

Применение метода межкадровой навигации для детектирования изображения проективно искажённого прямоугольного объекта в видеопотоке на мобильном устройстве

Т.В. Манжиков¹, И.Б. Мамай¹, Д.В. Полевой^{1, 2}

¹Национальный исследовательский технологический университет «МИСиС»

²Институт системного анализа ФИЦ УИ РАН

Проблема поиска границ проективно искажённого прямоугольного объекта в видеопотоке часто возникает при практическом решении прикладных задач оптического распознавания с помощью малоформатных цифровых камер мобильных устройств. Документы, вывески, этикетки и другие объекты часто имеют прямоугольную форму, но добиться положения камеры, при котором не будет проективных искажений практически невозможно. В данной работе решалась задача нахождения рамки телевизора на кадрах видеопотока, полученного с камеры мобильного устройства.

На каждом кадре видеопоследовательности требуется найти такой четырёхугольник, внутри которого заключён экран телевизора, а снаружи расположена рамка телевизора. Для решения используется межкадровая система координат, оценка геометрических характеристик рамки и предположение о монотонности цветовых характеристик рамки.

На первом шаге применяется существующий алгоритм детекции проективно искажённых прямоугольников в кадре [1] и для каждого изображения формируется набор гипотез четырёхугольников. Для каждой гипотезы вычисляется оценка качества на основе геометрических характеристик рамки и цветовых характеристик. Геометрическая оценка опирается на следующее утверждение: для произвольного строго выпуклого четырёхугольника Q , расположенного на кадре, существует единственный с точностью до гомотетии параллелограмм $P(Q)$, проективный образ которого совпадает с данным четырёхугольником Q [2], то есть можно оценить, насколько найденный четырёхугольник может быть изображением прямоугольника. Оценка цветовых характеристик изображения в некоторой окрестности этой гипотезы использует эвристику о том, что рамка телевизора имеет монотонный цвет. В целом, освещённость сцены может влиять на цвет рамки в различных её участках, но локальные изменения цвета невелики. Поэтому, если взять небольшую окрестность за пределами рамки и сравнить её с некоторой окрестностью в другой области рамки, то можно говорить и том, насколько наша гипотеза похожа на внутреннюю рамку телевизора

Затем с помощью детектора особых точек SURF [3] и метода оценки параметров модели RANSAC [4] строится межкадровая система координат, которая для каждой пары последовательных кадров определяет проективное отображение, переводящее один кадр в другой наилучшим образом. Это позволяет установить соответствие между теми гипотезами на соседних кадрах, которые являются изображениями одного и того же объекта. А также построить граф соответствия, вершинами которого являются гипотезы. Наилучший путь в таком графе состоит из гипотез, соответствующих, предположительно, одному и тому же объекту реального мира.

В рамках выполнения данной работы был разработан подход к решению задачи поиска рамки телевизора в видеопотоке. На основе данного подхода был реализован алгоритм на языке C++. На данном этапе корректность работы алгоритма проверена органолептически, на небольшом наборе тестовых данных. В дальнейшем планируется создать репрезентативную тестовую базу, методику и инструментарий оценки качества работы алгоритма для более точной проверки и тонкой настройки параметров.

Литература

- [1] *Skoryukina N., Nikolaev D., Sheshkus A., Polevoy D.* Real time rectangular document detection on mobile devices. – Seventh International Conference on Machine Vision (ICMV 2014), 2015, 94452A, pp. 1-5.
- [2] *Каченовский М.И., Колягин Ю.М., Кутасов А.Д.* Учебник для сред. спец. учеб. Заведений. – Москва: Наука, 1982. – 319 с.
- [3] *Bay H., Tuytelaars T., Van Gool L.* SURF: speeded up robust features. – ECCV, 2006.
- [4] *Li X., Liu Y., Wang Y., Yan D.* Computing homography with RANSAC algorithm: a novel method of registration. – Proc SPIE, 2005, 5637, pp. 109 –112