Extended Hierarchical Temporal Memory for Motion Anomaly Detection

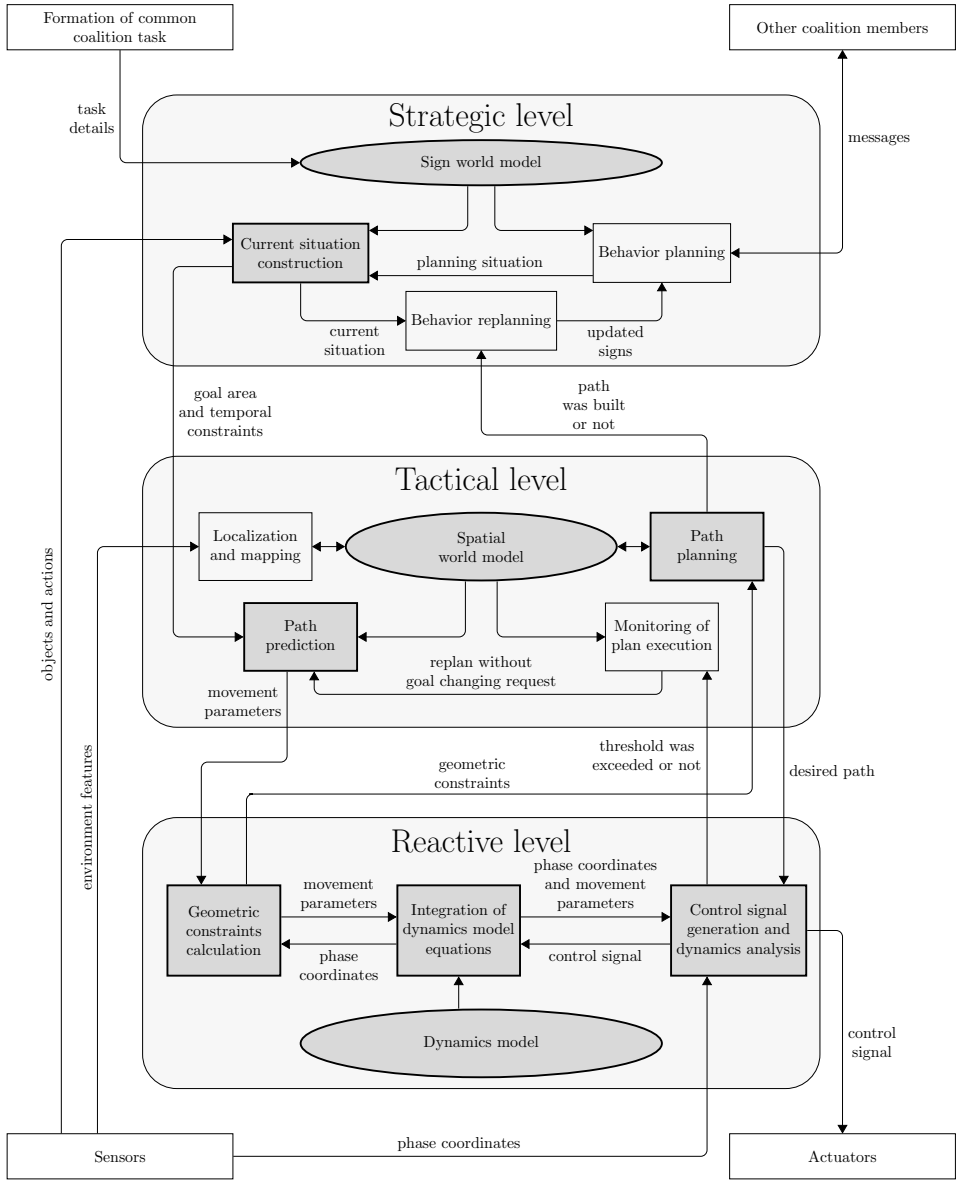
Ilya Daylidyonok, Anastasiya Frolenkova, Alelsandr I. Panov

Federal Research Center "Computer Science and Control"
Russian Academy of Sciences (RAS)
Moscow Institute of Physics and Technolgy
Moscow

August 24 – BICA 2018

STRL Architecture

- Cognitive functions modeling and construction of models that explain psychological phenomena.
- Algorithm of synthesizing the plan of behavior (algorithms MAP, MultiMAP, GoalMAP).
- Solving symbol grounding and symbol anchoring problems.
- Reconstruction of sign based world model of the actor based on texts.
- Text generation based on specific world models (virtual assistants).
- Multi-level architectures of control (robotic systems).



Sign based world model

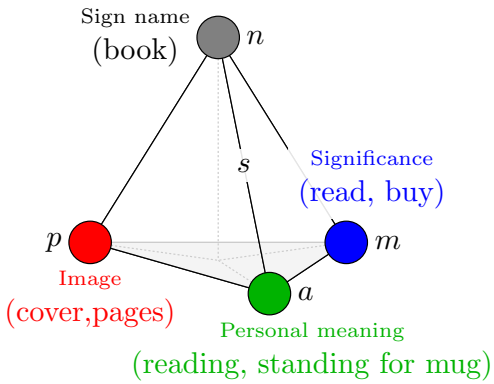
A component of knowledge representation is a sign:

- in sense of cultural-historical approach by L. Vygotsky,
- in sense of activity theory by A. Leontiev.

Sign based world model

A component of knowledge representation is a sign:

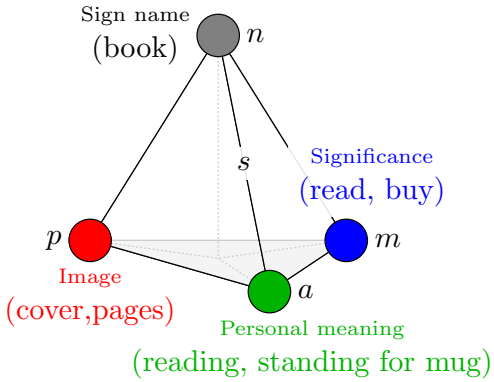
- in sense of cultural-historical approach by L. Vygotsky,
- in sense of activity theory by A. Leontiev.

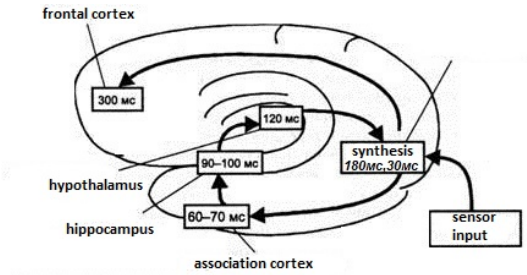


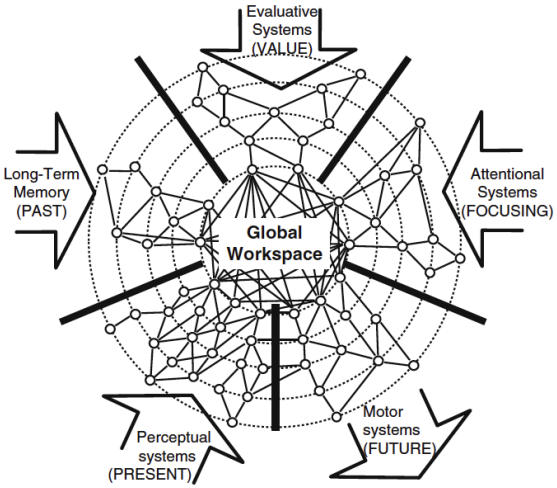
Sign based world model

A component of knowledge representation is a sign:

- in sense of cultural-historical approach by L. Vygotsky,
- in sense of activity theory by A. Leontiev.







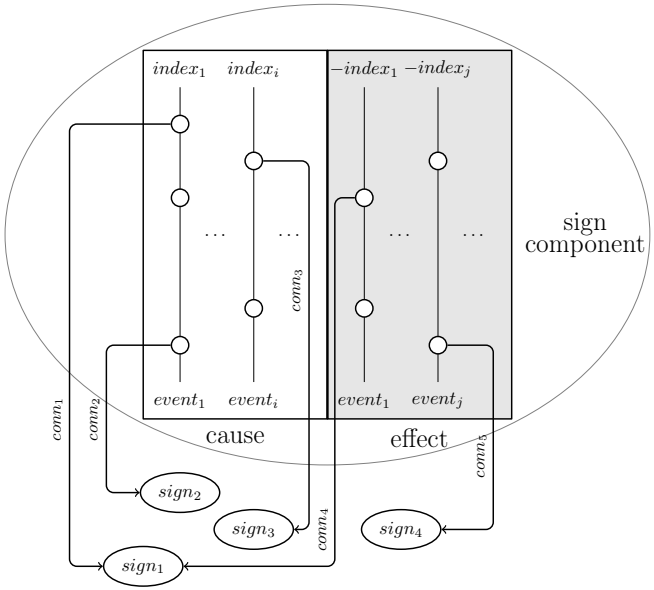
Supported ideas in psychology and biology:

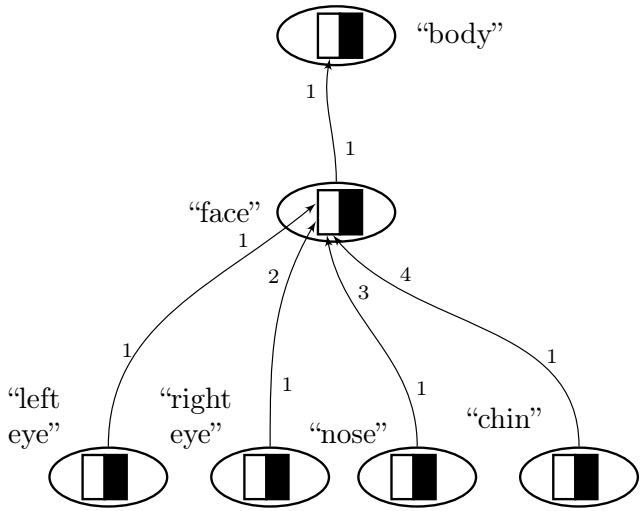
- neurophysiological data (Edelman, Ivanitsky, Mountcastle etc.),
- two and three levels psychological theories (Stanovich, Kahneman).

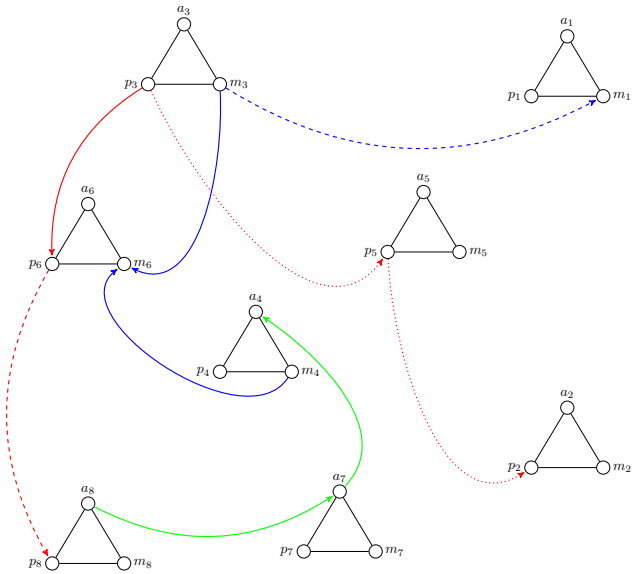
Osipov, G. S., A. I. Panov, and N. V. Chudova. "Behavior Control as a Function of Consciousness. II. Synthesis of a Behavior Plan". *Journal of Computer and Systems Sciences International*. 2015.

— ."Behavior control as a function of consciousness. I. World model and goal setting". *Journal of Computer and Systems Sciences International*. 2014.

Modeling of world model







Heterarchical causal network:

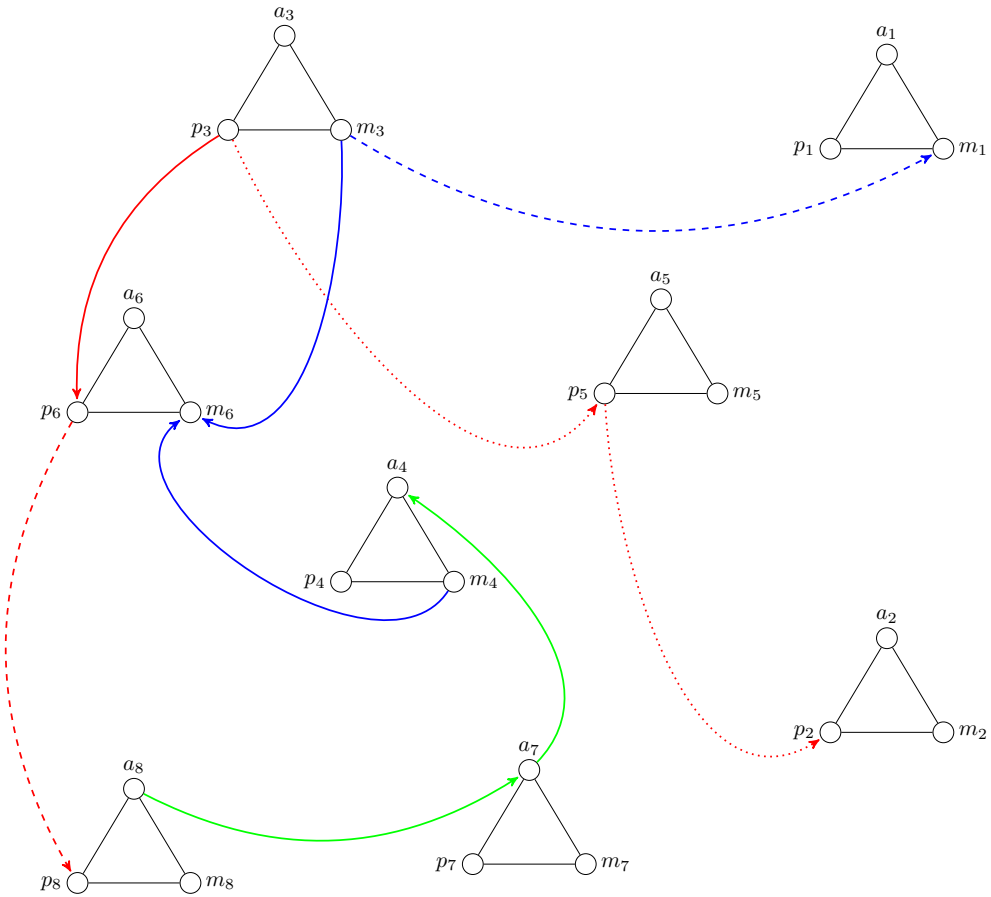
$$W_x = \langle V_x, E_x \rangle$$

$$v \to Z^x(s), x \in \{p, m, a\}$$

Panov, Aleksandr I. "Behavior Planning of Intelligent Agent with Sign World Model". *Biologically Inspired Cognitive Architectures*. 2017.

Osipov, Gennady S. "Signs-Based vs. Symbolic Models". *Advances in Artificial Intelligence and Soft Computing*.

Sign world model



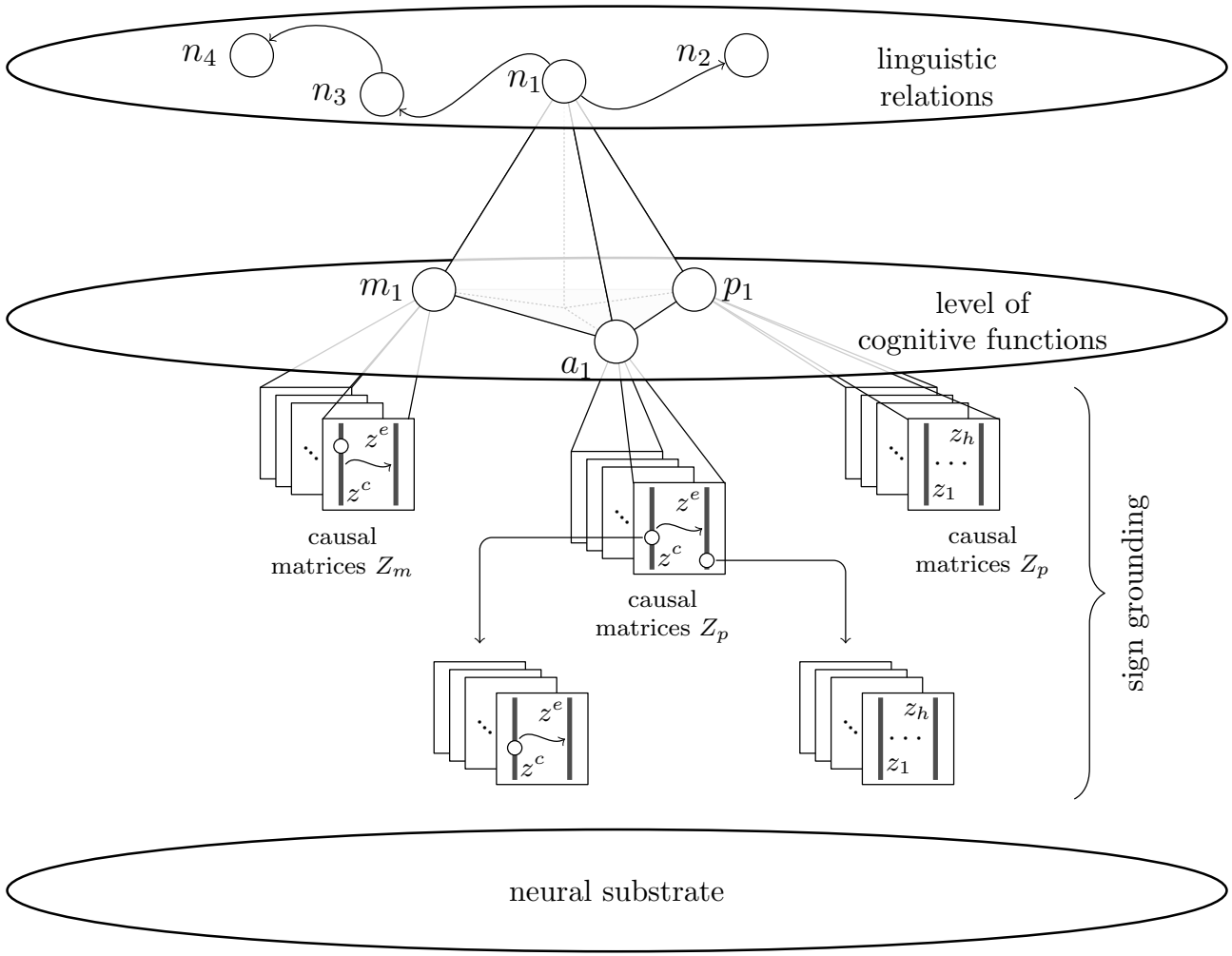
Semiotic network

$$\Omega = \langle W_p, W_m, W_a, R_n, \Theta \rangle$$

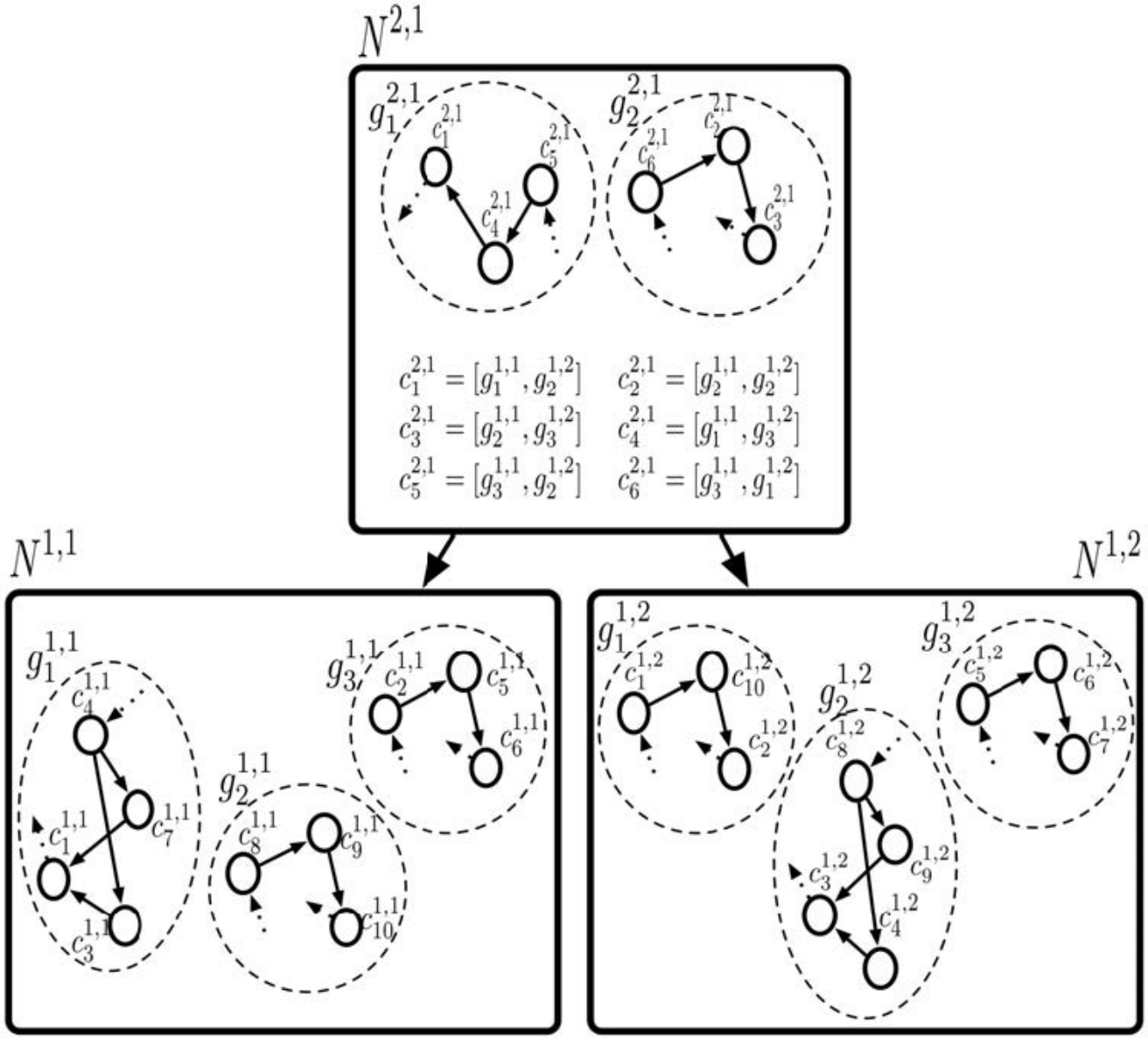
consisting of three causal network:

- $W_p = \langle 2^P, \mathfrak{R}_P \rangle$ – causal network on the set of sign images,
- $W_a = \langle 2^A, \mathfrak{R}_A \rangle$ – causal network on the set of sign meanings,
- $W_m = \langle 2^M, \mathfrak{R}_M \rangle$ – causal network on the set of sign significances.

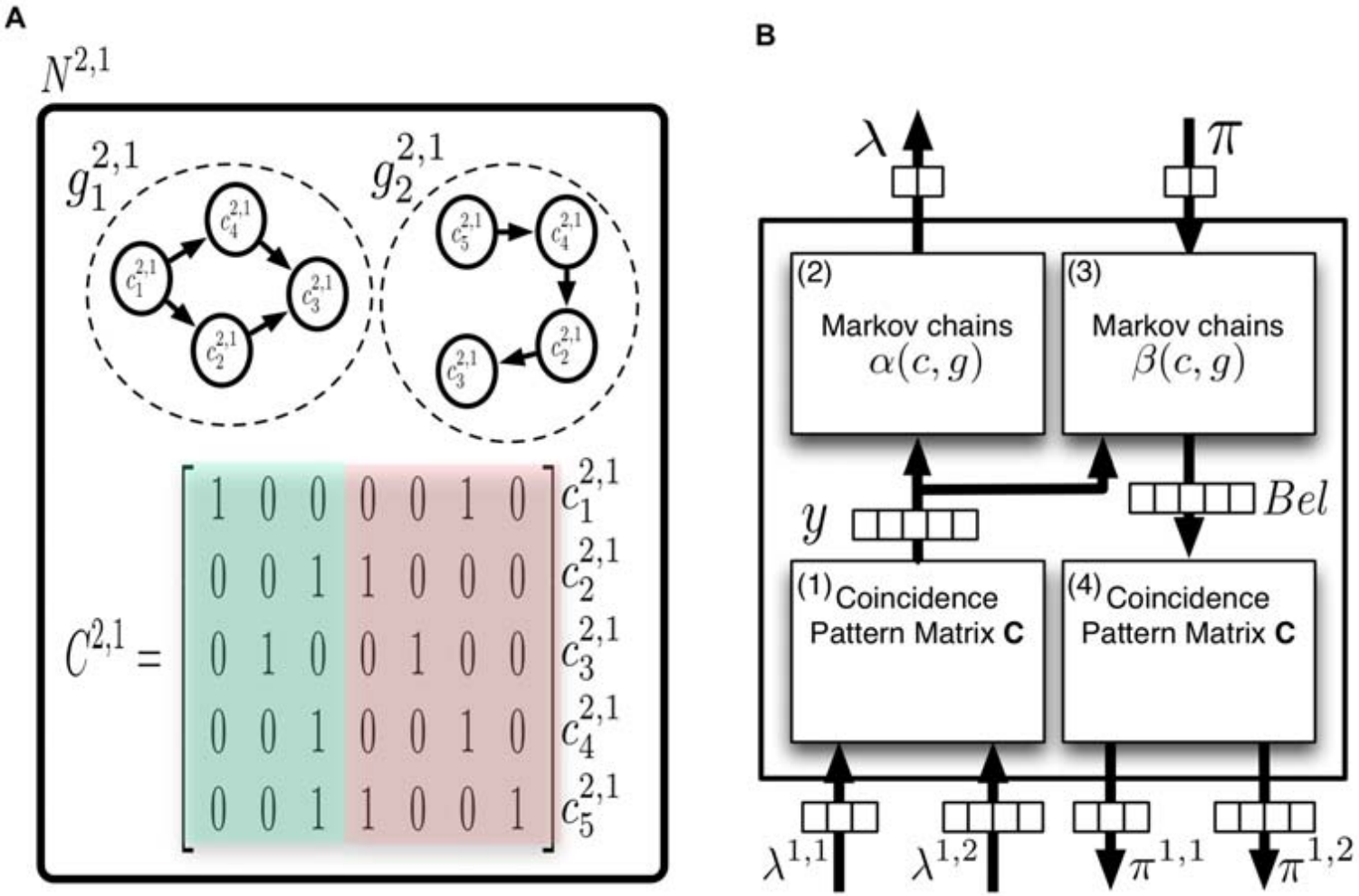
Sign world model



Formation of causal matrices

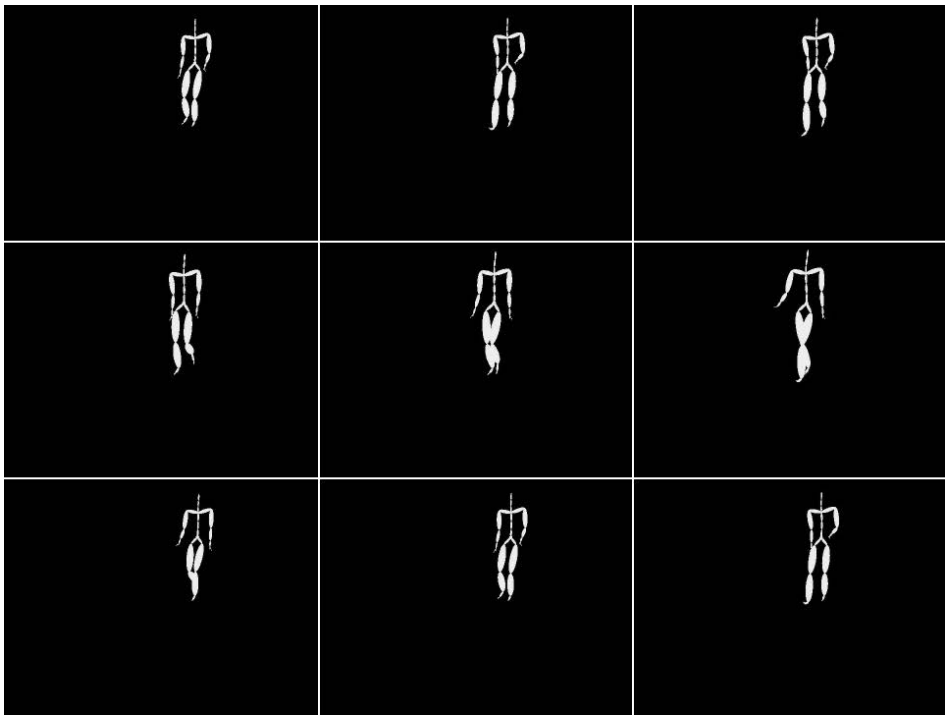


Formation of causal matrices



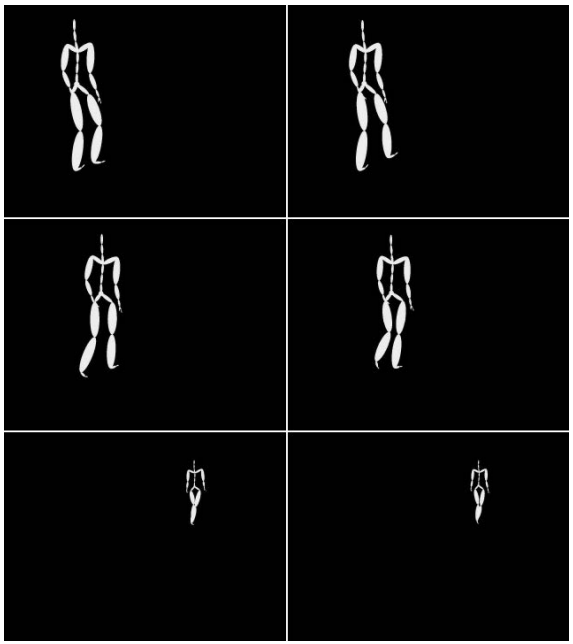
Anomaly Detection: Dataset

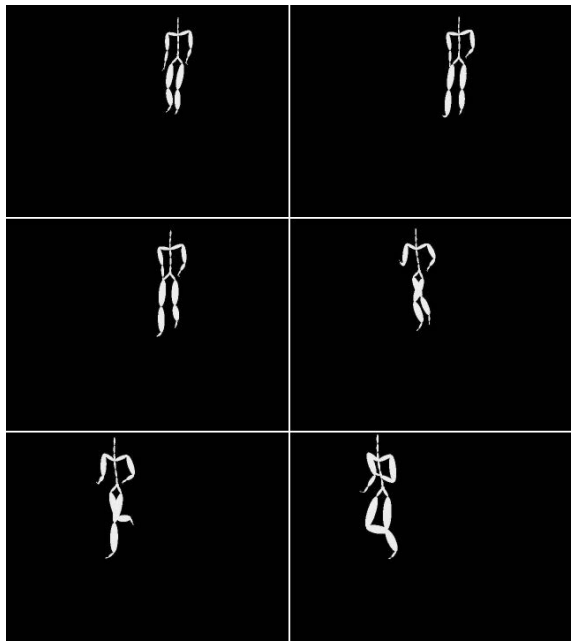
- Carnegie Mellon University Motion Capture Database.
- 2235 videos of 144 people performing different actions: walking, running, jumping and so on.
- 41 markers were positioned on a different actors' body parts and their positions were recorded during the action performance.



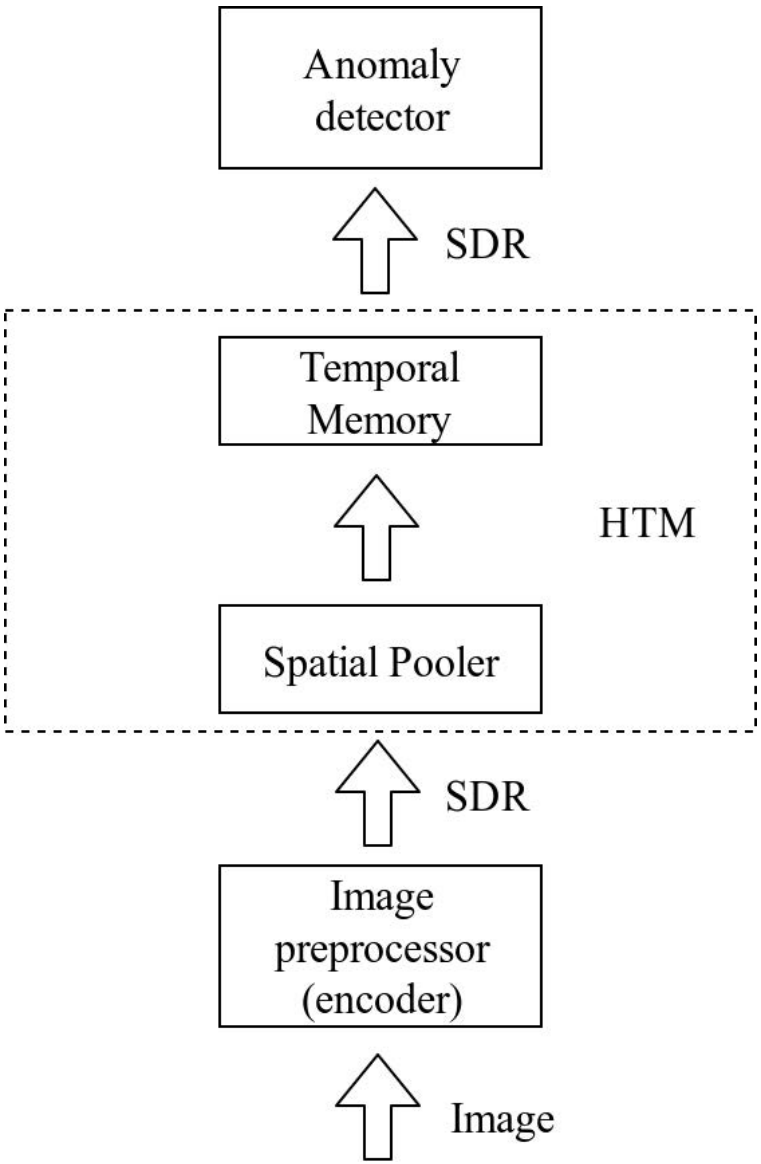
Anomaly Detection: Types

- Frames from other videos were inserted to produce an anomaly.
- Type 1: the object quickly changed the position.
- Type 2: the movements were similar, but different actions were performed.





Schema of the processing



Preprocessing:

- Image loading.
- Algorithms which control the sensor's movement around the image.

Score computing:

$$score = \frac{\sum_{i=0}^{n} \left(ac_{expected}^{(i)} - ac_{real}^{(i)} \right)^2}{n}$$

Experimental results

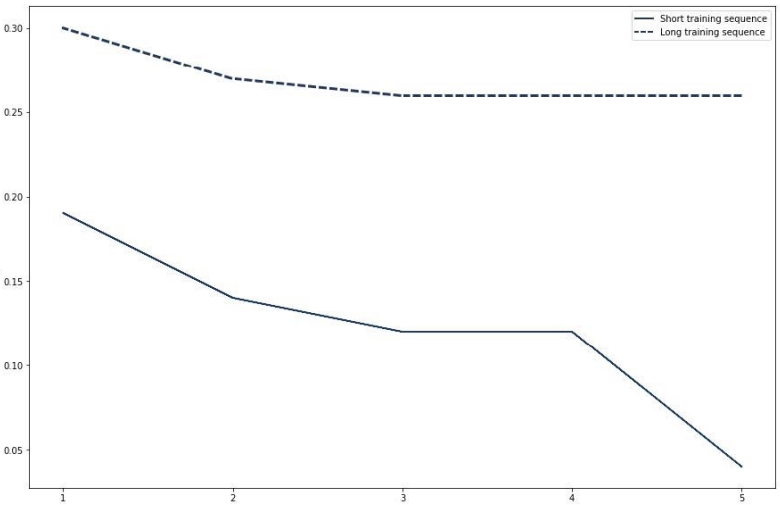
#	Original motion	# of frames	Anomaly motion	# of anomalous frames	# of detected anomalous frames	Score	Comments
1	walk	21	run	5	2	0.19	
2	walk	21	walk and turn right	12	5	0.19	
3	walk	21	jump	17	15	0.08	
4	walk	21	gorilla	12	2	0.40	wrong frames marked as anomaly
5	walk back and forth	136	scared walk	41	2	0.28	
6	walk back and forth	136	dance	31	2*	0.24	
7	walk back and forth	136	bad leg walk	22	2*	0.14	anomaly starts near the beginning
8	jump	30	walk back and forth	15	16	0.03	motion and location change
9	walk back and forth	136	walk back and forth	-	2*	-	test for sudden location change

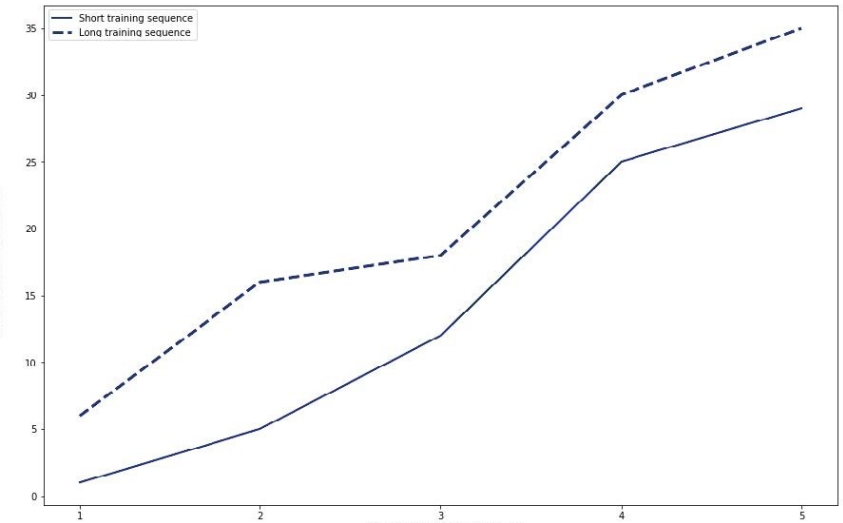
* Only first and last frames of anomaly were detected

Experimental results

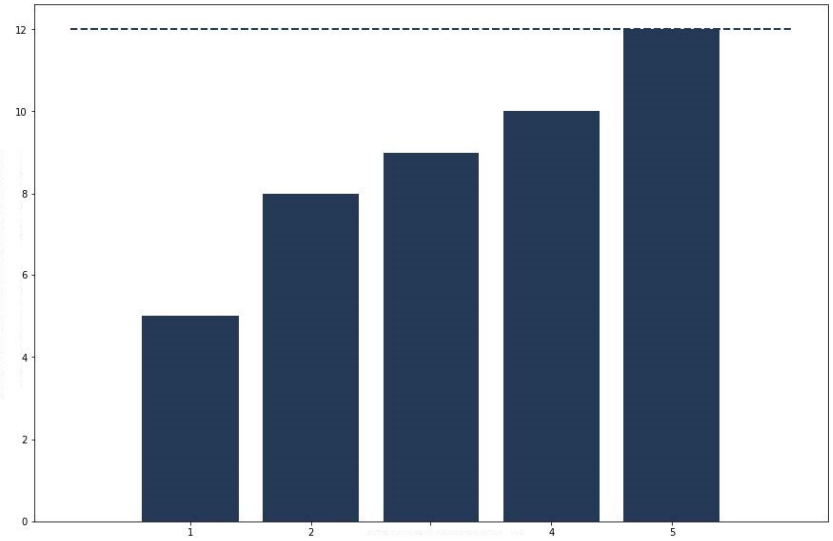
Extensions:

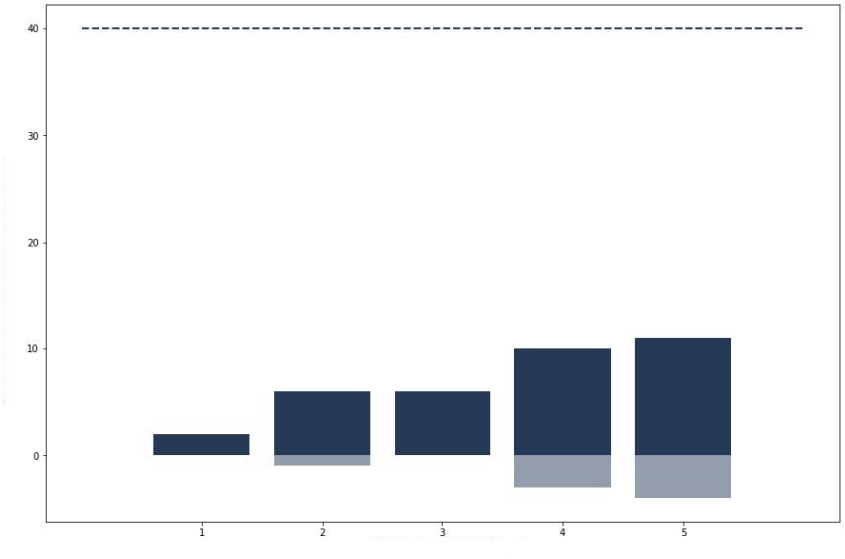
- A preprocessing step between layers: performing a convolution of temporal memory region's output with some function (kernel).
- Feedback: producing commands to a sensor's view which will move it accordingly, therefore focusing on different parts of the image. It will have an SDR of combined information as input: the output of higher level's temporal memory, the current frame of the video and reward.





Experimental results





Thank you for your attention!

panov.ai@mipt.ru