

19. Verkhagen K., Deyn R., Grun F. i dr. Raspoznavanie obrazov. Sostoyanie i perspektivy [Pattern recognition. State and prospects]. Moscow: Radio i svyaz', 1985, 104 p.
20. Kovalev V.V. Sergeev N.E. Metody resheniya problemy pereobucheniya neyronnykh setey v zadache obnaruzheniya malorazmernykh ob'ektov na izobrazheniyakh [Methods for solving the problem of re-training neural networks in the task of detecting small-sized objects in images], *Mezhdunarodnyy kongress po intellektual'nym sistemam i informatsionnym tekhnologiyam «IS&IT'23»* [International Congress on Intelligent Systems and Information Technologies "IS&IT'23"], 2023, No. 1, pp. 38-45.
21. Padilla R., Netto S.L., Silva E.A.B. A Survey on Performance Metrics for Object-Detection Algorithms, *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Niteroi, Brazil, 2020, pp. 237-242.

Статью рекомендовал к опубликованию д.т.н., профессор А.В. Боженюк.

**Ковалев Владислав Владимирович** – Научно-конструкторское бюро вычислительных систем, e-mail: vlad.kovalev94@mail.ru; г. Таганрог, Россия; тел.: +79525864492; программист.

**Сергеев Николай Евгеньевич** – Южный федеральный университет; e-mail: nesergeev@sfedu.ru; г. Таганрог, Россия; тел.: +79281742585; кафедра вычислительной техники; д.т.н.; профессор.

**Kovalev Vladislav Vladimirovich** – Scientific design bureau of computer systems, e-mail: vlad.kovalev94@mail.ru; Taganrog, Russia; phone: +79525864492; programmer.

**Sergeev Nikolay Evgenievich** – Southern Federal University; e-mail: nesergeev@sfedu.ru; Taganrog, Russia; phone: +79281742585; the department of Computer Science; dr. of eng. sc.; professor.

УДК 004.932.2

DOI 10.18522/2311-3103-2024-1-276-285

**К.И. Морев, П.А. Ледерер**

### **ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА ПОГРЕШНОСТЕЙ ВОССТАНОВЛЕНИЯ СТРУКТУРЫ НАБЛЮДАЕМОЙ СЦЕНЫ ИЗ СЕРИИ СНИМКОВ ДВИЖУЩЕЙСЯ КАМЕРЫ**

*Статья посвящена исследованию влияния использования различных математических моделей камер, следовательно, моделей формирования изображения сцены, при восстановлении 3-Д структуры сцены из набора 2-Д снимков в процессе движения камеры (восстановлении структуры из движения, далее – СИД). Сравнительная оценка проводится для двух моделей камер: классической модели камеры центральной проекции и, сравнительно новой, модели всенаправленной камеры. В статье дано краткое описание математической модели всенаправленной камеры, описанная модель используется в ходе экспериментов, также описаны способы представления изображений от всенаправленных камер. Дополнительно дается описание математической модели классической камеры центральной проекции. Описанная модель также используется в ходе проведения экспериментов. Используемые при решении задачи восстановления структуры из движения аналитические выкладки кратко упоминаются в статье. Также дается описание алгоритма получения 3-Д координат точек наблюдаемой сцены из последовательности снимков в движении. Проведенные в рамках исследования эксперименты подробно описаны в данной статье. Раскрыт процесс установки визуальных ориентиров, определения их истинных 3-Д координат. Описаны действия по формированию наборов данных для получения сравнительных оценок. В заключении работы дан анализ результатов экспериментов, выделены модели, позволяющие уменьшить погрешности восстановления 3-Д координат наблюдаемых визуальных ориентиров.*

*Математические модели камер; восстановление структуры из движения; визуальные ориентиры; всенаправленная камера; эпиполярная геометрия.*

K.I. Morev, P.A. Lederer

**EXPERIMENTAL ESTIMATION OF ERRORS IN RECONSTRUCTING  
THE STRUCTURE OF THE OBSERVED SCENE FROM A SERIES  
OF IMAGES BY VARIOUS CAMERAS**

*The article is devoted to the study of the influence of using various mathematical models of cameras, and therefore models of scene image formation, when restoring the 3-D structure of a scene from a set of 2-D images during camera movement (restoring the structure from motion, hereinafter referred to as LEDs). A comparative assessment is carried out for two camera models: the classic central projection camera model and the relatively new omnidirectional camera model. The article provides a brief description of the mathematical model of an omnidirectional camera, the described model is used during experiments, and also describes ways to represent images from omnidirectional cameras. Additionally, a description of the mathematical model of the classical camera of the central projection is given. The described model is also used during experiments. The analytical calculations used in solving the problem of restoring structure from motion are briefly mentioned in the article. An algorithm for obtaining 3-D coordinates of the points of the observed scene from a sequence of images in motion is also described. The experiments carried out as part of the study are described in detail in this article. The process of setting visual landmarks and determining their true 3-D coordinates is revealed. The steps for the formation of data sets for obtaining comparative estimates are described. At the end of the work, an analysis of the experimental results is given, models are identified that reduce the errors in restoring the 3-D coordinates of the observed visual landmarks.*

*Mathematical models of cameras; restoration of structure from motion; visual landmarks; omnidirectional camera; epipolar geometry.*

**Введение.** Задача восстановления структуры из движения является одной из множества задач сферы компьютерного зрения и как научная задача известна с конца прошлого века [1]. Однако, несмотря на давнюю известность и достаточно глубокий уровень изучения, в рамках задачи восстановления СИД продолжают появляться новые приложения. Так, ставшие популярными в последние годы всенаправленные камеры открывают новые возможности при решении задачи СИД. В практике компьютерного зрения всенаправленные камеры впервые были использованы в фотограмметрии, когда изображения всенаправленной камеры, как правило, регистрировалось путем вращения обычной камеры вокруг центра проекции или сшивкой панорамы из множества изображений от классических камер. В последние годы эффективность использования всенаправленных камер была значительно повышена благодаря прогрессу в области светочувствительных сенсоров и методов обработки изображений, например, появлению цифровых камер высокого разрешения и высокоточных алгоритмов сшивки панорам. В отличие от классических камер, использующих объективы с ограниченным полем зрения, всенаправленные камеры не ограничиваются полем зрения одного объектива, а захватывают всю окружающую сцену из одной точки пространства. Существуют разные способы получения такого панорамного изображения. Например, оно может быть сделано с помощью стандартной камеры и специализированного оборудования, такого как моторизованные панорамные комплексы (например, [2–5]) или специализированное автоматическое программное обеспечение (например, [3, 6]), или, что значительно удобнее, с помощью программно-аппаратных комплексов, которые автоматизируют все шаги, такие как Civetta от Weiss AG [7] или Sphero-Cam от SpheronVR [8].

Используемые модели камер. В этом разделе приведены используемые в расчетах математические модели камер. Первой описывается модель классической камеры центральной проекции (рис. 1). Изображение камеры формируется проекцией через центральную точку камеры всех наблюдаемых точек мира на плоскость изображения [9]. Камера определяется центром  $S$ , локальной прямоугольной системой координат и плоскостью изображения. Трехмерная точка  $M$  имеет координаты  $M_c = (X_{M_c}, Y_{M_c}, Z_{M_c})$  в системе координат камеры. Для получения изображе-

ния точки  $M$ , точка  $M_c$  проецируется на плоскость изображения в точку  $m = M_c / Z_{M_c}$ . Координата  $Z$  в системе камеры определяет расстояние от центра камеры до проецируемой точки [10].

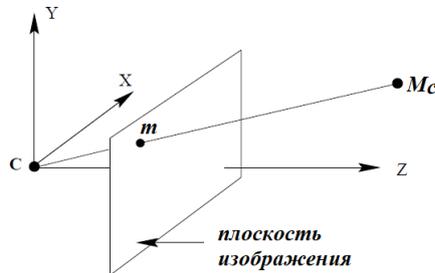


Рис. 6. Модель камеры центральной проекции

Изображение камеры центральной проекции представляет собой плоский снимок сцены, ограниченной шириной поля зрения камеры.

Ниже представлена модель всенаправленной камеры, используемая в дальнейших экспериментах. Изображение, полученное всенаправленной камерой, аппроксимируется центральной проекцией всех наблюдаемых точек пространства на сферическую поверхность – поверхность изображения. Модель всенаправленной камеры, аналогично камере центральной проекции, определяется как показано на рис. 2, ее центром  $C$ , локальной прямоугольной системой координат камеры и поверхностью изображения – поверхностью сферы [11]. Для получения изображения точки  $M$ , точка  $M_c$  проецируется на поверхность сферы в точку  $m = M_c / \|M_c\|$  (Вектор нормировали, чтобы получить единичную длину).

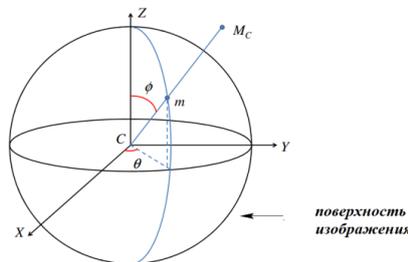


Рис. 7. Модель всенаправленной камеры

Таким образом, точки изображения всенаправленной камеры могут быть описаны единичными трехмерными векторами, подобно N-векторам в [12]. Единичные трехмерные вектора описывают положение точки в прямоугольной системе координат камеры, однако положение точек изображения удобнее выразить в сферических координатах. Сферические координаты  $(r, \theta, \phi)$  связаны с декартовыми координатами  $(X, Y, Z)$  через выражение

$$r = \sqrt{X^2 + Y^2 + Z^2}, \theta = \operatorname{tg}^{-1}\left(\frac{Y}{X}\right), \phi = \cos^{-1}\left(\frac{Z}{r}\right). \quad (1)$$

Изображение всенаправленной камеры представляет сферическую панораму сцены  $360^\circ \times 180^\circ$  ( $360^\circ$  по горизонтали и  $180^\circ$  по вертикали) вокруг центра камеры. Полученные всенаправленной камерой изображения будут представлены в виде равноугольной проекции сферы на плоскость. Формат именно равноугольной проекции наиболее широко применяется при обработке изображений, включая хор-

шо известные пакеты программного обеспечения с открытым исходным кодом и коммерческие пакеты программного обеспечения, например, OpenMVG [13], Agisoft Metashape [14] и Pix4Dmapper [15]. На рис. 3 показан пример изображения от всенаправленной камеры, развернутого в равноугольной проекции.

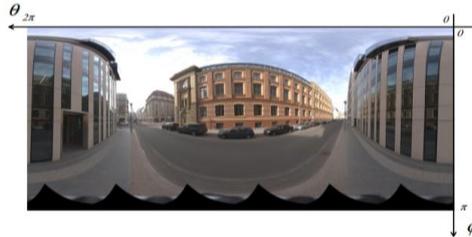


Рис. 8. Пример сферической панорамы в равноугольной проекции

Для упомянутых выше моделей камер предполагается существование глобальной системы координат (далее – ГКС), в которой может быть определено положение камеры (положение и ориентация). Если точка  $M$  имеет декартовы координаты  $M_w$  в мировой системе координат, то координаты  $M_c$  определяются поворотом  $R$  и сдвигом  $t$  таким образом, что:

$$M_c = RM_w + t. \quad (2)$$

Алгоритмы, применяемые при решении задачи СИД. Ниже описывается общая для обеих моделей камер схема действий для решения задачи СИД. Основная идея состоит в том, чтобы вычислить относительное движение между двумя первоначальными камерами, используя соответствия между точками двух изображений и уравнения эпиполярной геометрии [16]. Затем получить трехмерные координаты этих точек (начальная структура), решая триангуляционную задачу для всех сопоставленные между первоначальными камерами точек. После этой инициализации для каждого имеющегося следующего изображения последовательно выполняются следующие шаги: (1) добавляется одна новая камера в набор, (2) вычисляется 2D-3D соответствия из новой камеры с уже восстановленными триангуляцией точками, (3) вычисляется положение новой камеры, с использованием упомянутых 2D-3D соответствий и (4) восстанавливаются 3D координаты всех оставшихся точек по вычисленным параметрам связи между камерами. После добавления каждой новой камеры выполняется уточнение положений камер и структуры 3D-точек по вновь вычисленным координатам.

Эпиполярная геометрия для двух камер описана в [17]. Однако, для наглядности, ниже дано краткое описание основных принципов взаимосвязи между двумя камерами. Пусть  $C_1$  и  $C_2$  – центры двух камер. Без потери общности мы можем предположить, что система координат первой камеры является мировой системой координат, т.е.  $C_1 = 0$ , а  $R$  и  $t$  определяют поворот и смещение для второй камеры. Точка  $M$  проецируется в  $m_1$  и  $m_2$  в первой и второй камерах соответственно (рис. 4).

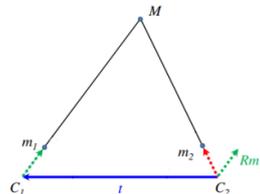


Рис. 9. Элементы эпиполярной геометрии для двух изображений произвольных камер

В общем случае точки  $M$ ,  $m_1$ ,  $C_1$ ,  $m_2$  и  $C_2$  лежат в одной плоскости – эпиллярной плоскости. И это можно выразить, указав, что векторы  $m_2$ ,  $Rm_1$  и  $t$  являются копланарными. Также существование эпиллярной плоскости позволяет определить основное ограничение эпиллярной геометрии:

$$m_2^T E m_1 = 0, E = [t]_{\times} R. \quad (3)$$

Хорошо известно, что матрица  $E$  – существенная матрица, может быть вычислена с использованием восьми пар соответствий [18]. Поскольку уравнение 2 использует только предположение о копланарности, без учета ориентации лучей, разложение  $E$  на  $R$  и  $t$  обычно имеет 4 решения [19]. В камерах плоской проекции правильное решение ищется через знание о положительной глубине сцены ( $z$ -значение), а, следовательно, положительного расстояния от камеры до трехмерных точек. Для всенаправленных камер применяется аналогичное ограничение, но проверяется не знак глубины, а знак произведения  $m^T M c$ .

**Проведение эксперимента.** Сбор данных является первым шагом в наших экспериментах. Исходя из принципа конструкции, сферические камеры можно разделить на три основные категории: диоптрические камеры, катадиоптрические камеры и полидиоптрические камеры [20, 21]. Для проведения эксперимента использовалась камера Samsung Gear360, относящаяся к полидиоптрическим камерам. Ниже на рис. 5 приведено фото используемой в экспериментах камеры.

В качестве экспериментальной площадки выбрано небольшое помещение. В ходе эксперимента требуется восстановить структуру комнаты размерами 5х4м, а именно: вычислить 3-Д координаты визуальных ориентиров – размещенных на стенах агисо-маркеров. Агисо-маркеры были размещены на одной высоте относительно уровня пола, расстояние между маркерами было определено и зафиксировано. Следует отметить, что расположение ориентиров на одной высоте обеспечивает нахождение их в одной плоскости. Расстояние между ориентирами определено в метрах. Схема расположения визуальных ориентиров приведена ниже на рис. 6.



Рис. 10. Фото камеры, используемой в экспериментах

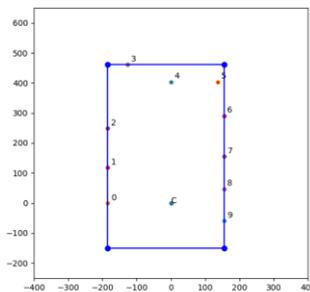


Рис. 11. Схема расположения визуальных ориентиров

На рис. 6 цифрами обозначены номера ориентиров, а крестами их положение. Также на рис. 6 отображена точка С – точка начала системы координат в которой определены координаты ориентиров. В табл. 1 указаны координаты визуальных ориентиров.

Таблица 6

**Истинные координаты визуальных ориентиров**

Номер ориентира	Координаты (X,Y,Z)
0	(-185; 0; 0)
1	(-185; 117,5; 0)
2	(-185; 248; 0)
3	(-126; 461; 0)
4	( 0; 403,5; 0)
5	( 137,5; 403,5; 0)
6	( 155; 290,5; 0)
7	( 155; 154,5; 0)
8	( 155; 47; 0)
9	( 155; -60; 0)

Упомянутая ранее камера Samsung Gear360 была установлена на штатив фиксированной высоты так, что визуальные ориентиры и оптический центр камеры находятся в одной плоскости. Ниже на рис. 7 представлен внешний вид стенда для выверки положения визуальных ориентиров и непосредственно ориентиры – агисо-маркеры.



Рис. 12. Внешний вид целевых визуальных ориентиров

Движение штатива с камерой выполнялось вручную. Каждое новое положение штатива определялось случайным образом, исходя из возможностей перемещения. Основной задачей при перемещении считался проход по помещению между противоположными углами для получения снимков как можно большего количества характерных для комнаты деталей.

Таким образом, было накоплено 10 наборов снимков всенаправленной камерой. Ниже на рис. 8 представлен пример снимка из набора.



Рис. 13. Пример исходного изображения от всенаправленной камеры

Как упоминалось ранее, обработка изображений от всенаправленной камеры будет производиться с изображениями, представленными в равноугольной проекции. Пример панорамы в равноугольной проекции, полученной из изображения входного набора данных приведен ниже на рис. 9.



Рис. 14. Изображение от всенаправленной камеры в равноугольной проекции

После преобразования исходного набора данных к виду равноугольных проекций, было произведено восстановление структуры помещения. Результат восстановления структуры приведен ниже на рис. 10.

Далее, для оценки координат ориентиров с использованием модели классической камеры центральной проекции из исходных снимков всенаправленной камеры (см. рис. 8) были вырезаны области разрешением 1920x1080. Таким образом имитировалась работа камеры с шириной поля зрения 90 градусов по горизонтали. На рис. 11 приведен пример полученного имитирующего изображения.



Рис. 15. Пример восстановленной структуры помещения



Рис. 16. Имитация снимка классической камеры центральной проекции

В конечном итоге, мы имеем 10 наборов снимков от всенаправленной камеры в движении по комнате с визуальными ориентирами (см. рис. 8). И 10 наборов снимков от классической камеры, аналогично двигающейся по комнате (см. рис. 11). Для каждого набора произведена последовательность действий по восстановлению структуры наблюдаемой сцены и получены 3-Д координаты наблюдаемых визуальных ориентиров.

Далее были вычислены среднее значение, средноквадратическое отклонение (далее – СКО), минимальное и максимальное значение разницы между истинным и восстановленным положением визуального ориентира. Указанные значения были вычислены для координат по каждой из осей отдельно. Выражение для вычисления СКО:

$$\sigma = \sqrt{\frac{\sum_{i=0}^n (x_i - \bar{x})^2}{n}}, \quad (4)$$

где  $x_i$  – значение оцененных координат визуального ориентира,  $\bar{x}$  – среднее значение координат для всех экспериментов,  $n$  – количество экспериментов. Ниже в табл. 2 и 3 содержатся статистические характеристики для величины разницы между истинными и вычисленными 3-Д координатами визуальных ориентиров. Результаты приведены для всех 10-ти наборов снимков от всенаправленной и классической камер, соответственно.

Для табл. 2 и 3 названия столбцов:  $\bar{x}$  – среднее значение разницы с истинным значением положения ориентира по 10-ти наборам снимков,  $\sigma$  – СКО значения разницы,  $min$  и  $max$  – минимальное и максимально значение ошибки, соответственно. Ниже на рис. 12 изображены вычисленные координаты визуальных ориентиров для обеих моделей камер. Эллипсами изображены пределы  $3\sigma$  для вычисленных координат каждого ориентира в отдельности. Координаты центра эллипса соответствуют среднему значению для вычисленных координат ориентира.

Таблица 7

Результаты оценки погрешностей для сферической модели камеры

Номер ориентира	$\bar{x}$	$\sigma$	$min$	$max$
0	(3.82; -6.31; -3.01)	(4.56; 12.85; 10.2)	(0.61 ; 3.53; 0.48)	(12.47 ; 8.96 ; 14.26)
1	(-0.25; -0.21; 4.89)	(8.35; 12.81; 8.61)	(1.64; 1.785; 1.47)	(8.88; 18.46; 15.86)
2	(3.16; 0.39; -1.53)	(4.59; 3.46; 8.86)	(4.40; 7.83; 1.46)	(10.50; 3.92; 16.6)
3	(3.56; 1.4; -3.17)	(7.84; 13.28; 12.78)	(1.11; 2.23; 3.51)	(18.31; 23.17; 10.52)
4	(0.06 ; 2.30; 4.08)	(6.67 ; 8.30; 7.61)	(3.56; 1.37; 0.86)	(15.15; 17.38; 15.17)
5	(1.62 ; 0.66; -1.46)	(4.98; 4.57; 5.85)	(0.78 ; 1.13 ; 1.29)	(9.28; 9.16; 7.21)
6	(0.24 ; -2.72 ; -5.35)	(5.16; 11.91; 13.23)	(0.81; 1.42; 2.67)	(8.15; 21.28; 19.81)
7	(1.39; 0.08; 0.51)	(10.86; 3.78; 5.13)	(1.32 ; 0.77; 0.34)	(16.1; 6.13; 9.9)
8	(7.20 ; 2.48 ; -3.73)	(3.52; 3.21; 3.44)	(1.64; 1.75; 1.27)	(18.78; 19.21; 14.55)
9	(-1.58; 0.11 ; -3.11)	(3.26; 5.65; 8.76)	(0.775; 0.768; 1.87)	(3.65; 10.03; 7.24)

Таблица 8

Результаты оценки погрешностей для классической модели камеры

Номер ориентира	$\bar{x}$	$\sigma$	$min$	$max$
0	(2.333; -7.66; 9.44)	(3.99; 14.01; 21.98)	(3.17; 2.94; 1.52)	(9.47; 15.30; 36.71)
1	(-1.89; 6.00; 5.05)	(8.94; 15.30; 10.54)	(5.13; 7.63; 1.61)	(12.41; 29.55; 25.42)
2	(2.83; 9.95; 9.01)	(14.59; 13.88; 16.12)	(5.93; 5.61; 6.28)	(25.62; 39.10; 36.57)
3	(-1.82; -0.87; -1.20)	(5.73; 18.36; 19.40)	(9.65; 3.08; 6.73)	(10.23; 38.80; 42.61)
4	(6.97; 2.99; 7.52)	(15.18; 9.13; 20.15)	(1.26; 1.15; 3.27)	(27.71; 19.36; 36.15)
5	(3.86; 1.50; 5.26)	(19.37; 12.15; 25.14)	(8.31; 9.79; 2.56)	(29.23; 24.45; 36.84)
6	(2.80; -0.25; -0.99)	(16.21; 6.84; 20.29)	(7.2; 9.69; 7.13)	(20.97; 12.93; 39.12)
7	(1.87; 0.19; 0.75)	(5.75; 18.75; 21.15)	(6.48; 3.89; 5.90)	(12.25; 32.60; 30.11)
8	(1.81; -6.39; -10.13)	(10.4; 14.13; 20.32)	(5.17; 4.0; 45.94)	(24.20; 18.00; 18.35)
9	(-2.91; -4.50; 17.78)	(17.75; 17.87; 26.75)	(4.51; 3.08; 7.81)	(33.06; 24.17; 55.96)

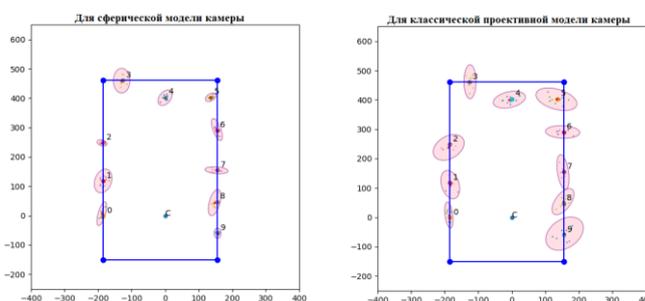


Рис. 17. Положения вычисленных 3-Д координат ориентиров

На рис. 12 продемонстрирована разница в точности решения задачи восстановления 3-Д координат заданных ориентиров. В левой части рисунка 12 приведены результаты вычисления координат при использовании сферической модели камеры, в правой для классической модели центральной проекции. Радиусы эллипсов слева, отображающих погрешности для сферической модели камеры, заметно меньше радиусов справа. А это значит что СКО ошибки относительно истинных координат меньше при использовании именно сферической модели камеры. Также можно обратить внимание на данные в табл. 2 и 3. Значения среднего  $\bar{x}$  сопоставимы при использовании обеих моделей камер, однако значения  $\sigma$  и  $max$  заметно больше при использовании модели проективной камеры, что опять же указывает на меньшую точность по результатам эксперимента.

**Заключение.** Результаты эксперимента позволяют сделать вывод о том, что использование всенаправленных камер дает возможность повысить точность решения задачи СИД при работе в закрытых помещениях. Более высокая точность определения координат наблюдаемых ориентиров при использовании сферической модели камеры может быть объяснена использованием сферической панорамы – относительно большого изображения, содержащего большее количество деталей и объектов сцены, позволяющего производить более точное сопоставление между двумя снимками. Также к преимуществам всенаправленных камер можно отнести более быстрый процесс формирования необходимого количества данных (снимков сцены) для запуска конвейера восстановления СИД.

В будущем предполагается исследование влияния выбранной модели ошибок эпиполярной геометрии на точность восстановленной структуры и скорость ее вычисления.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Xiong B., Jancosek M., Elberink S.O., Vosselman G. Flexible building primitives for 3d building modeling, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, Vol. 101, pp. 275-290.
2. Murtiyoso A., Grussenmeyer P. Documentation of heritage buildings using close-range uav images: dense matching issues, comparison and case studies, *The Photogrammetric Record*, 2017, Vol. 32, pp. 206-229.
3. Liao J., Yue Y., Zhang D., Tu W., Cao R., Zou Q., Li Q. Automatic tunnel crack inspection using an efficient mobile imaging module and a lightweight cnn, *IEEE Transactions on Intelligent Transportation Systems*, 2022.
4. Zhang C., Cui Y., Zhu Z., Jiang S., Jiang W. Building height extraction from gf-7 satellite images based on roof contour constrained stereo matching, *Remote Sensing*, 2022, Vol. 14, pp. 1566.
5. da Silveira T.L., Jung C.R. Dense 3d scene reconstruction from multiple spherical images for 3-dof+ vr applications, in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, 2019, pp. 9-18.
6. Leutenegger S., Chli M., Siegwart R. BRISK: Binary Robust Invariant Scalable Keypoints, *Computer Vision (ICCV)*, 2011, pp. 2548-2555.
7. Calonder M., Lepetit V., Strecha C., Fua P. BRIEF: Binary Robust Independent Elementary Features, *11th European Conference on Computer Vision (ECCV)*, 2010, pp. 778-792.
8. Wen W., Zhou Y., Zhang G., Fahandezh-Saadi S., Bai X., Zhan W., Tomizuka M., Hsu L.-T. Urbanloco: A full sensor suite dataset for mapping and localization in urban scenes, in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 2310-2316.
9. Pagani A., Stricker D. Structure from motion using full spherical panoramic cameras, in *2011 IEEE International Conference on Computer Vision Workshops*, 2011, pp. 375-382.
10. da Silveira T.L., Pinto P.G., Murrugarra-Llerena J., Jung C.R. 3d scene geometry estimation from 360 imagery: A survey, *ACM Computing Surveys*, 2022, Vol. 55, pp. 1-39.
11. Angelov D., Dulong C., Filip D., Frueh C., Lafon S., Lyon R., Ogale A., Vincent L., Weaver J. Google street view: Capturing the world at street level, *Computer*, 2010, Vol. 43, pp. 32-38.
12. Zhang X., Zhao P., Hu Q., Ai M., Hu D., Li J. A uav-based panoramic oblique photogrammetry (pop) approach using spherical projection, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, Vol. 159, pp. 198-219.
13. Lai P.K., Xie S., Lang J., Laganiere R. Real-time panoramic depth maps from omni-directional stereo images for 6 dof videos in virtual reality, in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019, pp. 405-412.

14. Hartmann W., Havlena M., Schindler K. Recent developments in large-scale tie-point matching, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, Vol. 115, pp. 47-62.
15. Zhao Q., Feng W., Wan L., Zhang J. Sphorb: A fast and robust binary feature on the sphere, *International journal of computer vision*, 2015, Vol. 113, pp. 143-159.
16. Torii A., Havlena M., Pajdla T. From google street view to 3d city models, in *2009 IEEE 12th international conference on computer vision workshops, ICCV Workshops*, 2009, pp. 2188-2195.
17. Tian Y., Fan B., Wu F. L2-net: Deep learning of discriminative patch descriptor in euclidean space, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 661-669.
18. Murrugarra-Llerena J., da Silveira T.L., Jung C.R. Pose estimation for two-view panoramas based on keypoint matching: A comparative study and critical analysis, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5202-5211.
19. Jiang S., Jiang C., Jiang W. Efficient structure from motion for large-scale UAV images: A review and a comparison of sfm tools, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, Vol. 167, pp. 230-251.
20. Mouragnon E., Lhuillier M., Dhome M., Dekeyser F., Sayd P. Real time localization and 3d reconstruction, in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, Vol. 1, pp. 363-370.
21. Scaramuzza D., Martinelli A., Siegwart R. A flexible technique for accurate omnidirectional camera calibration and structure from motion, in *Fourth IEEE International Conference on Computer Vision Systems*, 2006, pp. 45-45.

Статью рекомендовал к опубликованию д.т.н., профессор А.В. Боженюк.

**Морев Кирилл Иванович** – Акционерное общество Научно-конструкторское бюро вычислительных систем (АО НКБ ВС); e-mail: morev-ki@ya.ru; г. Таганрог, Россия; ; программист.

**Ледерер Петр Алексеевич** – e-mail: plederer25@gmail.com; программист.

**Morev Kirill Ivanovich** – Joint Stock Company "Scientific Design Bureau of Computing Systems" (JSC SDB CS); e-mail: morev-ki@ya.ru; Taganrog, Russia; programmer.

**Lederer Pyotr Alekseevich** – e-mail: plederer25@gmail.com; programmer.

УДК 28.23.24: 28.23.27: 004.82

DOI 10.18522/2311-3103-2024-1-285-299

**С.М. Соколов**

## **АНАЛИЗ СЦЕНЫ В ИНФОРМАЦИОННЫХ СИСТЕМАХ ПОДВИЖНЫХ РОБОТОТЕХНИЧЕСКИХ КОМПЛЕКСОВ**

Современные роботы способны выполнять все более сложные задачи, которые обычно требуют высокой степени взаимодействия с окружающей средой, в которой им приходится работать. Как следствие, роботизированные системы должны обладать глубокими и конкретными знаниями о своих рабочих пространствах, которые выходят далеко за рамки простого представления показателей, которое роботизированная система может создать с помощью приёмов обработки только зрительных данных, например в задаче одновременной локализации и картографирования (SLAM). Анализ сцены является связующим звеном между распознаванием объектов и знанием об окружающем мире и в том или ином виде присутствует в процессе извлечения из зрительных данных информации, необходимой для решения конкретной задачи. В статье представляется систематизированный подход к обеспечению анализа сцены бортовыми СТЗ. Рассматриваются технологии анализа сцены как составной части повышения степени автономности подвижных РТК. Ряд технологий только предстоит освоить и воплотить в жизнь, но общая структура позволяет постепенно углублять анализ сцены на борту РТК, тем самым повышая степень автономности без коренной переделки бортовой информационно-управляющей системы и СТЗ, как ключевой части информационного обеспечения. Информация, извле-