

Sim2Real для автономного вождения с использованием скрытого пространства

Сазанович Никита Валерьевич, МПА191С

Научный руководитель:
заведующий центром анализа данных и машинного обучения,
Шпильман Алексей Александрович



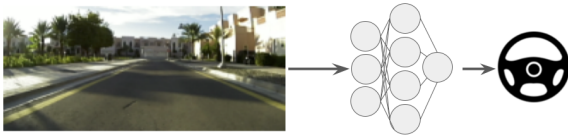
ВЫСШАЯ ШКОЛА ЭКОНОМИКИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

7 июня 2021

Предметная область

Два подхода к разработке алгоритма автономного вождения:

- последовательный конвейер
- обработка от начала до конца



Имеем набор данных $D = \{x_i, y_i\}_{i=1}^{|D|}$.

Необходимо задать **политику вождения**: $f_\theta : X \rightarrow Y$ и найти оптимальные относительно среднего абсолютного (САО) или квадратичного (СКО) отклонения параметры θ .

Область Sim2Real для автономного вождения

Трудности обучения алгоритмов в реальном мире:

1. трудоемкость сбора данных
2. политика вождения определена до начала сбора данных
3. проблемы этического характера при сборе экстремальных ситуаций
4. высокая размерность входных данных и факторов, которые задают их распределение

Цель области

Как, используя данные из симулятора, обучить политику вождения для реального мира?

Подходы Sim2Real можно объединить в следующие категории:

- метод рандомизации характеристик симуляционных доменов
- перенос высокоуровневых политик
 - использование модульности и абстракций¹
- установление соответствия между входными данными симулятора и реального мира
 - метод трансляции изображений²

Наблюдение

В литературе не исследованы Sim2Real методы использующие ограничения, основанные на темпоральном порядке данных.

¹Driving Policy Transfer via Modularity and Abstraction. / M. Müller [и др.] // CoRL. 2018.

²Learning to Drive from Simulation without Real World Labels. / A. Bewley [и др.] // 2019 International Conference on Robotics and Automation (ICRA). 2019. С. 4818—4824.

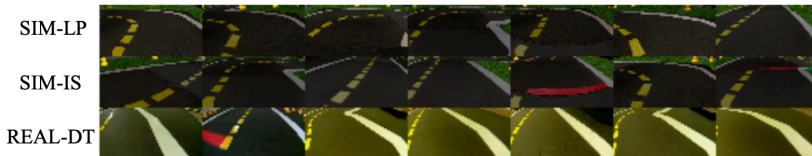
Цель

Исследовать методы Sim2Real, которые вводят дополнительные темпоральные ограничения на представления в скрытом пространстве, на примере платформы Duckietown.

Задачи:

1. собрать симуляционные и реальные наборы данных
2. реализовать методы Sim2Real для переноса политики
3. разработать дополнительные темпоральные ограничения
4. сравнить результаты методов и симуляционных окружений для задачи переноса
5. провести анализ полученных моделей

Сбор данных



Платформа Duckietown для сбора данных: изображения имеют размер 640x480, а управление задается напряжением привода на левом и правом колесе.

Для экспериментов было собрано 3 набора данных с помощью:

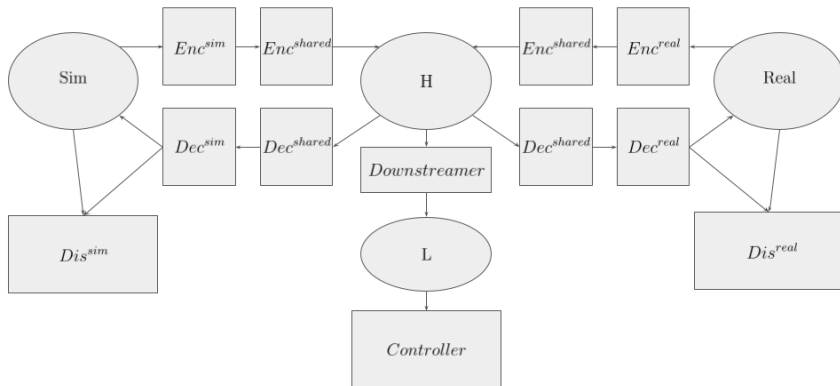
- ПД регулятора в симуляционном дорожном цикле (SIM-LP, 66560 пар) и дороге с перекрестками (SIM-IS, 66560 пар)
- людей-экспертов в реальном мире (REAL-DT, 66515 пар)

Предобработка обрезает верхнюю треть изображения, и сжимает его до 64x32.

Данные собраны в эпизоды: $D = \sqcup_{j=1}^{n_D} e_j$, где $e_j = \sqcup_{i=1}^{n_{e_j}} \{x_i, y_i\}$.

Архитектура метода Sim2Real

Всего в архитектуре 10 модулей, 4 компоненты (генераторы, дискриминаторы, сжимающий модуль и контроллер) и 4 явно выделенных пространства (Sim, Real, H, L).



Функции потерь

- генераторов: $L_{generator} = \alpha_1 L_{ll_dir} + \alpha_2 L_{ll_cyc} + \alpha_3 L_{kl_dir} + \alpha_4 L_{kl_cyc} + \alpha_5 L_{recon} + \alpha_6 L_{gen_gan} + \alpha_7 L_{control}$
- дискриминаторов: $L_{discriminator} = \alpha_6 L_{dis_gan}$
- контроллера: $L_{controller} = \alpha_7 L_{control}$

В качестве L_{ll_*} и L_{recon} взято L1 расстояние, L_{kl_*} есть КЛ-дивергенция с $\mathcal{N}(0, 1)$, L_{*_gan} представлены бинарной кросс-энтропией, $L_{control}$ есть средняя квадратичная ошибка.

Обучение происходит совместно на эпизодах из симулятора и реального мира.

Возможно добавлять дополнительные ограничения.

Ограничение Triplet Loss

Основано на идее из компьютерного зрения³.

Имея кадр x_i^s , его скрытое представление в L x_i^{shl} , представление следующего за ним кадра x_{i+1}^{shl} и представление случайного кадра x_j^{dhl} :



$$TL = \max(d(x_i^{shl}, x_{i+1}^{shl}) - d(x_i^{shl}, x_j^{dhl}) + M_{TL}, 0)$$

где d – функция расстояния, M_{TL} – отступ.

³*Schroff F., Kalenichenko D., Philbin J.* FaceNet: A unified embedding for face recognition and clustering. // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. С. 815—823.

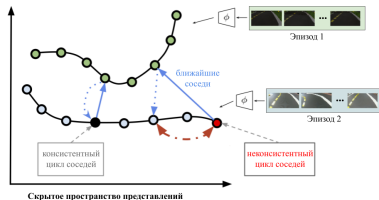
Ограничение Temporal Cycle-Consistency

Мотивировано задачей темпорального выравнивания видео⁴.

Имея представления для двух эпизодов $u_i = x_{a_i}^{dhl}$ и $v_j = x_{b_j}^{dhl}$, я нахожу непрерывного ближайшего соседа u_i :

$$\tilde{v}_i = \sum_j^{n_{cycles}} \alpha_j v_j, \text{ где } \alpha_j = \frac{e^{-||u_i - v_j||_2^2}}{\sum_{k=1}^{n_{cycles}} e^{-||u_i - v_k||_2^2}}$$

$$TCC = - \sum_{k=1}^{n_{cycles}} z_{ik} \log(\tilde{z}_{ik}), \text{ где } z_{ik} = \mathbb{1}[k = i] \text{ и } \tilde{z}_{ik} = \text{softmax}(-||\tilde{v}_i - u_k||_2^2)$$



⁴Temporal Cycle-Consistency Learning. / D. Dwibedi [и др.] // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2019.

Сравнение и анализ результатов

Метрики: среднее абсолютное (CAO) и квадратичное (CKO) отклонения.

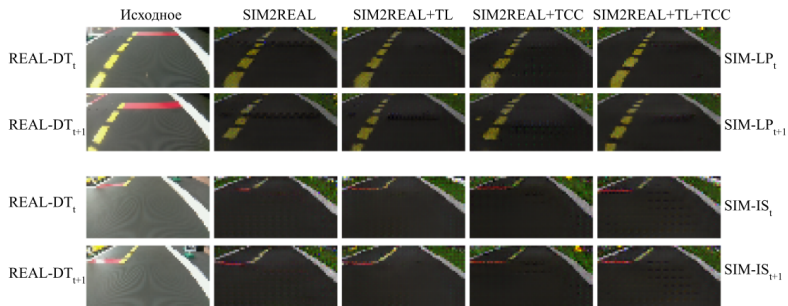
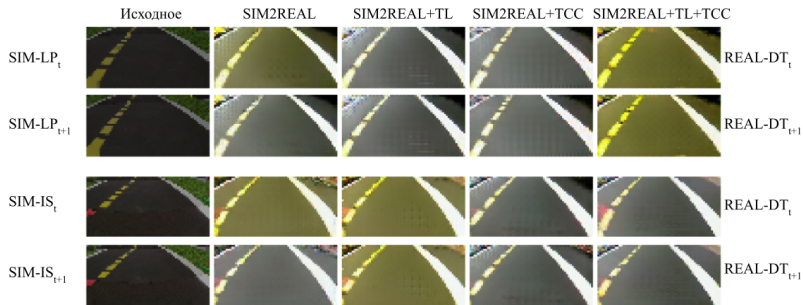
Метод	SIM-LP → REAL-DT		SIM-IS → REAL-DT	
	CAO	CKO	CAO	CKO
SIM	0.1182	0.0248	0.0935	0.0162
SIM2REAL	0.0657	0.0084	0.0624	0.0078
SIM2REAL+TL	0.0643	0.0082	0.0590	0.0070
SIM2REAL+TCC	0.0629	0.0075	0.0600	0.0069
SIM2REAL+TL+TCC	0.0631	0.0078	0.0587	0.0066
REAL	0.0446	0.0043	0.0446	0.0043

Таблица 1: Результаты тестовых метрик в реальном мире

Метод	SIM-LP → REAL-DT		SIM-IS → REAL-DT	
	CAO	CKO	CAO	CKO
SIM	100.00%	100.00%	100.00%	100.00%
SIM2REAL	28.67%	20.00%	36.40%	29.41%
SIM2REAL+TL	26.77%	19.02%	29.45%	22.69%
SIM2REAL+TCC	24.86%	15.61%	31.49%	21.85%
SIM2REAL+TL+TCC	25.14%	17.07%	28.83%	19.33%
REAL	0.00%	0.00%	0.00%	0.00%

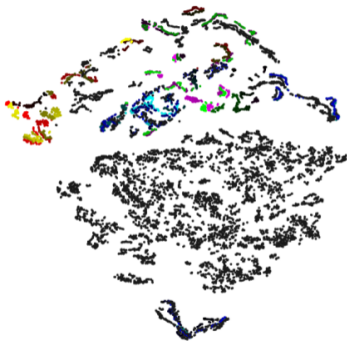
Таблица 2: Относительное доменное несоответствие в реальном мире

Трансляции изображений

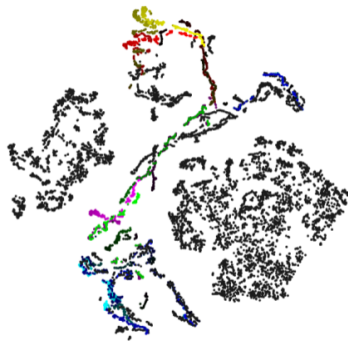


Визуализация представлений

Результаты стохастического вложения соседей с t -распределением для скрытых представлений кадров в контексте трансфера SIM-LP \rightarrow REAL-DT:



SIM2REAL



SIM2REAL+TL+TCC

Часть работы⁵ опубликована на Workshop on Artificial Intelligence for Autonomous Driving, ICML 2020.

Выводы:

- Перенос политики из симулятора в реальный мир требует техник Sim2Real
- Трансляция изображений в качестве Sim2Real метода сокращает доменное несоответствие в среднем до 32.54% для CAO и до 24.71% для СКО
- Использование темпоральных ограничений позволяет дальше сократить доменное несоответствие в среднем до 26.85% для CAO и до 17.47% для СКО

⁵Imitation Learning Approach for AI Driving Olympics Trained on Real-world and Simulation Data Simultaneously. / M. Sazanovich [и др.]. 2020.

Продолжения работы

- Рассмотрение других вариаций симулятора
- Выделение явно обозначенных условий реальности
- Проведение онлайн тестирования методов
- Исследование эффективности и скорости вывода модели для работы на ограниченных вычислительных ресурсах

Метод Drive Straight

Метод	SIM-LP → REAL-DT		SIM-IS → REAL-DT	
	CAO	CKO	CAO	CKO
DRIVE STRAIGHT	0.1256	0.0201	0.1256	0.0201
SIM	0.1182	0.0248	0.0935	0.0162
SIM2REAL	0.0657	0.0084	0.0624	0.0078
SIM2REAL+TL	0.0643	0.0082	0.0590	0.0070
SIM2REAL+TCC	0.0629	0.0075	0.0600	0.0069
SIM2REAL+TL+TCC	0.0631	0.0078	0.0587	0.0066
REAL	0.0446	0.0043	0.0446	0.0043

Таблица 3: Результаты тестовых метрик в реальном мире в сравнении с методом Drive Straight

Трансфер SIM-LP → REAL-DT

Метод	CAO		CKO	
	сред.	откл.	сред.	откл.
SIM	0.1182	0.0190	0.0248	0.0062
SIM2REAL	0.0657	0.0020	0.0084	0.0008
SIM2REAL+TL	0.0643	0.0016	0.0082	0.0003
SIM2REAL+TCC	0.0629	0.0018	0.0075	0.0010
SIM2REAL+TL+TCC	0.0631	0.0059	0.0078	0.0016
REAL	0.0446	0.0025	0.0043	0.0004

Таблица 4: Результаты тестовых метрик в реальном мире для трансфера из SIM-LP в REAL-DT

p-value для t-критерия Стьюдента на равенство средних метрики:

- CAO = 0.024198
- CKO = 0.077358

Трансфер SIM-IS → REAL-DT

Метод	CAO		CKO	
	сред.	откл.	сред.	откл.
SIM	0.0935	0.0083	0.0162	0.0021
SIM2REAL	0.0624	0.0019	0.0078	0.0004
SIM2REAL+TL	0.0590	0.0015	0.0070	0.0005
SIM2REAL+TCC	0.0600	0.0023	0.0069	0.0002
SIM2REAL+TL+TCC	0.0587	0.0009	0.0066	0.0001
REAL	0.0446	0.0025	0.0043	0.0004

Таблица 5: Результаты тестовых метрик в реальном мире для трансфера из SIM-IS в REAL-DT

p-value для t-критерия Стьюдента на равенство средних метрики:

- CAO = 0.002162
- CKO = 0.000093