

Анализ изображений и видео. Практикум 8.

В качестве отчета по каждому пункту должен быть приведен листинг кода на Python и, если требуется, скриншоты полученных результатов или ответы на вопросы.

Визуальный поиск

В этой лабораторной работе вам предстоит реализовать простейшую систему визуального поиска на основе дескрипторов Bag-Of-Words и геометрической верификации с помощью RANSAC. Работа состоит из нескольких этапов: обучение словаря визуальных слов, вычисление дескрипторов и собственно поиск с использованием геометрической верификации. Поисковая база изображения находится в папке “base”.

Обучение словаря визуальных слов

Извлеките SIFT-дескрипторы с изображений поисковой базы (“cv2.SIFT()”). На полученном множестве запустите кластеризацию методом k-средних (`kmeans = sklearn.cluster.KMeans()`). Для хорошего поиска размер словаря должен быть не меньше 1000. Множество центроидов (`kmeans.cluster_centers_`), полученное в ходе кластеризации, и будет являться словарем визуальных слов. Процедура обучения словаря долгая, поэтому имеет смысл сохранить обученный словарь, чтобы не пересчитывать его при отладке дальнейших экспериментов.

Вычисление дескрипторов Bag-Of-Words

Для каждого изображения из базы извлеките SIFT-дескрипторы. Для каждого дескриптора определите ближайший элемент словаря, то есть визуальное слово, к которому он относится

(можно делать на основе `sklearn.neighbors.NearestNeighbors()`). Посчитайте гистограмму встречаемости различных визуальных слов на изображении - это и есть дескриптор Bag-Of-Words. Также рекомендуется к полученным дескрипторам применить поэлементное извлечение квадратного корня и $L2$ -нормализацию.

Поиск

В качестве запросов нужно использовать изображения из папки “queries”. В качестве результатов поиска возвращаются те изображения, чьи дескрипторы оказались ближайшими к дескриптору запроса. Визуализируйте поисковую выдачу любым образом.

Поиск с RANSAC

(Перед выполнением этой части стоит перечитать раздел про RANSAC из лекции №4.) Примените к top-10 результатам из поисковой выдачи геометрическую валидацию на основе RANSAC. Для этого необходимо с помощью `cv2.BFMatcher(cv2.NORM_L2, crossCheck=True)` (как в 4-ой лабораторной) найти пары соответствующих дескрипторов с картинки-запроса и картинки-результата. Затем пары соответствующих точек необходимо передать в функцию `cv2.findHomography(src_pts, dst_pts, cv2.RANSAC, 10.0)`, которая возвращает пару (`H, mask`). `H` — матрица наилучшего перспективного преобразования, которое может перевести одно изображение в другое, `mask` — бинарный массив, причем `i`-ое значение массива равно единице, если `i`-ое соответствие оказалось инлайером, и нулю — если аутлайером. Соответственно, чем больше в массиве единиц, тем более сопоставимыми являются изображения. Переранжируйте top-10 результатов так, чтобы более сопоставимые по RANSAC изображения возвращались раньше. Визуализируйте новую выдачу. Убедитесь в улучшении результатов.