

ПОСТРОЕНИЕ КАРТЫ ГЛУБИНЫ ПО НЕРЕКТИФИЦИРОВАННЫМ ИЗОБРАЖЕНИЯМ

О. Ф. Ковалёв, М. М Важник

Белорусский государственный университет, Минск, Беларусь

E-mail: oleg-kovalev-2021@mail.ru

Метод Speeded Up Robust Features (SURF) применён к стереопарам с целью нахождению на изображениях точек, находящихся на эпиполярных линиях. Это позволит выровнять изображения стереопары так, чтобы эпиполярные линии на них были параллельны горизонтали, что в свою очередь позволит значительно ускорить построение карты глубины при помощи корреляционного анализа стереоизображений, так как идентичные фрагменты можно искать лишь на эпиполярных линиях снимков.

Ключевые слова: карта глубины; стереопара; корреляционный анализ; SURF; эпиполярная линия.

При растущих требованиях к автономности техники всё большее значение приобретает её способность ориентироваться в пространстве. Наиболее точно определить расстояния до окружающих объектов можно при помощи лазерных датчиков. Однако ввиду когерентности лазерного излучения, сканирование окружающего пространство требует значительных временных затрат. Помимо этого, сложные измерительные системы требуют и значительных финансовых затрат.

Дешёвой альтернативой является использование фототехники. В этом случае решается фундаментальная задача стереозрения: преобразование плоских изображений в трехмерное с восстановлением при этом информации об удалённости запечатлённых на фото областей от точки съёмки. Для этого используют две неподвижные камеры или одну, но с возможностью перемещения, для получения стереопары, - пары фотографий, изображающих некую область пространства с разных ракурсов (рис. 1).

Фундаментальная задача стереозрения, таким образом, сводится к установлению соответствия между левым и правым изображением стереопары путём вычисления диспаратета, - расстояния между одной и той же областью на левом и правом снимках (рис. 1). Существующие методы по вычислению карты диспаратета основаны на принципе «скользящего окна» [1] (рис. 2). Значения расстояния от объектива камеры до области пространства обратно пропорционально величине смещения [2] (обозначения с рис. 1):

$$(T - d)/(Z - f) = T/Z \rightarrow Z = fT/d . \quad (1)$$

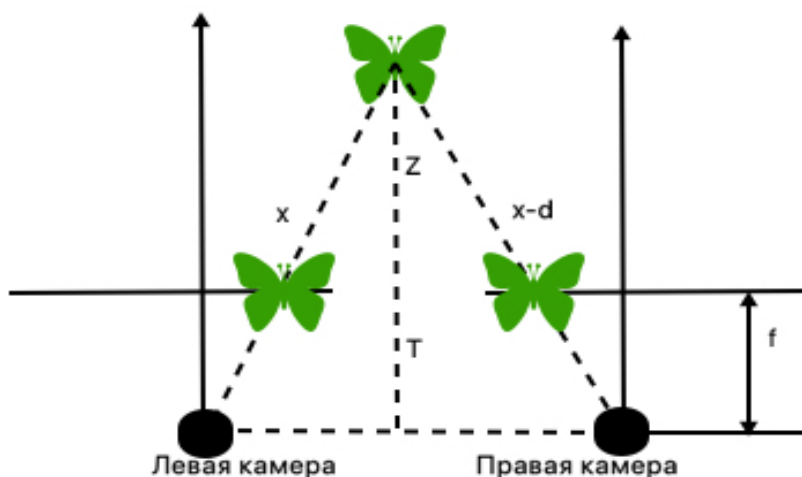


Рис. 1. Съёмка сцены двумя камерами: f - фокусное расстояние; T - расстояние между камерами; d - смещение объекта на изображении, полученном правой камерой относительно положения объекта на картинке левой камеры (x); Z - расстояние от линии камер до объекта

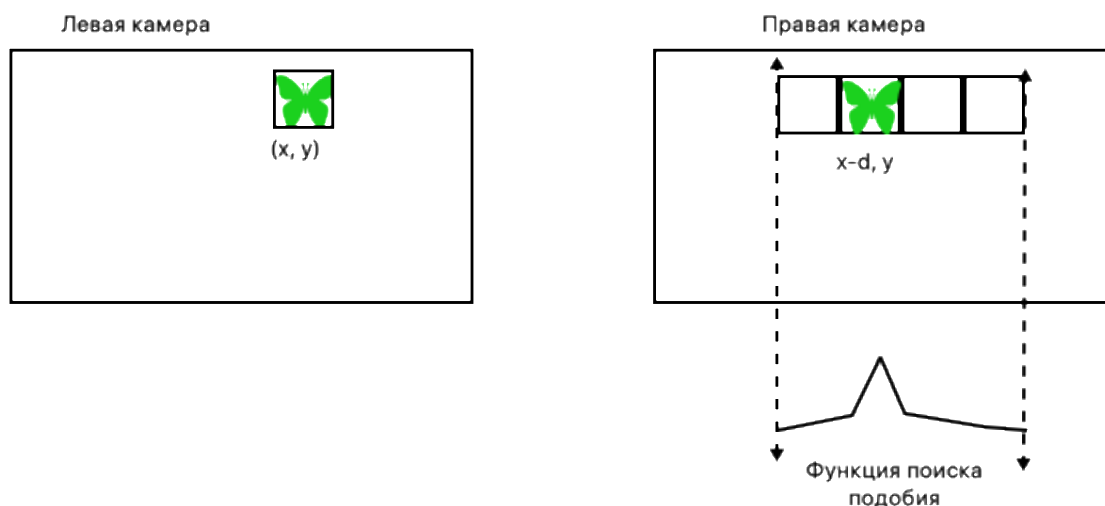


Рис. 2. Выбор окном сканирования области на левом снимке и поиск аналогичной области на правом снимке путём определения максимума функции поиска подобия

Так при наличии камер S и S' (рис. 3), некая точка пространства X проецируется в x на плоскость изображения левой камеры и в x' на плоскость изображения правой камеры. Прообразом точки x на изображении левой камеры является луч xX . Этот луч проецируется на плоскость второй камеры в прямую l' , называемую эпиполярной линией. Образ точки X на плоскости изображения второй камеры обязательно лежит на эпиполярной линии l' . Из этого выходит, что искать выбранную окном сканирования область на левом снимке стереопары можно искать на эпиполярной линии правого снимка, а не на всей фотографии.

Однако в случае использования, к примеру, мобильного телефона, при помощи которого делают снимки некой сцены с разных ракурсов, добиться получения идеально выровненных (ректифицированных) сте-

реопар невозможно. В этом случае требуется проведение ректификации - перенос двух плоскостей изображений в одно таким образом, чтобы все эпиполярные линии были бы параллельны оси абсцисс, а их точки имели бы одинаковые ординаты. В эксперименте использовался метод SURF (Speeded Up Robust Features), применяемый для определения местоположения и распознавания объектов, реконструкции трехмерных сцен и отслеживания особых точек [3].

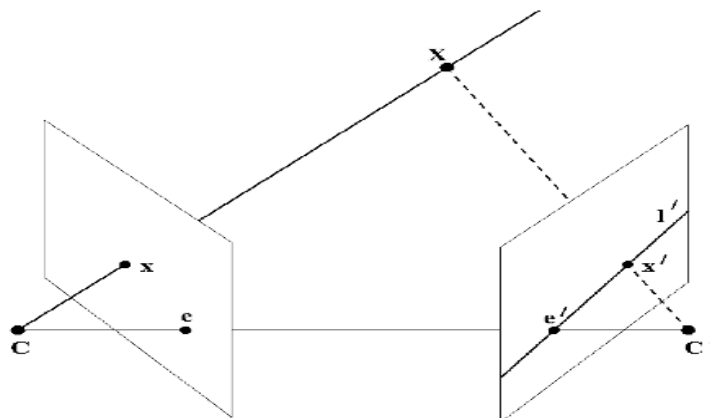


Рис. 3. Эпиполярная геометрия

Эти особые точки на изображениях стереопар инвариантны к масштабированию и вращению, а обнаруживаются в местах резкого перепада градиента. При помощи SURF можно найти эти точки, а затем определить дескрипторы - вектора, описывающие локальную окрестность вокруг этих точек. Это позволяет сопоставить эти точки на разных снимках, где их координаты различны.

Особые точки ищутся при помощи матрицы Гессе, определитель которой в этих точках достигает экстремума:

$$H(f(x, y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}, \det(H) = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2, \quad (2)$$

где H — матрица Гессе, $f(x, y)$ — функция изменения градиента яркости.

После нахождения ключевых точек, метод SURF формирует их дескрипторы, представляющие собой набор из 64 либо 128 чисел для каждой ключевой точки. Эти числа отображают флуктуации градиента вокруг ключевой точки. Поскольку ключевая точка представляет собой максимум гессиана, этим гарантируется, что в окрестности точки будут участки с разными градиентами, чем обеспечивается дисперсия дескрипторов для разных ключевых точек и достигается инвариантность дескриптора относительно поворота. Каждый дескриптор левого изоб-

ражения сравнивается со всеми дескрипторами правого, и так находятся соответствующие пары. При том, что каждая точка на левом изображении стереопары определяет эпиполярную линию на другом изображении (рис. 3), зная координаты соответствующих точек, можно связать их при помощи фундаментальной матрицы, представляющей собой линейный оператор, сопоставляющий каждой точке соответствующую ей эпиполярную линию. Так существует такая матрица F размера 3×3 , что пара точек x и x' являются соответствующими точками на снимках стереопары тогда и только тогда, когда:

$$x'^T \cdot F \cdot x = 0. \quad (3)$$

С помощью фундаментальной матрицы вычисляются уравнения эпиполярных линий. Так для точек x и x' уравнения эпиполярных линий будут иметь вид:

$$l' = F \cdot x, \quad l = F^T \cdot x'. \quad (4)$$

Ректификация изображения далее будет представлять процесс преобразования такой, что эпиполярные линии обоих снимков станут параллельны оси абсцисс.

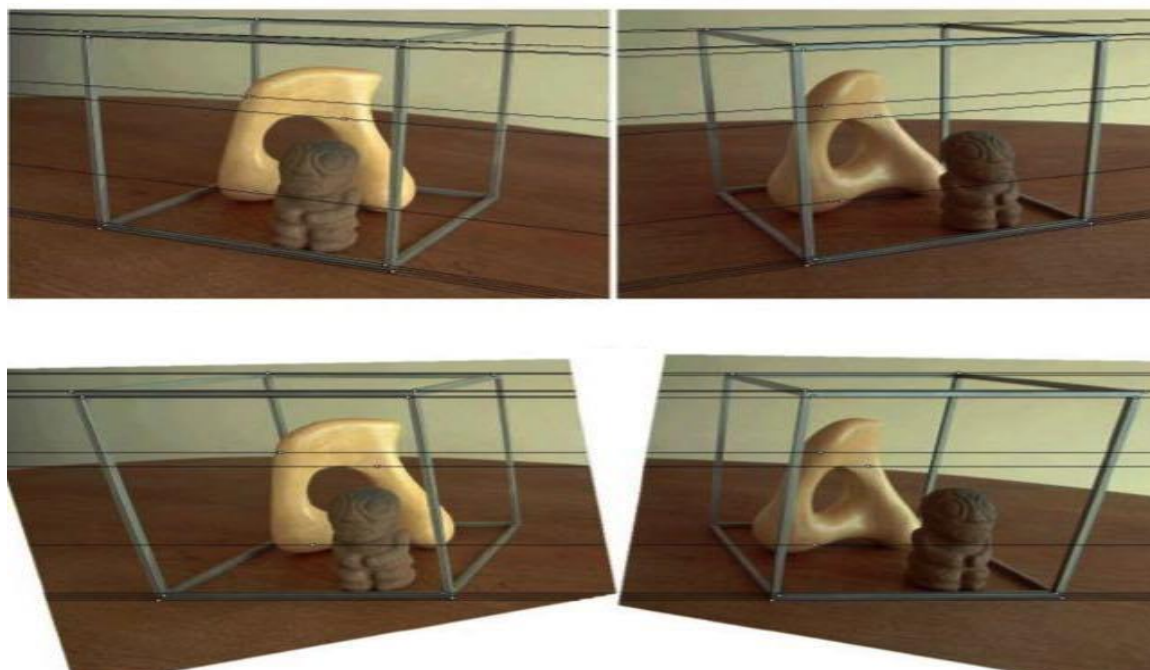


Рис. 4. Стереопара (вверху) и её ректификация (внизу)

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Kanade T., Okutomi M. A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment // Proceedings of the 1991 IEEE International Conference on Robotics and Automation (ICRA '91). 1991. P. 1088–1095. DOI: [10.1109/ROBOT.1991.131738](https://doi.org/10.1109/ROBOT.1991.131738)

2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. Москва: Техносфера. 2005.
3. Robust interest point detector and descriptor [Electronic resource]: pat. US 2009238460. Funayama R., Yanagihara H., Gool L., Tuytelaars T., Bay H. // Publ. date 24.09.2009. URL: <https://patents.google.com/patent/US8165401B2/en>

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ЭМОЦИЙ В РЕЧЕВОМ СИГНАЛЕ

К. Д. Политай, Е. И. Козлова

*Белорусский государственный университет, Минск, Беларусь
E-mail: earnur97@gmail.com*

Описана связь эмоционального состояния человека и параметров речевого сигнала, а также предложена модель автоматического распознавания эмоций в речи. Реализованы системы, использующие в качестве входных данных мел-частотные кепстральные коэффициенты и различные алгоритмы классификации. Проанализированы результаты распознавания эмоций для контрольных данных и проведен сравнительный анализ эффективности и точности различных подходов.

Ключевые слова: обработка речи; распознавание эмоций; кепстральный анализ; машинное обучение.

Распознавание эмоций в речевом сигнале является относительно новой областью исследований. Однако у систем автоматического распознавания эмоций из речи существует большое количество возможных применений, например, определение стрессоустойчивости сотрудников, выявление проблем, связанных с эмоциональным состоянием. С развитием технологий машинного обучения системы искусственного интеллекта обучаются интерпретировать эмоциональное состояние людей и адаптировать свою работу, соответствующим образом реагируя на проявление человеческих эмоций [1].

Каждое эмоциональное состояние имеет уникальные свойства, позволяющие его распознать [1]. Из-за биологических изменений, вызванных эмоциями, свойства акустического сигнала, полученного при произнесении одной и той же фразы, изменяются, что позволяет распознать эмоции при помощи автоматических методов.

Автоматическое распознавание часто используется для определения эмоций из определенного заданного набора их классов. Обработка речи в этом случае происходит в три основных этапа, включающих в себя предварительную обработку, выделение признаков и классификацию на основе полученных признаков [2]. Обработка в системах распознавания применяется для решения таких задач, как фильтрация речи и шумопо-