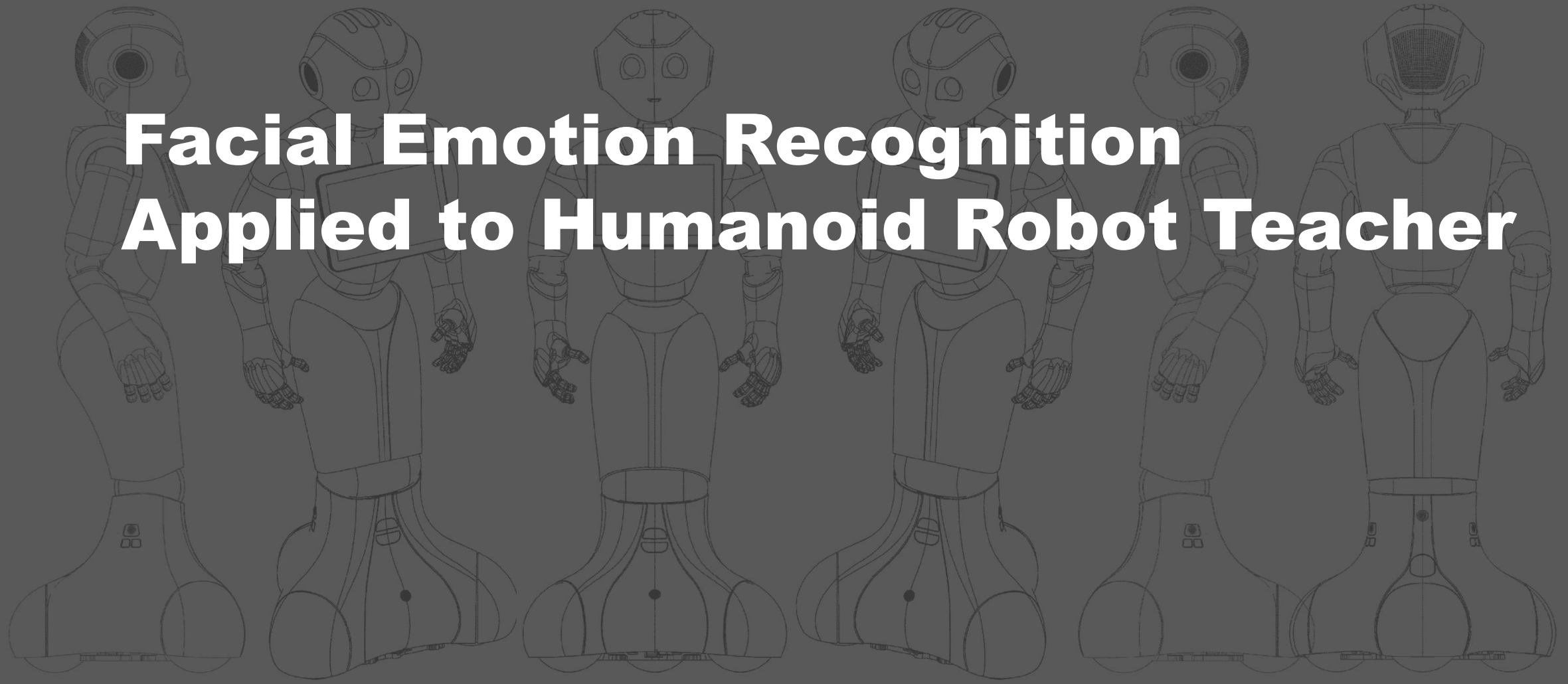
Queen Mary
University of London

北京郵電大學
Beijing University of Posts and Telecommunications



Facial Emotion Recognition Applied to Humanoid Robot Teacher

PROBLEM STATEMENT

Nowadays, more and more robots are introduced to the classroom.

(Softbank, 2019)

Students' engagement and concentration are not being observed.

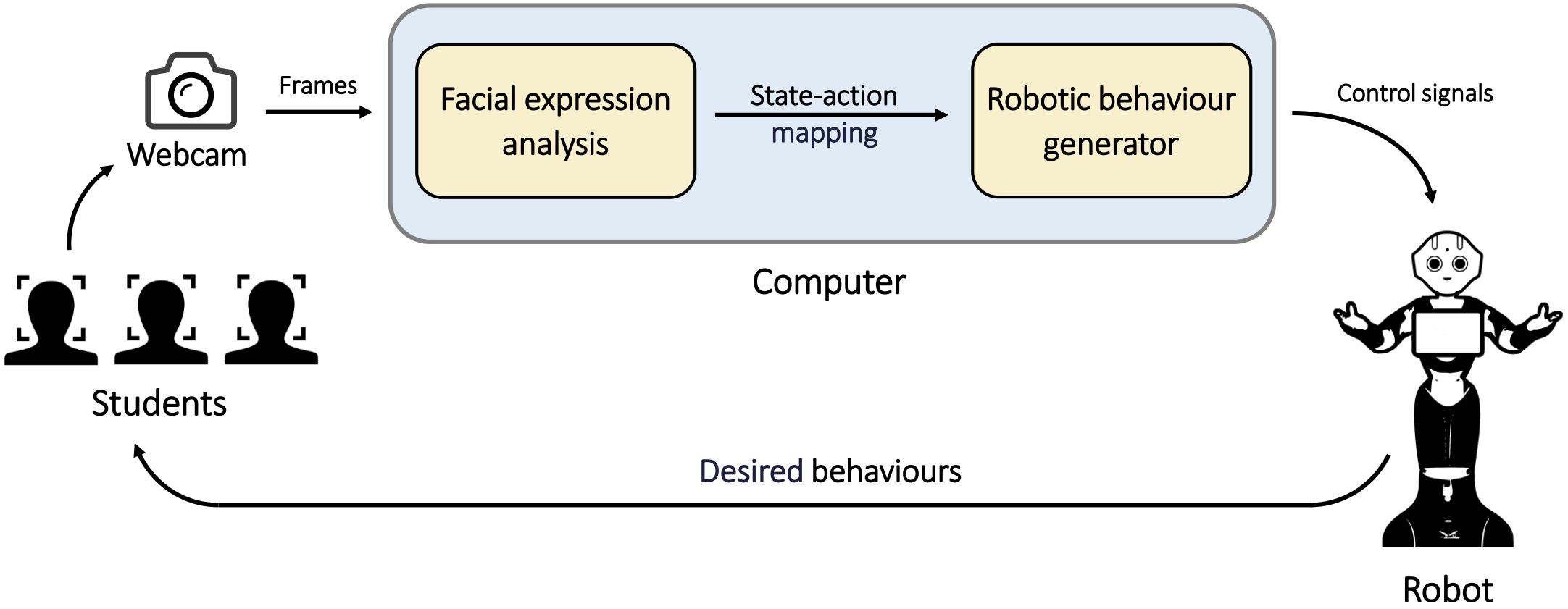
↓

Effective learning may not be guaranteed.

↓

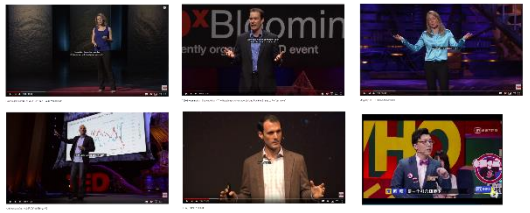
An emotional and sociable humanoid robot is needed.

SYSTEM ARCHITECTURE



DATA COLLECTION

Image dataset	CK+ dataset	Posed behaviour
		

Video dataset		Real class	Simulated scene
	students		
	lecturer		

EXPRESSION ANALYSIS

Face API – Cloud computing service provided by Microsoft

- Face detection: return 1 Face ID + 8 emotion intensities + 27 feature points

Image Input:

(Lucey, P., 2010)

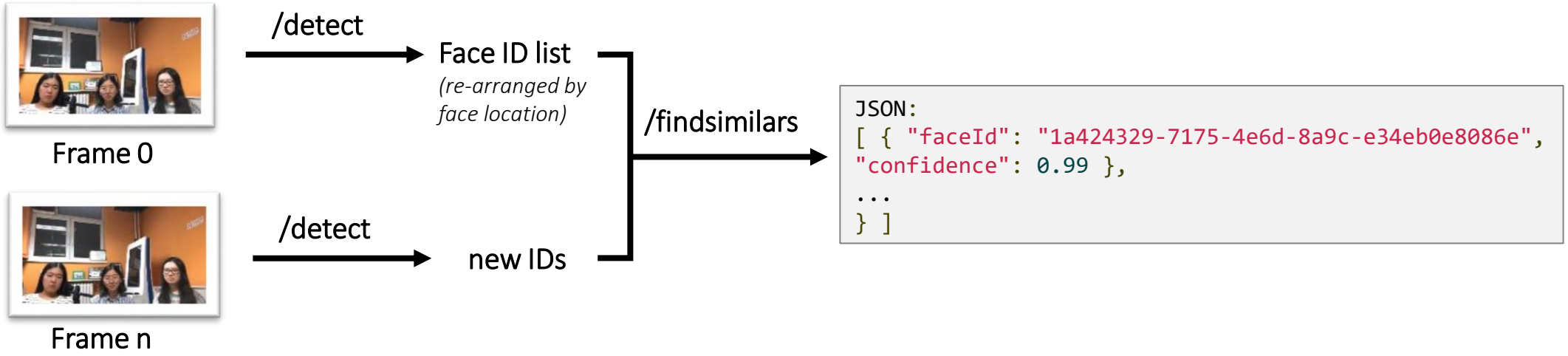
Detection result:


```
JSON:
[ { "faceId": "1a424329-7175-4e6d-8a9c-e34eb0e8086e",
"faceRectangle": { "top": 146, "left": 266, "width": 259, "height": 259 },
"faceAttributes": {
"emotion": { "anger": 0.0, "contempt": 0.0, "disgust": 0.0, "fear": 0.0,
"happiness": 0.975, "neutral": 0.025, "sadness": 0.0, "surprise": 0.0 },
"faceLandmarks": { "pupilLeft": { "x": 334.5, "y": 222.9 }, "pupilRight":
{ "x": 452.5, "y": 218.7 }, "eyeLeftOuter": { "x": 318.2, "y": 234.2 },
"eyeLeftInner": { "x": 353.6, "y": 225.9 }, "eyeRightInner": { "x": 433.9,
"y": 223.1 }, "eyeRightOuter": { "x": 472.5, "y": 220.6 },
...
} ]
```

EXPRESSION ANALYSIS

Group emotion recognition—tracking each face

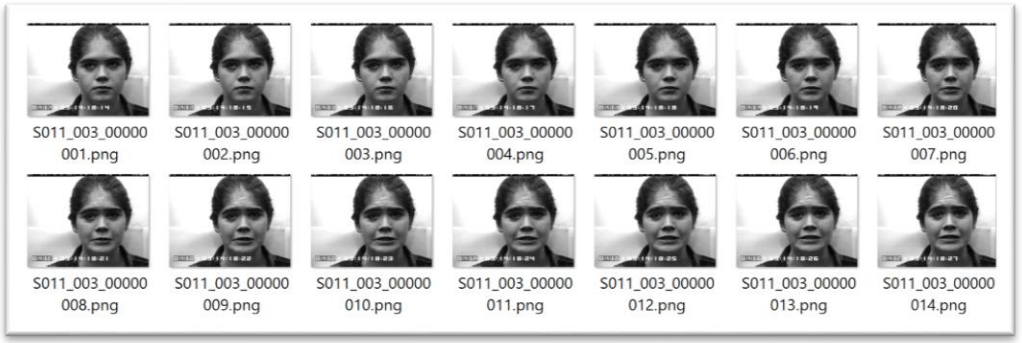
- Face verification: return a candidate Face ID



EXPERIMENT I

On CK+ dataset: 593 sequences across 123 subjects. All sequences are from the neutral face to the peak expression.

"disgust"



"fear"

The results has shown the drawbacks of current system:

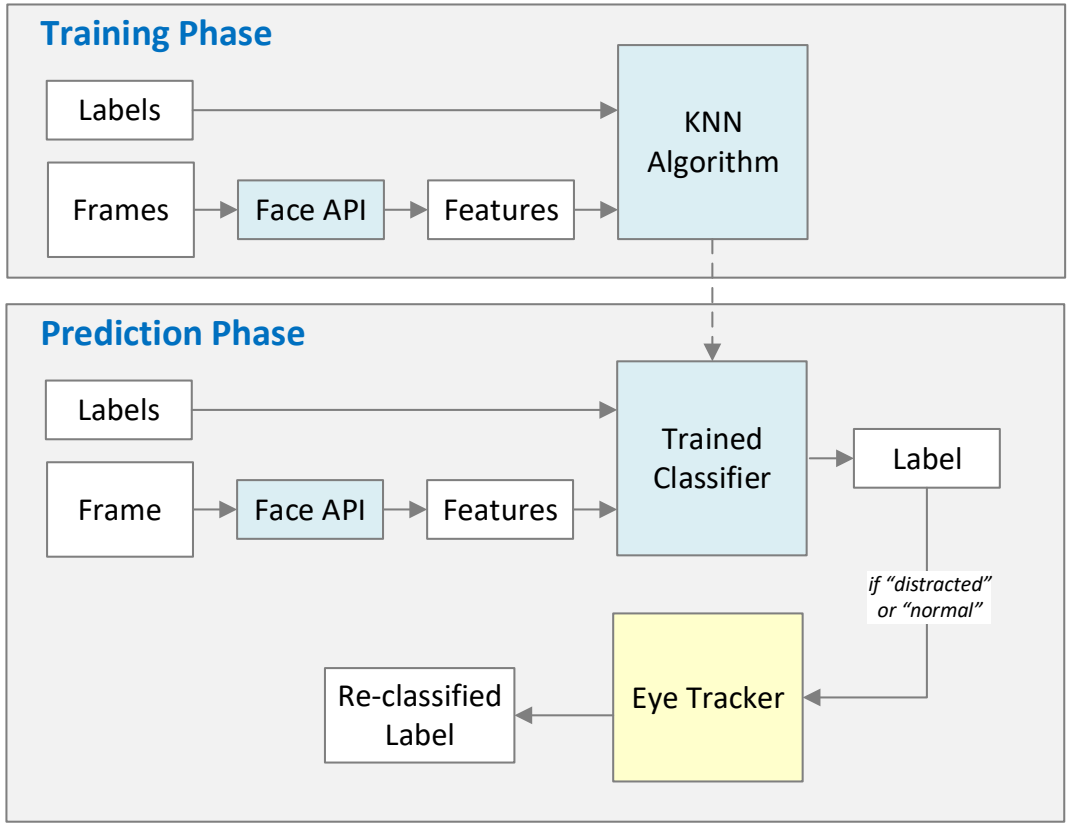
	anger	contempt	disgust	fear	happiness	neutral	sadness	surprise
1	0.004	0.003	**0**	0	0	**0.962**	0.03	0
2	0.006	0.005	**0**	0	0	**0.965**	0.024	0
3	0.014	0.011	**0.001**	0	0	**0.955**	0.018	0
4	0.037	0.01	**0.002**	0	0	**0.947**	0.004	0
5	0.345	0.004	**0.512**	0	0	**0.139**	0.001	0
6	0.28	0.001	**0.716**	0	0	**0.004**	0	0
7	0.271	0	**0.728**	0	0	**0**	0	0
8	0.107	0	**0.893**	0	0	**0**	0	0
9	0.129	0	**0.87**	0	0	**0.001**	0	0
10	0.129	0	**0.871**	0	0	**0**	0	0
11	0.084	0	**0.915**	0	0	**0**	0	0

	anger	contempt	disgust	fear	happiness	neutral	sadness	surprise
1	0	0	0	**0**	0	0.999	**0.001**	0
2	0	0	0	**0**	0	0.999	**0.001**	0
3	0	0	0	**0**	0	0.999	**0.001**	0
4	0	0	0	**0**	0	0.998	**0.002**	0
5	0	0	0	**0**	0	0.998	**0.002**	0
6	0	0	0	**0**	0	0.994	**0.006**	0
7	0	0	0	**0**	0	0.974	**0.026**	0
8	0	0.001	0	**0.001**	0	0.647	**0.351**	0
9	0	0	0	**0.007**	0	0.289	**0.701**	0.002
10	0	0	0	**0.03**	0	0.074	**0.894**	0.001
11	0	0	0	**0.042**	0	0.085	**0.869**	0.003
12	0	0	0	**0.008**	0	0.016	**0.976**	0
13	0	0	0	**0.015**	0	0.017	**0.967**	0
14	0	0	0	**0.009**	0	0.012	**0.978**	0

REFINEMENT

Convert basic emotion into “state”: confused, interested, distracted, and normal

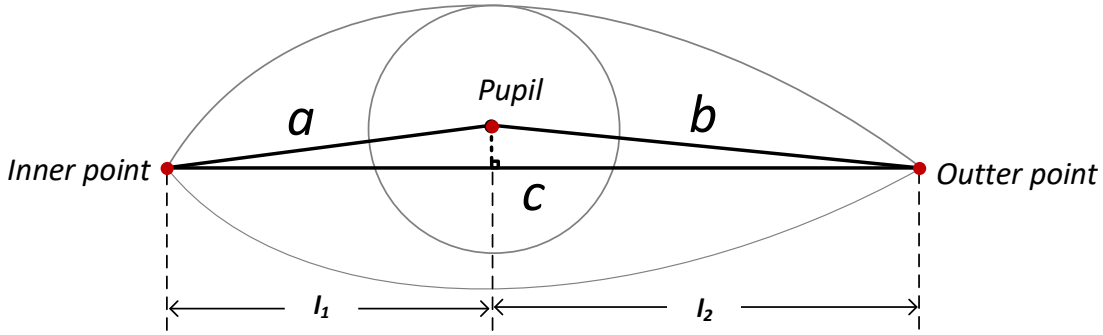
Classification: K-nearest neighbours algorithm + eye movement (by face feature points)



Calculation of the eye movement:

$$ratio = \frac{l_1}{l_2} = \frac{a^2 + c^2 - b^2}{b^2 + c^2 - a^2}$$

ranges in[0.4, 1.5] when looking straight at camera.



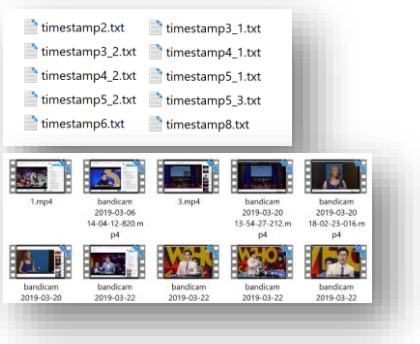
EXPERIMENT II

Relating the students' emotion to the lecturer's behaviour – to mark the time stamps when students are showing interest.

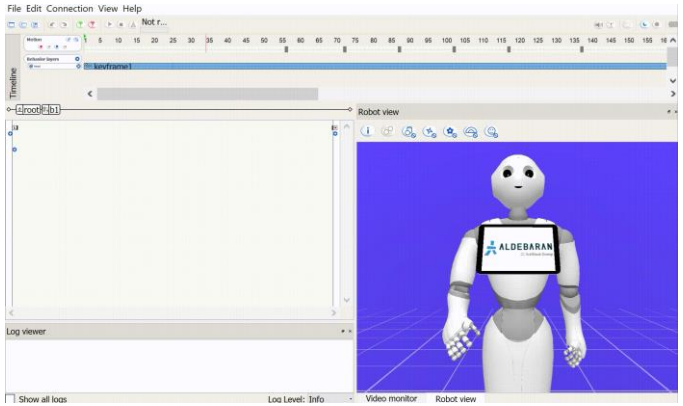
Desired behaviours are summarized as 15 different types.

time	anger	contempt	disgust	fear	happiness	neutral	sadness	surprise	state
0:20	0	0	0	0	0	0.988	0.012	0	3
0:21	0	0	0	0	0	0.994	0.005	0	3
0:22	0	0	0	0	1.0	0	0	0	1
0:23	0	0	0	0	0.992	0.008	0	0	1
0:24	0	0	0	0	0.965	0.035	0	0	1
0:25	0	0	0	0	0.994	0.006	0	0	1
0:26	0	0	0	0	0.998	0.002	0	0	1
0:27	0	0	0	0	0.652	0.348	0	0	1
0:28	0	0	0	0	0.624	0.375	0	0	1
0:29	0	0	0	0	0	0.995	0.004	0	3
0:30	0	0	0	0	0.016	0.982	0.002	0	0
0:31	0	0.001	0	0	0.229	0.749	0.015	0.006	0
0:32	0	0	0	0	0	0.99	0.007	0.003	3
0:33	0	0	0	0	0	0.986	0.002	0.012	3
0:34	0	0	0	0	0	0.984	0.003	0.012	3
0:35	0	0	0	0	0	0.996	0.004	0	3

0: confused
1: interested
2: distracted
3: normal



BEHAVIOUR DESIGN & IMPLEMENTATION



Simulation in Choregraphe

export motions

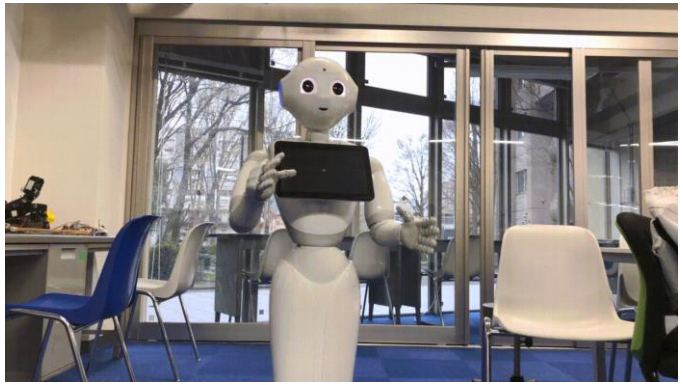
- Name of joint[ ]
- Time[ ]
- Angle[ ]

Control.py

```
from naoqi import ALProxy
ALProxy("ALMotion", ip, port)
motionProxy.post.angleInterpolation()
```

Programming with Naoqi SDK

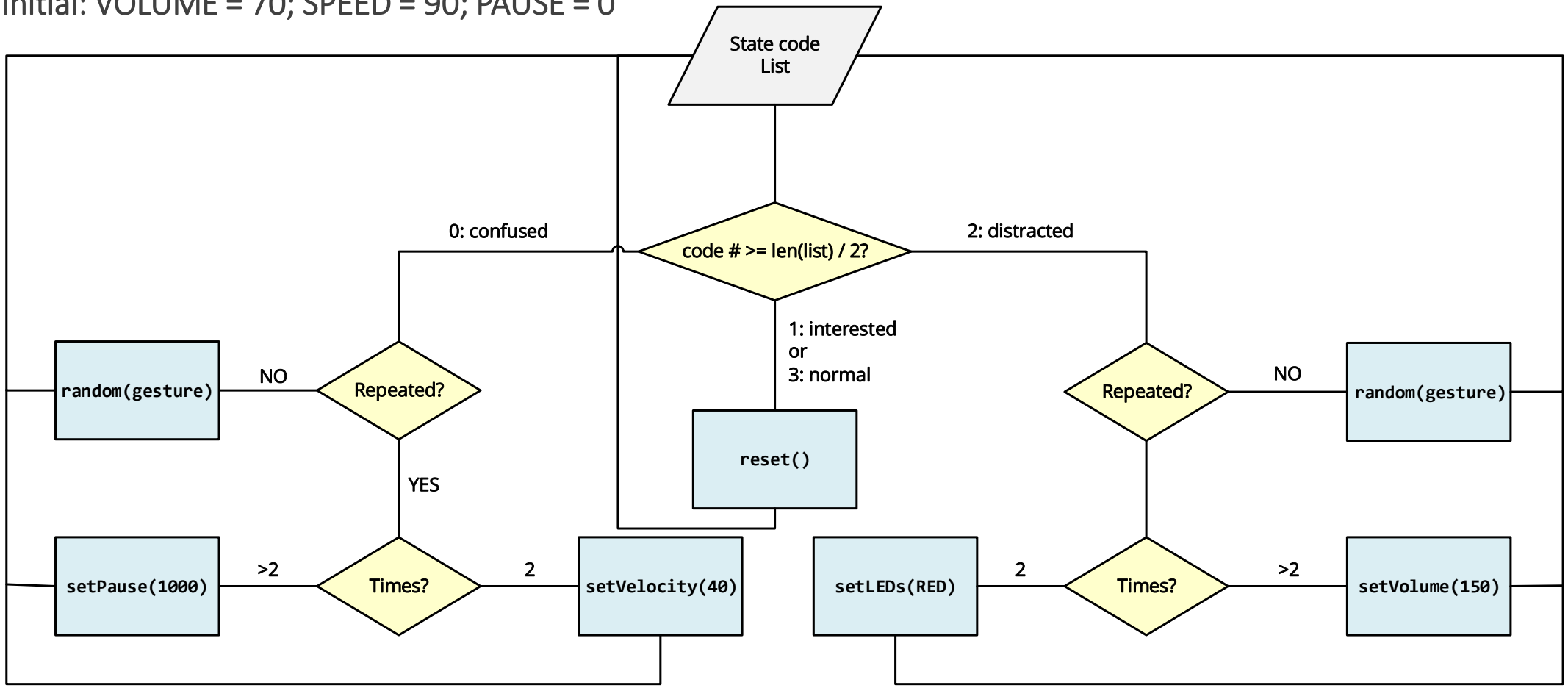
run



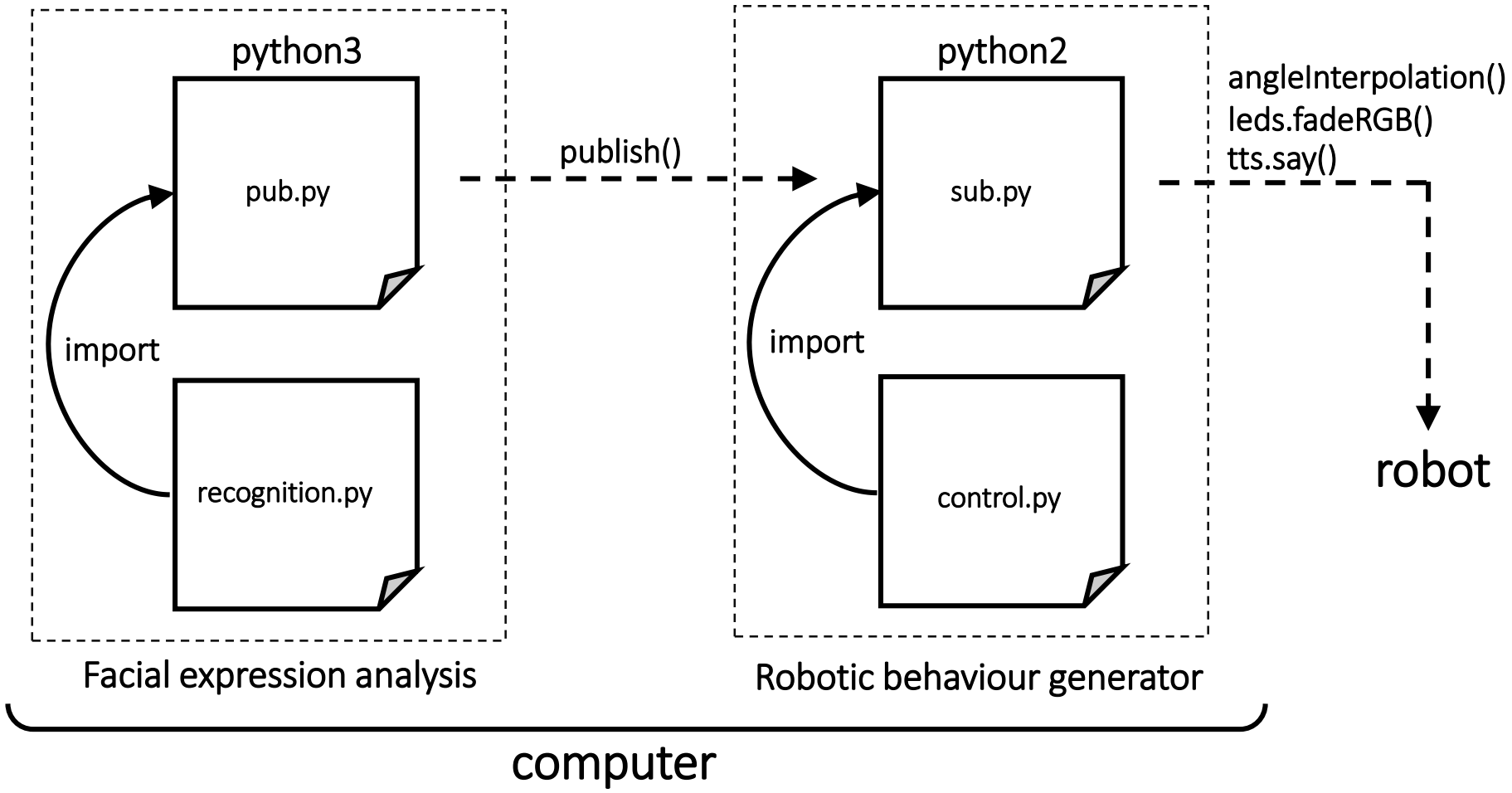
Implementation on the robot

MAPPING STRATEGY

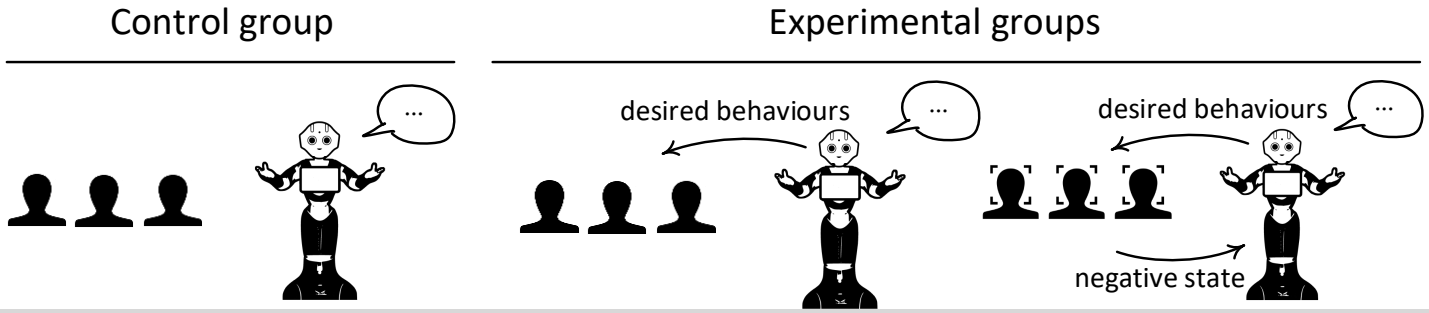
initial: VOLUME = 70; SPEED = 90; PAUSE = 0



SYSTEM INTEGRATION



SYSTEM EVALUATION



Independent variable: Connection with the HRI system	Without the HRI system	Connect to the behaviour generator	Connect to the whole HRI system
Questionnaires: filled up by 16 participants	↓	↓	↓
Dependent variable:			
Understanding of the lecture (%)	55.5%	48.3%	46.3%
Engagement in class (level 0 - 5)	0.188	2.687	3.125
Participants' preference	*2*	*6*	*8*

CONCLUSION & FURTHER WORK

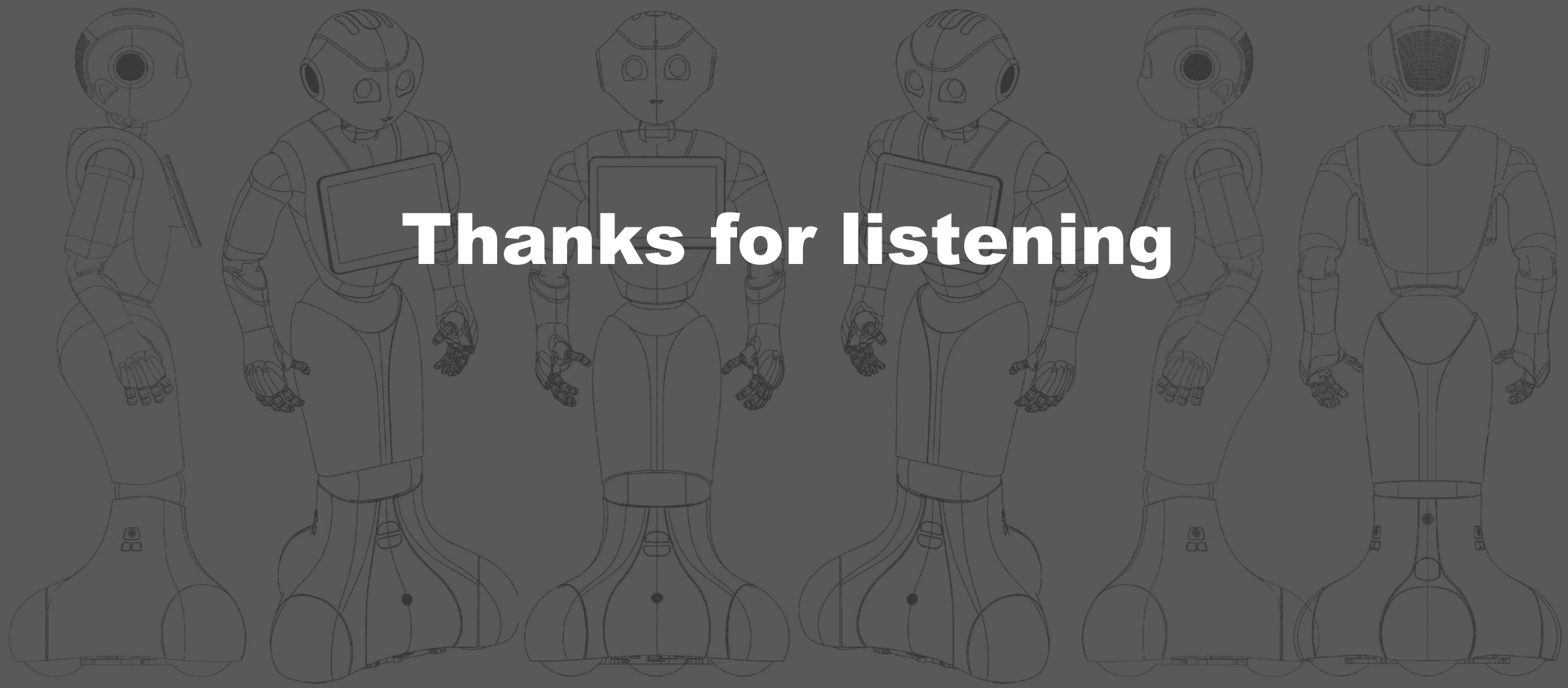
Proposed HRI system

- is an adaptive system
- excels other non-interactive systems in improving students' engagement
- visualises the current "state" with GUI
- is compatible with other robotic devices (e.g. robot arm, the NAO robot)
- is able to run without GPU

Further work: Robot avatar for the online lecture

Queen Mary
University of London

北京郵電大學
Beijing University of Posts and Telecommunications



Thanks for listening