

Сопоставление (совмещение) изображений - важная задача компьютерного зрения, которая заключается в приведении двух или более изображений к общей системе координат.

Большинство методов сопоставления изображений можно разделить на два класса: плотные и разреженные. Плотные алгоритмы сопоставления находят похожие области сопоставляемых изображений, и при нахождении преобразования анализируются все пиксели, входящие в эти области. Разреженные методы находят сопоставление анализируя множество отдельных точек, как правило, “разбросанных” по всему изображению. Метод сопоставления изображений с помощью сиамских нейронных сетей относится к разреженному классу (но об этом позже).

Разреженный подход, при котором принятие решения об искомом преобразовании осуществляется путем анализа и поиска соответствий между двумя наборами небольших участков изображений (“патчей”), состоит из трех основных шагов:

- выделение подмножества устойчивых точек, совокупность которых позволяет идентифицировать и/или сопоставить изображения;
- поиск пар устойчивых точек, соответствующих одной физической точке;
- нахождение геометрического преобразования, переводящего устойчивые точки одного изображения в соответствующие устойчивые точки другого.

Эти три шага показаны на рисунке 2

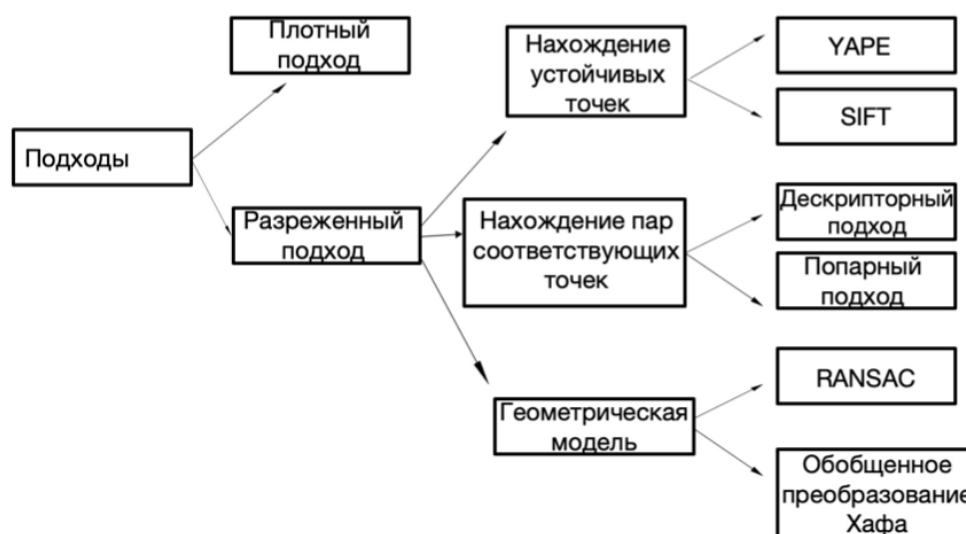


Рис. 1: Алгоритмы, используемые для сопоставления изображений

В качестве детектора особых точек можно использовать следующие методы: SIFT (Scale Invariant Feature Transform), SURF [20] (Speeded Up Robust Features) или YAPE (Yet Another Point Extractor) и так далее.

Для решения задачи сопоставления изображений разреженным методом, после нахождения устойчивых точек необходимо найти соответствие между устойчивыми точками двух изображений (то есть понять, какие пары устойчивых точек соответствуют одной физической точке). Обычно, для этого анализируются некоторые окрестности рассматриваемых точек. Такие окрестности называются “патчами”.

Алгоритмы, которые осуществляют поиск соответствий, можно разделить на два типа:

- “Попарный подход”. В этом случае алгоритм получает на вход два патча - по одному с каждой из картинок. Выходом такого алгоритма является численная оценка схожести двух входных патчей.
- “Дескрипторный подход”. В этом случае алгоритм получает на вход только один патч. Выходом алгоритма является некоторый вектор. Решение о том, какие патчи соответствуют друг другу, принимается путем сравнения дескрипторов некоторой легко вычисляемой метрикой (например, L1 или L2 норма).

В основе многих способов решения данной задачи с помощью нейронных сетей лежит сиамская нейронная сеть. Это такая сеть, которая принимает на вход 2 изображения, и обрабатывает их одинаковым способом (моделью с одинаковыми весами). На выходе получаем вектора, соответствующие входным изображениям. Такая сеть может быть применена как в “Попарном подходе” так и в “Дескрипторном подходе”:

- Вектора на выходе модели можно воспринимать как дескрипторы входных изображений и учить модель так, чтоб евклидово расстояние между дескрипторами для “похожих” изображений было малым, а для “непохожих” - большим (рисунок 1 а).
- Вектора на выходе модели можно подать на вход другой модели, которая будет решать задачу бинарной классификации: 1 - входные изображения, которым в соответствие ставятся дескрипторы, “похожи”, 0 - “непохожи” (рисунк 1 б).

Алгоритмы сопоставления получают на вход два патча, и выдают степень их схожести. Однако, при решении задачи сопоставления обычно число возможных ребер намного превышает число устойчивых точек. При этом дескрибируемые патчи - маленькие куски изображений - как правило имеют большую избыточность, и перед их сравнением, зачастую, эффективно снизить их размерность. Это позволит использовать вычислительно эффективный метод сопоставления патчей в пространстве уменьшенной размерности

Но в попарном подходе на вход второй (классифицирующей) модели можно подать дополнительную информацию о входных изображениях, которая может быть полезна для определения схожести. Например (как в модели hybrid approach на рисунке 1 б), можно на вход модели так же подать значение кросс-корреляции между входными изображениями.

Общий алгоритм сопоставления двух изображений с помощью сиамской нейронной сети:

На вход подаются два изображения I_1, I_2 , тип детектора, тип геометрической модели, семейство геометрических преобразований \mathcal{T} .

1. Подалгоритм детектирования устойчивых точек.

На вход: два изображения I_1, I_2 , тип детектора.

- С помощью заданного детектора находятся множества устойчивых точек F_1 и F_2 . Обозначим $N_1 = |F_1|$ и $N_2 = |F_2|$ - число найденных точек на оптическом и радиолокационном изображении соответственно.
- На изображениях выделяются подизображения (патчи) размера $L \times L$ с центром в найденных точках. Получаем множества патчей P_1 и P_2

На выходе: множества патчей двух изображений P_1 и P_2 .

2. Подалгоритм нахождения соответствующих пар точек.

На вход: множества патчей двух изображений P_1 и P_2 , множества устойчивых точек F_1 и F_2 .

- Из всевозможных пар патчей $Pairs = \{(\rho_1, \rho_2) | \rho_1 \in P_1, \rho_2 \in P_2\}$ оценивается подмножество соответствующих ($Corr_patches$) - такие патчи, центры которых соответствуют одной физической точке.
- Из найденных соответствующих пар патчей $Corr_patches$ и множеств устойчивых точек F_1 и F_2 находятся соответствующие пары устойчивых точек $Corr_pts$ - те точки, которые являются центрами патчей в $Corr_patches$.

На выходе: множество соответствующих пар точек $Corr_patches$.

3. Подалгоритм нахождения геометрического преобразования

На вход: множество соответствующих пар $Corr_patches$, семейство геометрических преобразований \mathcal{T} , тип геометрической модели.

- С помощью геометрической модели находится наиболее подходящее преобразование из указанного семейства преобразований \mathcal{T} .

На выходе: матрица преобразования из первого изображения во второе.

Итого, более точным алгоритмом сопоставления устойчивых точек является попарная нейронная сеть, но она также является самым вычислительно сложным методом из рассмотренных из-за использования нейросетевой метрики сравнения дескрипторов. Более быстрым, но чуть менее качественным является дескрипторная сиамская нейронная сеть - объединение базовых идей "классических" алгоритмов и нейронных сетей.

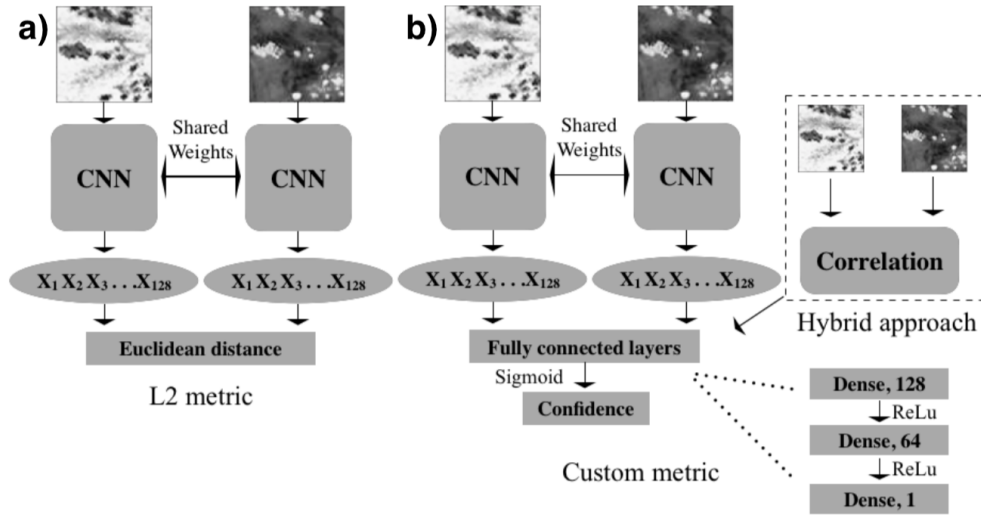


Рис. 2: Архитектуры нейронных сетей для сопоставления изображений. а) дескрипторная сиамская сеть, б) попарная сиамская сеть