



Human nonverbal behavior multi-sourced ontological annotation – ongoing research



Boris Knyazev, PhD Student, Engineer | Bauman Moscow State Technical University

bknyazev@bmstu.ru, biometric.bmstu.ru

General Problems

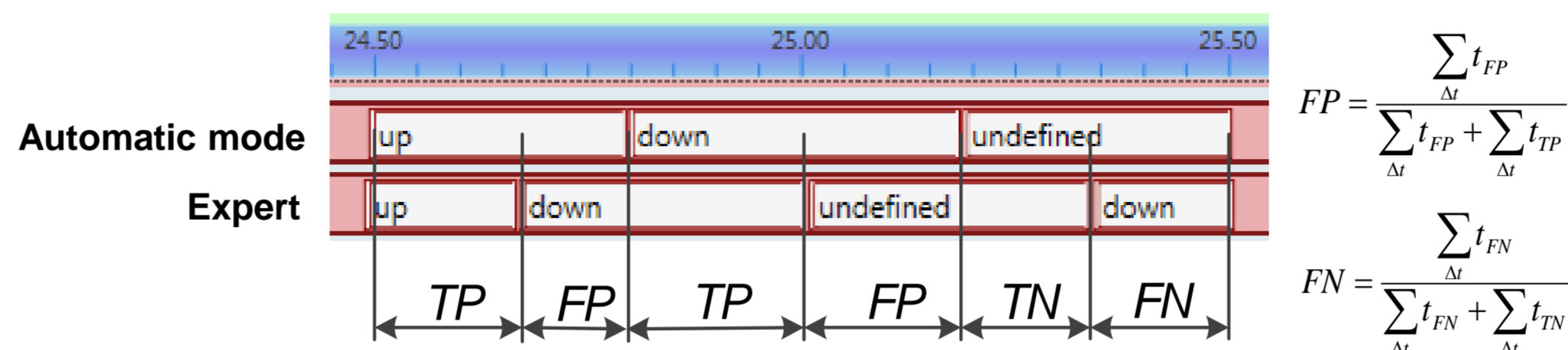
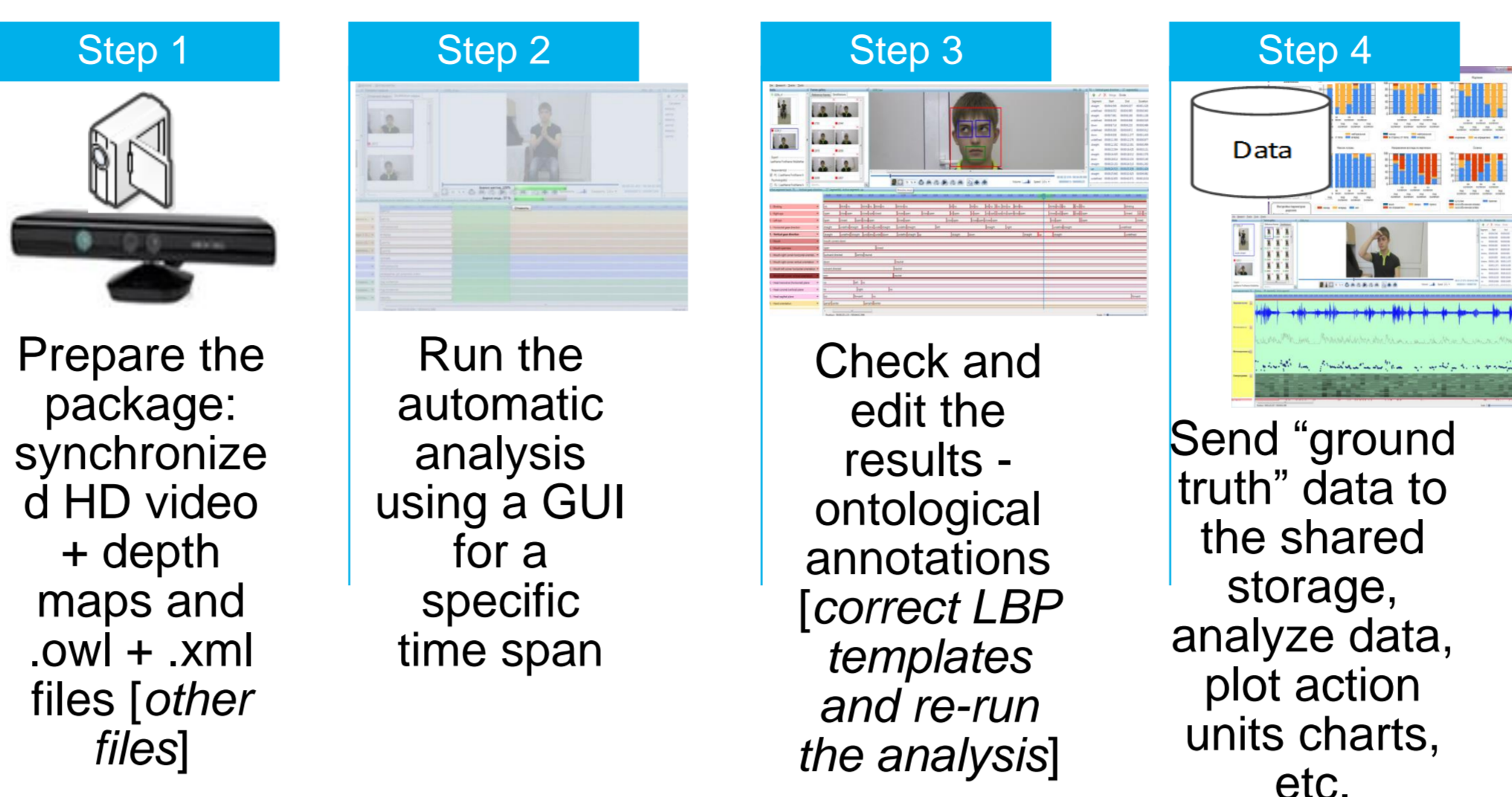
IF and HOW the nonverbal behavior and psycho-physiological state of a human are linked?

- Security
- Robotics
- Medical and Psychological diagnostics

Hypotheses

- They are linked in a unique, deeply concealed way, which might be visible through microexpressions and subtle body motions (motivated by [Ekman, P., 1969, Scherer, K.R., 1970])
- These micro activities can be registered and recognized by optical, IR (RGB-D) or other sensing devices coupled with the Computer Vision/Machine Learning techniques
- Semantic descriptions of nonverbal behaviors are more informative, provide effective reasoning and can increase the recognition rate (motivated by OWL, OWL-DL)

Data collection procedure



Experimental data

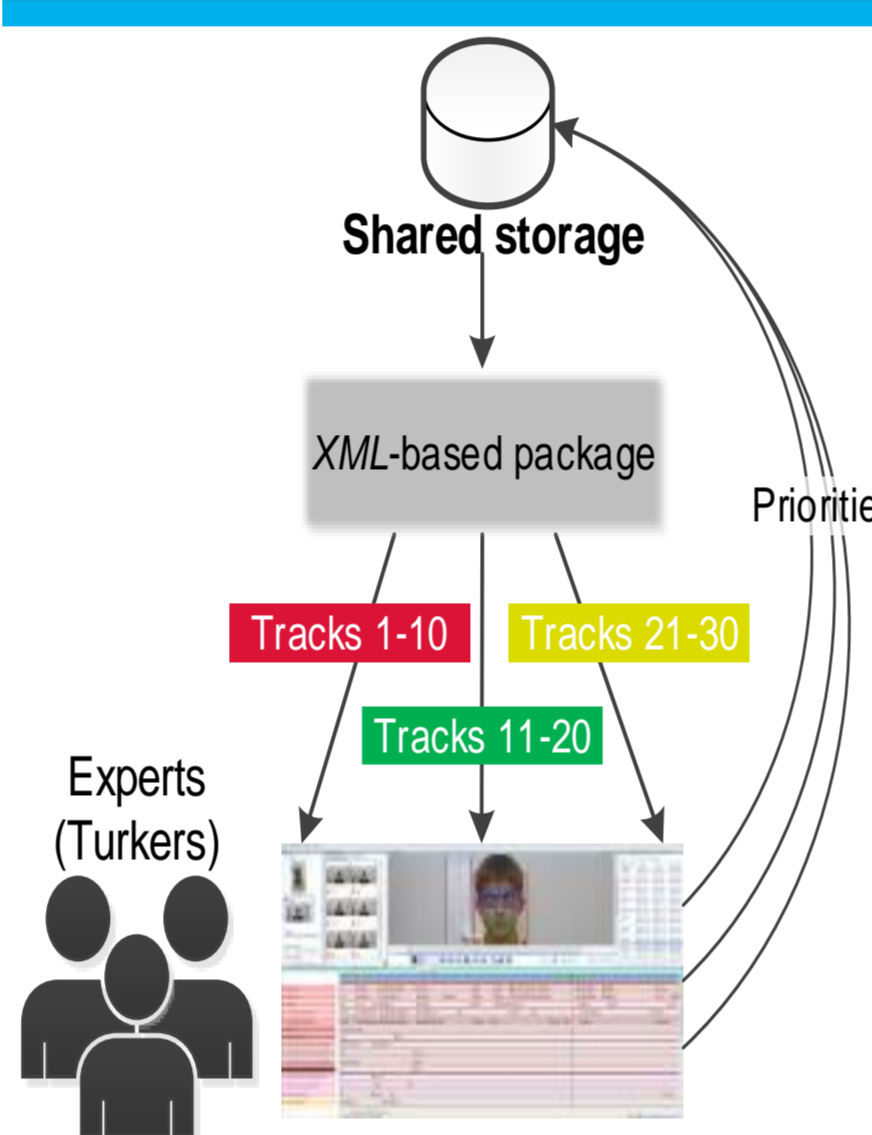
- For testing only **5!** packages with different subjects were available + 6 packages to train and test the ANN for knees positions
- Packages consisted of **5-7 min** HD facial and body (used only for display) videos with 25-30 fps + depth maps and other secondary files
- In average **1339** nonverbal events per package were recognized

Results grouped into 10 nonverbal categories

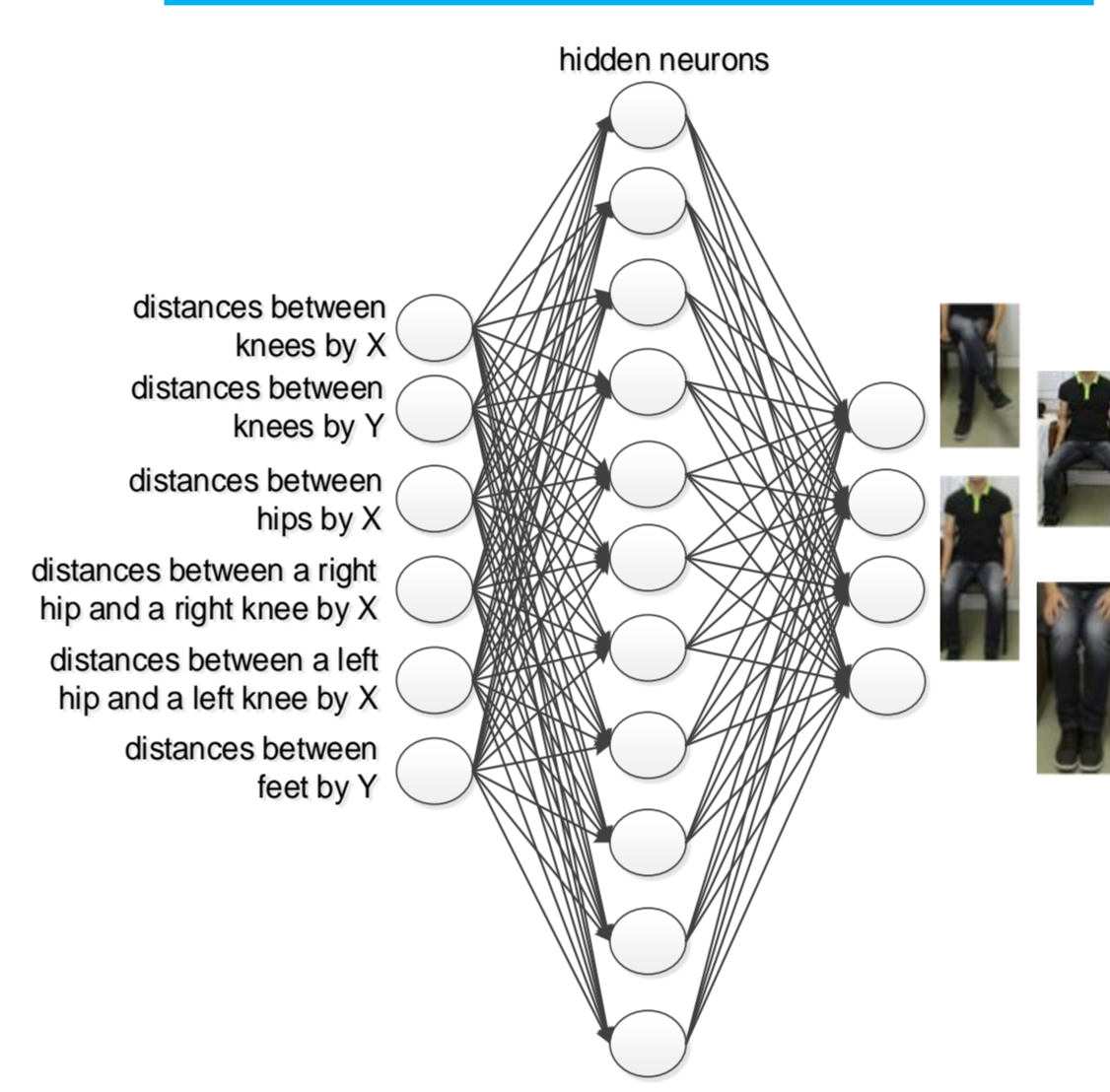
Nonverbal behavior group	N*	False positive	False negative
Eyes closed/opened states	178	0.35	0.39
Eyes blinking	115	0.30	0.42
Gaze directions	364	0.37	0.27
Eyebrows states	134	0.43	0.37
Lips and mouth corners states	164	0.24	0.31
Head position	36	0.40	0.29
Trunk position	68	0.39	0.30
Arms, hands and elbows position	74	0.34	0.35
Knees position	177	0.15	0.25
Feet and ankles position	29	0.30	0.38

*N is the average number of annotated segments of a nonverbal group per package

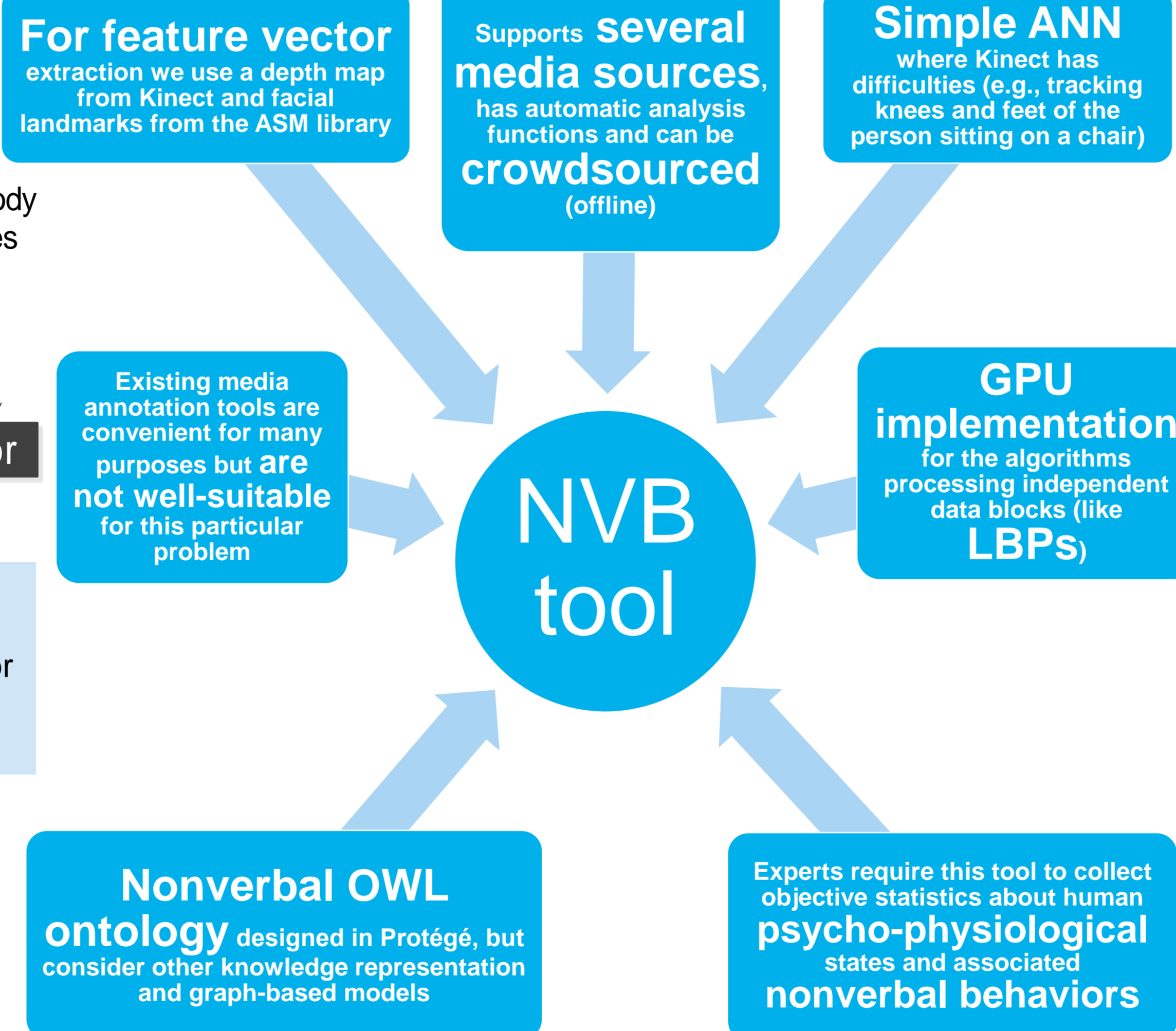
Offline crowdsourcing



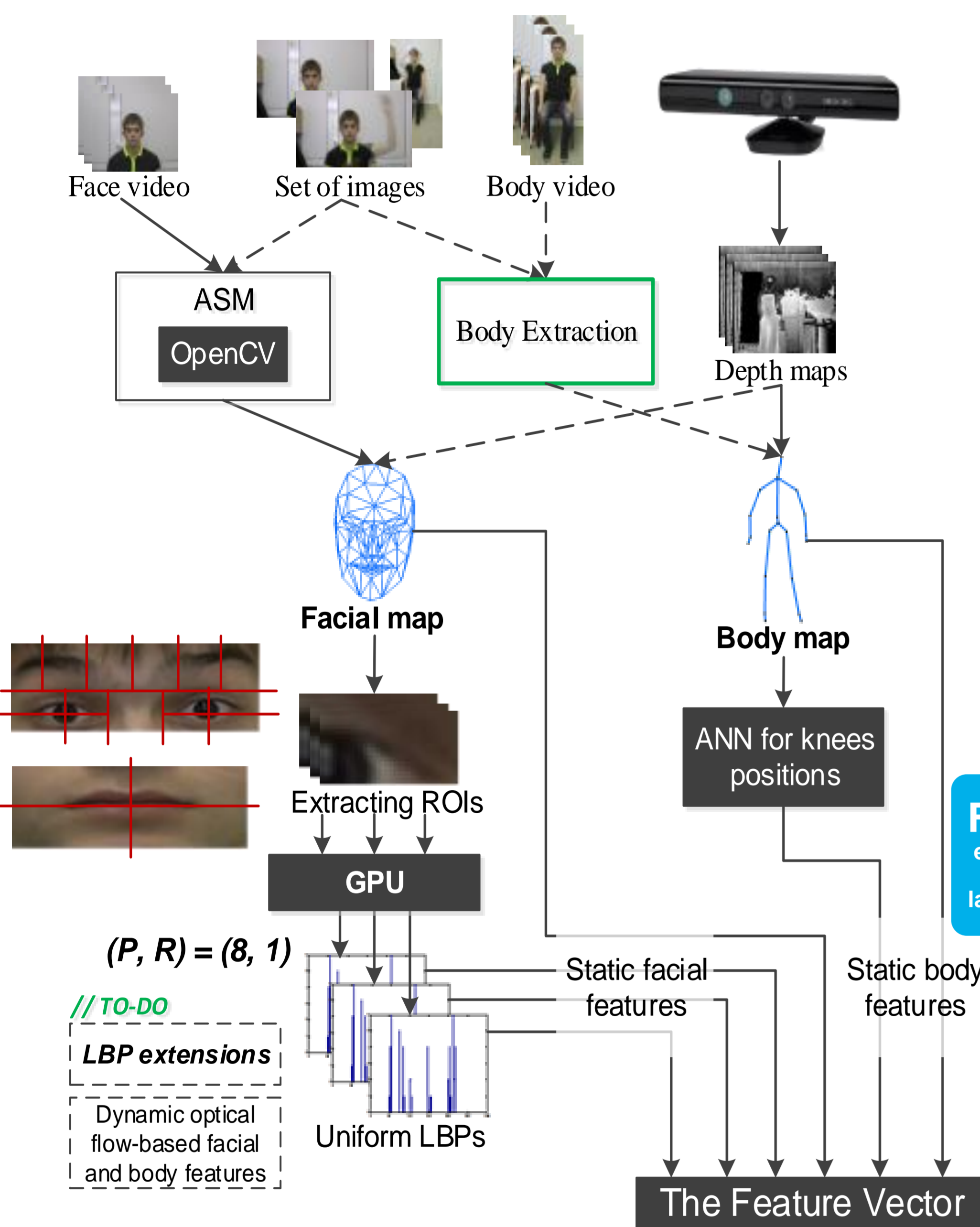
ANN



Project Features



NVB tool



The Feature Vector $S = \{S_{Holistic}, S_{Local}\} = \{S_{Kinect}, S_{LBPs}\}$

$S_{LBPs} = \{LBP_{Eyes}, LBP_{Eyebrows}, LBP_{Lips}\}$ – local component of the vector

sizeOf(uniform LBP) = 59 → N = 1082 – overall size of the vector

Algorithm 1. Pseudo code for LBP computation on a GPU

```

Require: TILE_W = 32, LBP_W = 3, indices = {0,1,2,5,8,7,6,3},
input, output, length
Result: uniform decimal codes for the input image
1. __global__ function declaration
2. begin
3. size = LBP_W^2
4. col, row, t = current thread X,Y and linear positions
5. if t > length then
6. return
7. __shared__ partialLBP[32x16] = read from the input
8. synchronize threads calling __syncthreads()
9. lbp_circle[size] = read current circle from partialLBP
10. threshold = central value of lbp_circle
11. if lbp_circle[0] >= threshold // first loop of cycle
12. dec_code += 0x80
13. for i = 1:size-1 // main loop
14. if lbp_circle[indices[i]] >= threshold then
15. dec_code += 0x80 >> i
16. if (bit value at i != bit value at (i-1)) then
17. ++transitions
18. if transitions > 2 then
19. dec_code = 0
20. break
21. synchronize threads calling __syncthreads()
22. write the dec_code value to the output
23.end

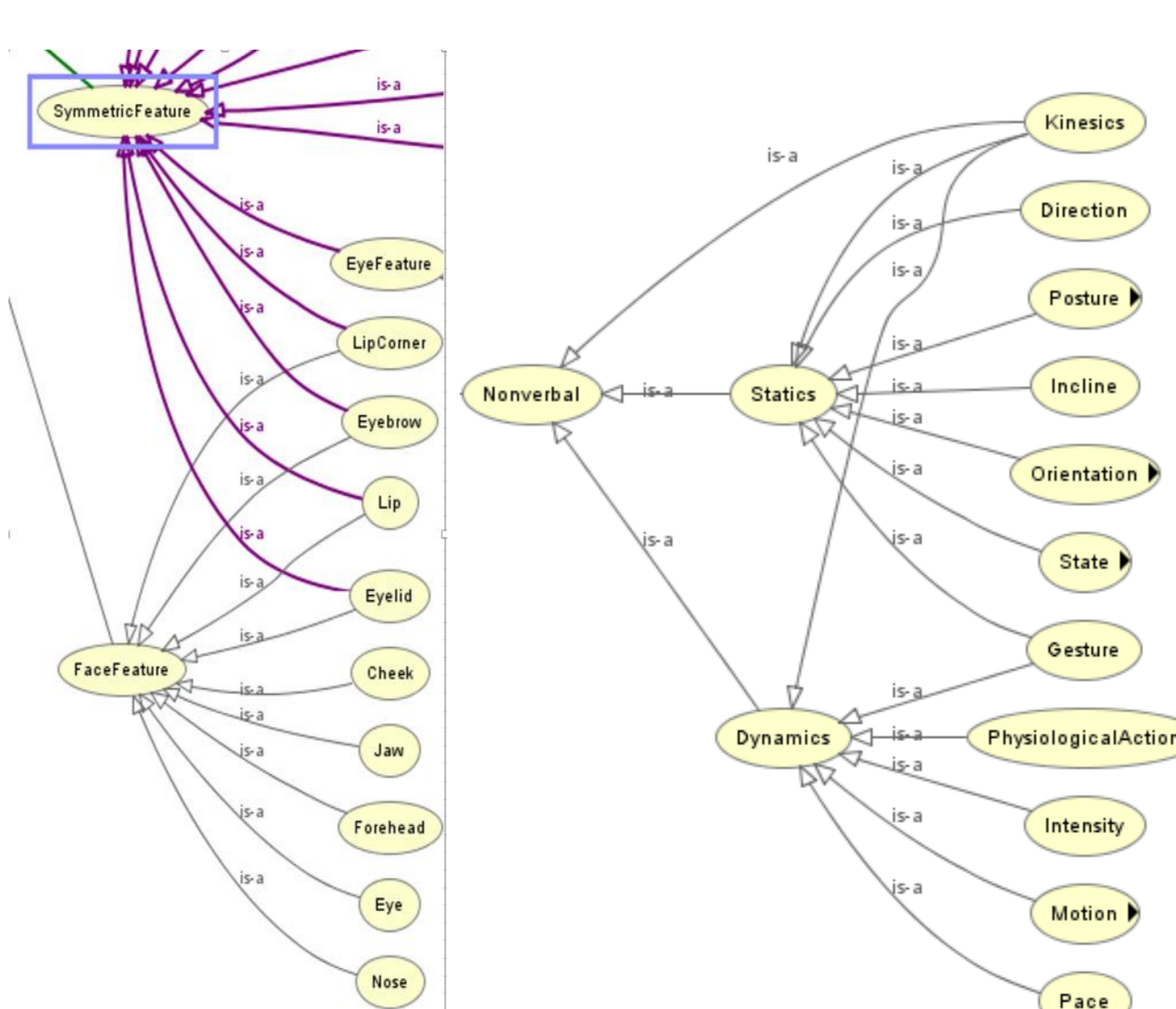
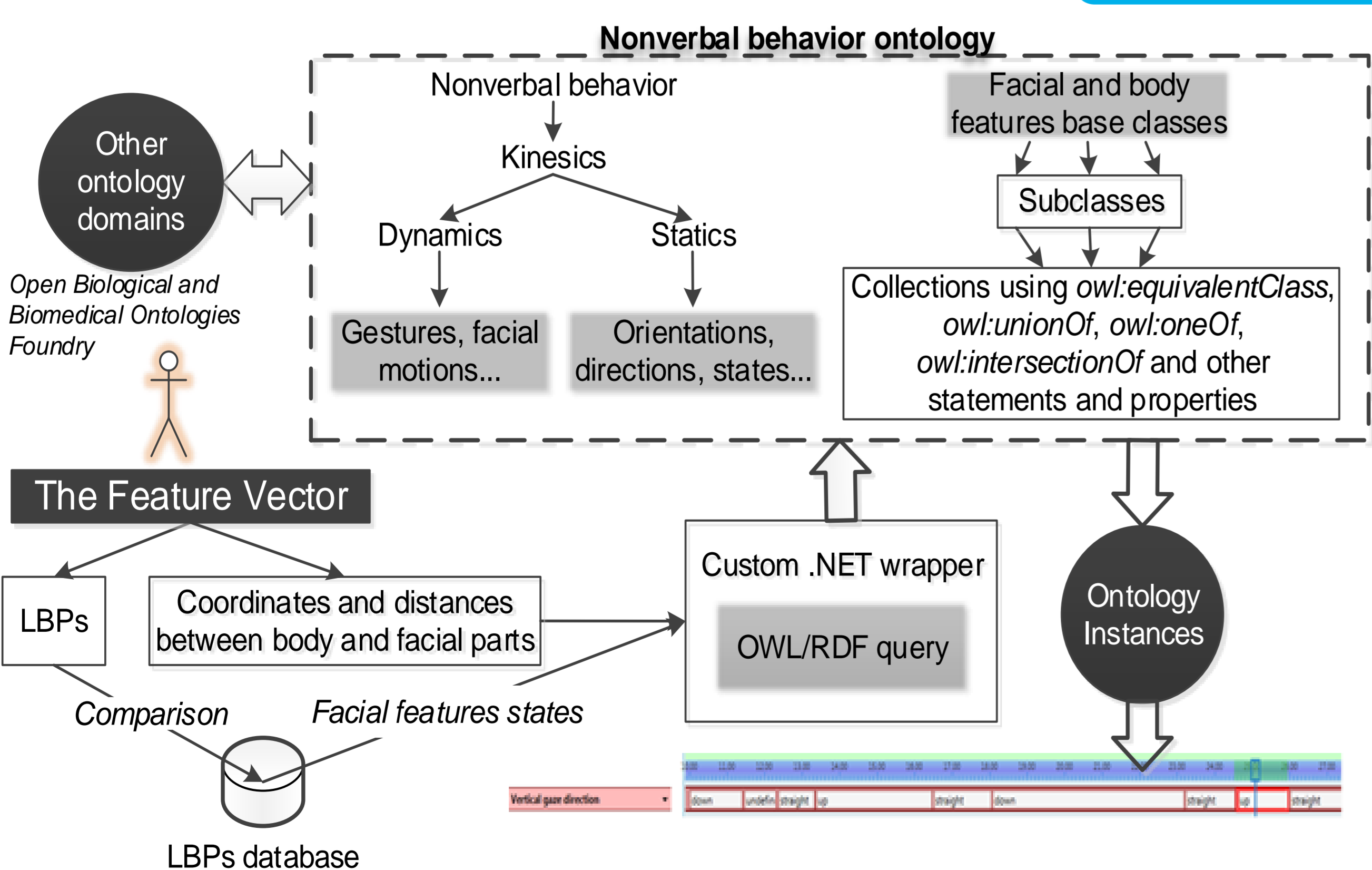
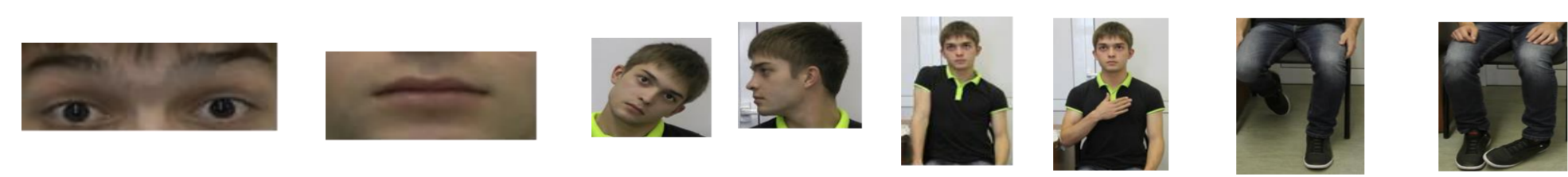
```

Conclusions

- Collected data *cannot* support or disprove these hypotheses at this moment because of the lack of more detailed "ground truth" data
- More powerful recording devices are required (new Kinect devices)
- Alternatively, new algorithms and local descriptors should be applied to video frames and regions of interests
- The results for publicly available data would be more reliable (ChaLearn datasets)
- Without a logic behavioral model our problem – the automatic generation of "ground truth" data about human facial and body activities – is difficult to be solved

Reference

Knyazev, B. Human nonverbal behavior multi-sourced ontological annotation. *Proceedings of the 2013 International Workshop on Video and Image Ground Truth in computer vision Applications*, St. Petersburg, Russia, 2013.



.OWL file