ISSN 2658-5014 (Print) СИИТ ISSN 2686-7044 (Online)
2025. Т. 7, № 2 (20). С. 3–29 http://siit.ugatu.su
СИСТЕМНАЯ ИНЖЕНЕРИЯ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
Научный обзор

УДК 004.65 DOI 10.54708/2658-5014-SIIT-2025-no2-p3 EDN AGHGVU

TRANSPORT-BY-THROWING — РОБОТОТЕХНИЧЕСКИЙ СПОСОБ ПЕРЕМЕЩЕНИЯ ПРЕДМЕТОВ ПЕРЕБРОСОМ: ЭКСПЕРИМЕНТЫ ПО НАБЛЮДЕНИЮ ЗА ТРАЕКТОРИЕЙ ОБЪЕКТА

К. В. Миронов

Аннотация. Перемещение предметов перебросом — это перспективный способ робототехнической транспортировки деталей в гибких производственных системах. Данный обзор посвящен исследованию точности наблюдения. Описанные эксперименты направлены на то, чтобы определить, как ошибка измерения повлияет на предсказание и как это влияние можно уменьшить. Для прогнозирования на основе обучения требуется набор примеров измерений траектории, этот набор формируется во время экспериментальных бросков. Приводится техническая информация об устройстве для метания, камерах и алгоритмах, которые позволяют получать трехмерную модель траектории на основе измерений камеры. На этой основе базируется исследование точности. Обсуждается точность позиционирования бросаемого мяча на одной паре изображений.

Ключевые слова: робототехническая система; транспортировка; переброс; захват; статистическая оценка баллистических кривых; трекинг движущихся объектов; алгоритм ближайших соседей.

Введение

Этот обзор является продолжением работ [Мир24а и Мир24б] и посвящён исследованию точности наблюдения за бросаемым предметом. Конечная оценка алгоритма прогнозирования полета тела производится на основе экспериментов по захвату бросаемых предметов. Процент успешных захватов показывает эффективность использования подхода в целом, но ничего не говорит о качестве отдельных частей системы: наблюдателя, предсказателя, планировщика путей и т. д. Эта информация может быть получена при численном моделировании процесса, если она более или менее точно отражает процесс, или путем сравнения результатов с имеющимися достоверными данными, то есть данными, которые измеряются с высокой точностью и не зависят от результатов прогнозирования. Таким образом, возникает необходимость информации о процессе полета и его наблюдении.

Описанные ниже эксперименты направлены на то, чтобы определить, как ошибка измерения повлияет на предсказание и как это влияние можно уменьшить. Само предсказание пока не рассматривается. Еще одним важным аспектом является описание наборов данных траекторий. Для прогнозирования на основе обучения требуется набор примеров измерений траектории. Этот набор приобретается во время экспериментальных бросков, рассматриваемых в этой работе.

Эксперименты по броскам описаны в разделе 1. Этот раздел в основном содержит техническую информацию об устройстве для метания, камерах и алгоритмах, которые позволяют получать трехмерную модель траектории на основе измерений камеры. Эта информация является основой для исследования точности в следующих разделах. В разделе 2 обсуждается точность позиционирования неподвижного мяча на одной паре изображений. В разделе 3 модель точности расширена для отслеживания летящих шаров.

1 Эксперименты с бросками

Данные эксперименты преследуют две основные цели. Они предоставляют набор данных выборочных траекторий для обучения предсказателей и позволяют исследовать свойства наблюдаемых траекторий, что наиболее важно для точности наблюдателя. Эксперименты проводились на основе имеющейся установки в Институте компьютерных технологий Технического университета Вены. Установка состоит из метательного устройства и двух цифровых камер (IDS uEye UI-3370CP с разрешением 2048×2048), интегрированных в стереопару. Обе камеры и метательное устройство управляются с настольного персонального компьютера (ПК). Управление камерами и метательным устройством осуществляется через универсальную последовательную шину (USB) и интерфейс RS-232 соответственно. Камеры запускаются одним генератором и синхронизируются. Броски были выполнены устройством, описанным в подразделе 1.1. Подразделы 1.2 и 1.3 описывают, как система визуального наблюдения отслеживает полет и как из изображений камеры извлекается ссылка на 3D-положение объекта. Подраздел 1.4 содержит обзор полученных траекторий выборки.

1.1 Бросок

Метательное устройство было сконструировано Мартином Понграцем и Гилбертом Маркумом [Pon24, Mar24]. Его фотография показана на рис. 1. Энергия броска обеспечивается натяжением двух пружин, которые тянут чашу катапульты с мячом. Натяжение обеспечивается электромеханическим приводом, который соединен с чашкой электромагнитным магнитом. Бросок происходит, если контроллер переключается на соленоид. Он освобождает чашу катапульты. Двигатель может регулировать положение запуска для достижения необходимых значений натяжения.

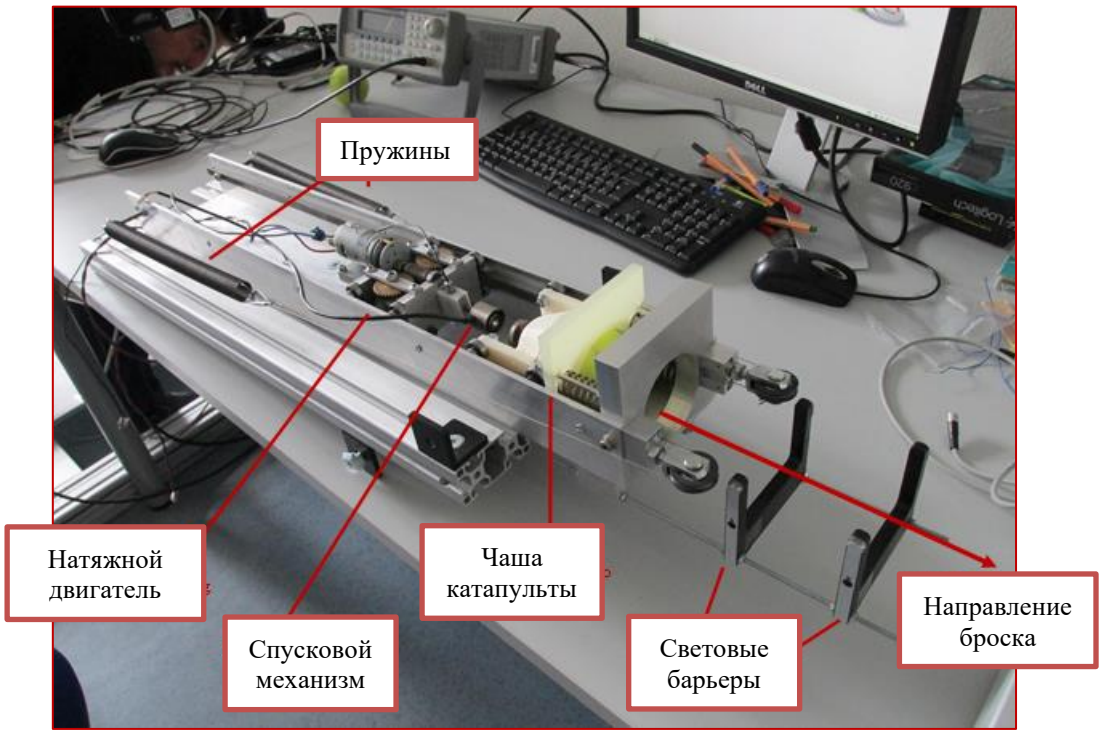


Рис. 1 Метательное устройство, используемое в экспериментах

Чтобы исследовать траекторию, полезно знать, какова была начальная скорость брошенного мяча. Если вектор скорости броска объекта каждый раз один и тот же, траектория объекта не будет меняться со временем в соответствии с моделями движения [Мир24б, уравнения (10) и (8)]. Реальное бросающее устройство отклоняется в скорости броска, поэтому фактическая

траектория может время от времени меняться. Изменение измеренной скорости запуска может дать представление об изменении траектории. Также этот вариант может быть использован в качестве основы для моделирования отклонений траектории в рамках моделирования процесса.

Начальная скорость, очевидно, зависит от натяжения пружины. В момент отпуска потенциальная энергия натянутой пружины E_p преобразуется в кинетическую энергию летящего объекта E_k. Взаимосвязь между энергией, натяжением пружины и скоростью объекта суммируется в уравнениях (1) и (2), которые следуют из простой школьной физики:

$$E_p = \frac{kd^2}{2}, \tag{1}$$

$$E_p = \frac{mv^2}{2}. \tag{2}$$

Здесь k – коэффициент восстановления; d – линейное удлинение пружины; F – сила натяжения; m – масса объекта; v – значение скорости объекта. Сила натяжения равна произведению линейного растяжения и коэффициента восстановления в соответствии с законом Гука:

$$F = kd. \tag{3}$$

Комбинация уравнений (1) и (3) соотношения энергии пружины и силы натяжения дает

$$E_p = \frac{F^2}{2k}. \tag{4}$$

Поскольку значения E_p и E_k равны, в соответствии с законом сохранения энергии известное значение силы позволяет оценить скорость броска:

$$v = \frac{F}{\sqrt{km}}. \tag{5}$$

В действительности значение v будет немного ниже из-за потерь энергии. В калибровочных экспериментах объект был брошен с определенной силой, и для измерения скорости запуска, соответствующей этой силе, использовались световые барьеры [Mar16]. В этих экспериментах была определена следующая эмпирическая зависимость между силой натяжения (в ньютонах) и скоростью запуска (в метрах в секунду) [Mar16]:

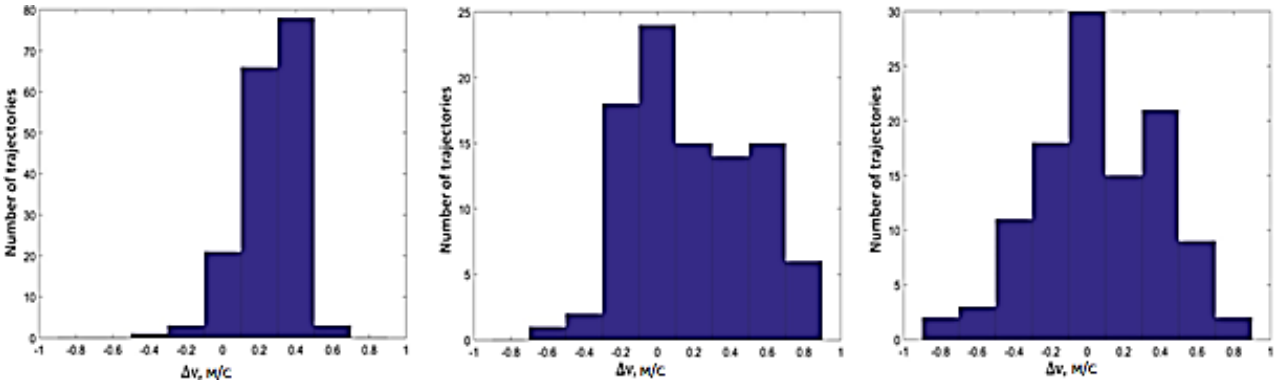
$$v = 0{,}1179\,F - 3.759. \tag{6}$$

Противоположное уравнение может быть использовано для определения силы натяжения по номинальной скорости [Mar16]:

$$F = 8.48\,v + 31.92. \tag{7}$$

Максимальная номинальная скорость броска установки (достигаемая при максимальном напряжении пружин) составляет около 10 м/с и соответствует 115 Н в соответствии с уравнением (7).

Два световых барьера расположены перед чашей в 10 см друг за другом. Они сигнализируют, когда объект проходит через них. Разница во времени между перехватами мяча световыми барьерами позволяет оценить фактическую скорость запуска объекта [Мир24б, уравнение (18)]. Сравнение значений скорости, измеренных системой световых барьеров, с номинальными значениями натяжения показывает, что отклонения в скорости запуска достигают 0.5 м/с, например, когда скорость номинальная была равна 4.5 м/с, фактические значения, измеренные световыми барьерами, были распределены в диапазоне от 4.2 до 5.1 м/с. Сравнение измеренной и номинальной скоростей броска, основанное на экспериментах по броскам, приведено на рис. 2, где представлены гистограммы, показывающие отклонение измеренной скорости от номинального значения, и в табл. 1, где приведены числовые значения для компонента систематической ошибки и стандартного отклонения.



Серия 1 *Серия 2* *Серия 3*

Рис. 2 Гистограмма отклонений броска для трех серий экспериментов

Таблица 1

Отклонения скоростей броска для трех серий экспериментов

Набор данных	Серия 1	Серия 2	Серия 3
Количество образцов, шт.	169	95	111
Систематическая ошибка, м/с	– 0.25	– 0.19	– 0.06
Стандартное отклонение, м/с	0.16	0.33	0.33

Эти результаты показывают, что и систематическая, и случайная составляющая ошибок достигает нескольких десятков сантиметров в секунду. Систематическая ошибка может быть связана с изменением упругих свойств пружины со временем и ошибками калибровки. Компонент систематической ошибки со временем становится меньше, но на самом деле это кажется случайным соответствием. Стандартное отклонение во второй и третьей сериях ухудшается. Причина может заключаться в том, что регулировка механических компонентов устройства со временем ухудшается (например, пружины со временем деформируются). Последний набор экспериментов, проведенных после создания всех наборов данных, показал, что, наконец, точность катапульты стала крайне низкой: фактическая скорость достигала более 6.5 м/с, в то время как номинальная была установлена на 5 м /с. Перед сбором каждого нового набора данных метательное устройство смазывалось и регулировалось. Обратите внимание, что эти данные показывают отклонение составляющей скорости в направлении, перпендикулярном плоскостям световых барьеров. Улучшение качества катапульты не является целью. Это, в частности, связано с тем, что при дальнейшем развитии системы ее можно было бы заменить робототехническим бросающим устройством. Эта информация полезна для понимания отклонений в доступных бросающих устройствах. Фронтальный световой барьер на этом устройстве используется для определения момента начала броска. Захват мяча с его помощью приводит в действие систему камер.

1.2 Наблюдение

Задача наблюдателя состоит в том, чтобы получить эталонные изображения мяча в движении и извлечь из этих изображений идеальные положения объекта в пространстве. На этапе обучения эти данные требуются в качестве образца траектории. При функционировании транспортной системы привязка 3D-координат в начале траектории используется в качестве входных данных для прогнозирования дальнейших значений. Процесс получения изображения входит в сферу действия этого подраздела, а извлечение 3D-координат рассматривается в следующем. Основными проблемами отслеживания являются необходимость обнаружения объекта на изображении и жесткие временные ограничения, например, текущий кадр должен быть обработан до поступления следующего.

Для наблюдения за полетом используются две монохромные камеры с разрешением 2048×2048 пикселей (IDS uEye UI-3370CP). Они подключаются к общему настольному персональному компьютеру (ПК) через USB-порт. В системе использовалось приложение для обработки изображений из [Goe15]. Обработка изображений размером 2048×2048 занимает некоторое время. Это ограничивает доступную частоту до 80 кадров в секунду [Goe15, с. 36]. Временные затраты обусловлены передачей данных по USB-кабелю и сложностью алгоритмов обработки изображений. Влияние обоих этих факторов тесно связано с размером изображения. Чтобы увеличить ограничения по частоте кадров, алгоритм обрабатывает только небольшую область необходимого (Area of Interest – AOI), а не все изображение.

В основном существуют два AOI: AIO камеры (Camera AOI – C-AIO) и алгоритм AOI (Algorithm AOI – A-AOI). C-AOI является частью датчика камеры, который используется для захвата изображения, а A-AOI – это часть полученного изображения, в которой выполняется поиск изображения мяча. Из-за технических ограничений камеры CAO I не может быть изменен во время отслеживания; его можно установить только до процедуры. Еще одно ограничение заключается в том, что границы C-API могут быть установлены только в одном измерении изображения. В другом случае он должен охватывать все 2048 пикселей. Размер C-API составляет 800×2048 пикселей [Goe15, с. 36]. Изображения с разрешением 800×2048 пикселей передаются с камеры на ПК с доступной частотой в 110 кадров в секунду.

A-AOI является частью C-API. Это подполье в изображении камеры, которое обрабатывается алгоритмом обнаружения окружности. Размер A-AOI составляет 300×300 пикселей [Goe15, с. 37]. Начальное положение A-AOI на изображении задано таким образом, чтобы оно закрывало область второго светового барьера, то есть мяч достигнет его сразу после начала броска. После обработки каждого полученного кадра положение AOI на изображении пересчитывается в соответствии с новым положением мяча. Такое иерархическое использование концепции AOI позволяет поддерживать частоту кадров на уровне 110 кадров в секунду. Иллюстрация C-API и движущегося A-AOI приведена на рис. 3.

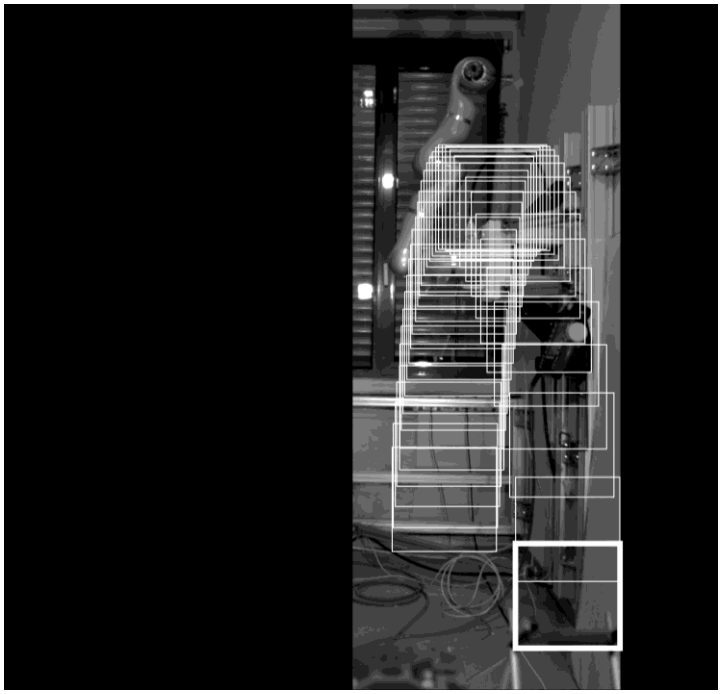


Рис. 3 Применение AOI

Черные области на рисунке соответствуют частям матрицы датчика камеры, которые не используются для получения изображения; прямоугольник, нарисованный жирными лини-

ями, показывает начальное положение A-AOI на изображении; он находится прямо над фронтальным световым барьером; прямоугольники, нарисованные тонкими линиями, показывают исходные положения AOI, пересчитанные во время полета, положение AOI на изображении сначала поднимается с восходящим движением шара, а затем опускается с нисходящим движением шара.

Новое положение A-AOI на изображении определяется на основе предыдущего. Для этого применяется быстрый алгоритм позиционирования мяча на изображении. Образец изображения сцены делается до броска. После получения нового кадра алгоритм быстрого отслеживания вычисляет разницу между последним изображением и той частью изображения образца, которая соответствует текущему A-AOI. Предполагается, что центр тяжести разностного изображения соответствует положению пикселя в центре мяча. Он вставляется в линейный фильтр Калмана, который определяет наиболее вероятное положение мяча на следующем изображении. Эта позиция установлена в качестве центра для нового A-AOI. В идеале изображение разницы равно внешнему виду объекта, но на самом деле это не так из-за мерцания системы освещения и влияния фоновых объектов. Различные объекты в сцене имеют разные значения яркости по сравнению с шаром. Когда мяч летит за объектами, имеющими аналогичный уровень яркости, работа алгоритма искажается. Алгоритм быстрого отслеживания достаточно точен, чтобы избежать выпадения мяча из AOI, но он не может предоставить точные данные о положении мяча на изображении.

Несмотря на это, вычитание фона является полезным этапом обработки изображений. Это уменьшает влияние фоновых объектов на точность позиционирования мяча. Изображение после вычитания фона показано на рис. 4.

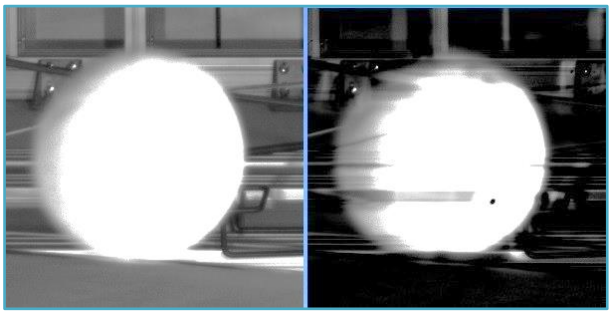


Рис. 4 Результаты вычитания фона

На рисунке исходное изображение находится слева, а изображение после вычитания – справа. Видно, что фоновые объекты все еще видны на изображении справа. Это эффект мигающей подсветки (снаружи изображения шара) и результат ее вычитания из монохромной поверхности шара (внутри изображения шара). Круглое изображение на самом деле представляет собой разницу между интенсивностью шара и фоном, покрытым шаром.

Точное позиционирование мяча на изображении осуществляется в два этапа. Прежде всего, алгоритм Кэнни обнаружения краев [Can86] используется для получения так называемого изображения края. Края – это точки на изображениях, где интенсивность (яркость пикселя) резко меняется. В простом изображении края каждый пиксель содержит ноль, если в точке нет края, или единицу, если в точке есть край.

В расширенных изображениях краев значение в пикселе края показывает направление края, которое должно быть нормальным в этой точке. Края могут соответствовать контуру объектов. В случае отслеживания мяча они показывают круг, соответствующий контуру мяча. Примеры изображений краев показаны на рис. 5. В первой строке результатов к входному изображению применяется функция обнаружения краев. Во втором ряду он применяется к разностному изображению. Можно видеть, что вычитание фона уменьшает количество краев

шума (то есть краев, которые не соответствуют контуру мяча), однако это добавляет шум в изображение внутри мяча.

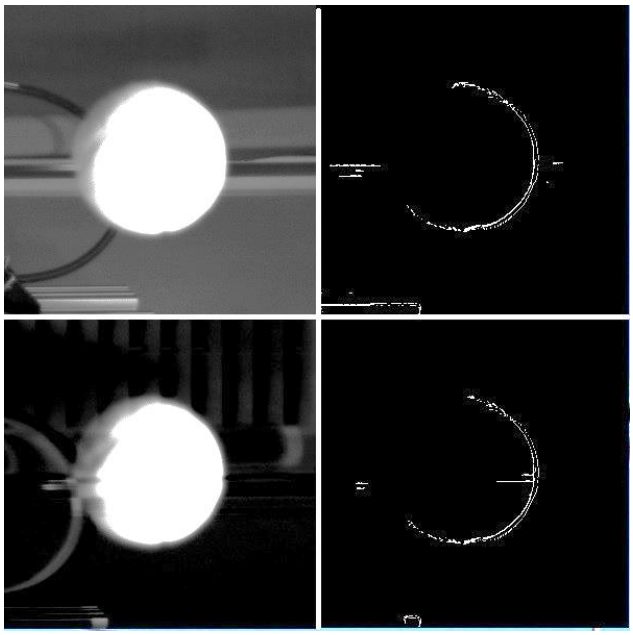


Рис. 5 Пример результатов определения границ
для одного и того же изображения
с (вверху) и без (внизу) вычитания фона

Второй шаг состоит в расположении центра шара на изображении края. Каждая точка ребра показывает возможную линию, которая может включать радиус окружности. Следовательно, трех точек идеально точного изображения края (то есть изображения края, которое точно представляет контур круга и ничего более) достаточно для определения положения центра круга на изображении. Реальное изображение края искажается краями фоновых объектов, текстурой мяча и тенями объекта. Специальные алгоритмы предназначены для извлечения центров окружностей из таких зашумленных изображений.

Преобразование Хафа, предложенное в [Hou62], является распространенным подходом для извлечения объектов из большого массива зашумленных данных (в частности, для извлечения фигур из изображений). При таком подходе требуемый объект (окружность) определяется минимальным набором параметров. Для окружности – это пиксельные координаты центра и радиуса

$$\left\{\begin{pmatrix} w_1(0) \\ w_2(0) \end{pmatrix}, r\right\}.$$

Эти три параметра формируют трехмерное пространство, которое называется аккумулятором [Hou62]. Аккумулятор дискретизирован, то есть он разделен на ячейки одного

$$\left\{\begin{pmatrix} w_1(1) \\ w_2(1) \end{pmatrix}, \begin{pmatrix} w_1(2) \\ w_2(2) \end{pmatrix}, \begin{pmatrix} w_1(3) \\ w_2(3) \end{pmatrix}\right\}$$

к кругу с конкретными значениями $\left\{\begin{pmatrix} w_1(0) \\ w_2(0) \end{pmatrix}, r\right\}$.

$$\left\{\begin{pmatrix} w_1(1) \\ w_2(1) \end{pmatrix}, \begin{pmatrix} w_1(2) \\ w_2(2) \end{pmatrix}, \begin{pmatrix} w_1(3) \\ w_2(3) \end{pmatrix}\right\} \rightarrow \left\{\begin{pmatrix} w_1(0) \\ w_2(0) \end{pmatrix}, r\right\} \quad (8)$$

Каждая возможная комбинация из трех очков голосует за определенную ячейку сетки. Это означает, что каждая ячейка имеет определенный рейтинг, который в начале равен нулю. Когда определенный набор

$$\left\{\begin{pmatrix} w_1(i) \\ w_2(i) \end{pmatrix}, \begin{pmatrix} w_1(j) \\ w_2(j) \end{pmatrix}, \begin{pmatrix} w_1(k) \\ w_2(k) \end{pmatrix}\right\} \rightarrow \left\{\begin{pmatrix} w_1(m) \\ w_2(m) \end{pmatrix}, r\right\},$$

номинальная мощность аккумуляторной батареи, включая точку

$$\left\{\begin{pmatrix} w_1(m) \\ w_2(m) \end{pmatrix}, r\right\},$$

увеличивается. После перебора всех возможных троек точек в центре ячейки с наибольшим рейтингом связывается с параметрами окружности. Если на изображении несколько кругов, выбираются ячейки с рейтингом выше указанного порогового значения.

Применение преобразования Хафа для отслеживания мяча, предложенное в [Kim75, Sca05] и примененное в [Bar09, Pon09], немного отличается от алгоритма, описанного выше. В нем используются три специфические особенности: предположение, что контур шара является единственным кругом на изображении или, по крайней мере, самым четким, тот факт, что значение r не требуется для дальнейшей обработки, и информация о нормали края, которая включена в изображение края. Каждая нормаль ребра соответствует одной возможной линии, на которой лежит центр шара. Накопитель выполнен в 2D, а не в 3D, а размер сетки ячеек равен размеру изображения края. На самом деле аккумулятор представляет собой изображение в оттенках серого. Значения пикселей накопителя в начале равны нулю. Затем все граничные пиксели обрабатываются итеративно. Для каждого пикселя края соответствующая нормаль края отображается на изображение аккумулятора. Пиксели изображения накопителя, которые покрыты нормалью, увеличивают свою интенсивность. После обработки всех краевых пикселей изображение сглаживается фильтром Гаусса. Это позволяет уменьшить влияние ошибок. Затем предполагается, что точка с наибольшей интенсивностью на изображении аккумулятора является центром мяча на изображении камеры. Пример последовательности соответствующего входного изображения, изображения-накопителя краевого изображения и проекции центра мяча для 2D-преобразования Хафа показан на рис. 6.

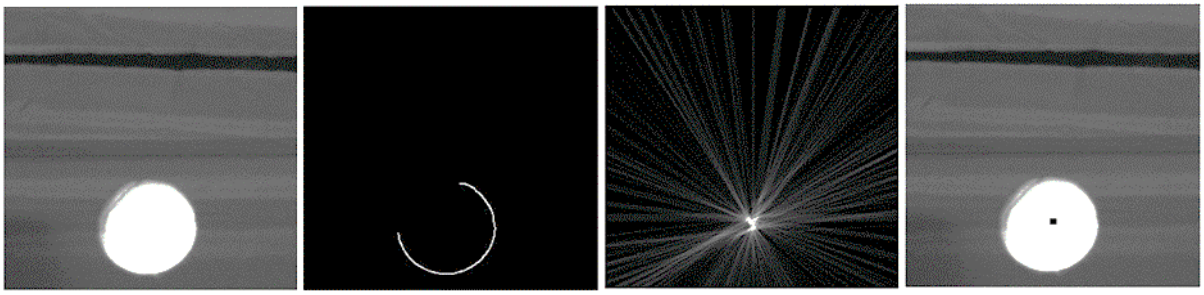


Рис. 6 Пример 2D распознавания кругов Хафа.
Слева направо показаны: базовое изображение, изображение края,
изображение аккумулятора и базовое изображение с центром проекции шара

Преимуществом 2D-обнаружения шара методом Хафа по сравнению с классическим кругом Хафа является меньший объем вычислений. Пространство накопителя также является 2D вместо 3D, и цикл включает обработку каждого краевого пикселя вместо каждого возможного триплета краевых пикселей. Но он менее надежен и не позволяет оценить радиус, хотя это можно сделать впоследствии. Из-за сложности сцены вместо 2D был применен метод 3D-Хаф.

RANSAC – это распространенный метод быстрой случайной подгонки в большом наборе данных. Метод был предложен в [Fis81]. Его применение для распознавания мячей также основано на сопоставлении, как и в уравнении (8). Здесь триплеты краевых точек выбираются

случайным образом, и гипотетический круг строится в соответствии с этим отображением. Все остальные точки ребер вписываются в этот круг. Если соответствие недостаточно, то описанную операцию повторяют до тех пор, пока не будет получен достаточный результат. Этот метод не гарантирует наилучшего соответствия, но на практике он обычно требует меньше вычислений, чем преобразование Хафа. Визуализация полного эталона распознавания мяча RANSAC приведена на рис. 7.

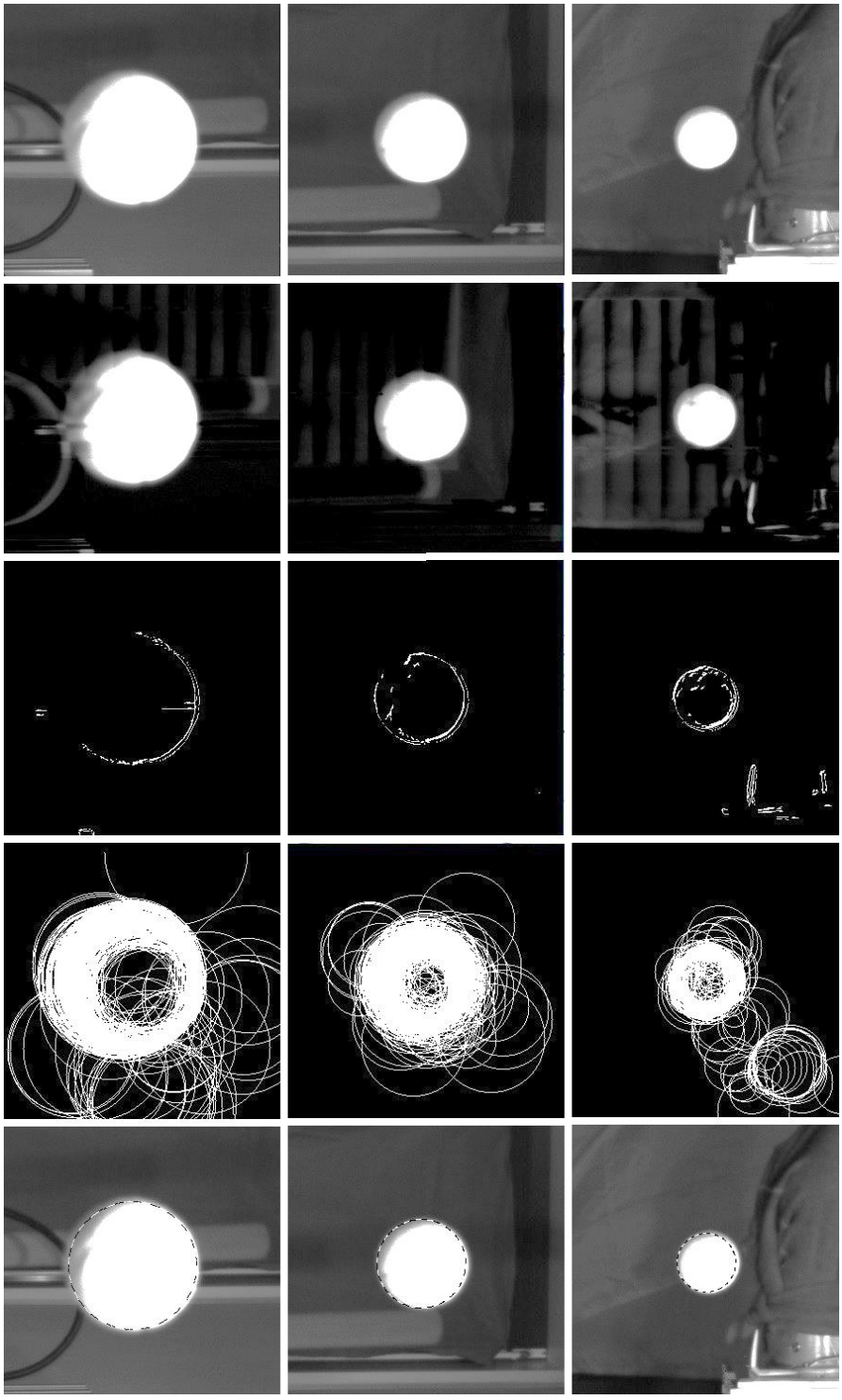


Рис. 7 Распознавание мячей RANSAC для трех образцов изображений

Каждый из трех столбцов представляет процесс извлечения центра из одного изображения: изображение, сделанное сразу после броска (левый столбец); изображение, сделанное в середине траектории (центральный столбец) и изображение, сделанное в области захвата (правый столбец). В первой строке представлены сами изображения, затем идут результаты вычитания фона (вторая строка), результаты обнаружения края (третья строка), набор гипотетических кругов, созданных RANSAC (четвертая строка). В последнем ряду выбранные круги проецируются на исходные изображения.

Как обнаружение краев, так и позиционирование центра шара (на основе преобразования Хафа или на основе RANSAC) являются дорогостоящими в вычислительном отношении, но их легко распараллелить; поэтому они реализованы в блоке графического процессора (Graphic Processing Unit – GPU) с использованием библиотеки C++ CUDA Гетцингера [Goe15]. В экспериментах, которые проверяют производительность, максимальное время выполнения последовательности операций обработки изображений (включая вычитание фона, обнаружение четких краев и обнаружение кругов Хафа или РАНСАКА) в большинстве случаев не исключает 10 мс [Goe15, с. 77–78]. Анализ характеристик точности алгоритма обработки изображений приведен далее в разделе 3.

1.3 Стереотриануляция

Когда координаты центра мяча извлечены из обоих изображений, применяется стереовидение для определения его положения в 3D. Операция определения пространственного положения объекта по его позициям пикселей на обоих изображениях называется стереотриангуляцией. В этой операции используются внутренние параметры камеры (фокусное расстояние, коэффициенты искажения и т.д.) и информация об относительном расположении камер (расстояние и углы поворота) в качестве необходимых входных данных для восстановления стерео. Значения этих параметров получаются до наблюдения во время процедуры калибровки. Стереотриангуляция – это обратная операция точечной проекции на изображении. Согласно модели камеры-обскуры из геометрической оптики, положение 3D-точки в плоскости изображения выражается следующей системой уравнений [Sze09, с. 59]:

$$u_1 = \frac{f}{x_3} x_1, \tag{9}$$

где f – фокусное расстояние камеры; $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ – вектор координат 3D-точки в системе координат камеры, а $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ – вектор плоских координат проекции точки на плоскость изображения (обратите внимание, что это не равно координатам точки в пикселях). В однородной системе координат это выражается следующим матричным уравнением [Sze09, с. 50]:

$$\begin{pmatrix} u_1 \\ u_2 \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} \frac{1}{x_3}. \tag{10}$$

Это означает, что если f и $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ точки известны, то могут быть найдены координаты проекции в плоскости изображения. Вычисление координат пикселей проекции точек производится с использованием приведенного ниже уравнения [Bou22].

$$\begin{pmatrix} u_1 \\ u_2 \\ 1 \end{pmatrix} = PFW \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{pmatrix}. \tag{11}$$

Здесь $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ является вектором координат точек в мировой трехмерной системе координат, P – матрица 3×3 для преобразования координат пикселей в геометрические координаты плоскости изображения. Значение P определяется линейным размером пиксельного элемента на датчике и коэффициентами, которые выражают радиальное искажение системы объективов камеры. F – матрица проекции 3×4, аналогичная матрице из уравнения (10), а W – матрица 4×4 для преобразования 3D-координат из системы, в которой они изначально определены, в систему, связанную с оптическим центром камеры. Эти матрицы определяются следующими параметрами [Bou22]:

– фокусное расстояние камеры f;

– положение начала координат пикселя в плоскости изображения $C = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$;

– линейный размер пиксельного элемента в датчике изображения $D = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$;

– вектор коэффициента, выражающий искажение системы объективов камеры $K = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \\ k_5 \end{pmatrix}$;

– угол α, выражающий наклон пиксельного элемента, если этот элемент является параллелограммом, а не прямоугольником.

Все эти параметры определяются с помощью процедуры калибровки камеры. Их влияние на матрицы преобразования здесь не рассматривается из-за его сложности. Следует только уточнить, что при знании этих параметров можно произвести расчет проекции точки на изображении [Bou22].

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, f, C, D, K, \alpha \right\} \rightarrow \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}. \quad (12)$$

Здесь $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = W$ представляет собой вектор координат пикселя проекции точки на изображении. Обратное отображение из известных $\{W, f, C, D, K, \alpha\}$ в $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ невозможно, так как конкретная точка на плоскости проекции соответствует бесконечному числу точек в трехмерном пространстве, лежащих на одном луче, которые начинаются от точки проекции и включают оптический центр камеры, как указано в [Мир24б, подраздел 3.4]. Если доступны координаты проекции точки на изображения с двух разных камер, то возможно следующее отображение [Bou22]:

$$\{W(1), f(1), C(1), D(1), K(1), \alpha(1), W(2), f(2), C(1), D(1), K(1), \alpha(2), T, O\} \rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \quad (13)$$

Здесь $T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$ – вектор перемещения, который выражает местоположение оптического центра второй камеры в системе координат, связанной с оптическим центром первой камеры; O – матрица 3×3, которая выражает вращение оптической оси второй камеры по сравнению с оптической осью первой камеры. Остальные элементы на левой стороне соответствуют соответствующим параметрам калибровки первой и второй камер. Значения T и O получаются во время калибровки стерео.

Диапазон в основном получается на основе уравнения несоответствия дальности [Liu06]:

$$x_3 = \frac{fb}{d}. \quad (14)$$

Здесь b – базовая линия системы камер, равная норме T; f – фокусное расстояние первой камеры; d – несоответствие между положениями точек на двух изображениях. Несоответствие в этом уравнении выражается в координатах плоскости изображения, то есть равно разнице между векторами точек в плоскостях двух изображений:

$$d = \sqrt{(u_1(1) - u_1(2))^2 + (u_2(1) - u_2(2))^2}. \quad (15)$$

Заметим, что

$$\begin{pmatrix} u_1(1) \\ u_2(1) \end{pmatrix}$$

в этом уравнении выражается в системе координат, связанной с первой камерой, а

$$\begin{pmatrix} u_1(2) \\ u_2(2) \end{pmatrix}$$

выражается в системе координат, связанной со второй камерой.

Калибровка системы камеры (то есть получение значений $W(1)$, $f(1)$, $C(1)$, $D(1)$, $K(1)$, $\alpha(1)$, $W(2)$, $f(2)$, $C(1)$, $D(1)$, $K(1)$, $\alpha(2)$, T, O) может быть выполнена в MATLAB с использованием набора инструментов для калибровки камеры Bouguet [Bou22] или калибратора камеры Mathworks [Cam24]. Приложение Mathworks работает быстрее и проще в использовании. Недостатком калибратора является то, что код вычислений скрыт. Исходный код бесплатный, поэтому он послужил основой для функции C++, реализующей стереотриангуляцию. Калибровка для экспериментов была выполнена с помощью калибратора Mathworks, а результаты были использованы в качестве параметров настройки для функции триангуляции C++.

1.4 Полученные наборы данных

Были проведены несколько отдельных серий экспериментов по метанию, чтобы получить траектории, изучить их свойства и возможности предсказателя. На основе проведенных экспериментов были собраны три основных набора данных. Эти три серии уже упоминались (см. табл. 1), поскольку они позволяли оценить точность бросающего устройства. Первый набор данных со 169 траекториями был получен с помощью камер, расположенных напротив устройства (рис. 8, *а*).

Номинальная скорость броска была 4.5 м/с, а измеренные значения варьировались от 4.3 до 5.1 м/с. Дальнейшее изучение точности показало, что ошибки позиционирования являются критическими, если камеры находятся далеко от объекта (подраздел 3.2). Поэтому в другой установке было проведено еще несколько экспериментов. Камеры были расположены таким образом, чтобы бросающее устройство находилось между ними (рис. 8, *б*). Оптические оси камер расположены почти горизонтально и почти параллельно друг другу.

Второй набор данных, основанный на второй установке, состоял из 95 траекторий. Траектории были записаны с несколькими различными скоростями запуска (4, 4.25, 4.5, 4.75 и 5 м/с). Предполагалось, что в экспериментах по захвату использовалась скорость 4.5 м/с, а остальные значения были только для обучения. Диапазон траекторий в обучающем наборе должен быть шире, чем в реальных условиях, чтобы избежать эффекта границы. Общий диапазон измеренных скоростей запуска составляет от 3.7 до 5.3 мс.

Третий набор данных был получен с той же настройкой. Для обработки изображений был применен более продвинутый алгоритм. Если во второй настройке изображения обрабатывались на основе алгоритма 2D преобразования Хафа, то второе изображение настройки обрабатывалось с помощью быстрой и более точной реализации GPU из работы [Goe15]. Исследование точности наблюдателя было проведено с использованием второго набора данных, но результаты этого исследования были затем заменены результатами из третьего набора данных. В последующих разделах третий набор данных рассмотрен в первую очередь.

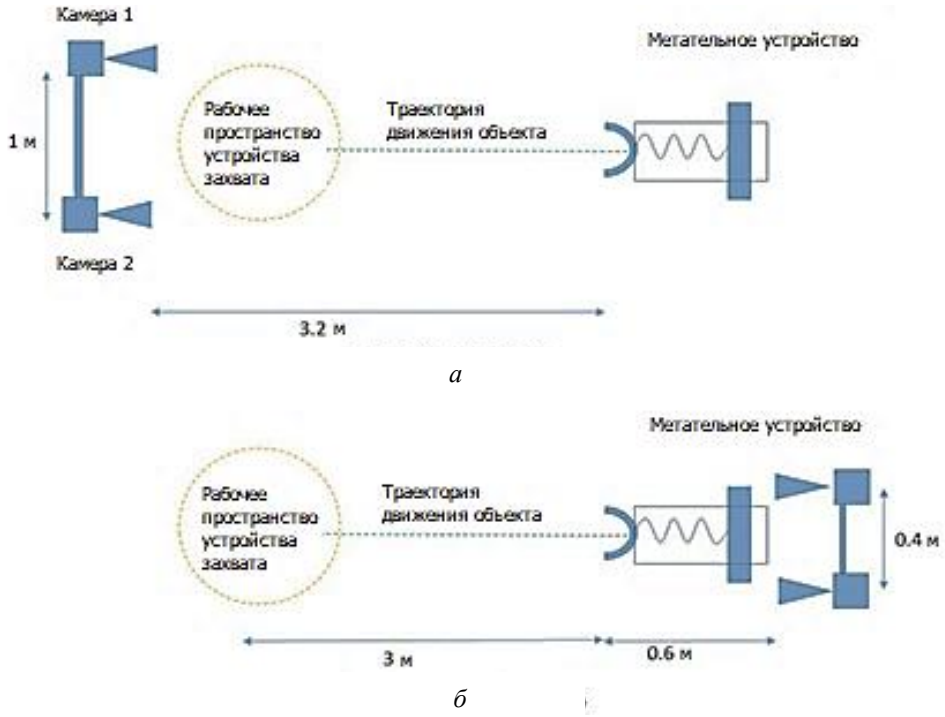


Рис. 8 Относительное расположение камер и метательного устройства

Количество траекторий, полученных для этого набора с различными номинальными скоростями, приведено в табл. 2.

Таблица 2

Параметры полученных наборов данных

Номинальная скорость броска, м/с	Количество попыток
4.00	22
4.25	22
4.50	22
4.75	23
5.00	22
Итого	111

Еще один набор данных был получен, чтобы определить, как движение робота в поле зрения влияет на захват траектории. Результаты этого эксперимента кратко показаны в начале раздела 3.

Представление траектории в базе данных состоит из нескольких блоков данных. Основной блок является эталоном координат объекта, измеренных в системе координат камеры. Это матрица, где 100 строк (каждая строка соответствует одному полученному кадру) и 3 столбца (каждый столбец соответствует одному пространственному измерению в системе координат, связанной с камерой). Если в кадре не было измерений координат (например, из-за того, что мяч в этот момент находился вне поля зрения), значения в этой строке были равны нулю. Дальнейшее отслеживание мяча после 100-го кадра не велось, поскольку в этих случаях мяч в тот момент уже находился вне рабочего пространства робота. На некоторых траекториях с низкими скоростями броска он уже отскочил от пола. Помимо привязки координат, в базе данных хранятся некоторые дополнительные параметры каждой траектории. Они в основном связаны

с предварительной обработкой траектории, например положения C-AOI и A-AOI в плоскости изображения, временные метки, координаты пикселей центра шара и т. д. Эти параметры также включают номинальные и измеренные значения скорости запуска.

2 Точность позиционирования неподвижного объекта

В этом и следующем разделе исследуется точность извлечения координат 3D-шара из изображений. Эта информация имеет решающее значение для дальнейшего развития алгоритма прогнозирования. Теперь эта точность будет проверена в ситуации, когда мяч неподвижен в пространстве. Раздел 3 расширяет модель точности до мяча в движении. Причина исследования статической точности заключается в том, что неподвижный шар может быть очень точно расположен в пространстве. Измеренное положение шара можно сравнить с его фактическим положением, установленным закрепленным оборудованием. Это сравнение позволяет оценить точность, но оно не учитывает конкретные факторы, влияющие на положение движущихся объектов, например, размытие при движении, ошибки при вычитании фона и т. д. Анализ этих факторов приведен в следующем разделе.

Исследование [Lee02] определяет следующие три источника ошибок в стереопозиционировании:

1) Ошибки калибровки (Calibration Errors – CE): они связаны с неправильной оценкой внутренних и внешних параметров во время калибровки. Например, если линейный размер опорного объекта, используемого при калибровке (размер сетки шахматной доски, используемой при калибровке [Bou22, Cam24]), измерен неточно, это может исказить линейные расстояния в реконструированной сцене.

2) Ошибки квантования (Quantization Errors – QE). Один пиксель соответствует определенной области изображения. Невозможно точно определить, где на самом деле находится точка в этой области. При триангуляции, предполагается, что положение точки соответствует центру пикселя. Разница между центрами пикселей показывает изменение координат точек.

3) Ошибки обработки изображений (Image Processing Errors – IPE). В [Lee02] ошибки в стереосогласовании, то есть в процедуре определения пар пикселей на левой и правой сторонах изображений, относятся к этой группе. Как упоминалось ранее, в алгоритме нет конкретного шага для сопоставления стереосигналов. Положение пикселей центра шара на обоих изображениях определяется с помощью экстрактора кругов Хафа или RANSAC. Алгоритмы обнаружения краев и выделения окружностей не обеспечивают 100 % точного расположения пикселей в центрах. Поэтому IPE считаются ошибками в этих алгоритмах.

Для определения точности системы видения были проведены ряд экспериментов [Pon15]. Целью эксперимента было создание среды, позволяющей сравнивать координаты объекта, измеренные системой камер, с информацией об их реальных значениях. Это обеспечивается mounting оборудованием, которое позволило точно расположить объект в пространстве.

Основание для крепления представляет собой горизонтальную плоскую пластину, изготовленную из стекла, и наклеенную на миллиметровую бумагу. Сферические объекты устанавливаются на подставках на определенной высоте (рис. 9, *а*). Эти опоры позволяют определить истинное положение сферы в вертикальном измерении. Точность определения истинности в двух горизонтальных измерениях обеспечивается размещением опор на миллиметровой бумаге (рис. 9, *б*). Небольшие трещины в основании опор позволяют регулировать его с помощью миллиметровой сетки. Миллиметровая пластина располагалась под областью полета объекта. Таким образом, настройка позволяет определить точность позиционирования объекта в зоне полета (рис. 10). Положения камер были аналогичны положениям, показанным на рис. 8, *б*.

Исходная точка системы координат стереокамеры (Stereo Camera Coordinate – SCC) была помещена в центр базовой линии. Перед оценкой точности позиционирования мяча было определено относительное расположение оборудования в системе координат камеры.



а *б*

Рис. 9 Установка для точного позиционирования сфер в пространстве

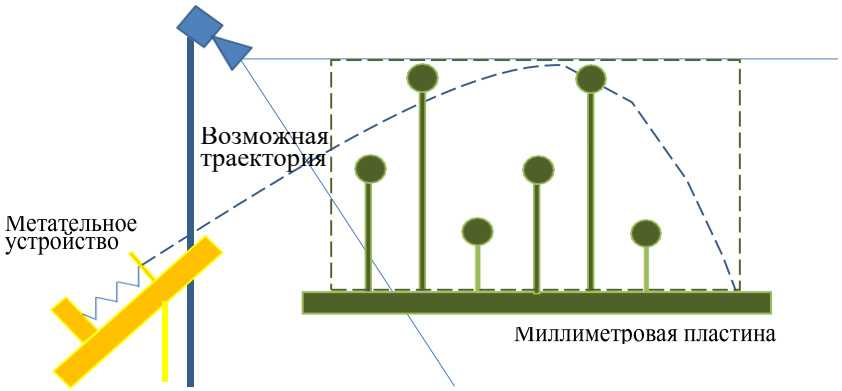


Рис. 10 Расположение сфер в области траектории наблюдения

Чтобы получить положение плоскости заземления с помощью миллиметровой бумаги в системе SCC, берется несколько точек на сетке. Положение пикселей на изображениях определяется вручную. Затем их SCC вычисляют методом стереотриангуляции. Плоскость, установленная в этих положениях, представляет положение миллиметровой пластины в SCC. Качество подгонки в первую очередь позволяет оценить влияние количественного смягчения. В общей сложности было выбрано 18 точек на расстоянии от 0.57 до 1.72 м от базовой линии. На самых дальних расстояниях узлы миллиметровой сетки не видны на изображениях, а на самом близком расстоянии нет точек, видимых обеим камерам. Расстояние от точек до установленной плоскости не превышает двух миллиметров, а стандартное отклонение достигает 0.7 мм.

Другими ошибками, извлеченными из этих данных, являются ошибки в оценке расстояний между интересующими точками с помощью стереотриангуляции. Точки были взяты на расстоянии 200 мм друг от друга для оценки плоскости. В первом вычислительном эксперименте было выявлено, что расстояние $\hat{d}$, оцененное методом стереотриангуляции, имеет систематическую составляющую ошибки по сравнению с реальным расстоянием d:

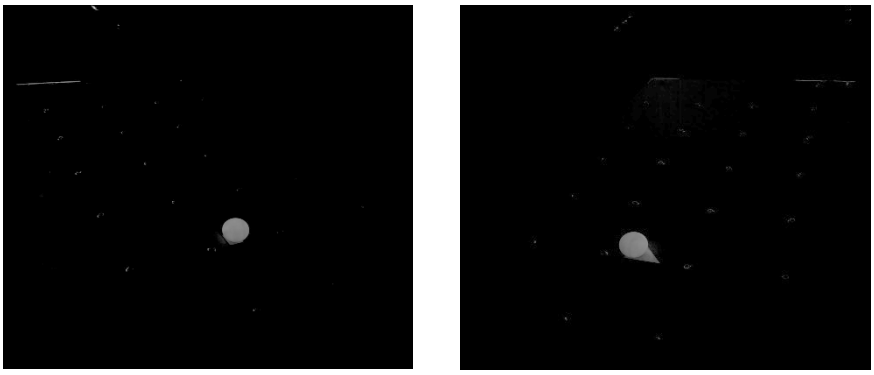
$$\hat{d} = 0.997d \ . \qquad (16)$$

Сферические объекты установлены на подставках с точной высотой (рис. 9, *а*). Пьедестал в правом нижнем углу рис. 9, *а* предназначен для самосветящейся сферы. Точное позиционирование в вертикальном измерении обеспечивается опорами, в то время как точное позиционирование в двух горизонтальных измерениях обеспечивается миллиметровой пластиной. На рис. 9, *б* показан пример точного позиционирования объекта в пространстве: подставка устанавливается в определенное положение на пластине. В других случаях предметы на различных пьедесталах располагались в разных положениях.

Другими словами, когда фактическое расстояние составляет 200 мм, триангуляция составляет 199.3 мм, для 600 мм – 597.9 и т. д. Анализ данных калибровки показал, что причина таких ошибок кроется в неправильном измерении размера сетки на калибровочной шахматной доске. Ошибка примерно в 50 микрон привела к этому искажению. После второго измерения шахматной доски с помощью высокоточного штангенциркуля при оценке расстояния не было обнаружено существенной составляющей систематической ошибки. Составляющая случайной ошибки достигает 0.5 мм, что, по-видимому, связано с QE.

После этого предварительного этапа были проведены эксперименты по позиционированию сферы. На первом этапе сферы освещались четырьмя галогенными прожекторами мощностью 500 Вт. Это были те же самые прожекторы, что использовались для освещения в экспериментах по метанию. Пример изображения, сделанного на этом этапе, показан на рис. 9. Стандартные отклонения положения шара с преобразованием Хафа достигли 3.6 мм. В основном погрешности составляли около 1.2 мм. Только в двух точках, которые лежат перед базовой линией, погрешность достигла значений более 5 мм. Возможно, что причиной такого большого значения могло быть неравномерное освещение, которое искажает результаты обнаружения краев и выделения окружности.

Чтобы определить влияние условий освещения, был проведен второй этап экспериментов. Здесь использовалась светящаяся сфера без внешнего освещения. Пример пары изображений, снятых в таких условиях, показан на рис. 11. Значение стандартного отклонения для светящихся сфер достигло 2.3 мм. За исключением набора внешнего освещения, вблизи базовой линии не было обнаружено резкого увеличения погрешности. Поскольку теннисный мяч не подсвечивается самостоятельно, необходимо использовать внешнее освещение. Чтобы избежать асимметрии освещения, рекомендуется, чтобы место для броска не располагалось ближе 0.6 метра к базовой линии, которая является расстоянием, на котором появляется асимметричное освещение, и использовать рассеянные источники света.



Левая камера Правая камера

Рис. 11 Изображения полученного положения светящейся сферы

Следовательно, расположение вычислительно извлеченных центров шаров является более ошибочным, чем расположение извлеченных вручную точек сетки. Это увеличение ошибок, очевидно, связано с ошибками обработки изображений на этапе выделения пиксельной координаты центра шара. Эти эксперименты игнорируют несколько факторов, например, размытие при движении, когда объект летит, ошибки вычитания фона при отслеживании движущегося

объекта и увеличение ошибок квантования на расстояниях более 1.72 м. Определить положение пикселей исходной точки на таких расстояниях не представлялось возможным. Изучение этих факторов и общей точности позиционирования летающих шаров приведено в разделе 3.

3 Точность определения местоположения летящего объекта

Эксперименты, описанные в предыдущем разделе, в основном касаются теоретического восприятия точности системы зрения. Далее будут проанализированы ошибки, обнаруженные в реальной ситуации. Данный анализ является сложной задачей, поскольку нет достоверных данных о реальном положении объекта в каждый момент времени. Большинство ошибок могут быть обнаружены, поскольку они искажают плавность кривой траектории. Эта кривая не может быть точно определена аналитически, но она плавная. Другой способ оценить ошибки – сопоставление точки данных с упрощенными моделями движения. Такое решение может быть не очень точным, поскольку эти модели не совсем точно передают движение объекта. Однако, если эти модели, как правило, более точны, чем система видения, качество подгонки может предоставить информацию о точности наблюдателя.

Ошибки отдельных этапов обработки могут быть обнаружены путем визуального анализа промежуточных изображений. Качество обнаружения краев может быть оценено путем сравнения изображения края с краями мяча на исходных изображениях. Качество извлечения кругов методами RANSAC или Хафа можно оценить, проецируя круги, которые были найдены на начальных изображениях. Например, визуальный анализ изображений с рис. 7 показывает, что алгоритм обнаружения границ имеет некоторый шум, но оценка RANSAC обеспечивает приемлемые результаты. Недостатком этого визуального анализа является то, что он выполняется человеком и не может предоставить объективную информацию. Тем не менее это позволяет обнаруживать некоторые очевидные ошибки отслеживания.

Еще один способ предоставления более точных данных для сравнения и проверки относится к RANSAC. Поскольку RANSAC не дает одинаковых результатов для разных запусков, несколько запусков дают возможность составить несколько предположений о центральном положении мяча. Правильная статистическая оценка, основанная на таких гипотезах, является более точной, чем результат одного прогона RANSAC. Результаты множественных измерений являются зашумленными и не поддерживаются моделью движения, основанной на истине, и предшествующими статистическими знаниями, например функцией плотности вероятности. Согласно [Kay93, Hla12, с. 11–12], в таких условиях следует использовать оценку наименьших квадратов. Оценка наименьших квадратов для статического параметра с неизвестным случайным шумом равна среднему значению результатов измерений [Hla12, с. 7]. В текущей оценке среднее значение заменяется медианой. Медианная и средняя оценка дают аналогичные результаты, но медиана более устойчива к отклонениям. При формировании базы данных использовалась медиана 1000 пробегов RANSAC, и любое дальнейшее увеличение числа пробегов не изменяет результаты оценки медианы. Использование такой оценки на этапе прогнозирования невозможно из-за большого объема вычислений. Графическому процессору, который способен выполнять извлечение одного круга RANSAC в режиме реального времени (то есть менее 9 мс для двух изображений и менее 1 секунды для всей траектории), потребовалось около 10 минут, чтобы запустить RANSAC 1000 раз.

Для движущихся объектов полная ошибка позиционирования может быть разделена на две составляющие. Полная ошибка e, указанная в миллиметрах или других единицах длины, представляет собой расстояние между расчетными координатами $\hat{X}(t)$ в момент времени t и реальным местоположением объекта в это время $X(t)$. Компонент пространственной ошибки e_s представляет собой расстояние между $\hat{X}(t)$ и $X(t+\tau)$, которое является ближайшей точкой к $\hat{X}(t)$ на реальной траектории объекта. Временная составляющая ошибки равна τ – это сдвиг во времени между t и моментом времени, когда реальное положение объекта ближе всего к $\hat{X}(t)$. Величина τ указывается в мс или других временных индексах. Временная ошибка также может быть определена как длина участка траектории между $X(t+\tau)$ и $X(t)$. Взаимосвязь между e_s, e

и τ проиллюстрирована на рис. 12. Разделение ошибок на временную и пространственную составляющие в основном важно для прогнозирования. Компонент временной ошибки менее важен для захвата, чем компонент пространственной ошибки.

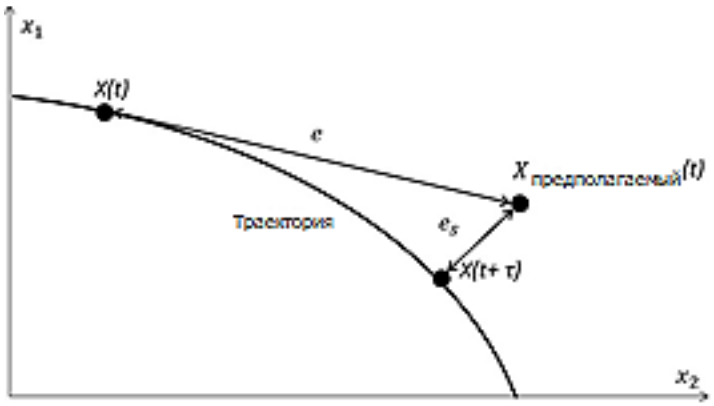


Рис. 12 Временные и пространственные составляющие ошибки позиционирования

Человеческому глазу на изображении не видно размытия при движении. Однако это не гарантирует, что размытие не повлияет на точность. Этот эффект возникает, когда объект значительно перемещается во время экспозиции. Время экспозиции камер в установке для метания было установлено равным 1 мс. Максимальная скорость мяча составляет 5 м/с. Это означает, что в момент экспозиции она достигает примерно 5 мм. Размытие происходит в основном из-за движения объекта в перпендикулярном направлении к плоскости изображения. Движение мяча вверх и вниз после броска является основным компонентом. Поскольку мяч брошен под углом $\frac{\pi}{3}$ радиана, скорость восходящего движения равна $v_a = v \sin\frac{\pi}{3} = 5\frac{\sqrt{3}}{2} = 4.33$ (м/с). Это значит, что шарик движется со скоростью 4 мм во время экспозиции в вертикальном направлении. Это намного меньше, чем размер шара, диаметр которого составляет более 65 мм [ITF15]. Поскольку размытие связано с движением объекта по траектории, влияние синего цвета движения является лишь частью временной составляющей ошибки. Это влияние ограничено временем экспозиции, которое в 10 раз меньше частоты кадров системы. Следовательно, специфическое влияние размытия при движении было проигнорировано, когда дело дошло до дальнейшего изучения ошибок.

Еще одним специфическим фактором, влияющим на качество отслеживания, является движение робота в зоне захвата. В четвертой серии экспериментов были получены ряд траекторий, когда робот двигался по зоне захвата. В этой ситуации вид робота на картинке не устраняется алгоритмом вычитания фона, и контур робота существует на краевом изображении. Испытания по отслеживанию с помощью движущегося робота показали, что конечный эффектор распознается как объект на большинстве изображений. Поэтому в текущей настройке отслеживание объекта на заключительном этапе траектории (после 70-го кадра) бесполезно, если робот движется. Следовательно, движение захвата, вдохновленное данными предсказателя, может не поддерживаться визуальной обратной связью, как только мяч достигнет зоны захвата.

Траектория пойманного мяча не используется для дополнительного обучения. Обучающие выборки обеспечиваются экспериментами по метанию без захвата.

3.1 Влияние вычитания фона

Анализ результатов отслеживания показал, что для правильного восстановления траектории необходимо вычитание фона. Если вычитание фона не применяется, отклонения измеренных значений при расстоянии более 1.5 м значительно возрастают. Эффект проиллюстрирован на рис. 13, где приведены графики для одной и той же траектории, извлеченные алгоритмом RANSAC с вычитанием фона и без него.

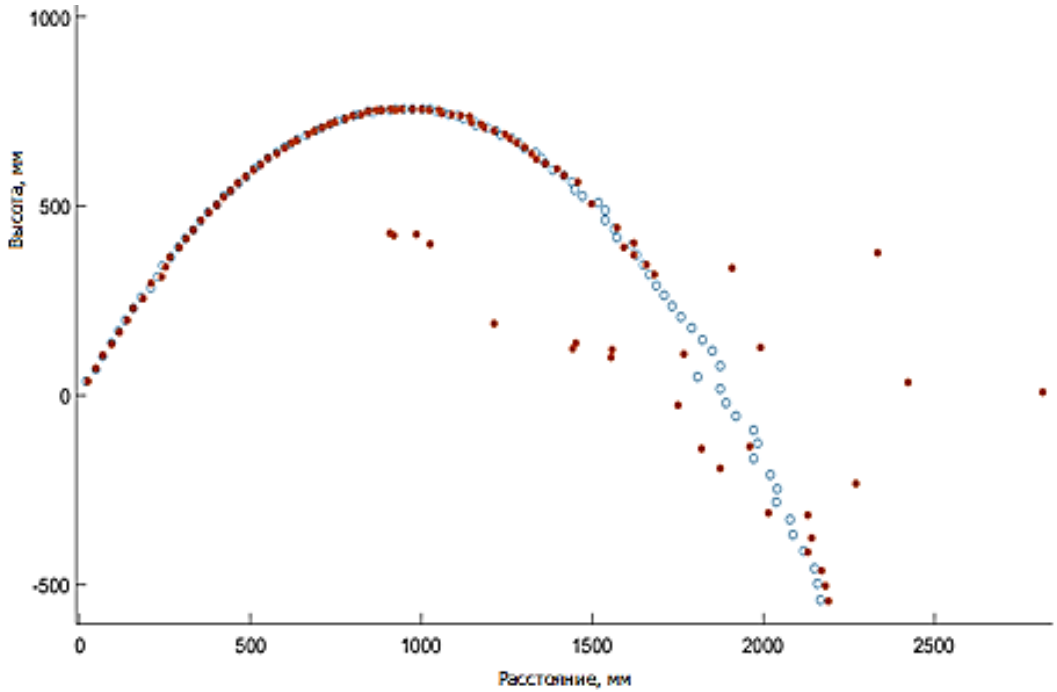


Рис. 13 График одной и той же траектории,
измеренной с учетом (синими кругами) и без учета (красные точки) фона

Видно, что на расстоянии примерно 1.5 метра измерения почти совпадают и выглядят как кривая второго порядка. Измерения с вычитанием фона (синие круги) сохраняют этот внешний вид, но внешний вид другого эталона (красные точки) становится хаотичным. Такое поведение типично для большинства траекторий в наборе данных. Численная оценка ошибок приведена в табл. 3.

Таблица 3

Сравнение различий между измеренными 3D-позициями, основанными на одном прогоне RANSAC, и медианой 1000 прогонов RANSAC для версий алгоритма, взятых с вычитанием фона и без него

Номер кадра	Стандартное отклонение без вычитания фона, мм	Стандартное отклонение с вычитанием фона, мм	Средняя погрешность без вычитания фона, мм	Средняя ошибка с вычитанием фона, мм
1...5	7.9	6.4	1.6	0.8
6...10	4.0	1.9	1.8	0.9
11...15	3.7	2.1	2.0	1.1
16...20	2.8	1.9	1.6	0.5
21...25	2.1	1.4	1.8	0.3
26...30	4.0	2.2	2.2	0.5
31...35	22.1	17.2	2.9	2.2
36...40	10.9	4.0	3.4	0.4
41...45	24.8	3.2	3.7	0.4
46...50	29.8	3.9	4.5	0.7
61...65	63.9	14.8	5.5	1.0
66...70	187.4	41.3	10.5	4.2
71...75	305.6	138.5	20.4	5.4
76...80	520.4	242.2	208.2	7.3
81...85	897.0	229.3	171.9	8.0
86...90	1361.6	197.5	163.9	8.4
91...95	1450.0	212.1	176.1	9.3
96...100	1272.6	106.63	146.8	10.1

Здесь координаты, извлеченные одним запуском RANSAC, сравниваются с результатами медианной оценки для 1000 запусков. Отклонения считаются ошибками. На самом деле эти цифры не приравниваются к ошибкам позиционирования, однако их можно использовать для отслеживания дисперсии измерений. На основе этих отклонений рассчитывается оценка стандартного отклонения для каждого кадра с момента броска мяча. В таблице кадры объединены в блоки по 5 кадров в целях экономии места. Стандартные отклонения суммируются на основе всех 111 траекторий, полученных в третьей серии экспериментов. Видно, что параметр начинает сильно расти после 65-го кадра, и этот рост гораздо более существенен для версии алгоритма без вычитания фона. Причина такой повышенной стабильности в начале заключается в том, что для начальных кадров размер шара больше и почти полностью покрывает изображение (необходимо сравнить первый и третий столбцы на рис. 7). Поэтому края фона вносят меньшие искажения в результаты обнаружения краев.

Видно, что даже для версии с вычитанием фона стандартное отклонение после 70-го кадра достигает очень высоких значений. Стандартное отклонение не лучший пример параметра, поскольку оно обладает низкой устойчивостью к отклонениям. Именно по этой причине статистики предпочитают отбрасывать отклонения из набора перед вычислением стандартного отклонения. Таким образом, в столбцах в правой части таблицы вместо этого приведены средние различия для одних и тех же блоков. Результаты медианы выглядят аналогично результатам для стандартных отклонений, но они более детализированы. Для алгоритма без вычитания фона средняя ошибка находится в интервале 3σ для неподвижных сфер, оцениваемая в 6.75 мм (см. подраздел 3.2), до 60-го кадра. Для алгоритма с вычитанием фона это свойство сохраняется до 80-го кадра. В версии без вычитания фона средняя разница более чем на 20 см достигается после 75-го кадра.

Это означает, что большинство кадров являются отклонениями в этой области. Таким образом, измерения, выполненные без вычитания фона, практически бесполезны.

Кроме того, плохое влияние вычитания фона возможно на начальной части траектории. Сразу после броска изображение мяча практически покрывает весь A-AOI. Из-за этого почти все фоновые объекты печатаются на изображении вычитания (рис. 14). Их края могут быть прочнее, чем края самого шара, и в этом случае они искажают процесс извлечения координат. Однако это искажение намного слабее, чем в предыдущем случае. Эффект прекращается после 4-го или 5-го кадра для измеренных траекторий. Проблему можно решить, применив вычитание фона только после 5-го кадра.

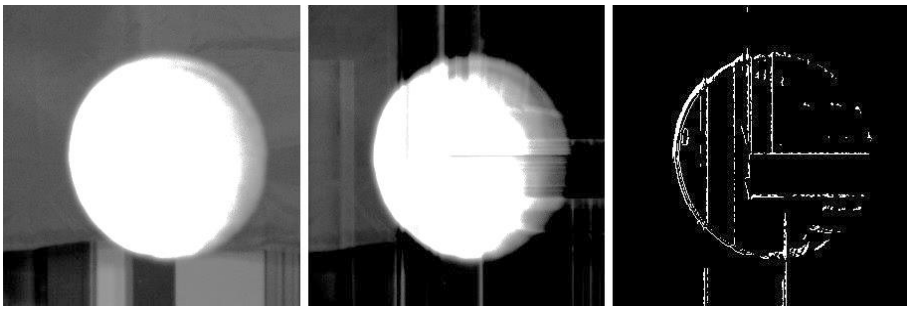


Рис. 14 Плохое влияние вычитания фона на начальном этапе полета:
шар практически покрывает все исходящее изображение, поэтому фоновые объекты печатаются на изображении вычитания, а их края видны на изображении края

3.2 Ошибки при измерении дальности больших расстояний

Значения стандартных отклонений и медианной разницы на заключительном участке траектории примерно после 65-го кадра выше, чем на предыдущем этапе. В этом подразделе проводится анализ источников этих ошибок и способов уменьшения их влияния. Измерения положения могут быть разделены на соответствующие значения и выбросы. Выбросы определяются как измерения, которые совершенно бесполезны, даже вредны для восстановления

траектории. Показатели могут быть ошибочными, но они помогают улучшить оценку. Очевидно, что невозможно со 100 % уверенностью решить, является ли измерение соответствующим или выбросом. Огромная разница между стандартным отклонением и средней ошибкой в конце траектории показывает, что выбросы оказывают значительное влияние. Задача обнаружения выбросов здесь пока не рассматривается, эта задача решается в рамках эталонных преобразований координат. В этом подразделе исследуются выбросы, измеренные с ошибками.

Построение траектории показывает специфическое свойство этих ошибок. На рис. 15 показаны три графика для примера траектории, например взаимосвязь между высотой объекта и расстоянием до камеры (верхний график), взаимосвязь между номером кадра и высотой (внизу слева) и взаимосвязь между номером кадра и расстоянием (внизу справа). Легко заметить, что первый и третий графики кажутся искаженными вправо, а привязка высоты сохранила вид гладкой кривой второго порядка. Другими словами, ошибки в основном локализуются в измерении диапазона.

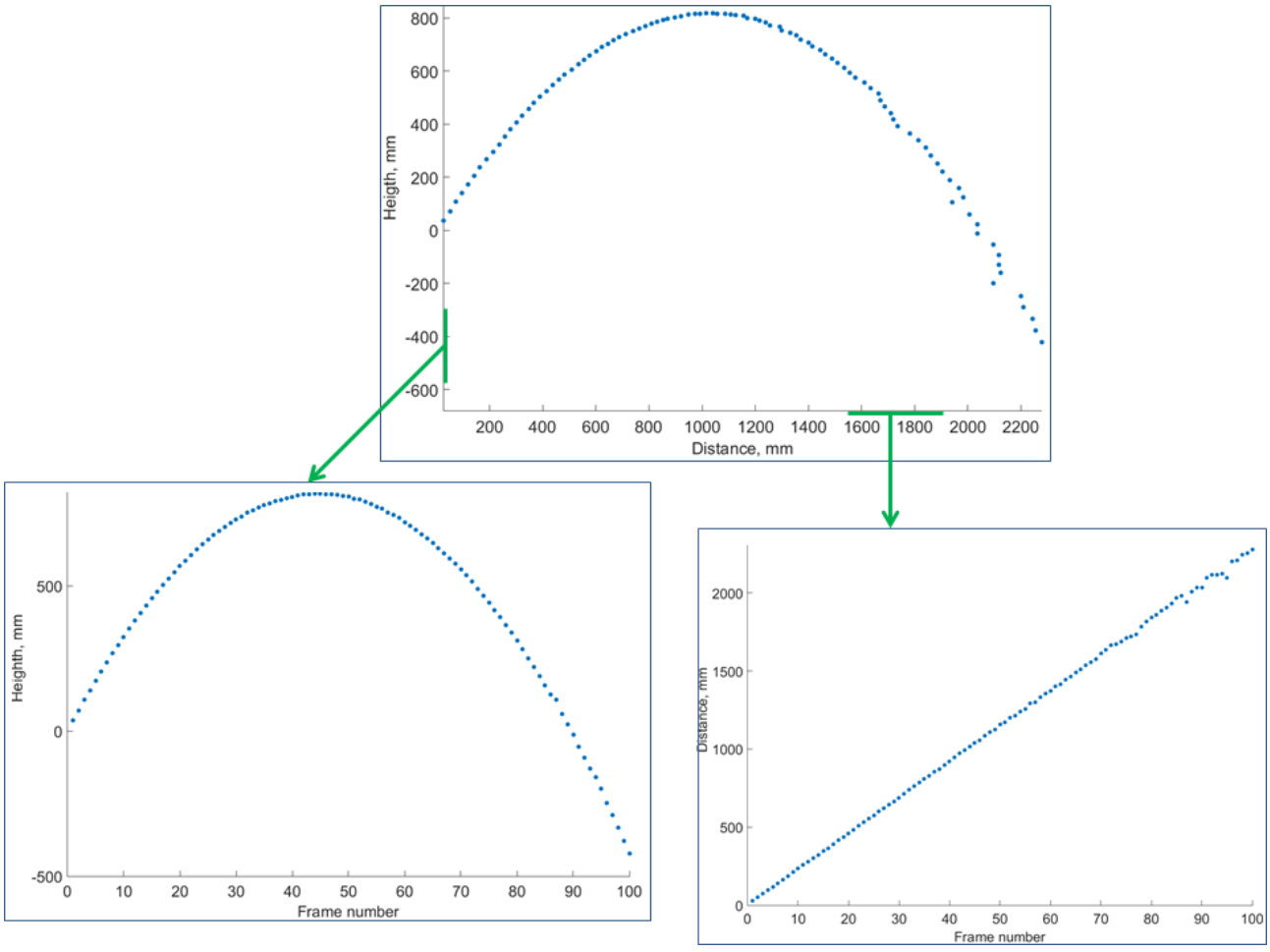


Рис. 15 Плавность измеренной траектории может повлиять на точность позиционирования. По большей части траектория плавная, но в заключительной части позиции колеблются. Соответствие между положением шара в определенном измерении и номером кадра показывает, что колебания в основном локализованы в измерении диапазона

Причина локализации ошибок в одном измерении заключается в том, что, когда расстояние до объекта больше базовой линии, разница в одном пикселе намного более значительна в измерении диапазона. Иллюстрация этого свойства приведена на рис. 16.

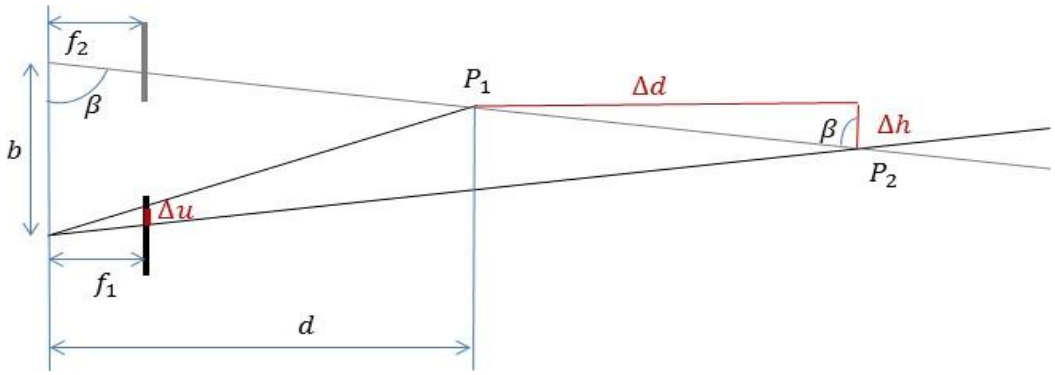


Рис. 16 Влияние ошибки пикселя δu на ошибки 3D-позиционирования в измерении диапазона (δd) и перпендикулярном измерении (δh)

Если значение полной ошибки равно длине P_1P_2, значения дальности и боковых составляющих ошибки ((δd) и (δh) соответственно) будут определяться углом β между базовой линией и лучом, который определяется положением пикселя точки на изображении со второй камеры.

$$\Delta h = P_1P_2 \cos\beta,$$
$$\Delta d = P_1P_2 \sin\beta. \qquad (17)$$

По мере увеличения расстояния от базовой линии значение β растет, то есть синус увеличивается, а косинус этого угла уменьшается. Поэтому на больших расстояниях погрешности измерения дальности достигают наиболее значительных значений.

Таблица 4

Сравнение разницы между измеренными 3D-позициями одного прогона RANSAC и медианы 1000 RANSAC для пространственных измерений в системе координат камеры

Номер кадра	Стандартное отклонение, мм			Средняя ошибка, мм		
	height	side	range	height	side	range
1...5	1.9	6.0	1.6	0.3	0.1	0.7
6...10	0.4	0.3	1.9	0.2	0.1	0.9
11...15	0.3	0.3	2.1	0.21	0.1	1.0
16...20	0.5	0.4	1.8	0.1	0.1	0.1
21...25	0.3	0.3	2.4	0.2	0.0	0.0
26...30	0.7	0.3	2.1	0.3	0.0	0.1
31...35	7.3	2.9	15.3	0.3	0.1	2.0
36...40	1.0	0.4	3.9	0.3	0.0	0.1
41...45	1.1	0.5	3.1	0.3	0.0	0.1
46...50	1.6	0.5	3.6	0.4	0.0	0.1
61...65	1.2	0.5	14.8	0.4	0.0	0.2
66...70	6.0	2.5	40.8	0.5	0.0	4.1
71...75	29.7	22.8	133.5	0.6	0.2	5.3
76...80	49.5	40.0	234.1	0.9	1.0	7.0
81...85	42.3	38.1	222.3	1.0	1.0	7.7
86...90	33.0	30.0	192.6	1.0	0.9	8.2
96...100	13.7	11.7	207.5	1.2	1.0	9.1

В связи с этим эксперименты показывают, что величина ошибки на больших расстояниях является существенной для отслеживания с точки зрения измерения дальности. Единственный способ решить эту проблему – увеличить количество камер, используемых для отслеживания, но это может быть дорогостоящим. Экспериментальная установка в Институте компьютерных

технологий включает в себя одну пару камер. В этой работе предполагается отслеживание объекта с помощью одной стереопары, а проблема ошибочного измерения решается с помощью алгоритмических средств и специфического выравнивания камеры.

Камеры должны быть выровнены таким образом, чтобы свести к минимуму влияние ошибок на большие расстояния на качество работы системы. Необходимо ответить на следующий вопрос: «На каком участке траектории точное позиционирование наиболее важно?» В первой установке камеры располагались напротив метательного устройства. В этой ситуации первые кадры являются ошибочными. Точно расположить мяч можно было только на 10-м или 12-м фреймах. В дальнейших экспериментах камеры располагались сбоку от метательного устройства. При таком выравнивании позиционирование на начальной части траектории довольно хорошее, но конечная часть траектории измеряется с более высокими ошибками. Высокая точность на начальном участке траектории на начальном этапе и более низкая точность на конечном этапе кажутся предпочтительнее, чем наоборот, в соответствии со следующим фактором. Начальная часть траектории используется для прогнозирования, сделанного в режиме реального времени. Заключительная часть траектории не обрабатывается в режиме реального времени. В реальных условиях транспортировки мяч уже находился бы в рабочей области захвата. Это используется только для обучения; поэтому его точность может быть повышена с помощью автономного статистического метода, например голосования 1000 RANSAC или подгонки модели к данным. Точное начальное позиционирование требуется для измерения параметров запуска: скорости, угла броска, положения на первом кадре и т. д. Еще один фактор заключается в ловле мяча, измерение положения мяча в конечной зоне в любом случае не будет точным. Робот, движущийся в поле зрения, сильно искажает работу алгоритма. Из-за этих факторов расположение камеры сбоку от метателя более вероятно, чем наоборот.

Также возможно расположить камеры другим способом, то есть с большей базовой линией или не параллельно направлению траектории. Вследствие этих изменений высокая погрешность не локализуется в измерении, совпадающем с направлением движения объекта. На самом деле эта локализация очень полезна для исправления ошибок. Объект движется с почти постоянной скоростью, и движение может быть аппроксимировано полиномом второго порядка (см. заключительный абзац этого раздела). Движение в этом направлении обеспечивает большее расстояние, чем в других измерениях, поэтому измерения расстояния имеют высокие значения по сравнению с размером ошибок. Благодаря этому сохраняется ориентация камеры, которая почти параллельна оптической оси направления движения, и короткая базовая линия сохраняется.

Увеличение ошибок дальше от камер связано с аспектами квантования, но выбросы в этой области в основном связаны с ошибками в обработке изображений. В основном эти ошибки связаны с влиянием фоновых объектов на алгоритм обнаружения границ. Некоторые примеры показаны на рис. 17.

Данный аспект мотивирует использование специального покрытия для фоновых объектов, чтобы уменьшить их яркость на изображениях. Детали робота, в частности, были покрыты черным материалом. Однако в реальной заводской среде фон может быть более контрастным, и это необходимо принимать во внимание. Кроме того, робот не может быть покрыт до такой степени, чтобы он не мог двигаться. Предложенные действия уменьшают влияние ошибок, но плохое позиционирование и обнаружение выбросов все еще остаются. Пример ошибочных траекторий, извлеченных с помощью 1000 запусков RANSAC, показан на рис. 18. Эти ошибочные случаи нетипичны. Большинство траекторий имеют более плавный, базовый вид, но ошибочные случаи все еще имеются. Многократные запуски RANSAC не могут устранить такие ошибки, поскольку RANSAC основан на детекторе Canny Edge, который является детерминированным алгоритмом. Если детектор границ работает ошибочно, во многих случаях эти ошибки не могут быть устранены RANSAC.

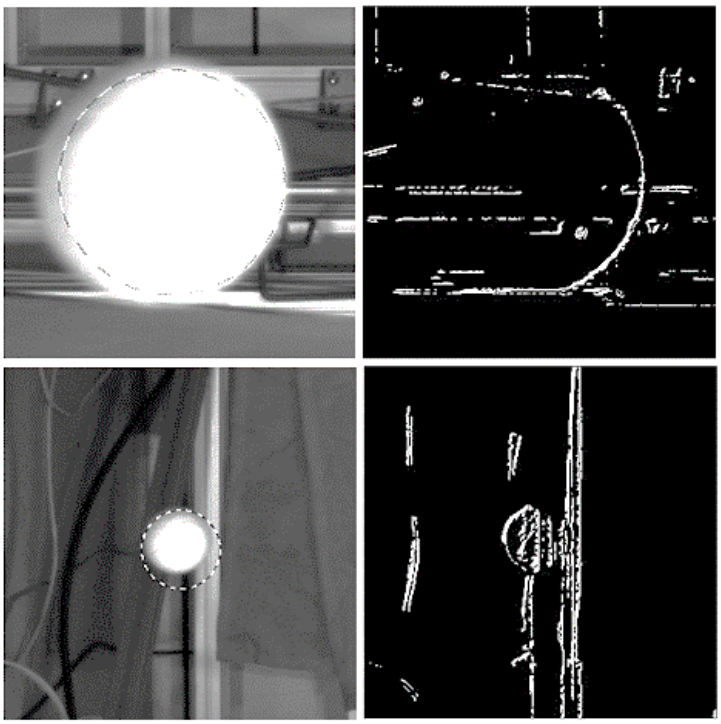


Рис. 17 Неправильное определение края может повлиять на извлечение окружности. В левом столбце извлеченные круги отображаются на входные изображения. В правой колонке показаны соответствующие изображения краев

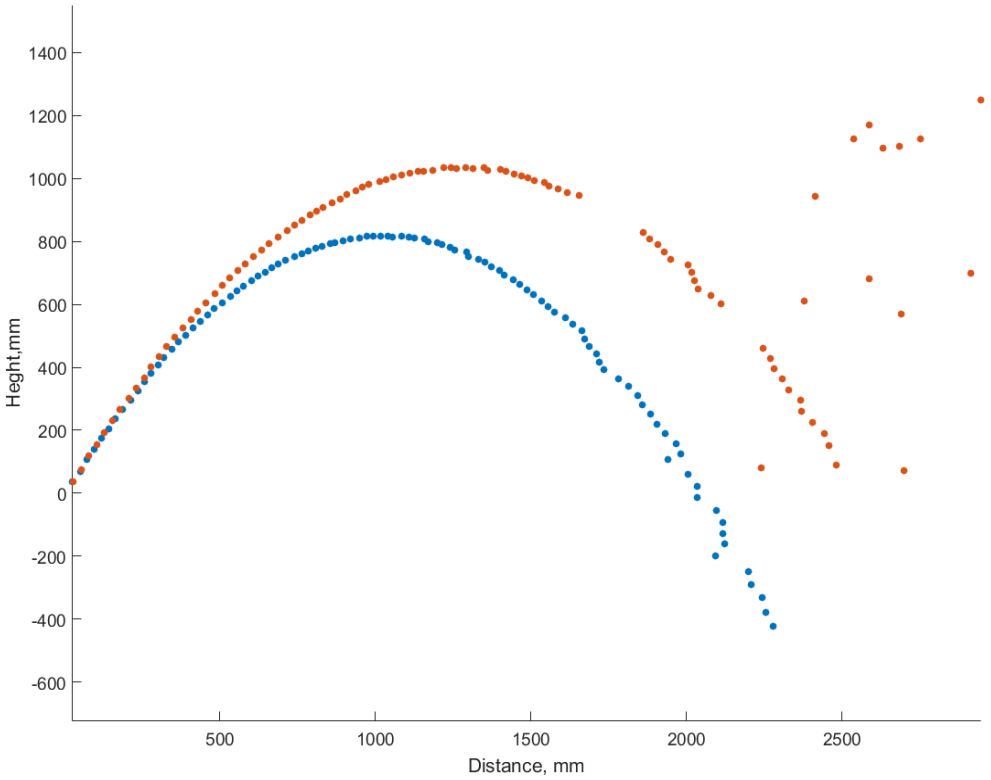


Рис. 18 Примерные графики траекторий, измеренных 1000 RANSAC, выполняются ошибочно. Красные точки представляют собой траекторию с выбросами, а синие точки – траекторию с плохой плавностью

Это влияние может быть уменьшено путем подгонки кривой к данным траектории. Для подгонки требуется точная модель полета объекта, поэтому предлагается тщательное использование подгонки кривой. Единственной целью подгонки является повышение точности данных на заключительной части траектории, которая используется для формирования обучающего набора данных. Таким образом, результат подгонки заменяет результат измерения только в заключительной части, начиная с 60-го кадра. Подгонка производится только для измерения диапазона; отдельная обработка каждого измерения уже рассматривалась в работе [Pon09]). Кривая второго порядка соответствует данным в измерении диапазона:

$$x_3 = p_1 n^2 + p_2 n p_3, \tag{18}$$

где n – номер кадра p_1; p_2 и p_3 – коэффициенты, полученные в результате операции подгонки. График измеренных и установленных эталонов диапазона объектов показан на рис. 19. С визуальной точки зрения такое соответствие кажется вероятным.

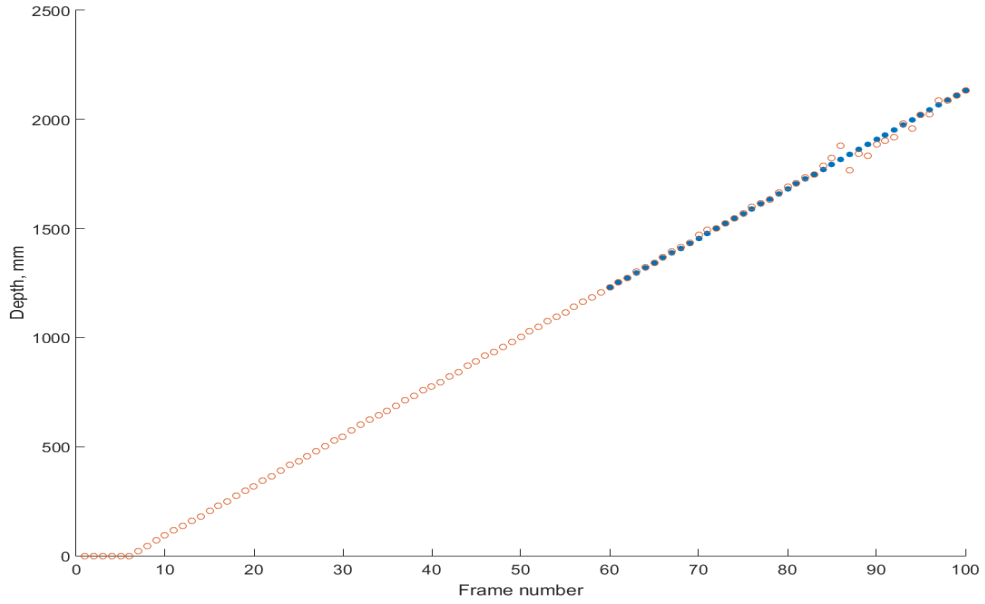


Рис. 19 Измеренные (красные кружки) и установленные (синие точки) значения диапазона объектов

3.3 Измерение скорости объекта

В предыдущих подразделах рассматривалась точность измерения положения объекта в пространстве. Измеренные положения могут быть использованы для оценки скорости и ускорения летящего тела. Вектор скорости V объекта в кадре с номером n оценивается по следующей формуле:

$$V(n) = \frac{X(n) - X(n-1)}{\tau} = \big(X(n) - X(n-1)\big) f, \tag{19}$$

где τ – межкадровый период, а f – частота кадров. Как межкадровый период, так и частота кадров являются постоянными параметрами, используемыми для выражения скорости в международной системе единиц, поэтому можно сказать, что скорость равна разности координат объекта на соседних кадрах.

Абсолютная погрешность измерения этих погрешностей имеет одинаковый масштаб, но относительные погрешности намного больше. Обсудим ошибки оценки скорости на основе ошибок оценки положения, определенных в предыдущих подразделах. Положение объекта со стандартными отклонениями можно увидеть в табл. 3. Для большинства кадров от 1 до 60 стандартное отклонение варьируется от 1.4 до 6.4 мм. Примем погрешность 5 мм для лучшей интерпретируемости (цель этого раздела – показать, что измерения скорости недостаточно

точны, поэтому правильно принимать значения точности немного лучше, чем в действительности). Если скорость объекта составляет 5 м/с, а частота кадров составляет 100 кадров в секунду, объект перемещается на 5 см между двумя кадрами. Абсолютная погрешность 5 мм на расстоянии 5 см соответствует относительной погрешности 10 % или абсолютной погрешности 0.5 м/с при оценке скорости объекта. Согласно хорошо известному 3σ-правилу [Hla12], отклонение не превышает трехкратного стандартного отклонения с достоверностью 99 %. В рассматриваемом случае это означает, что истинное значение скорости лежит где-то между 3.5 и 6.5 м/с.

Этот интервал полностью покрывает интервал скоростей броска от 4.2 до 5.1 м/с, измеренный в разделе 3.1. Таким образом, точность измерения скорости низкая из-за более высокой чувствительности к ошибкам измерения. Одни и те же абсолютные ошибки в низком масштабе оказывают более сильное влияние, чем в высоком. Это согласуется с хорошо известным правилом: производные более чувствительны к ошибкам измерения, чем их первообразные функции. Ускорение, которое является производной скорости, еще более чувствительно. Поэтому при разработке алгоритма, сделанного в следующих главах, кратковременные скорости и ускорения не учитываются.

Заключение

На основании изложенного можно сделать следующие выводы.

1. Предпочтительно расположение камеры сбоку от метательного устройства с оптической осью, сонаправленной с номинальной траекторией.
2. Вычитание фона необходимо для точного отслеживания.
3. Конечная часть траектории в этой установке измеряется с ошибками, которые превышают ошибки статического позиционирования. Эти ошибки локализованы в диапазоне пространственных измерений.
4. При формировании базы данных влияние ошибок уменьшается путем вычисления медианы 1000 запусков RANSAC и подгонки полиномиальной модели к позициям в измерении диапазона.

Вместо преобразования Хафа выбран экстрактор круга RANSAC. Сравнение в [Goe15, с. 70–82] показало, что RANSAC в большинстве случаев работает быстрее, а его точность не хуже, чем для Windows. Анализ данных траектории показал, что преобразование Хафа имеет те же проблемы с зашумленными краевыми изображениями, что и RANSAC. В исследовании выбран RANSAC, потому что он позволяет находить несколько выходов для одного и того же входного сигнала, которые можно сравнить, чтобы получить более точное значение.

Благодарности

Автор выражает признательность профессору Дитмару Дитриху (Dietmar Dietrich) и коллеге Мартину Понграцу (M. Pongratz) – лидерам исследовательского проекта «Транспортировка бросанием» в Техническом университете Вены, а также профессорам Л. Р. Черняховской, В. В. Васильеву, М. Б. Гузаирову, Р. А. Мунасыпову и Г. Р. Шахмаметовой (Уфимский университет науки и технологий) за полезные замечания и поддержку.

Список литературы / References

[Мир24а] Миронов К. В. Transport-by-Throwing – робототехнический способ перемещения предметов перебросом: обсуждение научно-технической задачи // СИИТ. 2024. Т. 6. № 1(16). С. 43–53. EDN QGFZBW.

[Мир24б] Миронов К. В. Transport-by-Throwing – робототехнический способ перемещения предметов перебросом: обзор используемых методов // СИИТ. 2024. Т. 6. № 3(18). С. 3–48. EDN FUUPEN.

[Bar09] Barteit D., Frank H., Pongratz M., Kupzog F. Measuring the Intersection of a Thrown Object with a Vertical Plane // IEEE International Conference on Industrial Informatics. Cardiff. UK. Pp. 680 to 685. June 2009.

[Bou22] Bouguet J.-Y. (2022). Camera Calibration Toolbox for Matlab. (1.0). CaltechDATA. https://doi.org/10.22002/D1.20164.

[Cam24] Camera Calibration – MATLAB & Simulink. http://mathworks.com/help/vision/cameracalibration. html, visited on September 28, 2024.

[Can86] Canny J. A Computational Approach to Edge Detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1986. Vol. 8. No. 6. November.

[Fis81] Fischler M. A., Bolles R. C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography // Communications of the Association for Computing Machinery. 1981. Vol. 24. No. 6. Pp. 381 to 395. June.

[Goe15] Goetzinger M. Object Detection and Flightpath Prediction. Diploma Thesis, Faculty of Electrical Engineering, Vienna University of Technology, June 2015.

[Hla12] Hlawatsch F. Parameter Estimation Methods: Lecture Notes, Grafisches Zentrum HTU GmbH. Vienna, Austria, March 2012.

[Hou62] Hough P. A method and means for recognizing complex patterns. U.S. Patent No. 3,069,654, December 1962.

[ITF15] ITF Tennis – Technical, http://www.itftennis.com/technical/balls/overview.aspx, visited on September 30, 2015.

[Kay93] Kay S. M. Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall, Englewood Cliffs, USA. March 1993.

[Kim75] Kimme C., Ballard D., Sklansky J. Finding circles by an array of accumulators // Communications of the ACM. 1975. Vol. 18. No. 2. Pp. 120 to 122. February.

[Lee02] Lee J. H., Akiyama T., Hashimoto H. Study on Optimal Camera Arrangement for Positioning People in Intelligent Space, IEEE/RSJ international conference on intelligent robots and systems, Lausanne, Switzerland. Pp. 220 to 225. October 2002.

[Liu06] Liu J., Zhang Y., Li Z. Selection of Cameras Setup Geometry Parameters in Binocular Stereovision, IEEE Conference on Robotics, Automation and Mechatronics. Bangkok, Thailand. Pp. 1 to 6. June 2006.

[Mar16] Markum G. Object Touchdown Position Prediction, Bachelor Thesis, Faculty of Electrical Engineering, Vienna University of Technology. March 2016.

[Mar24] TISS – Markum, Gilbert Harald: https://tiss.tuwien.ac.at/adressbuch/adressbuch/person/231464. visited on November 6. 2024.

[Pon09] Pongratz M. Object Touchdown Position Prediction, Diploma Thesis, Faculty of Electrical Engineering, Vienna University of Technology September 2009.

[Pon15] Pongratz M., Mironov K. V. Accuracy of Positioning Spherical Objects with Stereo Camera System // IEEE International Conference on Industrial Technology. Seville, Spain. Pp. 1608 to 1612. March 2015.

[Pon24] Martin Pongratz | Institute of Computer Technology. Wien. https://www.ict.tuwien.ac.at/en/users/pongratz, visited on November 6. 2024.

[Sca05] Scaramuzza D., Pagnotelli S., Valligi P. Ball Detection and Predictive Ball Following Based on a Stereoscopic Vision System // IEEE International Conference on Robotics and Automation. Barcelona, Spain. Pp. 1573–1578. April 2005.

[Sze09] Szeliski R. Computer Vision: Algorithms and Applications. Springer. September 2010.

Поступила в редакцию 6 декабря 2024 г.

МЕТАДАННЫЕ / METADATA

Title: Transport-by-Throwing — a robotic method of moving objects by through: overview of the methods used.

Abstract: Throwing is a promising method for robotic handling of parts in flexible manufacturing systems. This review focuses on the study of observation accuracy. The experiments described are aimed at determining how measurement error affects the prediction and how this effect can be reduced. Learning-based prediction requires a set of trajectory measurements, which is generated during experimental throws. Technical information is provided on the throwing device, cameras, and algorithms that allow obtaining a 3D trajectory model based on camera measurements. This is the basis for the study of accuracy. The accuracy of positioning the thrown ball in one pair of images is discussed.

Key words: robotic system; statistical evaluation of ballistic curves; tracking of moving objects; nearest neighbor algorithm.

Язык статьи / Language: Русский / Russian.

Об авторе / About the author:

МИРОНОВ Константин Валерьевич
Уфимский университет науки и технологий, Россия.
Доц. каф. вычислительной техники и защиты информации. Дипл. спец. по защите инф-и (Уфимск. гос. авиац. техн. ун-т, 2012). PhD (Техн. ун-т Вены, 2016). Иссл. в обл. робототехники, применения искусственного интеллекта в техн. системах.
E-mail: mironovconst@gmail.com
ORCID: http://orcid.org/0000-0002-4828-1345
URL: https://elibrary.ru/author_profile.asp?authorid= 939814

MIRONOV Konstantin Valeryevich
Ufa University of Science and Technology, Russia.
Assoc. Prof., Department of Computer Science and Information Security. Information Security Specialist (Ufa State Aviation Tech. Univ., 2012). PhD (Vienna University of Technology, 2016). Research in the field of robotics and intelligent control.
E-mail: mironovconst@gmail.com
ORCID: http://orcid.org/0000-0002-4828-1345
URL: https://elibrary.ru/author_profile.asp?authorid= 939814