

Сравнение качества построения карты глубины при использовании различных калибровочных моделей сверхширокоугольных объективов

Пантелеев М.

Санкт-Петербургский Политехнический Университет Петра Великого
`panteleev.md@edu.spbstu.ru`

Аннотация Пишу здесь на всякий случай, потому что прошлый раз мой комментарий на сайте Вы, кажется, не увидели. Мы с Виктором Витальевичем исследовали модели камер в этом семестре и решили опубликовать про это статью, которую я и сюда мог сдать (ну и статья с первого семестра, в общем-то, в тупике). Но у меня возникли проблемы с некоторыми моделями, у него со свободным временем, ещё и дедлайны как-то раньше ожидаемого случились, в итоге эксперименты проведены, но написать пока не успели. Этот документ - мой ранний черновик этой статьи. Здесь есть проблемы, картинки надо будет менять, эксперимент в конце переделывать, но на данный момент это выглядит так.

Камеры "рыбий глаз" имеют большой потенциал для применения в робототехнике. Одним из таких применений может быть поиск глубины и построение карт для задач автономной навигации. Однако чтобы реализовать этот потенциал, должна быть использована точная модель камеры. В этой статье исследуются актуальные калибровочные модели сверхширокоугольных камер и их применимости в системах стереозрения. Приведены оценки точности построения карты глубины для каждой модели. Проведённые эксперименты позволили определить самую оптимальную по нескольким параметрам модель.

1 Введение

За последние годы был достигнут существенный прогресс в доступности и точности сенсоров, позволяющих мобильным роботам проводить оценку окружающего пространства. Такие информационно-измерительные устройства как лидары, сонары и стереокамеры стали основным источником информации для алгоритмов автономной навигации и локализации. Тем не менее в роботах по-прежнему присутствуют телевизионные системы, так как они дают наиболее легко воспринимаемую информацию для оператора в случаях, когда его вмешательство необходимо. "Обычные" камеры, описываемые перспективной проекцией, дают изображение, понятное для восприятия и обработки, но покрывают зачастую слишком маленькую область пространства. Однако такие изображения в меньшей степени пригодны для применения в алгоритмах автономной навигации и локализации в виду того,

что видят те или иные особенности окружающего пространства в среднем более короткие промежутки времени [3].

Для достижения большего поля зрения могут быть использованы катадиоптрические системы, состоящие из выгнутого зеркала и перспективной камеры. Однако такой метод не всегда применим, так как система получается громоздкой и имеет "мёртвую зону" посередине кадра, как видно на рисунке 1 (а). Наконец, можно использовать камеры "рыбий глаз" (англ. fisheye-camera), позволяющие с помощью специальной системы линз одним кадром покрыть угловое поле свыше 180° . В сравнении с катадиоптрическими они обладают большей полезной областью кадра, как видно на рисунке 1 (б).

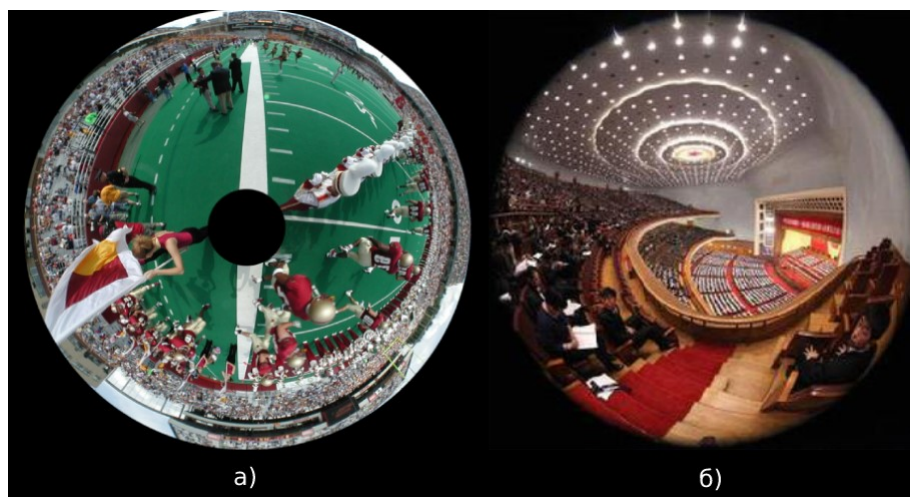


Рис. 1: Сравнение изображений с разных типов камер

При этом, как видно по изображению, сверхширокоугольные объективы накладывают на изображение заметные искажения. Обычно принято выделять два вида искажений: тангенциальные и радиальные, но в данном случае тангенциальные пренебрежимы по сравнению с радиальными и далее рассматриваться не будут. Устранение этих искажений является важной задачей, так как её решение позволяет нивелировать недостатки сверхширокоугольных объективов и применять их для чувствительных к точности передачи формы объектов в кадре задач.

Одной из таких задач является стереозрение. Несмотря на существование методов, позволяющих оценивать глубину по полным fisheye-снямкам [7], классические методы, требующие ректификации, всё ещё остаются самыми доступными и производительными [1]. Построив систему стереозрения на основе ортогонально расположенных сверхширокоугольных камер,

можно получить ряд преимуществ перед традиционными камерами. Для этого сначала нужно выбрать модель искажений, наиболее точно описывающую данные линзы.

Разработку и первоначальные испытания алгоритма стереозрения целесообразно проводить в виртуальной среде. Оптическая система была смоделирована в виртуальной среде с применением игрового движка Unity и плагина ZybrVR Dome Tools [9]. Благодаря этому появляется возможность моделировать оптические системы разной конфигурации и иметь точную информацию о геометрии в кадре.

1.1 Модели камер "рыбий глаз"

Сверхширокоугольные линзы изготавливают, закладывая разные виды проекций [2], но реальные линзы не всегда в точности соответствуют им, поэтому для более точного описания принято использовать модели дисторсии на основе других функций. Определим основные обозначения. Модель проекции для камеры это функция (обычно обозначаемая $\pi_c(\cdot)$), которая моделирует преобразование из точки трёхмерного пространства ($P = [x_c, y_c, z_c]^T$) в поле зрения камеры в точку на плоскости изображения ($p = [u, v]^T$). Единичная полусфера S с центром в точке O_c описывает поле зрения. На ней также лежит точка P_c , являющаяся результатом обратной проекции $\pi_c^{-1}(p)$. Угол θ является углом падения для рассматриваемой точки, а угол ϕ откладывается между положительным направлением оси u и $O_i p$.

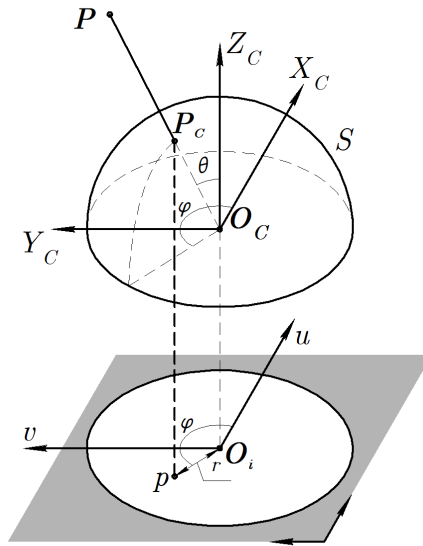


Рис. 2: Схема проекции точки трёхмерного пространства в точку на изображении

Основными критериями выбора моделей были: пригодность для моделирования сверхширокоугольных линз, простота калибровки и распространённость. Существует несколько программных пакетов, позволяющих автоматически калибровать камеры. Это kalibr (модели Канналы-Брандта, Скарамуззы, Двух Сфер), OdoCamCalib (модели Мея и Канналы-Брандта), OpenCV (модели Канналы-Брандта и Брауна) и MATLAB (модели Скарамуззы и Брауны). Таким образом, далее будут рассмотрены основные модели камер, отобранные для сравнения.

Модель Канналы-Брандта Модель Канналы и Брандта [5] реализована в частности OpenCV и описывает радиальные искажения через угол падения луча света на линзу, а не расстояние от центра изображения до места падения, как это делалось в более ранних моделях. Авторы посчитали, что для описания типичных искажений достаточно пяти членов полинома. Таким образом, указанную модель можно записать следующими уравнениями:

$$\begin{aligned}\pi(\mathbf{x}, \mathbf{i}) &= \begin{bmatrix} f_x d(\theta) \frac{x}{r} \\ f_y d(\theta) \frac{y}{r} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix} \\ r &= \sqrt{x^2 + y^2} \\ \theta &= \text{atan2}(y, x) \\ d(\theta) &= \theta + k_1\theta^3 + k_2\theta^5 + k_3\theta^7 + k_4\theta^9\end{aligned}\tag{1.1}$$

Калибровка камеры по этой модели произведена с помощью библиотеки OdoCamCalib, средняя ошибка перепроецирования составила 0.102 пикселя.

Модель Скарамуззы Также большое распространение получила модель Скарамуззы [8], которая легла в основу Matlab Omnidirectional Camera Calibration Toolbox, что позволяет быстро и удобно производить калибровку параметров модели. Она связывает точки на изображении с соответствующей им точкой в координатах камеры следующим образом:

$$\pi^{-1}(\mathbf{u}, \mathbf{i}) = \lambda \begin{pmatrix} u \\ v \\ a_0 + a_2 r^2 + a_3 r^3 + a_4 r^4 \end{pmatrix},\tag{1.2}$$

где r - расстояние от спроектированной точки до центра изображения; f - фокусное расстояние; λ - масштабный коэффициент.

Однако для устранения искажений необходима прямая проекция, которая не выражается в явном виде и находится методом Ньютона, что негативно сказывается на результате проекции.

Средняя ошибка перепроецирования составила 0.121 пикселей.

Модель Мея Модель Мея [6] изначально была создана для более эффективного моделирования катадиоптрических камер, но оказалась также весьма пригодной и для fisheye-камер.

$$\pi(\mathbf{x}, \mathbf{i}) = \begin{bmatrix} f_x \frac{x}{\alpha d + (1-\alpha)z} \\ f_y \frac{y}{\alpha d + (1-\alpha)z} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix} \quad (1.3)$$

Калибровка камеры по этой модели произведена с помощью библиотеки OdoCamCalib, средняя ошибка перепроецирования составила 0.099 пикселя.

Atan Эта модель представляет из себя идеальную эквидистантную fisheye-проекцию, которая заложена в использованную виртуальную широкоугольную камеру. Она добавлена к сравнению как эталонный способ устранения искажений из предположения, что с ней результаты стереосопоставления будут самыми лучшими.

$$\pi(\mathbf{x}, \mathbf{i}) = \begin{bmatrix} f_x \frac{\theta}{\sqrt{\frac{y^2}{x^2} + 1}} \\ f_y \frac{\theta}{\sqrt{\frac{x^2}{y^2} + 1}} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix} \quad (1.4)$$

Кроме этой проекции для сравнения в экспериментах присутствуют две идеальные камеры без искажений, направленные ровно в том же направлении, в котором будет производиться устранение искажений.

1.2 Построение стереопары

Наиболее сильные искажения на изображении "рыбий глаз" возникают по краям изображения, поэтому исследуемая виртуальная стереопара построена так, чтобы зона пересечения полей зрения камер располагалась на максимальном удалении от центра. Это соответствует ортогональному расположению камер с углом между принципиальными осями в 90° . При этом объективы камер имеют одинаковые параметры, поле зрения в 180° , а расстояние между ними составляет 20 см. Описанная конфигурация изображена на рисунке 3.

Здесь область пересечения полей зрения камер C_0 и C_1 обозначена красным. Эта область эквивалентна области пересечения полей зрения двух камер с полями зрения 90° (обозначены оранжевым), повернутых на $\beta_0 = \beta_1 = 45^\circ$ в сторону области интереса. Тогда B - база стереопары.

С помощью этой стереопары были получены снимки изображения шахматной доски, которые потом направлены в соответствующие программные пакеты для нахождения параметров каждой модели. Далее с помощью алгоритма [НИР] выполнялось устранение искажений для набора изображений, собранного виртуальными камерами. Пример входных и выходных снимков приведён на рисунке 4. По полученной паре изображений можно проводить стереосопоставление.

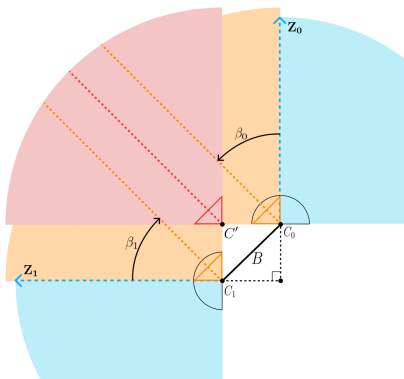


Рис. 3: Геометрическая модель бинокулярной системы стереозрения

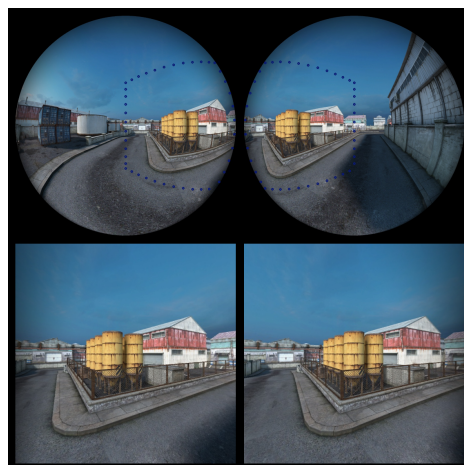


Рис. 4: Пример исходных изображений с отмеченной областью интереса и снимков виртуальной стереопары

1.3 Сравнение качества построения карт глубины

Оценка качества оценки глубины произведена на основе измерения среднеквадратичного отклонения и дисперсии полученных точек глубины от их предполагаемой позиции. Виртуальность эксперимента позволяет точно знать положение исследуемого объекта и, соответственно, точно определять ошибку. Кроме того, это позволяет изменять геометрические размеры исследуемого объекта в процессе исследования, чтобы его угловые размеры

в поле зрения камеры оставались постоянными. Таким образом возможно получить карты глубины примерно одинаковой плоскости на всём диапазоне дистанций. Эффективность стереосопоставления [4] сильно зависит от текстуры наблюдаемого объекта, поэтому для устранения влияния этого фактора эксперименты были проведены с различными текстурами. Примеры использованных текстур приведены на рисунке 5.

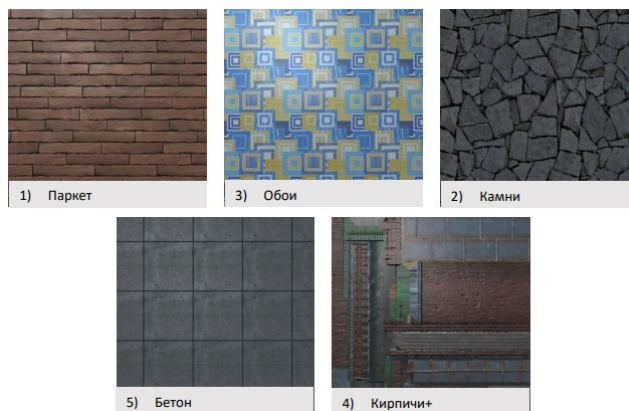


Рис. 5: Пример использованных текстур

Результаты оценки качества нахождения глубины приведены на рисунке 6. По оси абсцисс отложено расстояние до исследуемого объекта, по оси ординат среднеквадратичное отклонение. Каждый график соответствует своей модели.

Как можно заметить по графику, наилучший результат в оценке глубины логичным образом показывает эталонная стереопара. Далее следует идеальная модель, заложенная в виртуальную камеру, демонстрируя, что даже в случае идеального соответствии прямой и обратной проекции часть информации в изображении теряется или искажается. Из исследуемых моделей самую низкую ошибку демонстрирует модель Канналы-Брандта с результатом 0.06м на метр удаления.

Далее рассмотрены разностные изображения для всех моделей на выбранном расстоянии. Примеры таких изображений приведены на рисунке 7. Можно заметить, особенно на менее точных моделях, что различия проявляются не равномерно по всему изображению, а увеличиваются по мере приближения к краю оригинального изображения. Эта неравномерность может приводить к тому, что плоскости, параллельные плоскости снимка, на карте глубины будут изображаться с некоторым скосом.

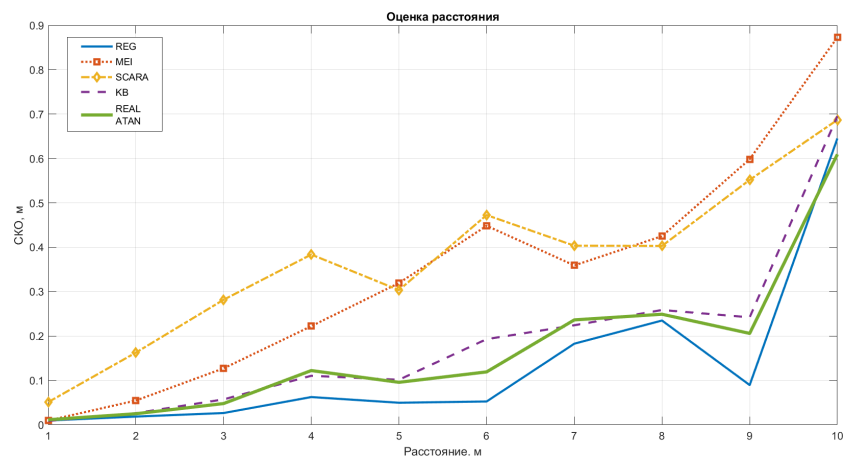


Рис. 6: Результаты оценки качества нахождения глубины

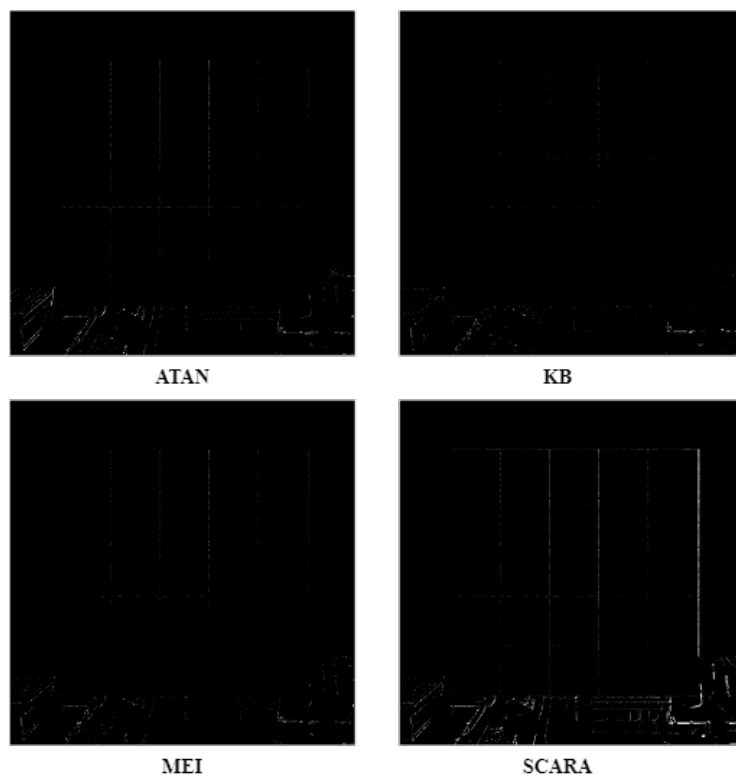


Рис. 7: Разностные изображения для моделей

Оценить описанный эффект можно аппроксимировав плоскость по полученному облаку точек и сравнив её нормаль с реальной. Аппроксимацию плоскости можно проводить методом RANSAC, а в качестве меры соответствия векторов нормалей использовать их скалярное произведение. Результаты исследования приведены на рисунке 8.

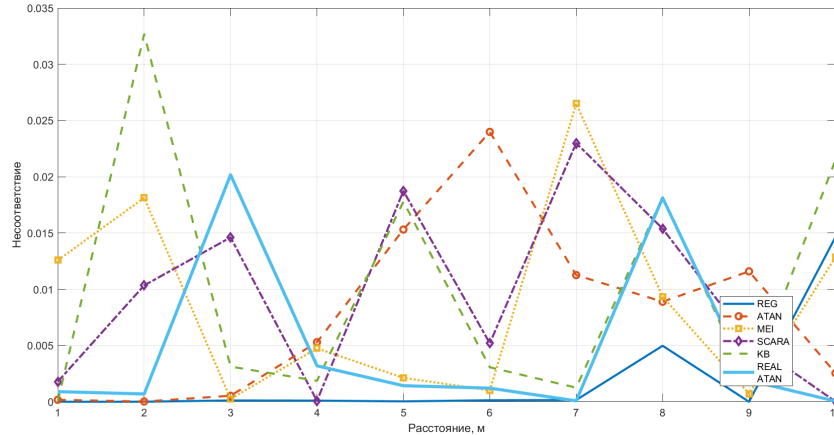


Рис. 8: Результаты исследования несоответствия плоскостей

Наиболее точные модели и тут показывают более хорошие результаты. Тем не менее, нельзя сказать, что непараллельность плоскостей систематически зависит от дистанции или модели.

2 Выводы

Проведённые исследования показали, что в применении к стереозрению наилучшие результаты демонстрирует модель Канналы и Брандта. Она демонстрирует наименьший прирост ошибки, который составляет в среднем 0.06м на метр удаления от камеры, что на 27% больше, чем для эталонной стереопары. Кроме того, модель является довольно доступной в плане автоматической калибровки и доступной в нескольких библиотеках технического зрения. Несмотря на меньшую точность, стереосистема, основанная на камерах со сверхширокоугольными объективами, может успешно применяться []. Полученные результаты мотивируют к дополнительным исследованиям в этой области.

Будущая работа может включать разработку модуля кругового стереозрения на основе четырёх сверхширокоугольных камер и тестирование разных моделей в системе локализации и построения карты на реальном роботе.

Список литературы

1. Brown, M., Burschka, D., Hager, G.: Advances in computational stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(8), 993–1008 (2003). <https://doi.org/10.1109/TPAMI.2003.1217603>
2. Clapa, J., Blasinski, H., Grabowski, K., Sekalski, P.: A fisheye distortion correction algorithm optimized for hardware implementations. pp. 415–419 (06 2014). <https://doi.org/10.1109/MIXDES.2014.6872232>
3. Herath, D., Kodagoda, S., Dissanayake, G.: Stereo Vision Based SLAM Issues and Solutions, p. 570 (06 2007). <https://doi.org/10.5772/5006>
4. Hirschmuller, H.: Accurate and efficient stereo processing by semi-global matching and mutual information. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). vol. 2, pp. 807–814 vol. 2 (2005). <https://doi.org/10.1109/CVPR.2005.56>
5. Kannala, J., Brandt, S.: A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. IEEE Transactions on Pattern Analysis and Machine Intelligence **28**(8), 1335–1340 (2006). <https://doi.org/10.1109/TPAMI.2006.153>
6. Mei, C., Rives, P.: Single view point omnidirectional camera calibration from planar grids. In: Proceedings 2007 IEEE International Conference on Robotics and Automation. pp. 3945–3950 (2007). <https://doi.org/10.1109/ROBOT.2007.364084>
7. Pan, B., Yu, L., Zhang, Q.: Review of single-camera stereo-digital image correlation techniques for full-field 3d shape and deformation measurement. Science China Technological Sciences **61**, 1–19 (09 2017). <https://doi.org/10.1007/s11431-017-9090-x>
8. 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 (ICVS'06). pp. 45–45 (2006). <https://doi.org/10.1109/ICVS.2006.3>
9. VR, Z.: A powerful, open-source plugin for the unity game engine that allows anyone to convert their content to the fulldome planetarium format. (2021), <https://zubr.co/work/dome-tools/>