

ВЫБОР МЕТОДОВ РАСЧЕТА ОБЪЕМОВ ПЕРЕВОЗИМЫХ ЛЕСОМАТЕРИАЛОВ С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ

THE CHOICE OF METHODS FOR CALCULATING THE VOLUME OF TRANSPORTED TIMBER USING A VISION SYSTEM

**D. Vegera
A. Zabavin
A. Novikova
I. Parkhomenko
E. Pokhvashchev**

Summary. The article presents the relevance of the problem of measuring the volumes of round timber transported by timber trucks, and considers a solution for automating the measurement process. Such mathematical methods for determining volumes based on determining the depth of points as triangulation, trifocal tensor and the Levenberg–Marquardt method were considered. On their basis, a software tool was developed and test calculations were carried out with an ideal load.

Keywords: triangulation, key point mapping, epipolar geometry, trifocal tensor, Levenberg–Marquardt method.

Вегера Денис Владимирович

старший преподаватель, ФГБОУ ВО «Тихоокеанский
государственный университет», г. Хабаровск
vegera79@mail.ru

Забавин Алексей Сергеевич

Преподаватель, ФГБОУ ВО «Тихоокеанский
государственный университет», г. Хабаровск
007913@pnu.edu.ru

Новикова Агата Алексеевна

Аспирант, ФГБОУ ВО «Тихоокеанский государственный
университет», г. Хабаровск,
agata2703@bk.ru

Пархоменко Игорь Сергеевич

Аспирант, ФГБОУ ВО «Тихоокеанский государственный
университет», г. Хабаровск,
011608@pnu.edu.ru

Похващев Егор Олегович

Ассистент, ФГБОУ ВО «Тихоокеанский государственный
университет», г. Хабаровск
79142102475@ua.ru

Аннотация. В статье представлена актуальность проблемы измерения объемов круглых лесоматериалов, перевозимых на лесовозах, и рассмотрено решение для автоматизации процесса измерения. Были рассмотрены такие математические методы определения объемов на основе определения глубины точек как триангуляция, трифокальный тензор и метод Левенберга–Марквардта. На их основе было разработано программное средство и проведены тестовые расчеты с идеальной нагрузкой.

Ключевые слова: триангуляция, сопоставление ключевых точек, эппиплярная геометрия, трифокальный тензор, метод Левенберга–Марквардта.

Введение

Проблема измерения объемов круглого леса, перевозимого на лесовозах, добавляет большие временные и логистические издержки для предприятий, занимающихся лесозаготовкой и лесопереработкой. Решение этой проблемы с помощью средств автоматизации определения количества вывозимого круглого леса позволяет снизить издержки и ускорить рабочие процессы, убрав неконтролируемые и субъективные факторы из процесса измерений.

На Дальнем Востоке богатый лесной массив вдоль ключевого участка от с. Лидога до д. п. Ванино трассы А-376 «Хабаровск–Лидога–Ванино–Комсомольск-на-Амуре» (далее — трасса) представляет интерес для коммерче-

ской деятельности по заготовке деловой древесины. В связи с этим, органы государственного лесного надзора и предприятия-лесозаготовители заинтересованы в разработке программно-аппаратного комплекса, устанавливаемого на пропускных пунктах трассы и позволяющего в автоматизированном режиме производить оценку объема в разрезе штабелей, находящийся на лесовозах, с учетом множества факторов.

На данный момент существует два основных способа замера объема груза, основанные на взвешивании грузовика и сканировании его кузова с помощью двух лидаров (продольного и поперечного). Как в одном, так и в другом случае измерения нужно производить два раза — на полном и пустом грузовике, чтобы сравнить изменение объема.

В данной статье будут рассмотрены математические методы определения объемов на основе определения глубины точек.

Триангуляция

Представление точки в трехмерном пространстве на двумерной плоскости изображено на рисунке 1.

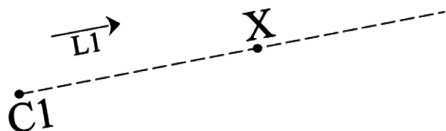


Рис. 1. Определение местоположения трехмерной точки

На рисунке 1 точки $C1$ и X находятся в трехмерном пространстве, а единичный вектор $L1$ показывает направление луча, идущего от точки $C1$ к точке X . Однако в данной модели невозможно решить уравнения, которые позволили бы нам определить глубину точки X . Поэтому модель нужно усовершенствовать [1].

Добавим ещё одну точку пространства — $C2$. Луч зрения, проведённый из этой точки, также проходит через точку X .

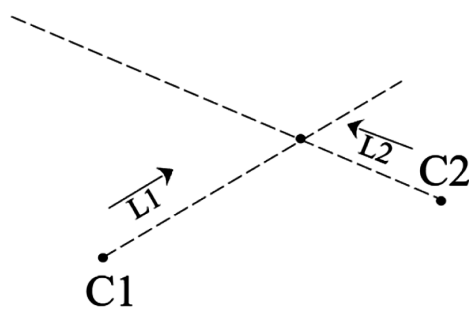


Рис. 2. Триангуляция

Адаптируем модель на использование двух камер и проекции объекта из 3D пространства на них.

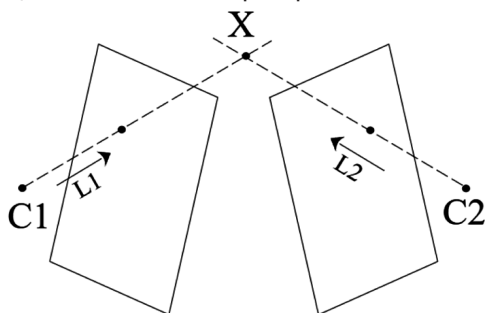


Рис. 3. Расширение концепции триангуляции

В представленной схеме (рисунок 3) отображены пространственные координаты $C1$ и $C2$, соответствующие положению левой и правой камер соответственно. Проекция трёхмерной точки X на плоскость изображения левой камеры обозначена как $x1$, тогда как проекция этой же точки на правую камеру обозначается как

$x2$. Эти две проекции, $x1$ и $x2$, классифицируются как соответствующие друг другу, поскольку они представляют собой результат преобразования одной и той же трёхмерной точки X .

На основе данных о положении $x1$ относительно $C1$ можно определить направление луча $L1$; аналогичным образом, зная положение $x2$ относительно $C2$, можно установить направление луча $L2$. Следовательно, используя метод триангуляции, можно точно определить координаты трёхмерной точки X в пространстве, как показано на схеме рисунка 2.

Из этого примера видно, что для проведения триангуляции трёхмерной точки на основе двух изображений, сделанных под разными углами, требуется следующее: расположение камер в реальной системе координат ($C1$ и $C2$), а также соответствие точек ($x1$ и $x2$) для каждой трёхмерной точки (X) в сцене, которая подлежит расчету [2, 3].

Главная трудность в оценке глубины на двумерных изображениях состоит в обнаружении и сравнении ключевых точек объекта ($x1$ и $x2$). Для решения данной задачи используется алгоритм ORB. Качество восстановления сцены напрямую связано с количеством совпавших точек: чем их меньше, тем менее точной будет полученная модель. Один из простых методов поиска соответствий основан на нахождении пикселей со схожими окружающими структурами. Тем не менее, наличие идентичных элементов или повторяющихся объектов на сцене может значительно усложнить эту задачу.

Для снижения числа ложных совпадений или ускорения поиска подходящих точек применяется эпиполярная геометрия — фундаментальный принцип стереоскопического зрения. Когда две камеры наблюдают трёхмерную сцену с различных ракурсов, существуют определенные геометрические взаимосвязи между 3D-точками и их проекциями на двумерные изображения, накладывающие ограничения на точки изображений. Эти соотношения основаны на допущении, что поведение камер можно моделировать как камеру-обскуру.

На рисунке 4 предполагается установка, схожая с той, что представлена на рисунке 3. Точка X в трёхмерном пространстве фиксируется камерами $C1$ и $C2$ в виде двумерных проекций $x1$ и $x2$ соответственно. Так как $x1$ является проекцией точки X , можно предположить, что если продлить луч $R1$ от камеры $C1$ через точку $x1$, он пройдет через саму точку X . Этот луч фиксируется как прямая $L2'$, а точка X отображается на изображении $i2$ как $x2$. Поскольку точка X находится на луче $R1$, то и её проекция $x2$ должна находиться на прямой $L2'$. Следовательно, возможные положения $x2$ ограничены

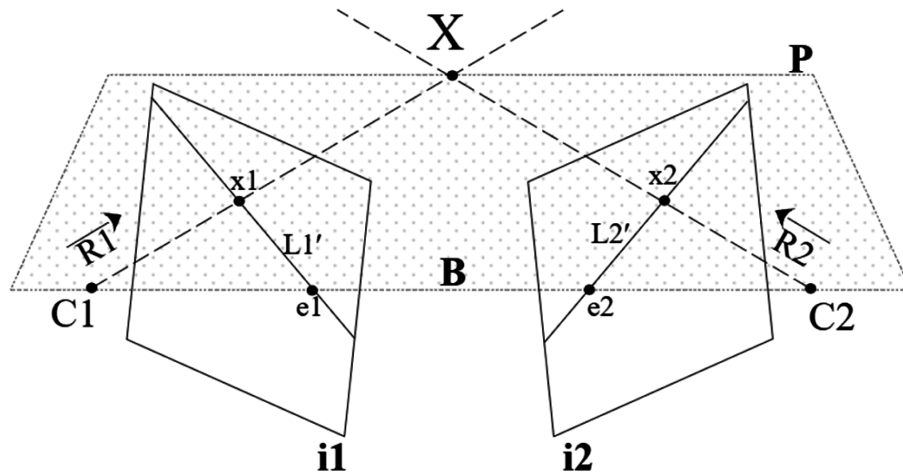


Рис. 4. Модель, описываемая эпполярной геометрией

этой линией, и таким образом, область поиска пикселей на изображении i_2 , соответствующих пикселям x_1 , сужается до единственной линии L_2' . Для нахождения L_2' используется эпполярная геометрия, что существенно уменьшает пространство поиска соответствий, описанное ранее [4].

Отметим, что:

- *эпполь* (e_1 и e_2) — изображение центра камеры одного ракурса в другом ракурсе
- *базовая линия* — линия, соединяющая центры двух камер (e_1 и e_2), то есть эпполь также можно определить как пересечение базовой линии с плоскостью изображения.
- *эпполярная плоскость* — плоскость P , которая содержит X , C_1 , x_1 , x_2 и C_2 , определенная R_1 и базовой линией.
- *эпполярная линия* — линия (L_2), полученная от пересечения эпполярной плоскости (P) и плоскости изображения.
- *эпполярное ограничение* — ограничение на поиск соответствующего пикселя сходя из эпполярной геометрии. Пространство поиска для пикселя в изображении (i_2), соответствующего пикселю (x_1), ограничено одной двумерной линией, которая является эпполярной линией (L_2) этого изображения.

Для разных значений X у нас будут разные эпполярные плоскости и, следовательно, разные эпполярные линии. Однако все эпполярные плоскости пересекаются на базовой линии, и все эпполярные линии пересекаются в эпполе. Все эти зависимости в совокупности и называются Эпполярной геометрией.

Эпполярные зависимости можно представить матрицей и для ее расчета достаточно двух изображений эпполярной системы. Такая матрица называется «фундаментальной».

Вычисление фундаментальной матрицы

На рисунке 3 предположим, что мы знаем матрицы проекции камеры для обеих камер, скажем, P_1 для камеры в C_1 и P_2 для камеры в C_2 .

Матрица проекции камеры определяет соотношение между координатами трехмерного мира и соответствующими пиксельными координатами при захвате камерой.

Точно так же, как P_1 проецирует трехмерные мировые координаты в координаты изображения, мы определяем P_1^{inv} , псевдообратный к P_1 , так что мы можем определить луч R_1 из C_1 , проходящий через x_1 и X как:

$$X(k) = P_1^{inv} \times x_1 + kC_1$$

k — параметр масштабирования, так как мы не знаем фактического расстояния X от C_1 . Нам нужно найти эпполярную линию Ln_2 , чтобы уменьшить пространство поиска для пикселя в i_2 , соответствующего пикселю x_1 в i_1 , поскольку мы знаем, что Ln_2 — это изображение луча R_1 , захваченное в i_2 . Следовательно, чтобы вычислить Ln_2 , мы сначала находим две точки на луче R_1 , спроецируем их на изображение i_2 с помощью P_2 и используем проецируемые изображения двух точек, чтобы найти Ln_2 .

Первая точка, которую мы можем рассмотреть на R_1 , — это C_1 , так как луч начинается с этой точки. Вторую точку можно рассчитать, оставив $k=0$. Следовательно, мы получаем точки как C_1 и $(P_1^{inv})(x_1)$. Используя матрицу проекции P_2 , мы получаем координаты этих точек на изображении i_2 как $P_2 * P_1^{inv} * x_1$ соответственно. Мы также заметили, что $P_2 * C_1$ — это в основном эпполь e_2 на изображении i_2 .

Прямая может быть определена в проективной геометрии с использованием двух точек p_1 и p_2 , просто найдя их векторное произведение $p_1 \times p_2$. Следовательно

$$Ln2 = P2 * C1 * P2 * P1inv * x1$$

$$e2 = P2 * C1$$

при

$$Ln2 = e2 \times P2 * P1inv * x1$$

Для фундаментальной матрицы

$$F = e2 \times P2 * P1inv$$

получим

$$Ln2 = F * x1$$

В проективной геометрии, если точка x лежит на прямой L , мы можем записать ее в виде уравнения

$$xT \times L = 0$$

Следовательно, поскольку $x2$ лежит на эпполярной прямой $Ln2$, получаем

$$x2T \times Ln2 = 0$$

Заменяя значение $Ln2$ из приведенного выше уравнения, мы получаем уравнение:

$$x2T \times F \times x1 = 0$$

Это необходимое условие для того, чтобы две точки $x1$ и $x2$ были соответствующими точками, и это также форма эпполярного ограничения. Таким образом, F представляет собой общую эпполярную геометрию системы с двумя точками зрения [5, 6].

Трифокальный тензор

Трифокальный тензор T — это тензор валентности 3 T_{jk}^i , с двумя контравариантными и одним ковариантным индексом. Он представляется однородным массивом $3 \times 3 \times 3$ (то есть 27 элементов) и имеет 18 степеней свободы.

Трифокальный тензор имеет аналогичную роль для вычисления глубины точки при захвате (проецировании) в трех разных видах (изображениях), что и фундаментальная матрица для двух видов. Он инкапсулирует все (проективные) геометрические отношения между тремя видами, которые не зависят от структуры сцены [7].

Тензор может использоваться для переноса точек или линий из соответствия двух видов в третий.

Трифокальный тензор имеет три принципиальных геометрических свойства:

- это гомографии между двумя видами, вызванные плоскостью, проецируемой обратно из линии в другом виде;
- отношения между соответствиями изображений для точек и линий, возникающих из инцидентов в 3-пространстве;
- извлечение фундаментальной и матриц камеры из этого тензора.

Тензорная система счисления

Точки и линии изображения представлены однородными столбцовыми и строчными 3-векторами соответственно, то есть $x = (x_1, x_2, x_3)^T$ и $l = (l_1, l_2, l_3)$. Элемент матрицы A с индексами ij обозначается как a_{ij} , где i — контравариантный индекс (индекс строки), а j — ковариантный (индекс столбца).

Определение трифокального тензора, используя тензорную нотацию

$$T_{jk}^i = a_j^i b_k^4 - a_k^4 b_j^i$$

Положение индексов в T_{jk}^i (два контраварианта и один ковариант) диктуются положением индексов справа от уравнения. Таким образом, трифокальный тензор является смешанным контравариантом-ковариантным тензором. В тензорной записи основное соотношение инцидентности принимает вид

$$l_i = l^j l^k T_{jk}^i$$

Краткое изложение трехфокусных тензорных отношений инцидентности — трехлинейностей

1. Соответствие линия — линия — линия

$$(l_r \epsilon^{ris}) l_j' l_k'' T_i^{jk} = 0^s$$

2. Соответствие точка — линия — линия

$$x^i l_j' l_k'' T_i^{jk} = 0$$

3. Соответствие точка — линия — точка

$$x^i l_j' (x''^k \epsilon_{kqs}) T_i^{jq} = 0_s$$

4. Соответствие точка — точка — линия

$$x^i (x'^j \epsilon_{jpr}) l_k'' T_i^{pk} = 0_r$$

5. Соответствие точка — точка — точка

$$x^i (x'^j \epsilon_{jpr}) (x''^k \epsilon_{kqs}) T_i^{pq} = 0_{rs}$$

Где x^i, x'^j и x''^k — точки, l_r, l_j', l_k'' — линии, ϵ^{ris} — тензор, T — трифокальный тензор, $\epsilon_{jpr}, \epsilon_{kqs}$ — контравариантный тензор [8, 9].

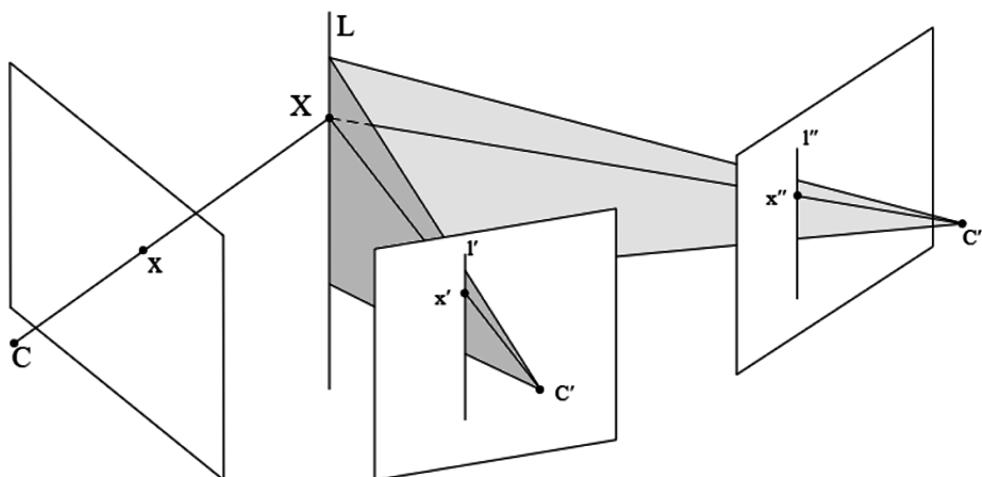


Рис. 5. Пример отношения инцидентности точка — линия — линия

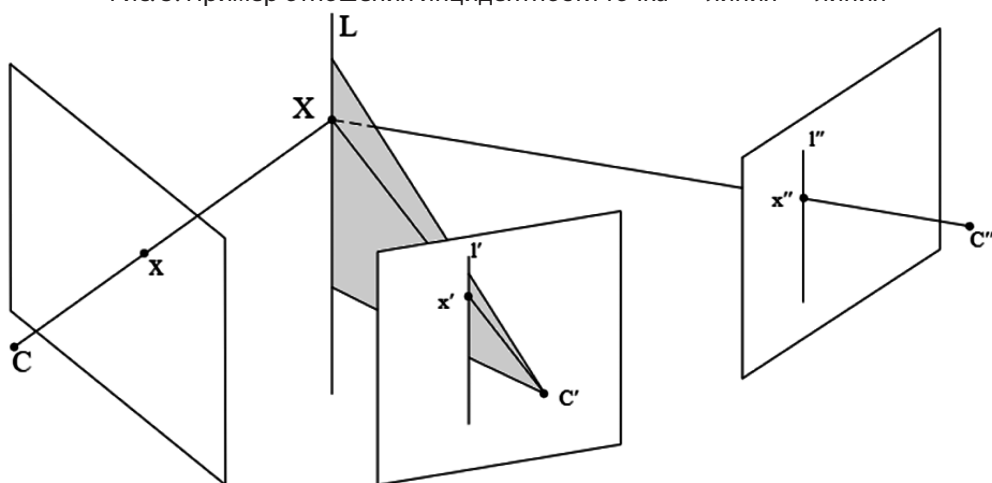


Рис. 6. Пример отношения инцидентности точка — линия — точка

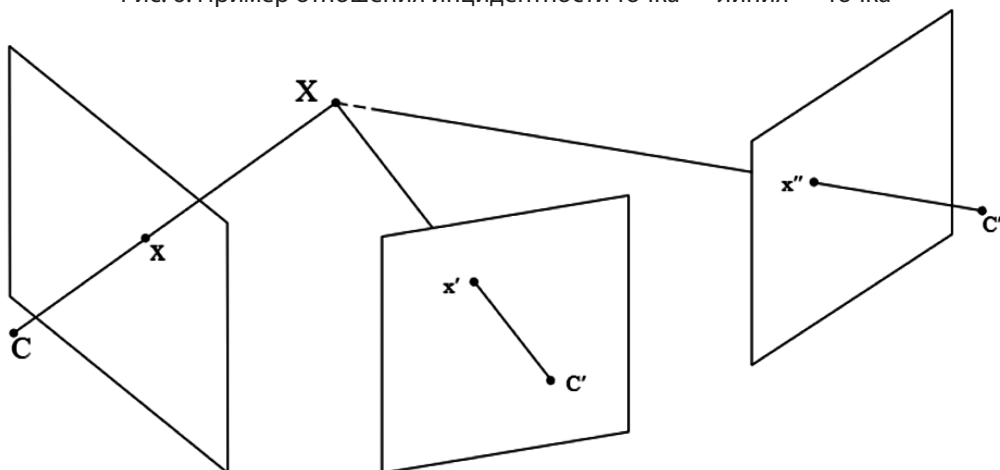


Рис. 7. Пример отношения инцидентности точка — точка — точка

Метод Левенберга–Марквардта

Матрица $H_k = f''(x^k)$ может не быть положительно определенной, даже при x^k сколь угодно близком к множеству решений. Поэтому глобализованный алгоритм должен включать в себя механизм выбора подходящего H_k , причем допускающий базовый выбор вблизи решений.

Алгоритм

Выбираем параметры $\rho_1, \rho_2 > 0, \tau_1 > 0, \tau_2 > 1, \sigma' > 0, q > 0, \varepsilon, \theta \in (0, 1)$. Выбираем $x^0 \in R^n$ и полагаем $k = 0$.

1. Если $f'(x^k) = 0$, стоп.

2. Полагаем $H_k = f''(x^k)$.

Если

$$\|H_k f'(x^k)\| \geq \rho_1 \|f'(x^k)\|^{\tau_1},$$

вычисляем p^k как решение уравнения при $\sigma_k = \min \{\sigma, \|f'(x^k)\|^q\}$. Иначе переходим к шагу 4.

3. Если

$$\{f'(x^k), p^k\} \leq -\rho_2 \|p^k\|^{\tau_2},$$

переходим к шагу 5.

4. Вычисляем симметричную $n \times n$ матрицу H_k и соответствующее решение p^k задачи, удовлетворяющие шагу 1 и 2.

5. Вычисляем $\alpha_k = \theta^j$, где j — наименьшее неотрицательное целое число, при котором выполняется неравенство Армихо

$$f(x^k + \theta p^k) \leq f(x^k) + \varepsilon \theta^j \{f'(x^k), p^k\}.$$

6. Полагаем $x^{k+1} = x^k + \alpha^k p^k$, увеличиваем k на 1, и переходим к шагу 1.

Ключевой вопрос состоит в том как реализовать шаг 4. Очевидно, что шаг 6 может выполняться для любого $\rho_1 > 0$, если H_k достаточно положительно определена, а именно, ее минимальное собственное значение не меньше ρ_1 . Более того, из шагов 4 и 5 следует, что шаг 2 также достигается достаточной положительной определенностью H_k . Последнего же можно добиться посредством модифицированного разложения Холецкого, или посредством модифицированного симметричного зна-конеопределенного разложения матрицы Гессе $f''(x^k)$. При этом на каждой итерации алгоритма потребуются решить не более двух систем линейных уравнений, как и в случае, когда для второй системы берется просто $H_k = \rho I$ с достаточно большим $\rho > 0$.

Альтернативным образом можно выбирать H_k последовательно, заменяя H_k на $H_k + \omega I$ с некоторым $\omega > 0$, до тех пор, пока шаги 1 и 2 не окажутся выполнены. Разумеется, при таком подходе итерация алгоритма может потребовать решения более двух систем линейных урав-

нений, но получаемое в результате направление может быть лучше, так как выбранная в итоге матрица H_k может быть ближе к истинной матрице Гессе функции f [10].

Применение описанных методов к решению исходной задачи

По итогу изучения данных методов была разработана программа на языке Python с использованием библиотек OpenCV и NumPy. Были проведены тестовые расчеты с идеальной нагрузкой на 2 лесовозах: в 21,3 т., и в 33т. Было проведено по 25 тестов для каждого метода с каждым лесовозом.

В таблице представлены средние значения тоннажа и минимальные и максимальные значения тоннажа по методам.

Средние значения тоннажа

Идеальная нагрузка		Метод Левенберга–Марквардта	Метод эпиполярной геометрии	Метод трифокального тензора
21,30	Среднее значение	20,22	20,96	21,12
	Минимальное значение	19,08	20,41	20,85
	Максимальное значение	21,5	21,48	21,38
33,00	Среднее значение	31,55	32,57	32,71
	Минимальное значение	30	31,53	32,19
	Максимальное значение	33,32	33,27	33,2

Заключение

Результаты расчетов показывают, что использование метода, основанного на трифокальном тензоре, дают наиболее достоверные результаты. Для получения более точных результатов необходимо увеличить сложность алгоритма и использовать более энергозатратные вычислительные средства.

ЛИТЕРАТУРА

1. Barrett E.B., Brill M.H., Haag N.N., Payton P.M. Invariant linear methods in photogrammetry and model-matching. In: Mundy J.L., Zisserman A. (Eds.). *Geometric invariance in computer vision*. Cambridge: MIT Press, 1992.
2. Гульванский В. В., Миненко М. В. Программно-аппаратный комплекс для измерения объема сыпучих материалов // Интеллектуальный пункт пропуска в России и мире: компетентностный подход к созданию : Сборник докладов Всероссийской практической конференции, Санкт-Петербург, 10–11 февраля 2022 года. Санкт-Петербург: Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), 2022. С. 37–38. EDN RWZCCW.
3. Hartley R.I., Sturm P. Triangulation // *Computer Vision and Image Understanding*. 1997. Vol. 68, No. 2. P. 146–157.
4. Тушев С.А., Суховилов Б.М. Алгоритм полиномиальной сложности для поиска соответствующих точек на основе эппиполярной геометрии // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2018. Т. 7, № 4. С. 83–104. DOI: 10.14529/cmse180406.
5. Рощупкина С.Н., Асташенкова Л.К., Булгакова В.С., Кудринская К.В. Построение карты диспаратности по неоткалиброванной паре изображений // Молодой ученый. 2019. № 26 (264). С. 43–46. URL: <https://moluch.ru/archive/264/61141/> (дата обращения: 05.12.2024).
6. Faugeras O.D. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. Cambridge: MIT Press, 1993.
7. Yang Y., Jia W., Bucher T., Zhang H., Sun M. Image-Based food portion size estimation using a smartphone without a fiducial marker // *Public Health Nutrition*. 2019. Vol. 22, No. 7. P. 1180–1191. DOI: 10.1017/S136898001800054X.
8. Hartley R., Zisserman A. *Multiple View Geometry in Computer Vision*. 2nd ed. Cambridge: Cambridge University Press, 2000, 2003.
9. Rother C., Carlsson S. Linear multi view reconstruction and camera recovery. In: *Proceedings of the 8th International Conference on Computer Vision*, Vancouver, Canada, 2001. P. 1–42–49.
10. Измаилов А.Ф., Куренной А.С., Стецюк П.И. Метод Левенберга–Марквардта для задач безусловной оптимизации // Вестник российских университетов. Математика. 2019. Т. 24. С. 125:60–74.

© Вегера Денис Владимирович (vegera79@mail.ru); Забавин Алексей Сергеевич (007913@pnu.edu.ru);

Новикова Агата Алексеевна (agata2703@bk.ru); Пархоменко Игорь Сергеевич (011608@pnu.edu.ru); Похвощев Егор Олегович (79142102475@ya.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»