

Определение 3D координат для облаков точек с помощью стереомэтчинга и суперпиксель сегментации¹

М.Г. Мозеров*, В.И. Кобер*,***, В.Н. Карнаухов*, Л.В. Зимина**

* Институт проблем передачи информации, Российская академия наук, Москва, 127051, Россия,

** Московский политехнический университет, Москва, 107023, Россия,

*** Центр научных исследований и высшего образования, Энсенада, 22860, Мексика

Поступила в редколлегию 21.03.2025 г. Принята 21.04.2025 г.

Аннотация—Важным элементом движения роботизированных систем является определение позиции объекта в координатах, связанных со сценой, а также вектор собственного движения автономного устройства в этой системе координат (матрица поворота и вектор смещения). Все эти параметры можно определить, если оценить значения 3D координат облака точек сцены, связанных с координатной системой бортовой видео камеры автономного роботизированного устройства. Процесс нахождения матрицы поворота и вектора смещения называется регистрацией трехмерных облаков точек и имеет большое значение в робототехнике и компьютерном зрении. Поэтому существует множество методов такой регистрации. Однако проблема заключается в том, как оценить значения 3D координат облака точек сцены. В этой работе мы предлагаем определять искомые координаты точек с помощью стереомэтчинга. Основное предположение классического стереомэтчинга заключается в утверждении, что совпадающие пиксели на стереоизображениях имеют одинаковые значения яркости. Однако это предположение в целом неверное, если учесть наличие шума на изображениях и различную освещенность левого и правого изображения в стереопаре. Кроме того, попиксельное сопоставление не работает в областях, где нет реальной текстуры. Методы, которые решают все проблемы плотного стереомэтчинга, достаточно сложные и требуют больших вычислительных затрат. Однако проблему можно решить более простым и менее затратным способом, учитывая, что облака, предназначенные для регистрации, состоят из небольшого количества точек, значительно меньшего, чем пикселей на изображении. Поэтому стереомэтчинг в нашем случае осуществляется, основываясь на некоторой окрестности сопоставляемых пикселей, называемых суперпикселем. Точность подобного мэтчинга превосходит использование близкого по принципу робастного сенсуса-признака. Кроме того, мы предложили оригинальный метод верификации надежности определения диспаратности, основанный на локальной автокорреляции, таким образом удаляя точки облака с ненадежной оценкой диспаратности. В результате наша оценка 3D координат "надежных" точек сцены сравнима с лучшими state-of-the-art алгоритмами глубокого обучения.

КЛЮЧЕВЫЕ СЛОВА: компьютерное зрение, облако точек, стереомэтчинг, суперпиксель-сегментация.

DOI: 10.53921/18195822_2025_25_1_46

1. ВВЕДЕНИЕ

Стереомэтчинг используется во многих приложениях и остается одной из самых сложных открытых проблем в компьютерном зрении [1, 2, 3]. Методы стерео играют ключевую роль в

¹ Работа выполнена при поддержке государственного задания ИППИ РАН: FFNU-2025-0031, утвержденного Минобрнауки России.

дополненной реальности, в управлении транспортными средствами, 3D-реконструкции и обработке медицинских изображений. Тем не менее, стерео согласование в практических приложениях часто зависит от многих факторов, таких как шум изображения, низкая текстура, окклюзия, различное освещение и экспозиция. Поэтому возникает необходимость разработки робастных локальных признаков.

Самый распространённый способ решения проблемы стереосогласования — минимизация энергии. Решение должно сходиться к максимальной апостериорной вероятности (МАН) или к минимуму энергии функционала условного случайного поля, определённого над графом изображения, где пиксели или участки изображения (суперпиксели) являются вершинами, а попарные отношения между пикселями обычно кодируются рёбрами графа [4]. Локально связанную модель рассматривали в [5, 6, 7, 8]. Позднее стали применять глобальную модель связанности [9]. В последнее время прогресс в задаче плотного стереомэтчинга (оценка диспаратности осуществляется в каждой точке-пикселе изображения) достиг небывалых результатов с применением нейросетевых технологий и глубокого обучения [10]. Однако такая технология требует специального оборудования и серьёзных вычислительных возможностей. Кроме того, в нашей задаче не требуется плотного стереомэтчинга, а предложенная в этой статье схема вычислений достигает тех же результатов, что и state-of-the-art методы глубокого обучения.

Модель минимизации энергии требует вычисления функции стоимости сопоставления или функции рассогласования. Формирование функции рассогласования — важный компонент любого алгоритма стереомэтчинга. Более того, функция рассогласования является основным элементом любого мэтчинга, включая задачу определения оптического потока. Однако в случае, когда необходимо определить параметры движения роботизированной системы с помощью регистрации облаков точек [11, 12], задача облегчается за счет того, что необходимо найти координаты только небольшого количества точек 3D сцены. В этом случае действует допущение, что у каждой такой точки можно найти окрестность, внутри которой функция глубины непрерывная. А значит, все точки окрестности на левом изображении должны совпадать по яркости с точками подобной окрестностью на правом изображении. Разрыв в функции глубины сцены обычно совпадает на изображении с границами яркости. Поэтому принципиально получить предварительно разбиение одного из стереоизображений на небольшие области, границы которых совпадают с границами яркости на этом изображении. Такое деление обычно называют суперпиксель сегментацией.

В последнее время суперпиксель-сегментация привлекла большой интерес, особенно в области компьютерного зрения, поскольку она обеспечивает удобный способ вычисления характеристик изображения и снижает сложность последующих задач обработки изображений. Для решения задачи избыточной сегментации было предложено множество алгоритмов [13, 14, 15, 16]. Для нашей задачи определения 3D координат для облаков точек такая сегментация предоставляет широкие возможности. Поэтому мы выбрали алгоритм [16] чтобы создать N окрестностей для N точек. Если проинтегрировать по таким окрестностям веса поточечной функции рассогласования, то полученная интегральная функция рассогласования, отнесенная к центру окрестности, будет представлять собой робастную функцию рассогласования, и координаты точек, вычисленные с помощью такого алгоритма, будут достаточно точными для целей одометрии. Экспериментальные результаты, основанные на реальных стереоизображениях, показали, что идея использовать суперпиксель-сегментацию в стереомэтчинге оказалась весьма плодотворной.

Статья организована следующим образом: в разделе 2 разъясняются принципы суперпиксель-сегментации для стерео изображений на основе геодезического расстояния, в разделе 3 предложен новый алгоритм стереомэтчинга, основанный на интегрировании поточечных весов функ-

ции рассогласования по окрестностям ранее полученных сегментов, в разделе 4 приведены экспериментальные результаты и, наконец, раздел заключение суммирует наши выводы.

2. ПРИНЦИПЫ СУПЕРПИКСЕЛЬ-СЕГМЕНТАЦИИ НА ОСНОВЕ ГЕОДЕЗИЧЕСКОГО РАССТОЯНИЯ

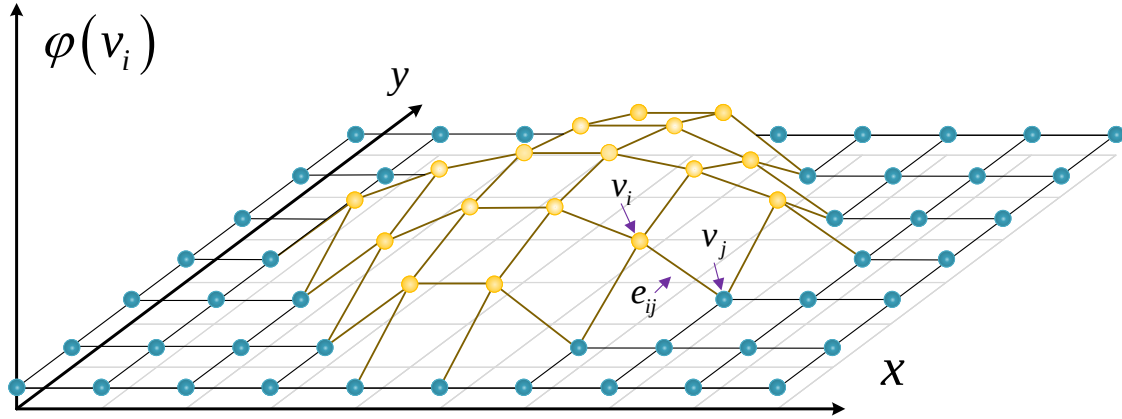


Рис. 1. Изображение как граф на регулярной четырёхсвязной решетке.

Представим изображение, как регулярный граф: $G = \{V, E\}$, где вершины $v_i \in V$ совпадают с пикселями изображения, а ребра $e_i \in E$ составляют регулярную четырёхсвязную решетку, как изображено на Рис. 1. Тогда расстояние между соседними вершинами v_i и v_j задается как функция ребра e_{ij} , равная квадрату частной производной функции яркости изображения $\varphi(v_i)$:

$$d(e_{ij}) = (\varphi(v_j) - \varphi(v_i))^2 + \delta, \quad (1)$$

где δ — усредненное значение расстояний на ребрах графа изображения.

$$\delta = \frac{1}{|E|} \sum_{e_{ij} \in E} e_{ij}. \quad (2)$$

Важным термином для аффинного пространства геодезического расстояния является понятие пути между двумя вершинами $p(v_i, v_j)$, который определяется как конечная последовательность вершин и ребер:

$$p(v_i, v_j) = \{v_i, e_{i,i+1}, v_{i+1}, \dots, v_{j-1}, e_{j-1,j}, v_j\}. \quad (3)$$

Важно отметить, что в рассматриваемом аффинном пространстве вершины и ребра могут встречаться только один раз. Число возможных путей между двумя вершинами пропорционально экспоненте от числа ребер, и длина произвольного пути $|p(v_i, v_j)|$ определяется как:

$$|p(v_i, v_j)| = \sum_{e_{k,l} \in p(v_i, v_j)} d(e_{k,l}). \quad (4)$$

Тогда расстояние между двумя произвольными вершинами $|v_i, v_j|$ для аффинного пространства геодезического расстояния определяется как минимальная длина пути между этими вершинами:

$$|v_i, v_j| = \min \{|p_n(v_i, v_j)|\}. \quad (5)$$

В общем виде задача сегментации изображения $\varphi(v_i)$ формулируется следующим образом: найти функцию $s(v_i)$, ставящую в соответствие каждому пикселю-вершине v_i некоторое натуральное число, означающее номер сегмента. При этом пиксель на изображении должен обладать некоторыми свойствами, по которым этот пиксель можно было бы отнести к тому или иному сегменту. Обычно это какое-то расстояние в пространстве близости. Для суперпиксель-

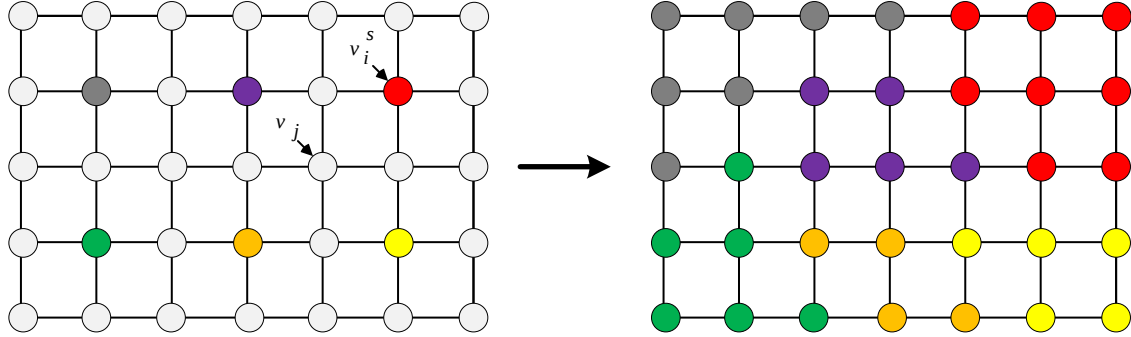


Рис. 2. Исходное распределение вершин-центров $v_i^s \in V^S$ (левая часть рисунка) и результат сегментации (правая часть рисунка).

сегментации добавляется еще два важных условия: а) все пиксели одного сегмента должны образовывать односвязное множество; б) площадь суперпикселей должна быть примерно одинакова, а карта сегментов должна представлять из себя подобие регулярной сетки. Поэтому предполагается, что существует подмножество вершин $v_i^s \in V^S$, называемые центрами суперпикселей, геодезическое расстояние до которых определяет принадлежность всех остальных вершин к тому или иному сегменту, как это показано на Рис. 2. Таким образом для суперпиксель-сегментации на основе геодезического расстояния задача формулируется следующим образом: найти для всех вершин графа-изображения геодезическое расстояние до всех вершин-центров и, выбрав наименьшее, присоединить каждую вершину к множеству ближайшей вершины-центра $v_j^s \in V^S$:

$$s(v_i) = \arg \min_{v_j^s} \{|v_i, v_j^s|\}. \quad (6)$$

Быстрый и точный алгоритм решения подобной задачи был предложен в статье [16] и в дальнейшем мы будем использовать именно этот алгоритм для формирования суперпиксель-сегментации одного из изображений стереопары. Результат такой сегментации представлен на Рис. 3.

Можно видеть, что границы полученных сегментов практически везде совпадают с границами объектов. При этом предполагается, что точки облака совпадают с геометрическими центрами суперпикселями (инициальными точками графа роста суперпиксель-сегмента). Таким образом, облако точек формируется на регулярной решетке левого изображения стереопары, что позволяет суммировать значения функции поточечного рассогласования по окрестностям ранее полученных сегментов. Таким образом, получая робастную функцию рассогласования, связанную именно с точкой центра сегмента. Действительно, в область суммирования попадают лишь те точки суммирования, которые принадлежат некоторому объекту, а значит, функция диспаратности в этой окрестности непрерывная.



Рис. 3. (a) — Результат суперпиксель-сегментации для стереоизображения Adirondack из Middlebury dataset; (б) — Результат суперпиксель-сегментации для стереоизображения Motorcycle из Middlebury dataset.

3. СТЕРЕОМЭТЧИНГ С ПОМОЩЬЮ ИНТЕГРАЛЬНОЙ ФУНКЦИИ РАССОГЛАСОВАНИЯ

Чтобы понять идею нашего метода, необходимо рассмотреть один из самых популярных локальных признаков мэтчинга на изображениях: сенсус-признак [17].

Пусть $I_{i_n \in N_{i_0}} = [i_0, i_1, \dots, i_n, i_{N-1}]$ значение сигнала на изображении в некоторой окрестности N пиксела p_0 (для сенсус-признака это обычно окно 8×8 пикселей) со значением яркости на изображении в этом пикселе i_0 . Тогда сенсус-признаком окрестности N на изображении называется бинарный вектор $B^N = [b_0, b_1, \dots, b_{N-1}]$ так что:

$$b_n = \begin{cases} 1 & \text{if } i_n > i_0 \\ 0 & \text{elsewhere} \end{cases} \quad (7)$$

Вычислительные преимущества такого признака очевидны: вместо сравнения двух окрестностей на разных изображениях, нам достаточно сравнить два битовых слова по метрике Хемминга:

$$D(B_p, B_q) = \sum_{k=0}^{N-1} b_k^p \oplus b_k^q, \quad (8)$$

где $\oplus$ — исключающее или. Преимущества такого признака показал в плане точности мэтчинга опыт работы с этим признаком многих стереоалгоритмов и методов. По сравнению с прямым мэтчингом двух окрестностей сенсус-признак не зависит от изменения освещенности на правом и левом изображениях стереопары. Также сенсус-признак более робастен по сравнению с полным значением сигнала и на границах областей с разными диспатностями. В самом деле, наличие окклюзий на границе двух областей с разными диспатностями приводит к тому, что одна половина окрестности полностью не совпадает с подобной на другом изображении.

Функция рассогласования точки $\mathbf{p}$ на левом изображении I_L относительно точки $\mathbf{q}$ на правом изображении I_R с использованием расстояния сенсус векторов в (8) выражается следующим образом:

$$c(\mathbf{p}, d) = D(B^{I_L}(\mathbf{p}), B^{I_R}(\mathbf{q})). \quad (9)$$

Здесь точки $\mathbf{p}$ и $\mathbf{q}$ связаны как: $\mathbf{p} = [p_x, p_y] \Rightarrow \mathbf{q} = [p_x - d, p_y]$. И искомая функция диспаратности находится по формуле:

$$d(\mathbf{p}) = \arg \min_d \{c(\mathbf{p}, d)\}. \quad (10)$$

Тем не менее, результат стереомэтчинга, основанный на поточечной функции рассогласования в (10), очень неточный (см. Рис. 4 (б)). Поэтому предлагается находить значение диспаратности на основе интегральной функции рассогласования, когда поточечные функции рассогласования суммируются в окрестности сегмента суперпикселя:

$$C(\mathbf{P}(s), d) = \sum_{\mathbf{p} \in s} c(\mathbf{p}, d) \quad (11)$$

и тогда значение диспаратности в точке центра сегмента $\mathbf{P}_s$ вычисляется как:

$$d(\mathbf{P}_s) = \arg \min_d \{C(\mathbf{P}(s), d)\}. \quad (12)$$

Как можно видеть на Рис. 4 (в) использование интегральной функции рассогласования значительно улучшает точность оценки диспаратности (а значит и вычисления значения глубины 3D точек облака). Однако результаты сравнения с самыми лучшими методами не позволяют говорить об state-of-the-art уровне нашего метода. Относительно плохие результаты (еще

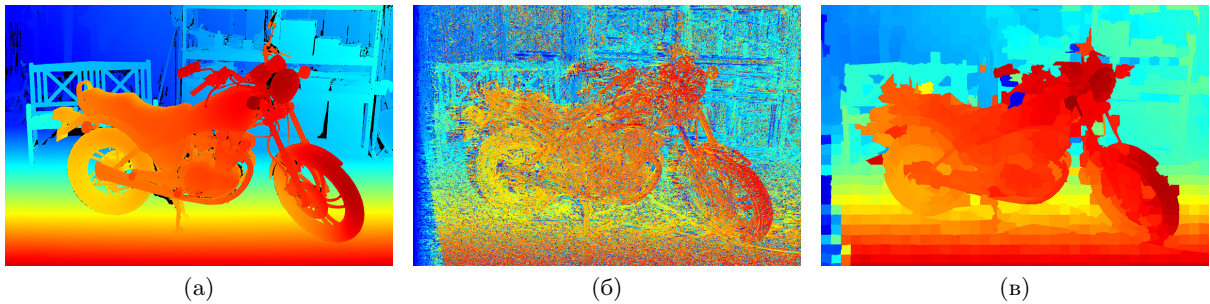


Рис. 4. (а) — Истинное значение карты диспаратности для стереоизображения Motorcycle из Middlebury dataset; (б) — Результат оценки карты диспаратности для стереоизображения Motorcycle из Middlebury dataset при использовании поточечной функции рассогласования. (в) — Результат оценки карты диспаратности для стереоизображения Motorcycle из Middlebury dataset при использовании интегральной функции рассогласования.

15 лет назад такие результаты были бы в топе сравнительной таблицы сайта Middlebury) можно объяснить наличием аутлаеров, точек, где по причине окклюзий или слабой текстуры определить точное значение диспаратности просто невозможно. Так как для получения вектора собственного движения системы, связанной с камерой, не столь важно, сколько точек в облаке — 1000 или 300, но более важна точность определения координат, то имеет смысл исключить из множества точек облака те, для которых определение диспаратности недостоверно. В этой работе нами предложен критерий выбраковки недостоверных точек из числа центров суперпиксель-сегментов, основанный на функции рассогласования между пикселями одного и того же изображения (обычно левого в стереопаре). То есть обе точки $\mathbf{p}$ и $\mathbf{q}$ в (9) принадлежат одному и тому же изображению, а значение диспаратности в этой функции авторассогласования $d = 1$. Если минимум функции интегрального рассогласования в точке $\mathbf{p}$

$$\min_d \{C(\mathbf{p}^{I_L} \rightarrow \mathbf{q}^{I_R}, d)\} > C(\mathbf{p}^{I_L} \rightarrow \mathbf{q}^{I_L}, 1), \quad (13)$$

то такая точка считается недостоверной.

На Рис. 5 (в) показаны только достоверные точки, а на Рис. 5 (б) результат оценки функции диспаратности во всех точках. Такой подход позволяет значительно увеличить точность оценки для всего облака в целом. Что и будет показано в экспериментальной секции этой статьи.

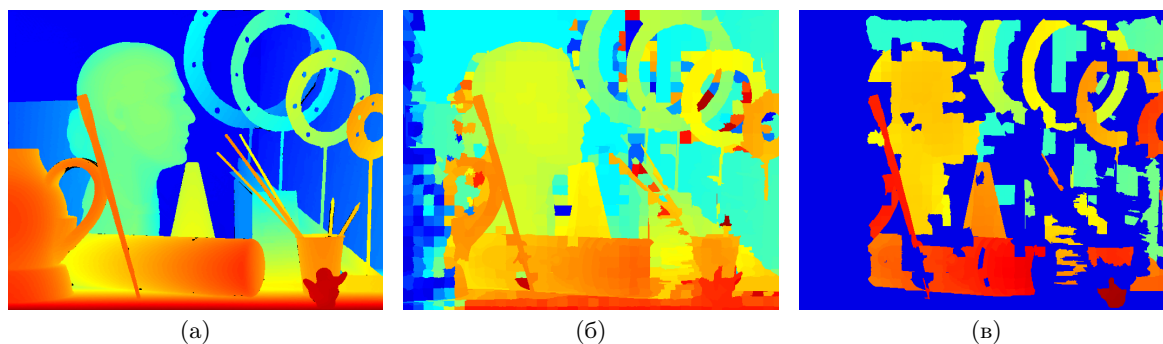


Рис. 5. (а) — Истинное значение карты диспаратности для стереоизображения ArtL из Middlebury dataset; (б) — Результат оценки карты диспаратности для стереоизображения ArtL из Middlebury dataset при использовании интегральной функции рассогласования; (в) — Результат оценки карты диспаратности для стереоизображения ArtL из Middlebury dataset при использовании интегральной функции рассогласования после удаления "ненадежных" точек.

Таблица 1. Численные результаты сравнения реконструкции карты диспаратности различных методов с истинными значениями по критерию MAE (верхняя часть таблицы) и критерию превышения порога неточности 2 пикселей в процентах (нижняя часть таблицы).

Images:	ArtL	Motorcycle	Piano	Pipes
Algorithms:				
MoCha [10]	0.421	0.754	0.570	1.021
Our Base	10.22	4.925	5.035	9.354
+Outliers check	0.506	0.843	0.524	0.945
MoCha [10]	1.162	3.071	4.085	6.426
Our Base	27.83	19.08	22.81	25.83
+Outliers check	1.118	3.045	3.238	2.931

4. ЭКСПЕРЕМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Экспериментальная часть была задумана таким образом, чтобы показать основные достоинства предлагаемого алгоритма, поэтому мы анализируем численные результаты применения нашего метода в случае при использовании интегральной функции рассогласования в сравнении с результатами после удаления "ненадежных" точек. Дополнительно мы сравниваем наш алгоритм с лучшим методом [10] из опубликованных статей. Здесь надо заметить, что наш алгоритм решает разные проблемы, чем метод с применением нейросетевых технологий и глубокого обучения [10]. Однако это не умоляет того факта, что наш достаточно простой и быстрый алгоритм достигает таких же результатов в сравнении с технологиями глубокого обучения, хотя последние технологии требуют гораздо больших вычислительных затрат и специального оборудования. Анализ численных результатов производится на основе открытой базы данных стереоизображений [18].

4.1. Эксперименты с численным анализом результатов

В этой части экспериментальной секции результат, полученный с помощью предложенного базового алгоритма (Our Base) и алгоритма с удалением значений ненадежных точек, сравнивается с методом MoCha [10]. Взятый для сравнения алгоритм на настоящее время

является лучшим из опубликованных по совокупности реконструкции глубин на тестовой выборке Middlebury dataset. В Таблице 1 представлены результаты такого сравнения. Базовый алгоритм определяет значение диспаратности в центре суперпиксель-сегмента. Число сегментов — это параметр алгоритма суперпиксель-сегментации, и в нашем случае этот параметр выбран как 1000 сегментов. По нашим оценкам, это оптимальная величина для стереоизображений. Большее количество суперпикселей приводит к уменьшению окрестности, связанной с точкой облака, что влечет неточность в определении величины диспаратности. С другой стороны, уменьшение числа сегментов ведет к увеличению окрестности и неминуемому включению пикселей, не относящихся к поверхности с одинаковым или близким значением диспаратности. Дополнительный алгоритм удаления "ненадежных" точек сокращает первоначальное число точек облака до 300–400, в зависимости от изображения. Такое количество все еще составляет значительную выборку для робастного определения значений матрицы вращения и вектора смещения.

Анализируя результат в Таблице 1, мы можем сказать, что предложенный в статье алгоритм в варианте +Outliers check значительно превосходит базовый и сравним с лучшим алгоритмом MoCha [10].

5. ЗАКЛЮЧЕНИЕ

В этой работе мы представили новый алгоритм для определения 3D координат точек облака с помощью стереомэтчинга. В отличие от методов плотного стереомэтчинга, требующих достаточно сложных алгоритмических и технических решений, а также больших вычислительных затрат, наш подход решает проблему более простым и менее затратным способом, учитывая, что облака, предназначенные для регистрации, состоят из небольшого количества точек, значительно меньшего, чем пикселей на изображении. Поэтому стереомэтчинг в нашем случае осуществляется, основываясь на некоторой окрестности сопоставляемых пикселей, называемой суперпикселем. Кроме того, мы предложили оригинальный метод верификации надежности определения диспаратности, основанный на локальной автокорреляции, таким образом, удаляя точки облака с ненадежной оценкой диспаратности. Мы показали, что в результате применения процедуры верификации наша оценка 3D координат точек сцены сравнима с лучшими state-of-the-art алгоритмами глубокого обучения.

СПИСОК ЛИТЕРАТУРЫ

1. Scharstein D., Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms // International journal of computer vision. 2002. Vol. 47, no. 1-3. P. 7–42.
2. Ershov E., Karnaukhov V., Mozerov M. Probabilistic choice between symmetric disparities in motion stereo matching for a lateral navigation system // Optical Engineering. 2016. Vol. 55, no. 2. P. 023101–023101.
3. Mozerov M., van de Weijer J. Accurate stereo matching by two-step energy minimization // IEEE Transactions on Image Processing. 2015. Vol. 24, no. 3. P. 1153–1163.
4. Kolmogorov V., Zabih R. What energy functions can be minimized via graph cuts? // IEEE transactions on pattern analysis and machine intelligence. 2004. Vol. 26, no. 2. P. 147–159.
5. Boykov Y., Veksler O., Zabih R. Fast approximate energy minimization via graph cuts // IEEE Trans. Pattern Analysis and Machine Intelligence. 2001. Vol. 23, no. 11. P. 1222–1239.
6. Kolmogorov V., Zabih R. Computing visual correspondence with occlusions using graph cuts // ICCV. 2001. P. 508–515.
7. Sun J., Zheng N.-N., Shum H.-Y. Stereo matching using belief propagation // IEEE Trans. Pattern Analysis and Machine Intelligence. 2003. Vol. 25, no. 7. P. 787–800.

8. Ihler A., Fisher J., Willsky A. Loopy belief propagation: Convergence and effects of message errors // J. Machine Learning Research. 2005. Vol. 6. P. 905–936.
9. Mozerov M. G., van de Weijer J. One-view occlusion detection for stereo matching with a fully connected crf model // IEEE Transactions on Image Processing. 2019. Vol. 28, no. 6. P. 2936–2947.
10. Chen Z., Long W., Yao H. et al. Mocha-stereo: Motif channel attention network for stereo matching // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024. P. 27768–27777.
11. Zheng Y., Li Y., Yang S., Lu H. Global-pbnet: A novel point cloud registration for autonomous driving // IEEE Transactions on Intelligent Transportation Systems. 2022. Vol. 23, no. 11. P. 22312–22319.
12. Makovetskii A. Y., Kober V., Voronin S. et al. Registration of point clouds in 3d space using soft alignment // Journal of Communications Technology and Electronics. 2024. P. 1–9.
13. Papon J., Abramov A., Schoeler M., Worgotter F. Voxel cloud connectivity segmentation-supervoxels for point clouds // Proceedings of the IEEE conference on computer vision and pattern recognition. 2013. P. 2027–2034.
14. Wang J., Wang X. Vcells: Simple and efficient superpixels using edge-weighted centroidal voronoi tessellations // IEEE Transactions on pattern analysis and machine intelligence. 2012. Vol. 34, no. 6. P. 1241–1247.
15. Giordano D., Murabito F., Palazzo S., Spampinato C. Superpixel-based video object segmentation using perceptual organization and location prior // Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. P. 4814–4822.
16. Mozerov M., Karnaukhov V., Kober V., Zimina L. Accurate and fast geodesic distance calculation algorithm for superpixel segmentation // Journal of Communications Technology and Electronics. 2023. Vol. 68, no. Suppl 2. P. S254–S262.
17. Zabih R., Woodfill J. Non-parametric local transforms for computing visual correspondence // ECCV. 1994. P. 151–158.
18. Scharstein D., Hirschmüller H., Kitajima Y. et al. High-resolution stereo datasets with subpixel-accurate ground truth // German conference on pattern recognition / Springer. 2014. P. 31–42.

3D point cloud coordinates estimation using stereo matching and superpixel segmentation

M.G. Mozerov, V.N. Karnaukhov, V.I. Kober, L.V. Zimina

An important element of the movement of robotic systems is the determination of the object's position in the coordinates associated with the scene, as well as the vector of the autonomous device's own movement in this coordinate system (rotation matrix and displacement vector). All these parameters can be determined by estimating the values of the 3D coordinates of the cloud of scene points associated with the coordinate system of the onboard video camera of an autonomous robotic device. The process of finding the rotation matrix and the displacement vector is called the registration of three-dimensional point clouds and is of great importance in robotics and computer vision. Therefore, there are many methods of such registration. However, the problem is how to estimate the values of the 3D coordinates of the point cloud of the scene. In this paper, we propose to determine the desired coordinates of points using stereomatching. The main assumption of classical stereomatching is that matching pixels in stereo images have the same brightness values. However, this assumption is generally incorrect, given the presence of noise in the images and the different illumination of the left and right images in the stereo pair. In addition, pixel-by-pixel mapping does not work in areas where there is no real texture. The methods that solve all the problems of dense stereo patching are quite complex and require high computational costs. However, the problem can be solved in a simpler and less costly way, given that the clouds intended for registration consist of a small number of points, significantly smaller than the pixels in the image. Therefore, stereomatching in our case is carried out

based on a certain neighborhood of the matched pixels, called a superpixel. The accuracy of such matching surpasses the use of a feature close to the principle of robust sensation. In addition, we have proposed an original method for verifying the reliability of the disparity determination based on local autocorrelation, thus removing cloud points with unreliable disparity estimates. As a result, our estimate of the 3D coordinates of "reliable" scene points is comparable to the best state-of-the-art deep learning algorithms.

Ключевые слова: Computer vision, point cloud, stereomatching, superpixel segmentation.