

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ "ЛЬВІВСЬКА ПОЛІТЕХНІКА"

ЛЕКЦІЯ 5. Застосування машинного навчання в комп'ютерному зорі

Львів -- 2025

Лекція зі штучного інтелекту 2025-05

Вступ

На цьому занятті ми розглянемо застосування машинного навчання в комп'ютерному зорі. Ми ознайомимось з основними задачами комп'ютерного зору, детально розглянемо класичні та сучасні підходи до задач зіставлення зображень (image matching) та структури з руху (Structure-from-Motion, SfM), включаючи побудову 3D-хмар точок та локалізацію камери.

Теми, що розглядаються

1. Вступ до комп'ютерного зору
2. Основні задачі комп'ютерного зору
3. Класичні методи зіставлення зображень
4. Структура з руху (Structure-from-Motion, SfM)
5. Класичні алгоритми для SfM
6. Глибоке навчання в задачах комп'ютерного зору
7. Сучасні підходи до задачі image matching та SfM
8. Практичні застосування та інструменти

Вступ до комп'ютерного зору

Комп'ютерний зір (Computer Vision) — це галузь штучного інтелекту, яка займається розробкою алгоритмів для автоматичного аналізу, інтерпретації та розуміння візуальної інформації з цифрових зображень або відео.

Основні задачі комп'ютерного зору

- Класифікація зображень
- Виявлення та локалізація об'єктів
- Сегментація зображень
- Відстеження об'єктів
- Зіставлення зображень (Image Matching)
- Відновлення 3D-структури (Structure-from-Motion, SfM)

Зіставлення зображень (Image Matching)

Зіставлення зображень — це задача знаходження відповідних точок або областей між двома або більше зображеннями однієї сцени. Це є ключовим етапом у багатьох задачах, таких як реконструкція 3D-сцен, локалізація камери, панорамне зшивання зображень тощо.

Класичні методи зіставлення зображень

Класичні методи базуються на виявленні та описі ключових точок (features):

- **Виявлення ключових точок:** SIFT, SURF, ORB
- **Опис ключових точок:** створення дескрипторів (SIFT, SURF, ORB)
- **Зіставлення дескрипторів:** алгоритми на основі найближчих сусідів (Nearest Neighbor, NN), FLANN, RANSAC для відкидання помилкових відповідностей

Алгоритм SIFT (Scale-Invariant Feature Transform)

SIFT — один з найпопулярніших класичних алгоритмів, що забезпечує інваріантність до масштабу, обертання та частково до освітлення.

Основні етапи SIFT:

1. Виявлення ключових точок (keypoints) за допомогою різниці гаусіанів (Difference of Gaussians, DoG)
2. Визначення орієнтації ключових точок
3. Побудова дескрипторів ключових точок
4. Зіставлення дескрипторів між зображеннями

Структура з руху (Structure-from-Motion, SfM)

SfM — це методика реконструкції 3D-структури сцени з набору 2D-зображень, отриманих з різних точок зору.

Основні етапи SfM:

1. **Зіставлення зображень:** знаходження відповідних точок між зображеннями
2. **Оцінка фундаментальної матриці:** визначення геометричних зв'язків між парами зображень
3. **Тріангуляція:** обчислення 3D-координат точок з відповідних точок на зображеннях
4. **Оптимізація:** Bundle Adjustment для уточнення параметрів камери та 3D-точок

Алгоритми, що використовуються в SfM:

- **RANSAC** (Random Sample Consensus): для оцінки фундаментальної матриці та відкидання помилкових відповідностей
- **Тріангуляція:** лінійна або нелінійна оптимізація для отримання 3D-координат точок
- **Bundle Adjustment:** глобальна оптимізація параметрів камери та 3D-точок

Сучасні підходи на основі глибокого навчання

З появою глибокого навчання з'явилися нові підходи до задачі зіставлення зображень та SfM.

Глибокі нейронні мережі для зіставлення зображень

- **SuperPoint:** глибока нейронна мережа для виявлення ключових точок та побудови дескрипторів

- **SuperGlue**: графова нейронна мережа для зіставлення дескрипторів, що враховує контекстні зв'язки між точками

Глибокі підходи до SfM

- **DeepSfM**: використання глибоких мереж для оцінки відповідностей та геометричних параметрів
- **NeRF (Neural Radiance Fields)**: нейронні мережі для реконструкції 3D-сцен з набору зображень

Практичні застосування та інструменти

Застосування

- Реконструкція 3D-моделей об'єктів та сцен
- Автономні транспортні засоби (локалізація та картографування)
- Доповнена реальність (AR)
- Робототехніка та навігація

Інструменти

- **OpenCV**: класичні алгоритми комп'ютерного зору
- **COLMAP**: SfM та реконструкція 3D-хмар точок
- **PyTorch3D**: бібліотека для роботи з 3D-даними та нейронними мережами
- **TensorFlow Graphics**: інструменти для 3D-реконструкції та візуалізації

Висновок

Машинне навчання значно розширило можливості комп'ютерного зору, дозволяючи ефективно вирішувати складні задачі, такі як зіставлення зображень та реконструкція 3D-структур. Сучасні підходи поєднують класичні методи з глибокими нейронними мережами, забезпечуючи високу точність та ефективність.