

Сопоставление разномодальных изображений (оптика и КТ-реконструкция)

В. Р. Коновалов

Московский физико-технический институт

Существует ряд исследований, посвящённых виртуальной развёртке документов на основании данных, полученных с помощью компьютерной томографии. Для анализа работы этих алгоритмов искусственно создаются образцы с характерным содержанием, деформируются и далее осуществляется их виртуальная развёртка. Соответственно встаёт задача сопоставления скана исходного образца и его виртуальной развёртки, в том числе для оценки качества работы алгоритма разворачивания. Данное сопоставление строится по ключевым точкам двух изображений.

Имеются автоматические методы выделения и сопоставления этих точек для изображений из оптического диапазона. Данная работа посвящена оценке качества результатов работы различных алгоритмов на наших образцах и сравнению их между собой.

Помимо разномодальности, специфика работы заключается в том, что искажения, которые претерпевает исходное изображение в нашем случае, являются нелинейными, то есть разные области картинки деформируются по-разному. Данное обстоятельство не позволяет построить единую модель преобразования для всего изображения, что осложняет проверку правильности соответствий.

В качестве решения этой проблемы оценка качества проводилась на изображениях с частичной разметкой, проставленной вручную, приблизительно 200 точек на одной картинке. Также мы предполагаем, что искажения можно считать локально афинными. Это позволяет для произвольной точки исходного изображения по эталонным точкам восстановить её образ на виртуальной развёртке. Делается это с помощью барицентрической системы координат и поиска ближайших эталонных точек. Был написан и протестирован соответствующий код. С его помощью далее оценивалась величина невязки между реальным образом точки и тем, который подобрал алгоритм. Также был написан код для визуализации отклонений и характеристик найденных особых точек.

Для построения соответствия нужны детектор, дескриптор и мэтчер особых точек. В ходе исследования были опробованы различные пары детектор-дескриптор, в то время как мэтчер везде брался один и тот же, стандартный Brute-Force Matcher. Подробнее про них см. в статьях [1] и [2]. При этом эти детекторы и дескрипторы имеют различные гиперпараметры, соответственно был проведён их подбор для максимизации качества. Правильными соответствиями считались те, величина невязки у которых не превышает трёх пикселей, так как это минимальное расстояние между размеченными вручную ключевыми точками. Доли таких соответствий для разных алгоритмов представлены на рис. 1.

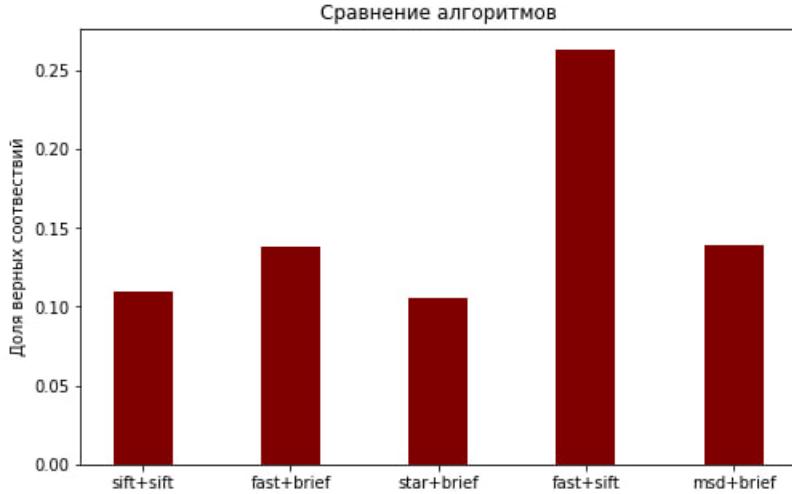


Рис. 1: Сравнение различных пар детекторов и дескрипторов

Выигрыш детектора FAST можно объяснить спецификой наших изображений, они включают в себя по большой части печатный текст и прямые линии, поэтому детектор углов, коим является FAST, справляется лучше остальных. Также необходимо отметить, что точки, выделенные с помощью SIFT и Star плохо подходят для ручной верификации. В этом плане FAST и MSD показали себя сильно лучше, то есть выделенные ими точки легко распознаются на картинке человеком. У FAST есть и недостатки, например отсутствие ориентации у найденных точек, MSD в этом плане более продвинутый, что может иметь значение при вращении изображения. Поэтому в итоге можно утверждать, что из рассмотренных алгоритмов SIFT как детектор работает в нашем случае хуже всего, а наиболее предпочтительными будут пары FAST+SIFT и MSD+BRIEF.

Реализации всех детекторов и дескрипторов взяты из библиотеки OpenCV и весь написанный код предназначен в первую очередь для работы с этой библиотекой. В будущем можно также протестировать алгоритмы не из этой библиотеки, в рамках данной работы на это не хватило времени.

Литература

- [1] Lin Chen, Franz Rottensteiner and Christian Heipke "Feature detection and description for image matching: from hand-crafted design to deep learning"
- [2] David G. Lowe "Distinctive Image Features from Scale-Invariant Keypoints"