

Сопоставление изображений при условии значительных смещений и изменений среды

Выступающий: М.Ю.Швец

Руководитель: проф., к.ф.-м.н. В.С.Лемпицкий

Соруководитель: проф., PhD L.Torresani

Московский физико-технический институт
(государственный университет)

Москва, 2017

Задача

Сопоставление пар изображений, отображающих одну трехмерную сцену из различных точек пространства при различных состояниях среды

Актуальность

Локализация автономной платформы, ориентация робота, задачи картографии



Задача

Сопоставление пар изображений, отображающих одну трехмерную сцену из различных точек пространства при различных состояниях среды

Актуальность

Локализация автономной платформы, ориентация робота, задачи картографии



- Се Liu, Jenny Yuen и Antonio Torralba. “Sift flow: Dense correspondence across scenes and its applications”. В: *IEEE transactions on pattern analysis and machine intelligence* 33.5 (2011), с. 978—994
- Alex Kendall, Matthew Grimes и Roberto Cipolla. “Convolutional networks for real-time 6-dof camera relocalization”. В: *Proceedings of the International Conference on Computer Vision (ICCV)*. 2015
- Philipp Fischer и др. “Flownet: Learning optical flow with convolutional networks”. В: *arXiv preprint arXiv:1504.06852* (2015)

Множество трехмерных признаков представлений: \mathcal{F}

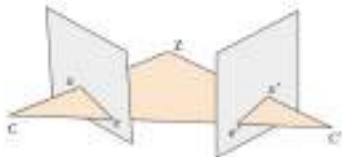
Изображения $X, Y \in \mathcal{X} = \{0, \dots, 255\}^{H \times W \times 3}$

Относительное расположение камеры $P \in \mathbb{R}^d$

Для двух изображений таких, что $\|P\| < \delta$, определены матрица сопоставимости M и матрица смещений $F \in \mathcal{F}$:

$$M_{i,j} = \begin{cases} 1, & \text{если } \exists Z : (i,j) = P_1(Z) \in X_1 \text{ и } P_2(Z) \in X_2; \\ 0, & \text{иначе.} \end{cases}$$

$$\forall (i,j) : M_{i,j} = 1 \rightarrow X_1(i,j) \sim X_2(i + F_{i,j}^1, j + F_{i,j}^0)$$



Обучающая выборка: $({}^\ell X_1, {}^\ell X_2, {}^\ell F, {}^\ell M, {}^\ell P)_{\ell=1}^L$

Предсказание $\tilde{F} = \mu(X_1, X_2)$, где $\mu \in \mathfrak{M}_F$

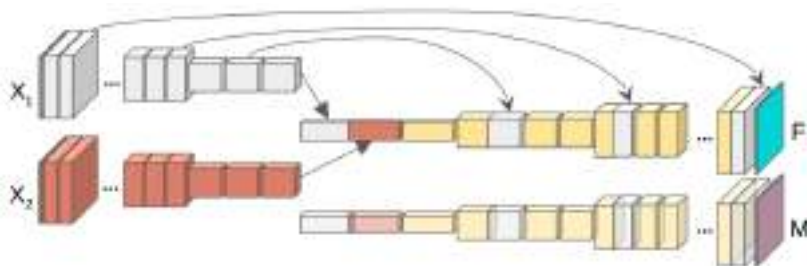
Выбор модели:

$$\hat{\mu} = \arg \min_{\mu \in \mathfrak{M}} \frac{1}{L} \sum_{\ell=1}^L \mathcal{L}_{\text{соотв}}(\mu({}^{\ell}X_1, {}^{\ell}X_2), {}^{\ell}F, {}^{\ell}M)$$

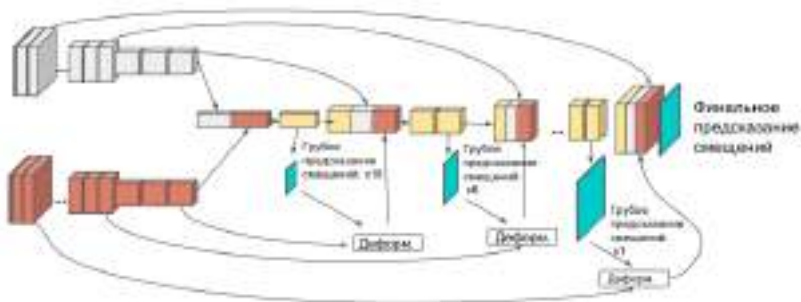
Функция потерь:

$$\mathcal{L}_{\text{соотв}}(\tilde{F}, F, M) = \frac{1}{N_{\text{соотв}}} \sum_{i,j: M_{i,j}=1} \iota(\tilde{F}_{i,j}, F_{i,j})$$

$$\iota(\tilde{F}_{i,j}, F_{i,j}) = \min(\|\tilde{F}_{i,j} - F_{i,j}\|_2^2, T^2)$$



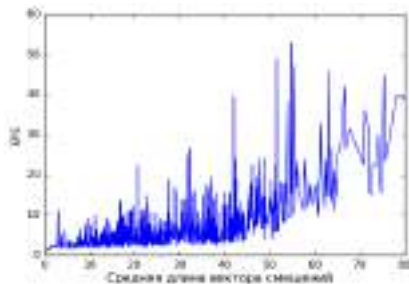
- Две ветви извлечения признаков
- Одновременное предсказание целевой переменной для всех пикселей
- Регуляризация признаков с помощью матрицы соответствий
- Повторное использование представлений с нижних уровней



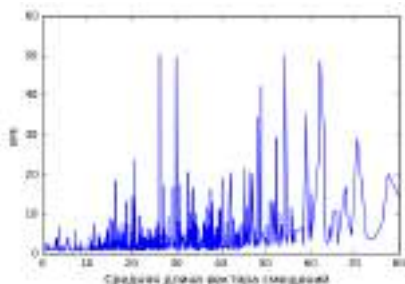
Деформация $W(X, F) = Y$: для $\begin{cases} i = i' + F_{i',j'}^1 \\ j = j' + F_{i',j'}^0 \end{cases}$ и $\begin{cases} \theta_i = i - \lfloor i \rfloor \\ \theta_j = j - \lfloor j \rfloor \end{cases}$

$$Y_{i',j'} = (1-\theta_i)(1-\theta_j)X_{\lfloor i \rfloor, \lfloor j \rfloor} + \theta_i(1-\theta_j)X_{\lceil i \rceil, \lfloor j \rfloor} + (1-\theta_i)\theta_j X_{\lfloor i \rfloor, \lceil j \rceil} + \theta_i\theta_j X_{\lceil i \rceil, \lceil j \rceil}$$

Сравнение базовой и расширенной моделей



Базовая модель



Расширенная модель

Сравнение базовой и расширенной моделей

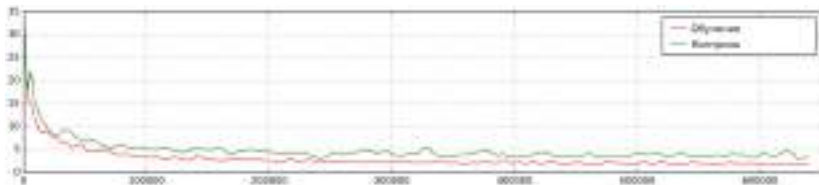


Базовая модель



Расширенная модель

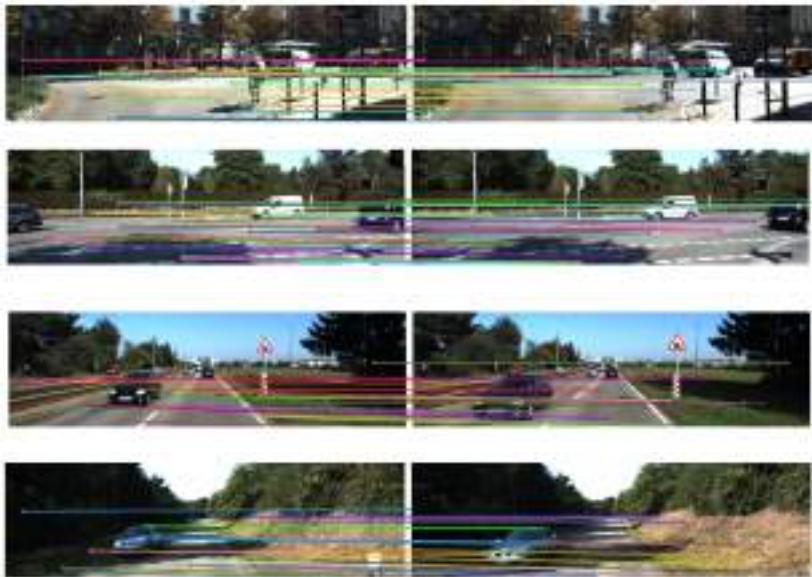
Результаты предсказания матрицы смещений



Зависимость функции потерь расширенной модели от количества итераций на наборе SYNTHIA

	SYNTHIA	NCLT	Chairs	KITTI
Расширенная модель	3.2	14.8	2.1	6.4
Базовая модель	6.0	17.0	3.2	8.0
FlowNetS	5.8	17.1	2.9	-
SIFT Flow	16.2	29.6	4.4	21.4
Предсказание среднего	29.2	44.3	10.7	31.3

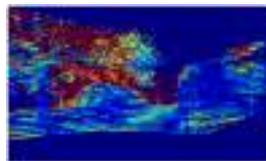
Результаты на наборе данных KITTI



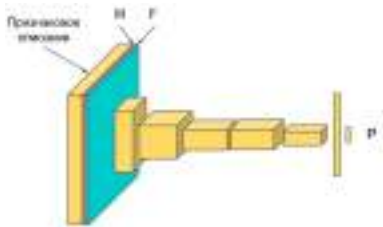
Результаты на наборе данных NCLT



Кодирование матрицы смещений



Оценка положения камеры

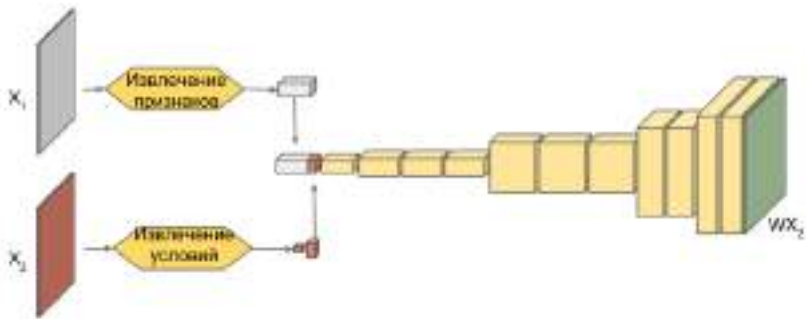


Предсказание $\tilde{P} = \mu(X_1, X_2)$

$$\mathcal{L}_{\text{поз.}}(\tilde{P}, P) = \mathcal{L}_{\text{смещ.}}(\tilde{P}, P) + \lambda \mathcal{L}_{\text{пов.}}(\tilde{P}, P)$$

Оценка положения по матрице смещений F :
извлечение фундаментальной матрицы ($x'^T \mathfrak{F} x = 0$),
из разложения существенной матрицы выделяется
матрица поворота R и смещение u (с точностью до константы)

	SYNTHIA	NCLT
Ошибка полярного угла, °	4.7	10.0
Ошибка угла рыскания, °	-	2.2
Ошибка длины, м	0.11	0.56
Ошибка смещения, м	0.20	0.76
Средняя длина, м	1.75	2.62



Состояние среды – $c : \mathcal{X} \rightarrow \mathbb{R}^d$.

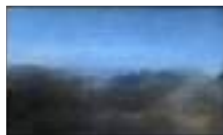
Для $X_1, X_2 \in \mathcal{X}$ требуется определить $r : \mathcal{X} \times \mathbb{R}^d \rightarrow \mathcal{X}$ такую, что

$$r(X_1, c(X_2)) \approx W(X_2, P_1)$$

Результаты на наборе данных SYNTHIA



Результаты на наборе Transient Attributes



- Усовершенствование модели переноса состояний среды путем изменения определения состояний в сторону более богатого семейства
- Разработка методов постобработки на основании полученных представлений
- Введение в модель семантической информации
- Ускорение модели для работы в режиме реального времени
- Составление эталонной выборки для задач межсезонной локализации автономной платформы

- Разработан подход к решению задачи сопоставления пары изображений при значительных смещениях и изменениях среды
- Предложена модель, сочетающая традиционный подход к решению задачи с идеями обучения глубоких нейронных сетей
- Предложен метод моделирования состояний среды, а также метод видоизменения изображения
- Получены результаты на наборах синтетических данных вождения, данных навигации робота, оптического потока
- Предложены методы совмещения моделей для задачи предсказания соответствий и задачи локализации
- Исследована задача классификации с совместной регрессией остатков