

# Методы распознавания места для локализации робота с использованием графа 3D-сцены

Цатурян Егор, 3 курс, кафедра «ПИР»

Научный руководитель:

Дмитрий Александрович Юдин

Ментор: Сергей Линок

Центр когнитивного моделирования МФТИ

20 мая, 2025г.

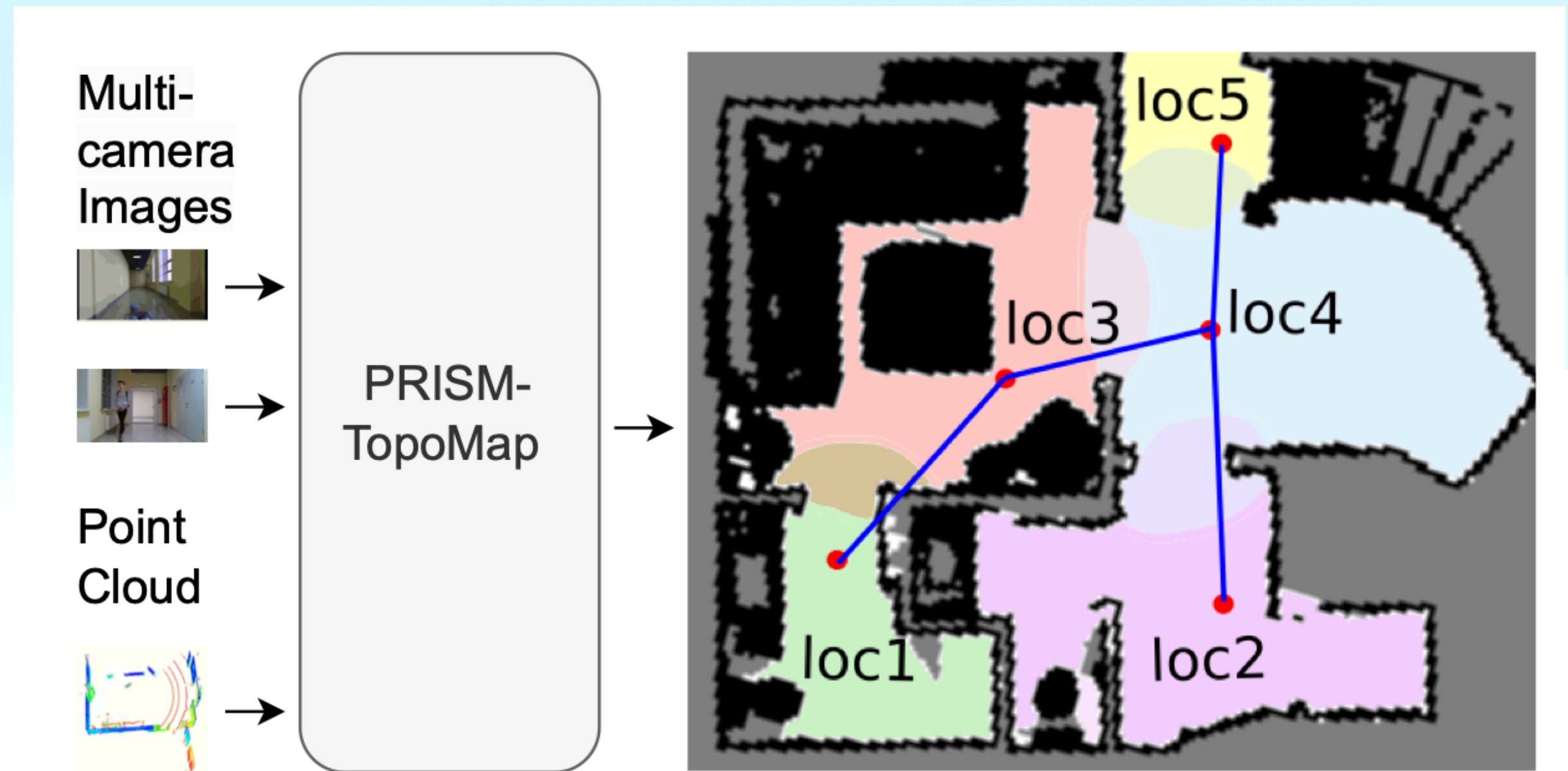
# Решаемая проблема и перспективы проекта:

- Локализация робота в помещениях и на открытой местности в условиях отсутствия спутниковых сигналов.
- Повышение качества распознавания места за счет учета взаимосвязей между объектами сцены.



# Моя задача

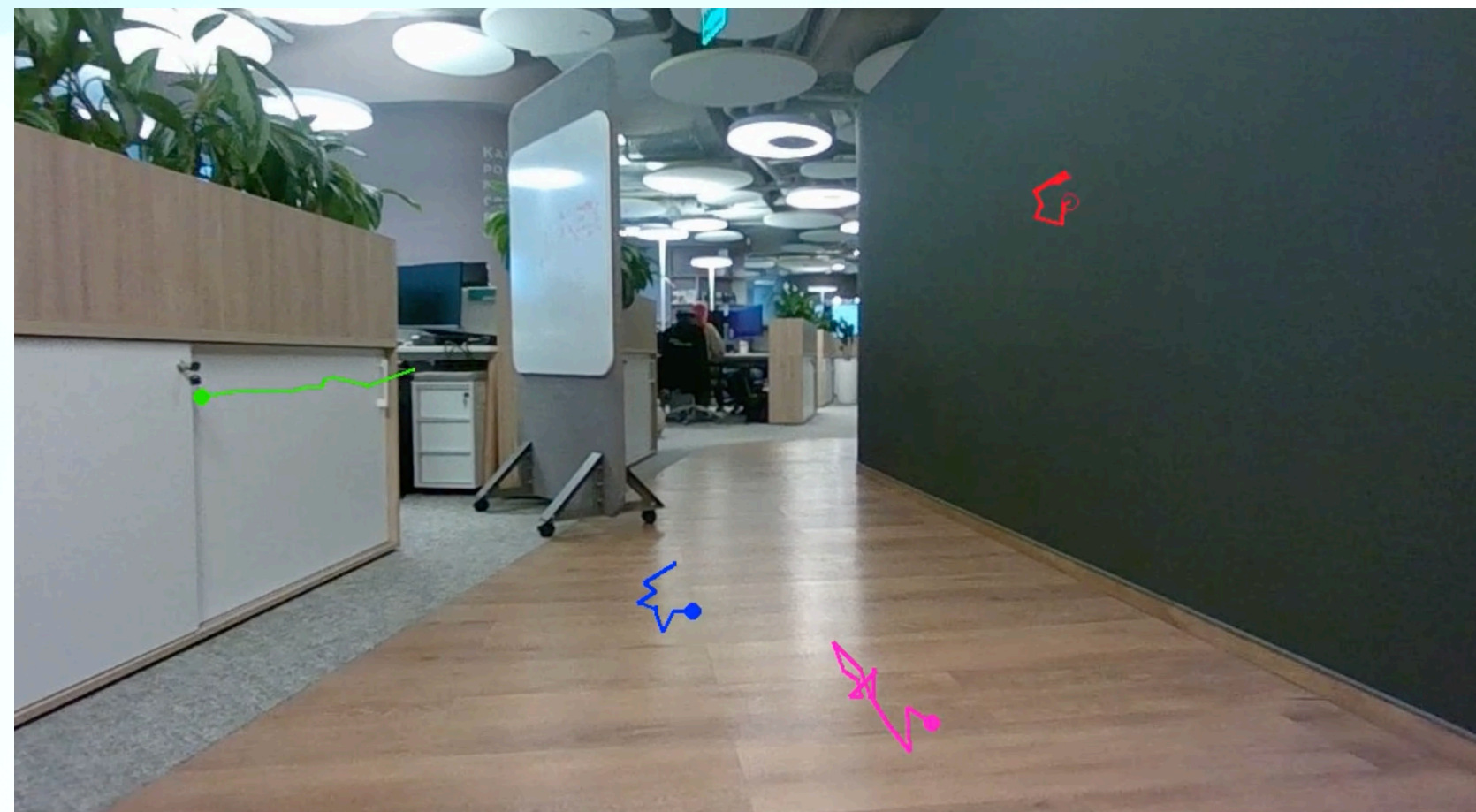
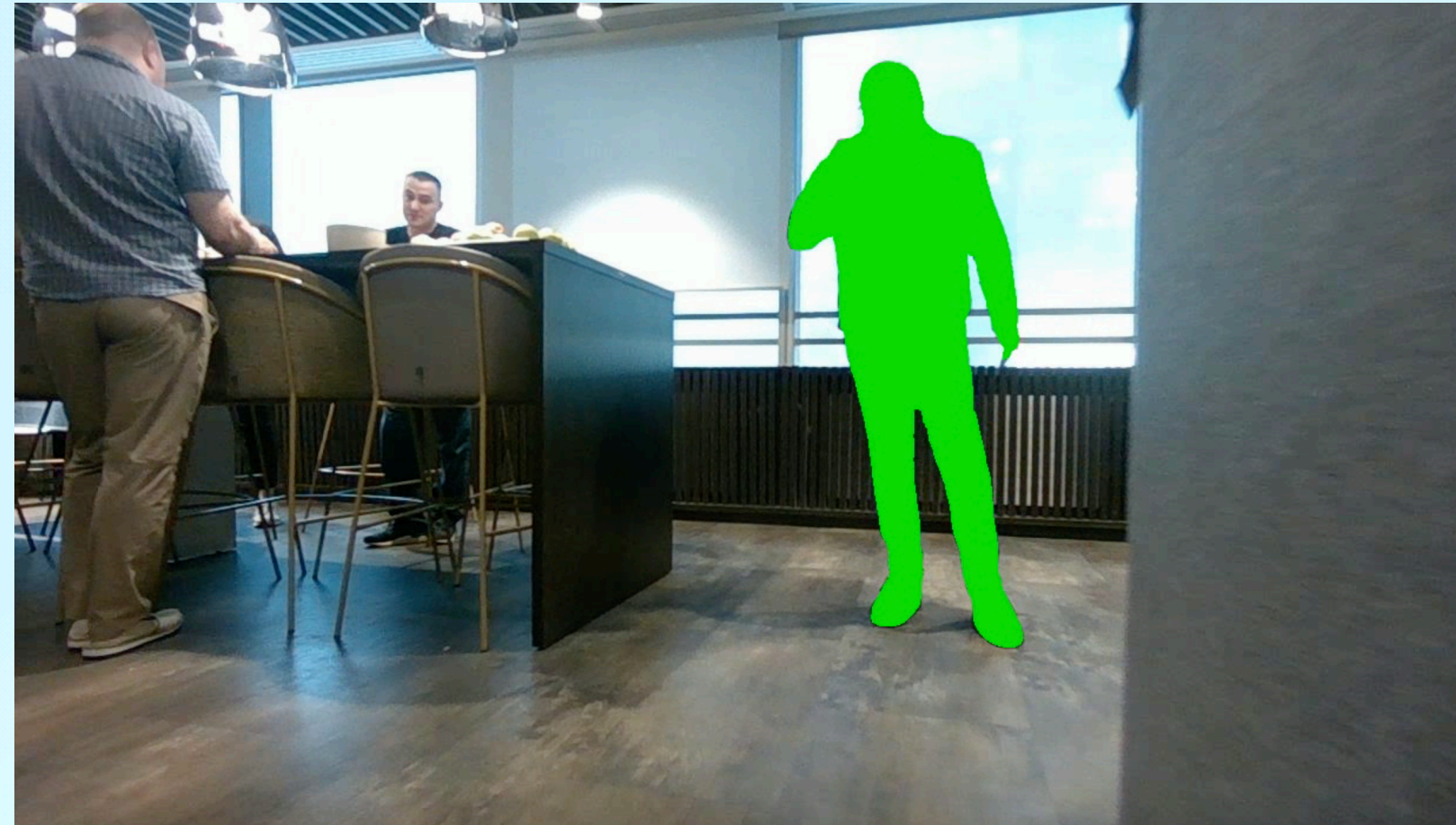
Разработка алгоритма для отслеживания перемещения объектов графа 3D-сцены по последовательности изображений и прогнозирование его изменения





# Существующие методы

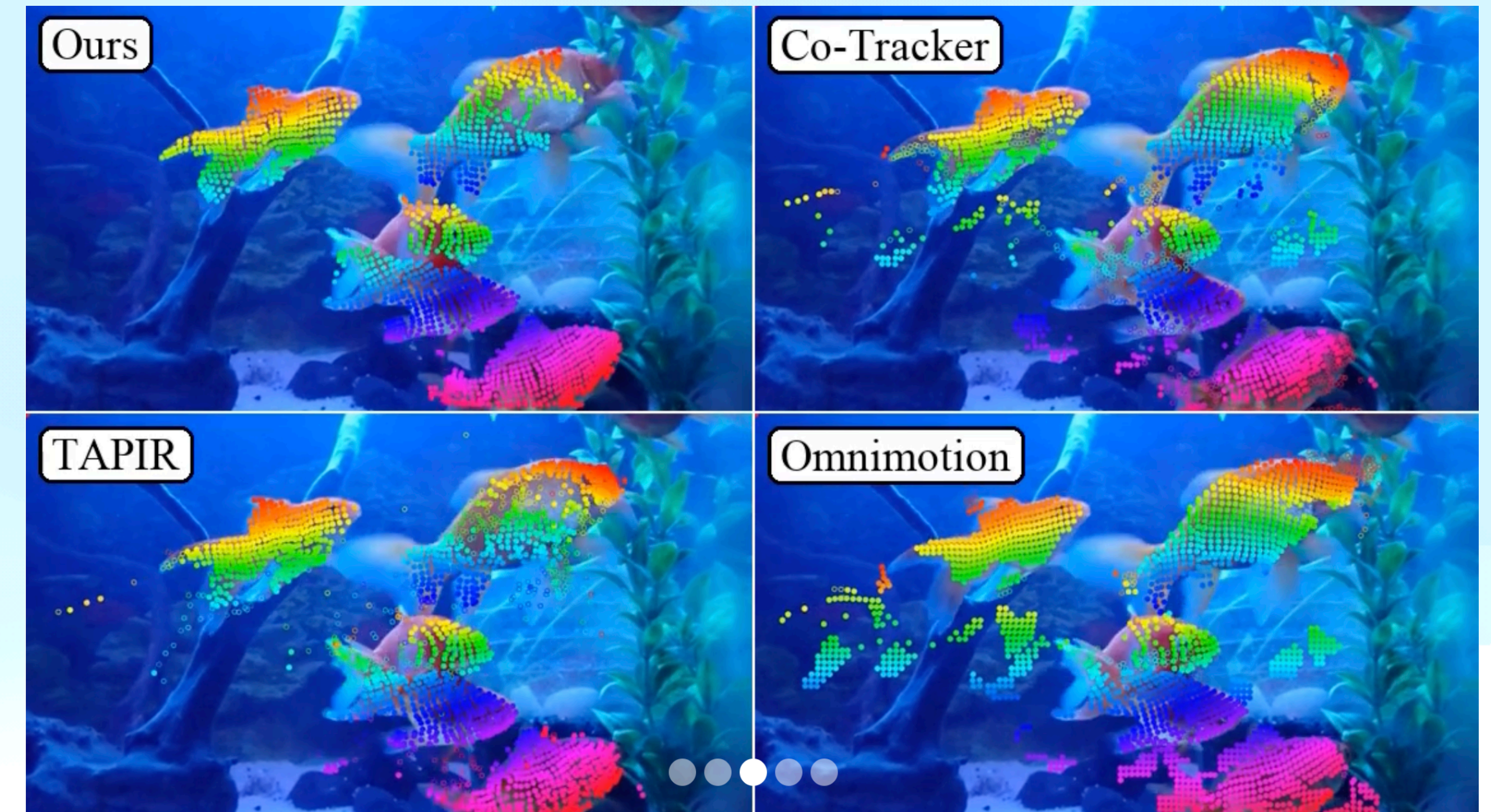
- SAM-2
- DINO-Tracker
- SAMURAI
- CoTracker3
- TAPIR





# Мотивация выбора базового метода

- CoTracker3 отслеживает отдельные точки объекта
- Оптимизированная архитектура
- Хорошо справляется с трекингом при большом occlusion rate (процент скрытия объекта за сторонними объектами)
- Поддерживает online и offline режим работы



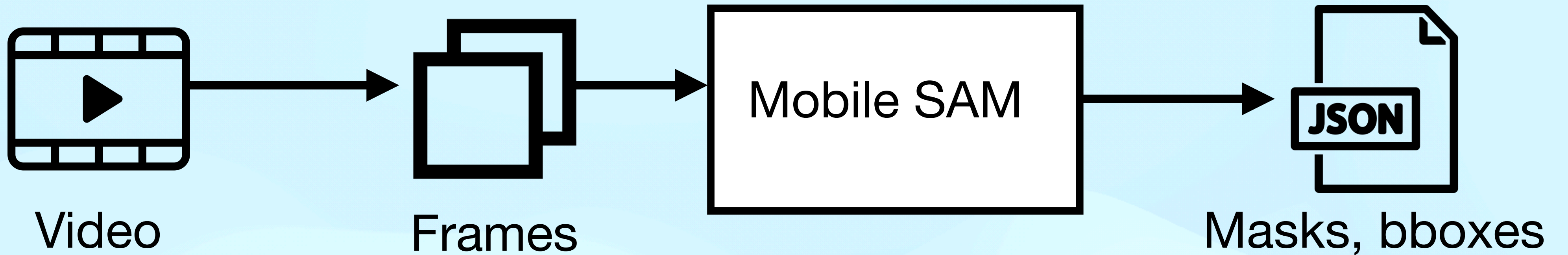
Method	Train	Kinetics			RGB-S			DAVIS			Mean
		AJ $\uparrow$	$\delta_{vis}^{avg} \uparrow$	OA $\uparrow$	AJ $\uparrow$	$\delta_{vis}^{avg} \uparrow$	OA $\uparrow$	AJ $\uparrow$	$\delta_{vis}^{avg} \uparrow$	OA $\uparrow$	$\delta_{vis}^{avg} \uparrow$
PIPs++ (Zheng et al., 2023)	PO	—	63.5	—	—	58.5	—	—	73.7	—	65.2
TAPIR (Doersch et al., 2023)	Kub	49.6	64.2	85.0	55.5	69.7	88.0	56.2	70.0	86.5	68.0
CoTracker (Karaev et al., 2024)	Kub	49.6	64.3	83.3	67.4	78.9	85.2	61.8	76.1	88.3	73.1
TAPTR (Li et al., 2024)	Kub	49.0	64.4	85.2	60.8	76.2	87.0	63.0	76.1	<u>91.1</u>	72.2
LocoTrack (Cho et al., 2024)	Kub	52.9	66.8	85.3	69.7	83.2	89.5	62.9	75.3	87.2	75.1
CoTracker3 (Ours, online)	Kub	54.1	66.6	87.1	71.1	81.9	90.3	<b>64.5</b>	<u>76.7</u>	89.7	75.1
CoTracker3 (Ours, offline)	Kub	53.5	66.5	86.4	<u>74.0</u>	<u>84.9</u>	90.5	63.3	76.2	88.0	75.9
BootsTAPIR (Doersch et al., 2024)	Kub+15M	54.6	68.4	86.5	70.8	83.0	89.9	61.4	73.6	88.7	75.0
CoTracker3 (Ours, online)	Kub+15k	<b>55.8</b>	<b>68.5</b>	<b>88.3</b>	71.7	83.6	<u>91.1</u>	63.8	76.3	90.2	<u>76.1</u>
CoTracker3 (Ours, offline)	Kub+15k	<u>54.7</u>	67.8	<u>87.4</u>	<b>74.3</b>	<b>85.2</b>	<b>92.4</b>	<u>64.4</u>	<b>76.9</b>	<b>91.2</b>	<b>76.6</b>

Table 1: **TAP-Vid benchmarks** CoTracker3 trained on synthetic Kubric shows strong performance compared to other models, while the online version fine-tuned on 15k additional real videos (Kub+15k) outperforms all the other methods, even BootsTAPIR trained on 1,000× more real videos. Training data: (Kub) Kubric (Greff et al., 2022), (PO) Point Odyssey (Zheng et al., 2023).

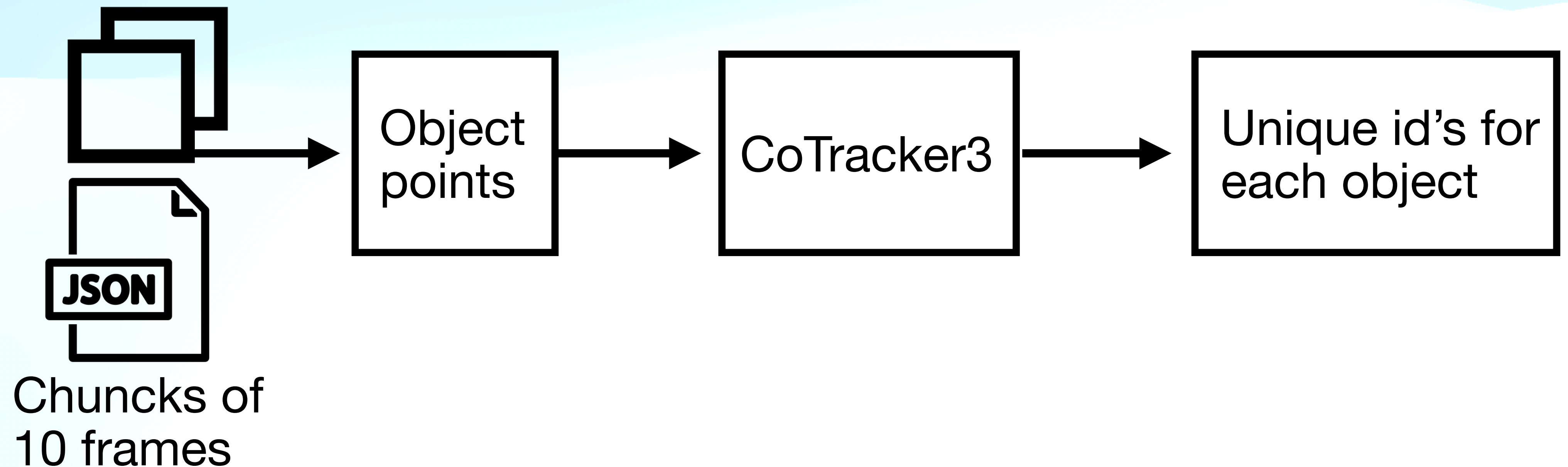


# Предложенный подход

Step 1:



Step 2:



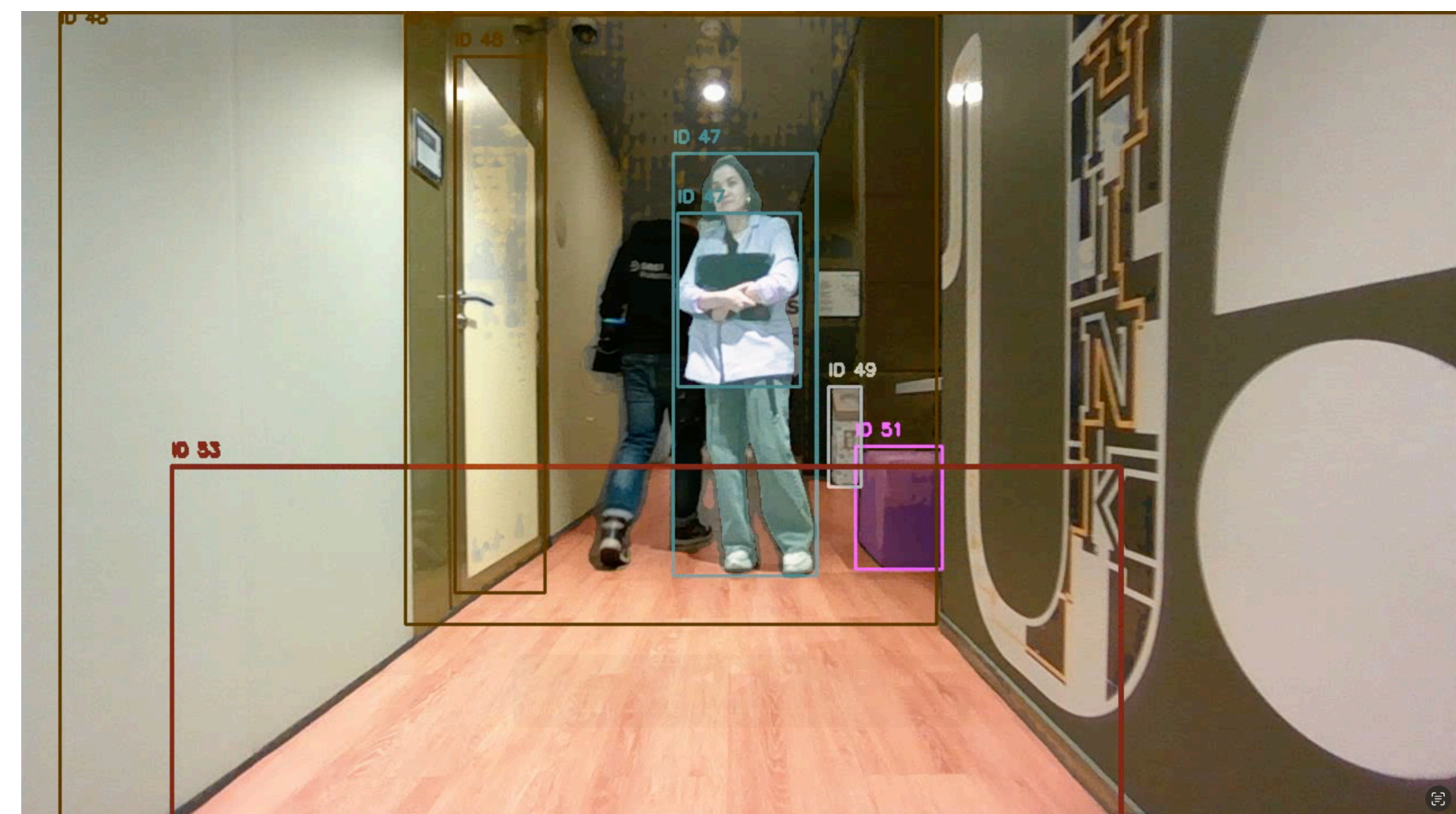
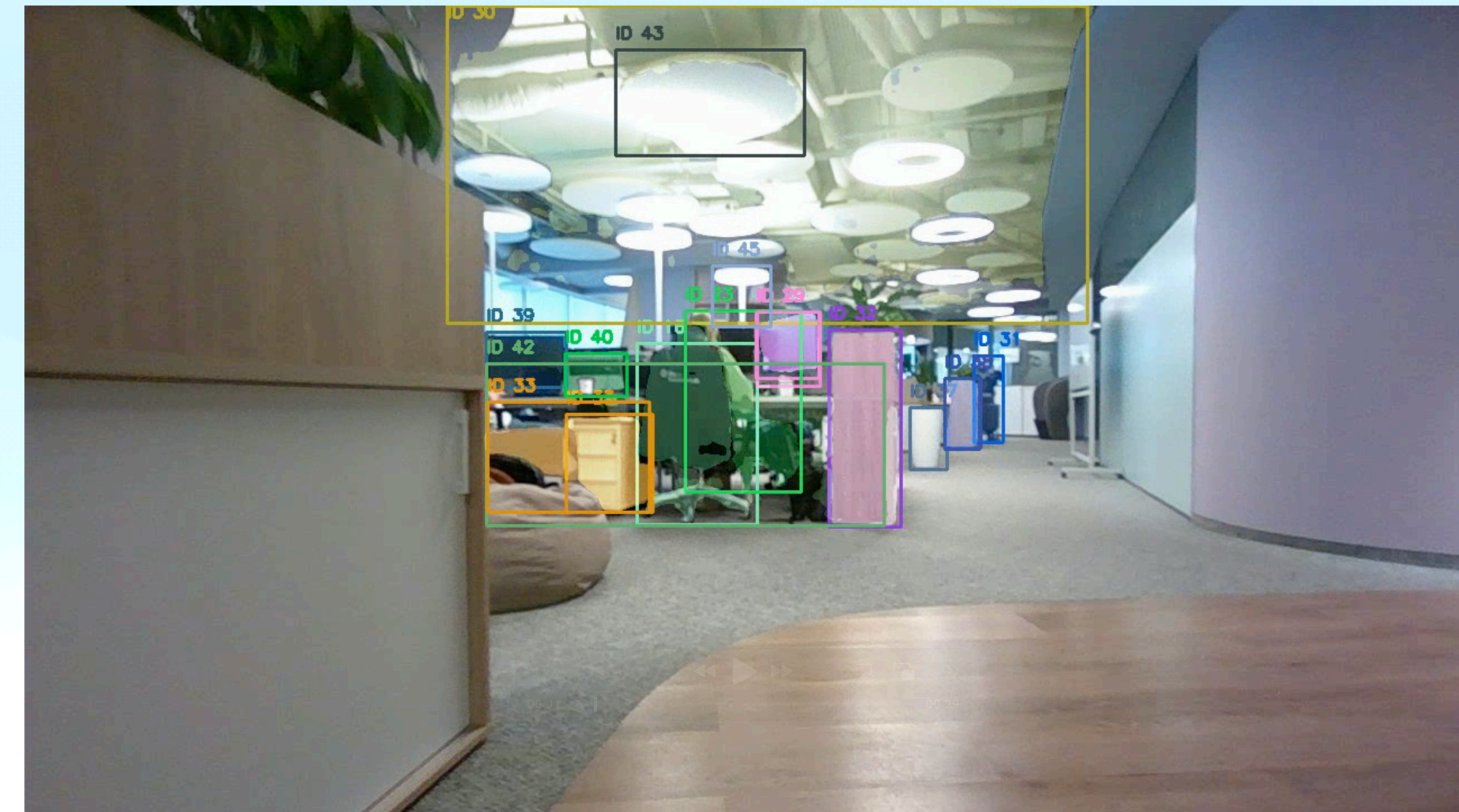


# Сравнение с существующими подходами

SAM-2: 19:31:43

My solution:

Chunk size	Time
5	11:04:24
10	10:09:71
15	09:39:15
20	08:33:45
25	07:41:23





# Вывод

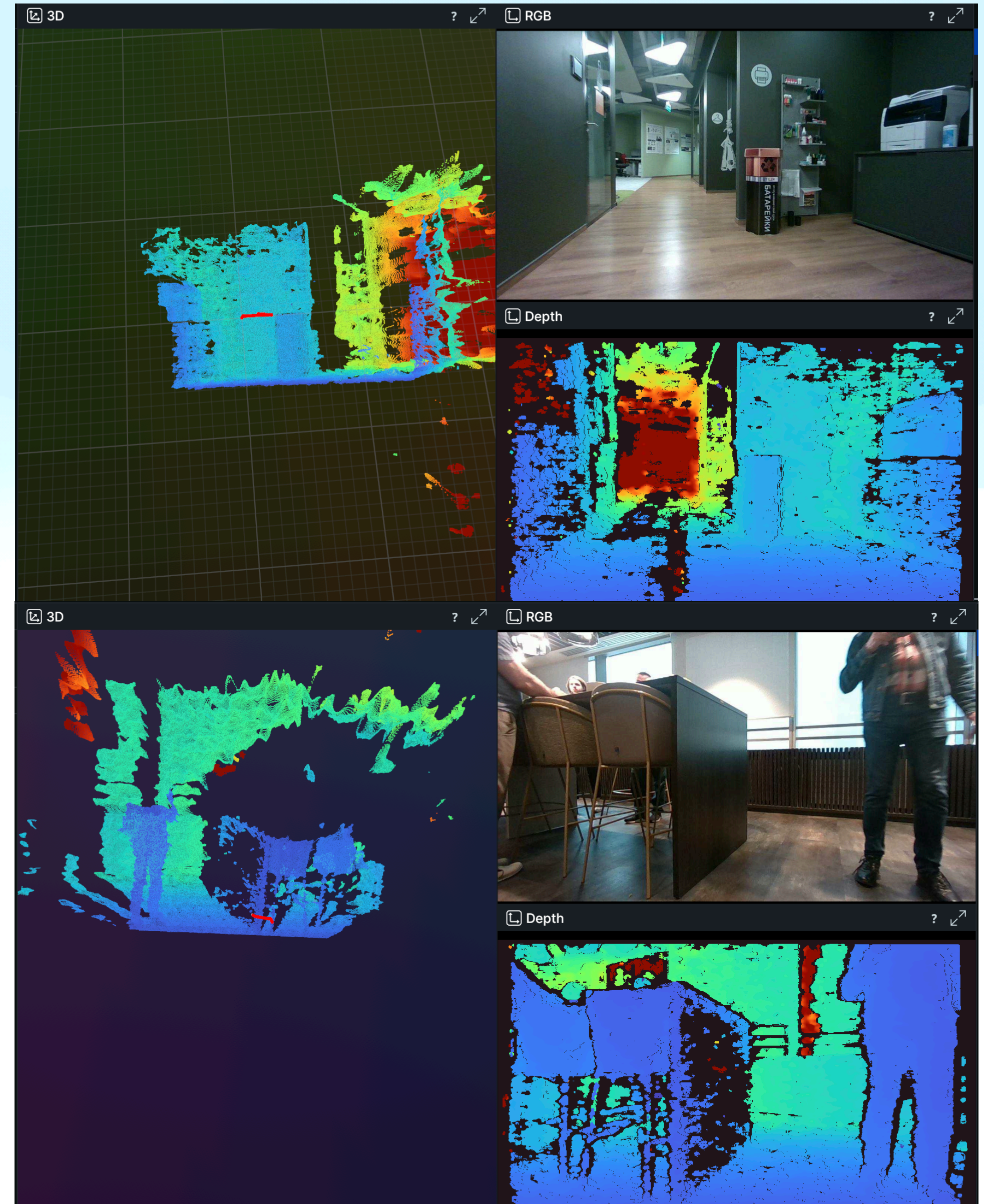
Разработан эффективный метод трекинга, используемый в едином фреймворке быстрых и надежных нейросетевых подходов распознавания места, который команда Центра когнитивного моделирования планирует внедрять на практике в партнерстве с Центром робототехники Сбера.





# Возможные улучшения

- Добавление нескольких точек на объекты, нахождение баланса «точность/время работы» метода
- Доработка визуализации результатов в 3Д версии





# References

[https://docs.google.com/spreadsheets/d/19zumUqsBUrXaEuGZgLbgWg6VOZLBxaVGoF\\_f-x2vyu4/edit?gid=0#gid=0](https://docs.google.com/spreadsheets/d/19zumUqsBUrXaEuGZgLbgWg6VOZLBxaVGoF_f-x2vyu4/edit?gid=0#gid=0)

<https://www.youtube.com/playlist?list=PLUjeNilwUr-GNTc1Mn5YRCTIg7ib1Ste5>