks and thyme leaves. Cook on a medium heat for 2-3 minutes. During this time, chop the zucchini, eggplant and mushrooms. The zucchini and eggplant should be cut into quarters (lengthways twice) and then diced.
4. When the water boils, add the pasta for 7-10mins depending on the type. Then drain them, drizzle them with olive oil and keep them aside.

4. Ratatouille



Recette du Pad Thai (ผัดไทย)

- 1/2 Oignons
- 2 à 4 sauges de calabasses grillées
- 2 à 4 sauges de coriandre
- 1 à 4 sauges de crevette séchées
- 2 à 4 sauges d'ail de bonnet
- Filés de tamarin acide
- Tofu
- Protéines de soja
- 1 échalote
- 1 citron
- Nouilles de riz
- Sauce poisson (nuoc mam)

Faites tremper les nouilles dans de l'eau chaude pour les ramollir.

Faites dissoudre une boucle de pâte de tamarin dans 100 ml d'eau chaude, mélangez et filtrez le mélange pour obtenir un jus de tamarin.

Coupez le tofu, en cubes puis réservez, idem avec l'échalote et l'oignon.

Faites chauffer l'huile dans une poêle et faites revenir le tofu jusqu'à ce qu'il soit doré et réservez de côté.

Ajoutez les nouilles égoutées, et laissez le cuire avec le jus, rapétez un peu d'eau si les nouilles ont tout absorbé.

5. Pad Thai

4 recipes

5 sessions of ~1h

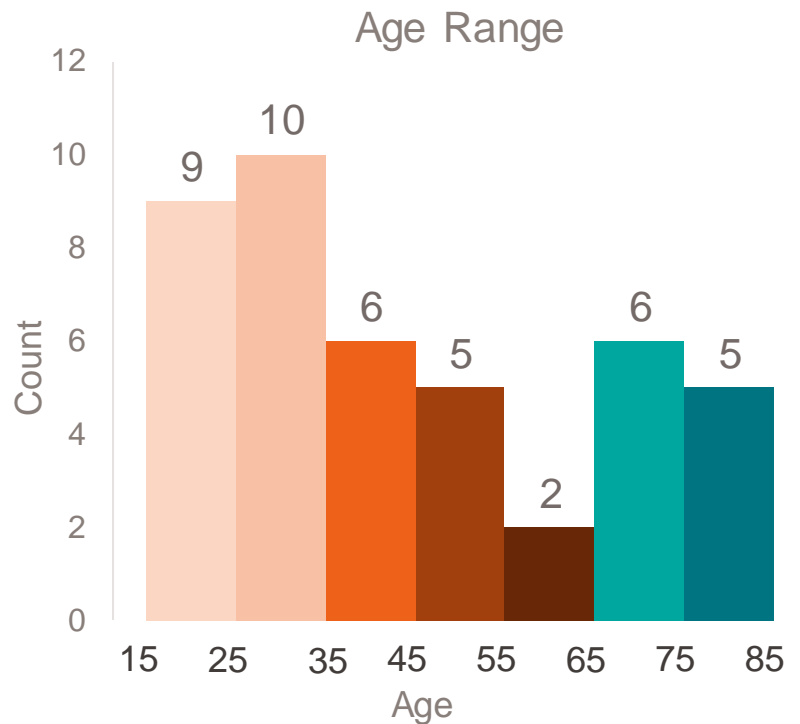
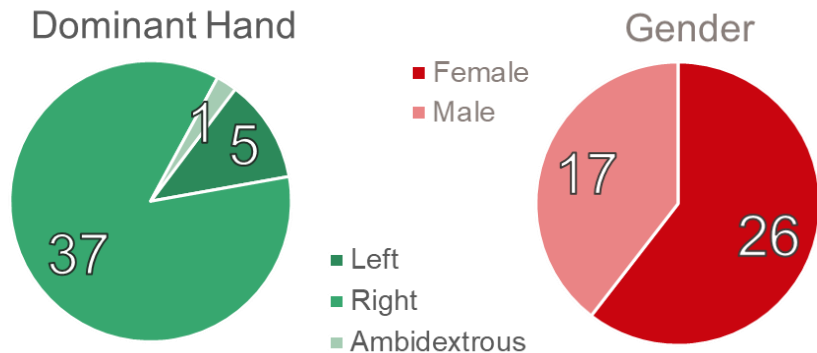
EPFL Data collection summary

Each participant cooks **5** times under 4 recipes

Number of participants: **43**

Number of sessions: **190** sessions for now

Total data memory usage: Around **200 TB**

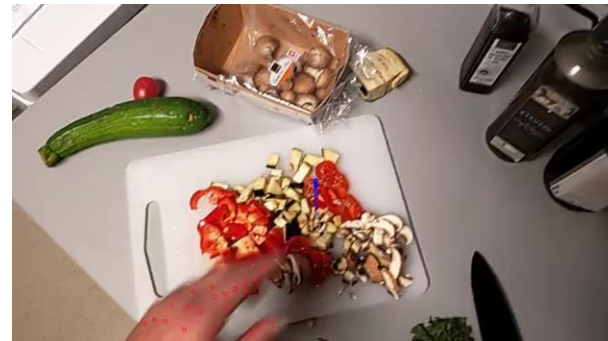




1. Find ingredients



2. Cut ingredients



3. Seasoning



4. Cooking



5. Make pasta

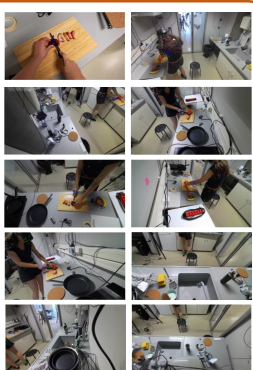


6. Clean the pot

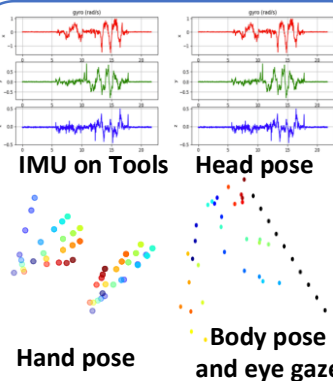
EPFL In summary, we have ...

Multi-modal
data types

Visual information



Kinematic information



Category information

Identity level

Age

Gender

Cooking
skills

Dominant
hands

Session level

Recipes

Time

Frame level

Fine-grained
action
annotations

Coarse actions

Reading Recipe
Fetch ingredients
Process ingredients
Cooking on the stoves
Cleaning

XClip

Rule

LLM
Summary

Multiple
cohorts



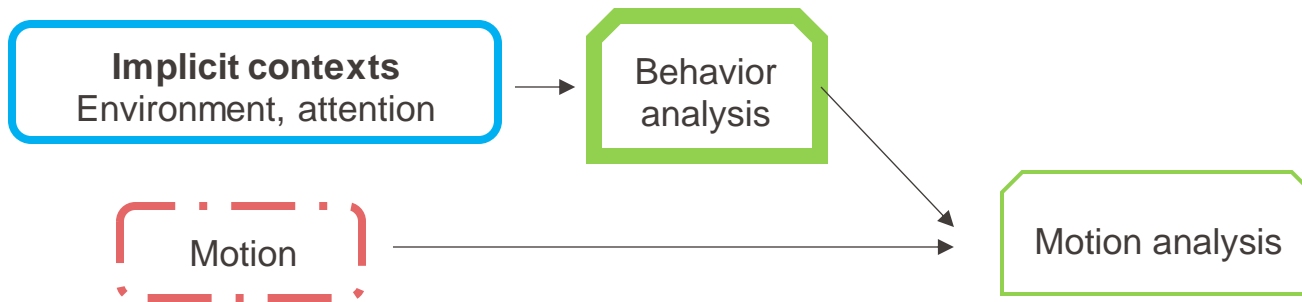
Professional chef data

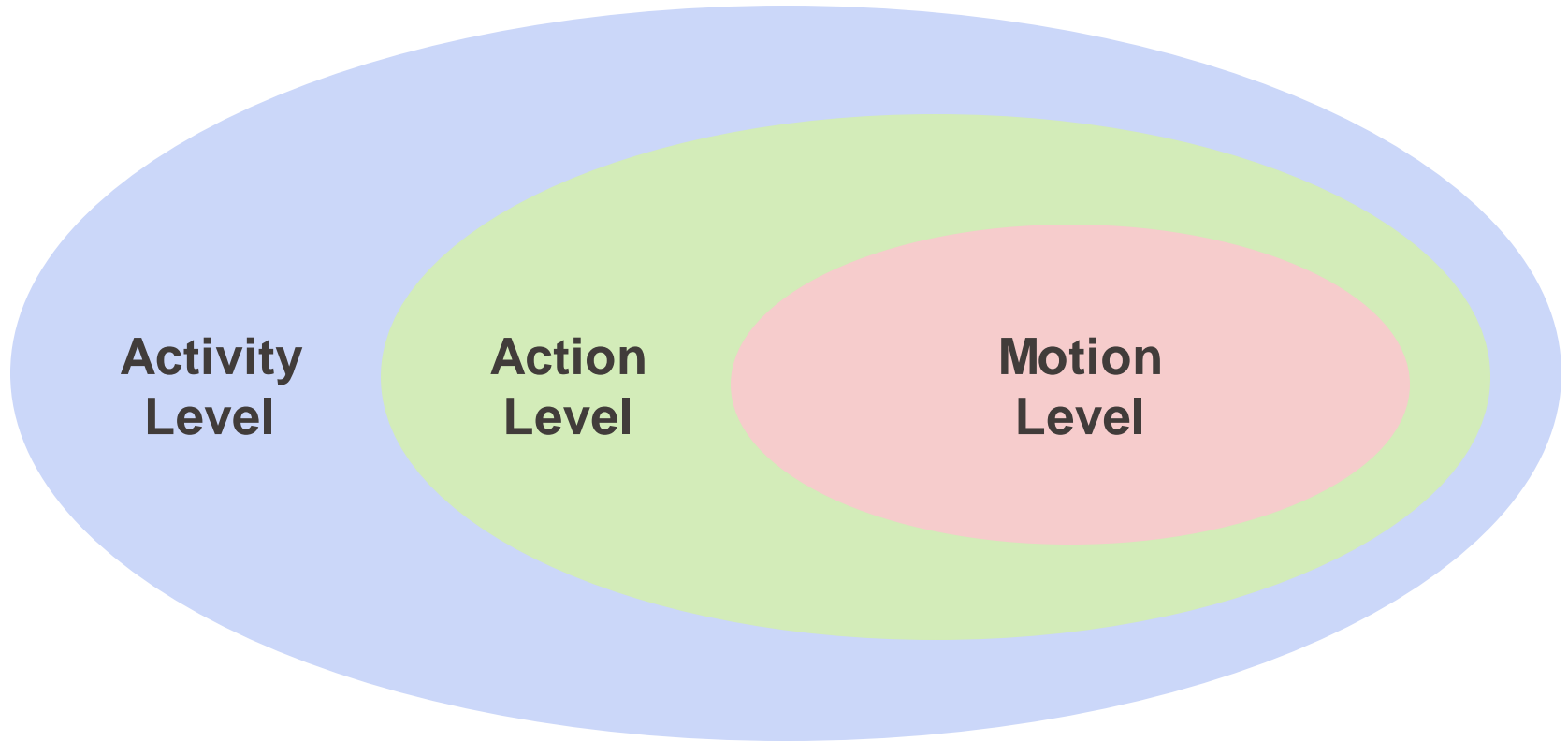


Patients data

Q.2: How do we quantify natural behavior?

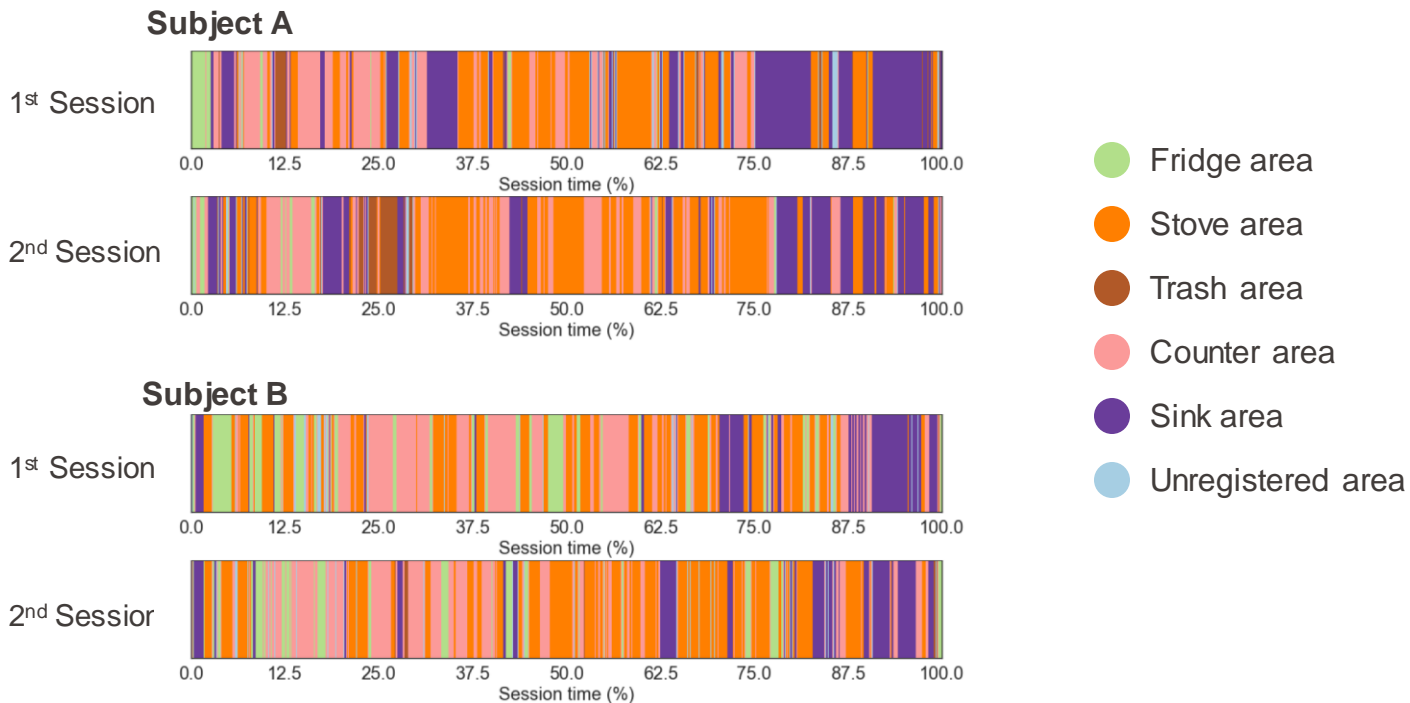
Unconstrained human motion analysis





EPFL Activity level analysis: Within-subject consistency

Despite different strategies across subjects, individuals are consistent

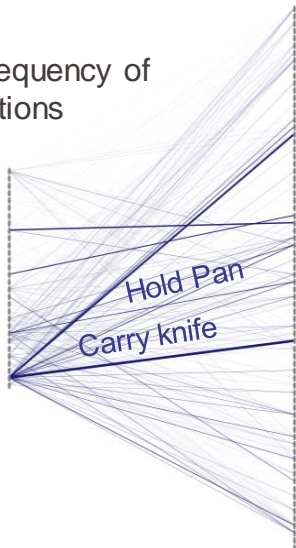


Cooking recipe dictates action repertoire.

- | | | | |
|---|---|--|--|
| Verbs | Food | Nouns | Objects |
| Open Put Carry
Read Tap Slide
Close Clean Peel
Add Taste Stir
Wash Move Cut
Switch Take off
Split Touch Pour
Press Sauté Wait
Throw Squat Dry
Put on Shake
Other action | Mushrooms Noodles
Radish Bean sprouts
Peanuts Salad Surimi
Frying Oil Stock Cube
Broccoli Sauce Pasta
Broth Tomatoes Rice
Fresh Coriander Lemon
Processed ingredient
Tamarind paste Tofu
Zucchini Carrots Eggs
Onions Water Butzler
Shrimps Cheese Salt
Seasoning Cucumber
Shallots Bell pepper
Eggplant Other food | Pan Knife Grater Sponge
Salt Box Recipe Pot Soap
Bottle Tissue Towel Whip
Fork Bowl Trash Button
Doser Glass Brush Glove
Cutting board Colander
Spatula Package Drawer
Salad bowl Spoon Trivet
Pot lid Hand Plate Peeler
Other object | Appliances
Stoves
Fridge Sink
Cupboard
Ventilation
Trash bin |
| | | | Recipes
Tomato salad
Green salad
Ratatouille
Pad Thai
Omelet Risotto |

Frequency of actions

Verbs



Nouns

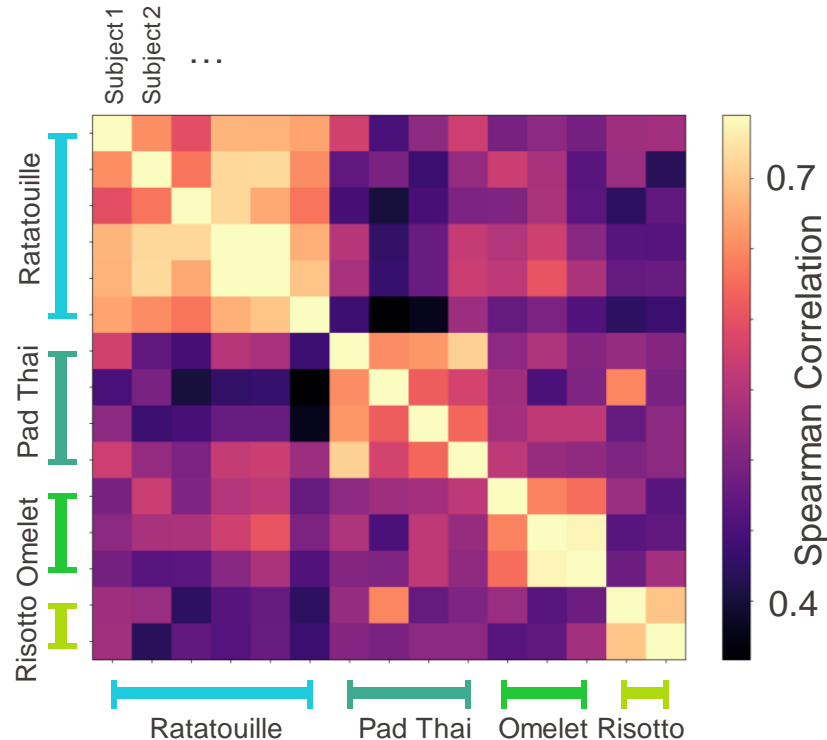
Graph definition

$$G = \{(w_{ij} \in \mathbb{R}^{N_v \times N_n})\}$$

Spearman correlation between graphs

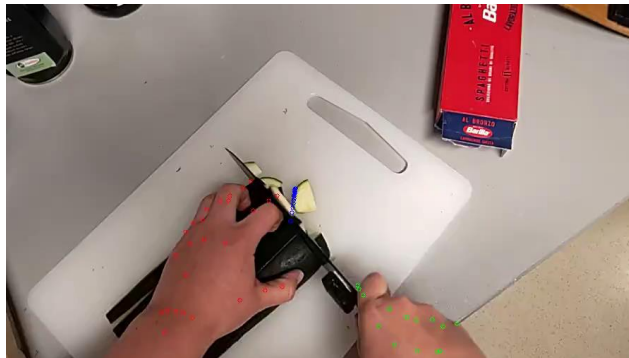
$$S = \rho(G_A, G_B)$$

$$\rho_{R(X), R(Y)} = \frac{\text{cov}(R(X), R(Y))}{\sigma_{R(X)} \sigma_{R(Y)}}$$



Characterize the distribution of motions for specific actions

$P(\text{kinematics} | \text{action}_t)$



$P(\text{kinematics} | \{\text{actions}, \text{activities}\}_{t-\tau, t+\tau})$



EPFL Smart Kitchen for unconstrained human motion analysis



- **Conclusion:**
 - Collected multi-modal dataset related to natural human motions.
 - Estimated accurate 3D poses and fine-grained actions. (*Novel action segmentation benchmark!*)
 - Analyzed cooking behavior at different temporal levels.
- **Ongoing:**
 - Scaling to all participants & analysis of motions across actions and their transitions
- **Future:**
 - Contrast with stroke patients (for functional assessment) and arm prosthetics users (for better motion control)

EPFL SMART
KITCHEN



JRC project team:

- Prof. Marc Pollefeys (Microsoft, ETH)
- Dr. Mahdi Rad (Microsoft)
- Prof. Alexander Mathis (EPFL)
- Prof. Dr. MD Friedhelm Hummel (EPFL)
- Prof. Silvestro Micera (EPFL)
- Dr. Solaiman Shokur (EPFL)

Students:

- Andy Bonnetto
- Franklin Leong
- Josephine Raugel

The whole Mathis Group