

# ИСПОЛЬЗОВАНИЕ ДЕТЕКТОРОВ И ДЕСКРИПТОРОВ ЛОКАЛЬНЫХ ПРИЗНАКОВ ИЗОБРАЖЕНИЙ В ЗАДАЧЕ ТРЕХМЕРНОЙ РЕКОНСТРУКЦИИ ОБЪЕКТОВ ПО ВИДЕОДАНЫМ

Головатая Е. А., Садов В. С.

Кафедра интеллектуальных систем, факультет радиофизики и компьютерных технологий, Белорусский  
государственный университет  
Минск, Республика Беларусь

E-mail: katerina-golovataya@gmail.com, sadov@bsu.by

*В работе рассматриваются алгоритмы извлечения и сравнения локальных признаков цифровых изображений. Исследуются методы поиска ключевых точек, а также вычисления векторов признаков изображений по ключевым точкам для их дальнейшего сопоставления с точками, найденными на других изображениях набора. Полученные пары ключевых точек описывают изменения положений объектов на разных изображениях и в дальнейшем используются для оценки положений видеокамер в трехмерном пространстве и пространственной поточечной трехмерной реконструкции.*

## ВВЕДЕНИЕ

Задача трехмерной реконструкции является одним из важнейших направлений исследований в области обработки визуальной информации и компьютерного зрения. В общем виде задачу можно сформулировать следующим образом: имеется реальный объект в пространстве и набор данных визуальных наблюдений этого объекта, обычно в виде двумерных изображений. Задача состоит в том, чтобы по этим данным восстановить как можно более точную трехмерную модель объекта. В рамках данной работы рассматриваются изображения видеоряда, полученного в ходе видеоэндоскопического обследования.

Одними из наиболее распространённых методов для проведения трехмерной реконструкции являются методы оценки структуры по сдвигу (structure from motion) [1]. Основная идея этих методов заключается в определении сдвига объектов между парами изображений; информация о глубине может быть оценена исходя из эффекта параллакса: при смещении камеры более близкие объекты получают больший сдвиг, а более дальние – меньший.

Основной задачей алгоритмов оценки структуры по сдвигу является поиск точек, принадлежащих одним и тем же частям объекта на разных изображениях. При этом можно выделить 2 этапа – непосредственно нахождение набора точек, достаточно полно описывающих ключевые характеристики и признаки объекта, и сопоставление найденных точек на двух изображениях для определения того, как изменилось положение объекта [2–3].

### I. АЛГОРИТМЫ ОБНАРУЖЕНИЯ КЛЮЧЕВЫХ ТОЧЕК

Ключевая точка – это такая окрестность на изображении, в которой присутствует некоторый

характерный признак. К основным характерным признакам можно отнести следующие [1, 3]:

- Границы и стыки – ключевые точки в местах перепада яркости на изображении. Такие точки можно найти исходя из того, что в них наблюдается большое значение модуля градиента яркости. Граница может быть произвольной формы и содержать разветвления. Основными алгоритмами обнаружения границ являются методы Собела и Кэнни.
- Углы и точечные структуры – ключевые точки в местах, где на изображении наблюдается признак, похожий по структуре на точку, а также в местах резкой смены направления границы. В таких точках, как правило, наблюдаются высокие значения кривизны градиента яркости. Наиболее популярными алгоритмами является детектор углов Ши-Томаси и алгоритм поиска ключевых точек FAST, а также дифференциальные методы - лапласиан гауссиана, разница гауссианов, и определитель матрицы Гессе.
- Совокупности точек и замкнутые регионы – в отличие от предыдущих признаков, совокупности точек и регионы описывают некую область на изображении. Алгоритмы поиска ключевых точек, как правило, сводят описание всей области к описанию одной точки этой области – например, её центра. Для поиска таких областей используются специализированные алгоритмы поиска пятен (blob detector), которые способны обнаружить признаки на сглаженных изображениях, для которых не сработают детекторы границ и углов. Для поиска этих признаков используются специализированные алгоритмы, которые условно можно разделить на методы оценки интенсивно-

сти (MSER) и методы оценки структуры (PCBR).

Некоторые алгоритмы обнаружения ключевых точек могут также вводить свои виды характерных признаков.

Основное требование к алгоритму поиска ключевых точек – воспроизводимость. Воспроизводимость означает, что алгоритм будет обнаруживать одни и те же точки объектов на различных изображениях инвариантно относительно их положения и различных видов аффинных преобразований – наклона, поворота, масштабирования и их совокупности.

Как правило, поиск ключевых точек осуществляется с использованием нескольких различных алгоритмов, результат работы которых комбинируется.

## II. АЛГОРИТМЫ ОПИСАНИЯ ЛОКАЛЬНЫХ ПРИЗНАКОВ

После нахождения ключевых точек необходимо осуществить их сопоставление. Для этого можно сравнивать совокупность локальных признаков в некоторой их окрестности. Описание локальных признаков по точкам осуществляется при помощи специальных алгоритмов, называемых дескрипторами. Дескрипторы позволяют вычислять для произвольной точки вектор признаков  $\vec{d}$  в некотором пространстве  $U_D$ , для которого определена метрика  $s(\vec{d}_i, \vec{d}_j)$ . Метрика  $s$  в этом пространстве служит мерой «похожести» окрестности пары точек [4].

Таким образом, алгоритм поиска соответствующих точек на двух изображениях можно описать в виде следующих шагов:

1. Определить наборы ключевых точек  $P_1 = \{p_{1i}\}$  и  $P_2 = \{p_{2j}\}$  для первого и второго изображения соответственно;
2. При помощи дескриптора  $D(p)$  с метрикой  $s(\vec{v}_1, \vec{v}_2)$  рассчитать для каждой ключевой точки  $p_{1i}$  и  $p_{2j}$  значения их векторов признаков  $\vec{d}_{1i} = D(p_{1i})$ ,  $\vec{d}_{2j} = D(p_{2j})$ ;
3. Для каждой точки  $p_{1i}$  исходного изображения найти номер точки на втором изображении  $m_i$ , такой, что  $s(\vec{d}_{1i}, \vec{d}_{2m_i}) = \min_j s(\vec{d}_{1i}, \vec{d}_{2j})$ ;
4. Из множества пар точек  $\{(p_{1i}, p_{2m_i})\}$  выбрать подмножество  $\{F_k\} = \{(p_{1k}, p_{2m_k}) | s(\vec{d}_{1k}, \vec{d}_{2m_k}) < T\}$ , где  $T$  - некое пороговое значение, которое можно определять экспериментально;

Различают 2 основных вида дескрипторов – гистограммные и бинарные.

К наиболее распространённым гистограммным детекторам относятся SIFT (Scale-Invariant Feature Transform), SURF (Speeded Up Robust Features) и GLOH (Gradient Location and Orientation Histogram).

Гистограммные дескрипторы статистически оценивают окрестности ключевых точек с ис-

пользованием различных видов преобразований и строят на их основании набор градиентных гистограмм. Метрика  $s$  для таких дескрипторов, как правило, вычисляется достаточно сложно и включает в себя нелинейные преобразования исходных дескрипторов [4-5].

Бинарные дескрипторы подразумевают вычисление некоторой битовой строки по окрестности точки, обычно только на основании яркости окружающих пикселей. К таким дескрипторам относятся BRIEF (Binary Robust Independent Features), ORB (Oriented FAST and rotated BRIEF), BRISK (Binary Robust Invariant Scalable Keypoints), FREAK (Fast Retina Keypoint) и другие.

Значение бинарного дескриптора (битовая строка) вычисляется в соответствии с определенной последовательностью сэмплирования. Основное преимущество бинарных дескрипторов заключается в намного более высокой производительности – в качестве метрики  $s$  сравнения значений дескрипторов используется расстояние Хэмминга (количество несовпадающих бит), вычисление которого значительно проще, чем гистограммный анализ [5].

Анализ алгоритмов обнаружения ключевых точек применительно к данным видеоэндоскопических исследований показал несостоятельность большинства методов вследствие таких особенностей данного типа изображений, как контрастность, недостаточная освещенность и отсутствие резких границ. После проведенной предварительной обработки наилучшие результаты показали алгоритмы, основанные на методах оценки интенсивности. Вследствие однородности самих видеоэндоскопических изображений и однотипности объектов на них рассматриваемые дескрипторы вызвали множество ошибок первого и второго рода. Наилучшие результаты по точности определения соответствующих точек показали гистограммные дескрипторы, но они имеют значительный проигрыш по производительности и времени работы относительно бинарных дескрипторов.

1. Do We Need Binary Features for 3D Reconstruction? / B. Fan [et al.] // Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE Conference – 2016.
2. Lindeberg, T. Scale-space / T. Lindeberg // Encyclopedia of Computer Science and Engineering (Benjamin Wah, ed), John Wiley and Sons, 2008–2009 – №4 – P. 2495–2504.
3. Mikolajczyk, K. Scale and affine invariant interest point detectors. / K. Mikolajczyk, C. Schmid // International Journal of Computer Vision, 2004 – P. 63–86.
4. Mikolajczyk, K. A performance evaluation of local descriptors. / K. Mikolajczyk, C. Schmid // Pattern Analysis and Machine Intelligence, 2005 – P. 1615–1630.
5. Rublee, E. ORB: an efficient alternative to SIFT or SURF / E. Rublee, V. Rabaud, K. Konolige, G. Bradski // IEEE International Conference on Computer Vision (ICCV). – 2011.