

УДК 355/359.07
DOI: 10.22213/2410-9304-2025-4-33-39

Определения расстояния до объекта с использованием технического зрения

А. В. Колузов, ВИТ «ЭРА», Анапа, Россия
К. С. Корнеев, ВИТ «ЭРА», Анапа, Россия
А. С. Гой, ВИТ «ЭРА», Анапа, Россия

Целью данной статьи является оценка эффективности и точности использования комплексного подхода для определения расстояния до объектов, основанного на модели камеры и применении библиотек машинного обучения, в контексте систем помощи водителю (ADAS) и автономного вождения.

В работе рассматривается, как параметры модели камеры (разрешение сенсора, фокусное расстояние, угол обзора и характеристики диафрагмы) влияют на качество входных данных для алгоритмов, реализованных через библиотеки машинного обучения и, как следствие, на общую точность оценки расстояния. Особое внимание уделяется практическому применению современных алгоритмов машинного обучения, в том числе глубоких нейронных сетей, для решения задачи монокулярной оценки глубины по одному изображению.

Исследование предполагает выявить потенциальную зависимость точности определения расстояния от дистанции до объекта. В частности, будет проанализировано, как увеличение дистанции может привести к непропорциональному росту погрешности в оценке, что позволит определить возможные ограничения текущих реализаций алгоритмов машинного обучения при работе с удаленными объектами. Будут рассмотрены факторы, способствующие снижению точности измерений на больших расстояниях, такие как усиление искажений в процессе захвата и обработки изображений, а также вопросы адаптации алгоритмов к особенностям данных от удаленных объектов.

Ключевые слова: параметры камеры, алгоритмы обработки изображений, точность определения расстояния, компьютерное зрение, стереозрение, фокусное расстояние, монокулярная глубина, диафрагма, выдержка, сегментация изображений, тестирование алгоритмов.

Введение

Определение дистанции до объектов, особенно транспортных средств, является фундаментальной проблемой в широком спектре современных технологий, к примеру системы интеллектуального видеонаблюдения.

Традиционно, системы оценки расстояния полагаются на комбинацию камер и алгоритмов обработки изображений. Однако эффективность и надежность таких систем тесно связаны с качеством исходных изображений, которые, в свою очередь, зависят от тщательно подобранных параметров камеры. Более того, выбор алгоритма обработки изображений также играет решающую роль. Неоптимальный выбор параметров камеры или использование неэффективных алгоритмов может приводить к значительным погрешностям в оценке расстояния, ставя под угрозу производительность и безопасность системы.

Данное исследование направлено на выявление и анализ ключевых факторов, влияющих на точность определения расстояния до автомобиля. Мы проводим систематическое изучение влияния параметров камеры на качество изображения и оцениваем эффективность различных алгоритмов обработки изображений, используемых для оценки расстояния.

На основе полученных результатов предлагаются рекомендации по выбору оптимальных настроек камеры и наиболее подходящих алгоритмов, позволяющих достичь максимальной точности и надежности определения расстояния до автомобиля в различных условиях [1].

Анализ параметров камеры и их влияние на изображение

Методы компьютерного зрения для определения расстояния до объектов критически зависят от качества входных изображений, формируемых камерой. Соответственно, качество изображения, являющееся определяющим фактором, непосредственно связано с параметрами камеры. В рамках данного раздела будет представлен детальный анализ основных параметров камеры и их влияния на получаемые изображения. Параметры камеры делятся на следующие группы:

1. Разрешение камеры.
2. Фокусное расстояние (f).
3. Размер сенсора.
4. Диафрагма.
5. Выдержка.
6. ISO (светочувствительность).
7. Искажение объектива.

Разрешение камеры определяет количество пикселей, составляющих изображение. Обычно это ширина и высота в пикселях. Высокое разрешение сенсора обеспечивает захват большего количества деталей, что способствует улучшению точности определения расстояния. Более точная идентификация и локализация ключевых точек (углов, границ) становятся возможными, минимизируя потенциальные ошибки в вычислениях, особенно применительно к удаленным объектам, где даже небольшие смещения пикселей оказывают существенное влияние на итоговый результат. Изображения с низким разрешением в свою очередь характеризуются ограниченной детализацией, что создает трудности при идентификации объектов и точном определении их кон-

туров. Данные недостатки негативно влияют на производительность алгоритмов, предназначенных для определения расстояния, снижая общую точность системы. В свою очередь выбор высокого разрешения влечет за собой увеличение вычислительной нагрузки на систему обработки изображений, а также повышение требований к объему оперативной и постоянной памяти. Поэтому при проектировании системы необходимо тщательно сбалансировать стремление к детализации с ограничениями, накладываемыми доступными вычислительными ресурсами.

Фокусное расстояние объектива определяет угол обзора камеры и степень увеличения объектов в кадре. Использование объективов с малым фокусным расстоянием, обеспечивающих широкий угол обзора, позволяет получить больше информации об окружающей среде и полезно в сценариях, требующих охвата большой области, например, при движении в ограниченных пространствах. Однако широкоугольные объективы приводят к уменьшению масштаба объектов и геометрическим искажениям (дисторсии), что может затруднить точное определение расстояния. Дисторсия особенно выражена на краях изображения, где происходит наибольшее отклонение от идеальной перспективы. Объективы с большим фокусным расстоянием обеспечивают узкий угол обзора и увеличение объектов, что делает их оптимальным выбором для наблюдения объектов на больших расстояниях с высокой детализацией. К недостаткам относится ограниченная видимость окружающих объектов, что может представлять неточности в полученных сценариях. В то время как подобные объективы менее подвержены дисторсии, они более чувствительны к вибрациям камеры, что может негативно сказаться на качестве изображения. Перспектива, отображаемая на изображении, напрямую зависит от выбранного фокусного расстояния. Объективы с малым фокусным расстоянием склонны к преувеличению расстояний, в то время как длиннофокусные объективы сжимают перспективу, уменьшая видимую глубину сцены. Следовательно, выбор оптимального фокусного расстояния должен учитывать как требуемую дистанцию обнаружения объектов, так и условия окружающей среды, а также потенциальное влияние перспективных искажений на интерпретацию сцены.

Формирование изображения происходит на светочувствительной поверхности сенсора, размер которой характеризуется физической площадью матрицы. Большие сенсоры характеризуются повышенной светочувствительностью, что приводит к улучшению качества изображений, снижению уровня шума и увеличению динамического диапазона. При одинаковом фокусном расстоянии обеспечивается более широкое поле зрения по сравнению с сенсорами меньшего размера. Малые сенсоры ограничивают захват света, что приводит к увеличению шума и уменьшению динамического диапазона изображения, особенно в условиях низкой освещенности. Выбор размера сенсора является важным этапом оптимизации системы обработки изображений. В то время как большие сенсоры обеспечивают улучшенные характеристики изображения, они также приводят к увеличению стоимости и габа-

ритов системы. Следовательно, проектирование эффективной и экономичной системы требует тщательного учета всех факторов, включая требуемые характеристики изображения, доступный бюджет и ограничения по размерам камеры [2].

Количество света, проходящего через объектив, регулируется диафрагмой, параметром, обозначаемым *f*-числом (например, *f*/2, *f*/8). Важно понимать, что меньшее значение *f*-числа указывает на более широкое отверстие диафрагмы и, соответственно, на больший проходящий световой поток. Выбор широкой диафрагмы (низкое *f*-число) обеспечивает возможность захвата большего количества света, что повышает качество изображения при слабом освещении. Однако необходимо учитывать, что это приводит к уменьшению глубины резкости, что может потребовать более точной фокусировки и приводить к размытию объектов, не находящихся в фокусе. Объективы с малым относительным отверстием, достигаемым при использовании узкой диафрагмы (большое *f*-число), пропускают меньше света, что может потребовать компенсации путем увеличения времени экспозиции или повышения чувствительности сенсора (ISO). Преимуществом использования узкой диафрагмы является увеличение глубины резкости, позволяющее получить четкое изображение объектов, находящихся на различных расстояниях от камеры. Дифракция света, усиливающаяся при сильном сужении диафрагмы, может снижать резкость изображения. Поэтому выбор оптимального значения диафрагмы – компромисс между глубиной резкости и дифракционным размытием, зависящий от конкретных условий съемки [3].

Выдержка представляет собой период времени, в течение которого светочувствительный элемент (сенсор) камеры открыт для воздействия света. Данный параметр измеряется в единицах времени, как правило, в секундах или их долях (например, 1/1000 секунды, 1 секунда), и является ключевым фактором, определяющим экспозицию. Применение коротких выдержек позволяет эффективно минимизировать смазывание изображения при съемке динамичных сцен. Это достигается за счет фиксации движения на коротком временном интервале. Недостаток заключается в необходимости компенсации уменьшенного светового потока, что может потребовать открытия диафрагмы или увеличения чувствительности ISO [4].

ISO управляет чувствительностью камеры к свету. Высокое значение ISO позволяет снимать в условиях недостаточной освещенности, но может привести к появлению шума на фотографии. Низкие значения ISO обеспечивают максимальное качество изображения с минимальным уровнем зернистости (шума). Данная настройка требует достаточного освещения сцены для достижения оптимальной экспозиции. Высокие значения ISO расширяют возможности съемки в условиях недостаточной освещенности. Однако это приводит к увеличению уровня шума на изображении, а также может негативно сказаться на детализации и динамическом диапазоне. Таким образом, для оптимального качества изображения всегда предпочтительнее использовать минимально возможное значение ISO [5].

Оптические системы объективов могут вносить в изображение различные aberrации, в частности геометрическую дисторсию, проявляющуюся в виде бочкообразного или подушкообразного искажения перспективы, и хроматические aberrации, которые выражаются в появлении цветных ореолов вокруг объектов. Геометрические искажения, вызванные дисторсией, могут существенно влиять на точность измерений на изображении. Искаженная перспектива затрудняет правильное определение размеров и координат объектов, что необходимо учитывать при расчете дистанции. Цветные окантовки, вызванные хроматическими aberrациями, вносят погрешности в определение границ объектов на снимке. Это может стать серьезной проблемой, особенно если требуется высокая точность измерения расстояний до этих объектов. В конечном итоге, чтобы получить максимально точные результаты, нужно выбирать объектив с минимальными искажениями, скорректировать искажения при обработке изображений и регулярно калибровать камеру. Эти шаги помогут минимизировать влияние оптических недостатков на измерения [6].

Анализ алгоритмов обработки изображений

Для автоматического определения расстояния до объектов на изображении существует ряд алгоритмов обработки. Далее будут рассмотрены наиболее распространенные методы, используемые в этой области.

Метод стереозрения основан на использовании нескольких камер, расположенных на известном расстоянии (базе) друг от друга, для получения изображений одной и той же сцены. Анализируя разницу в положении (параллаксе) идентичных объектов на изображениях, можно вычислить расстояние до этих объектов. Стереозрение обеспечивает высокую точность, однако требует тщательной калибровки камер и значительных вычислительных ресурсов для поиска соответствий между изображениями. Для решения этой задачи используются различные алгоритмы, такие как Block Matching, Semi-Global Matching (SGM), и Graph Cut [7].

Метод монокулярной глубины оценивает расстояние до объектов, анализируя единственное изображение. В качестве входных данных используются различные признаки изображения, включая размер объектов, текстурные градиенты, информацию о тенях, перспективные проекции и априорные знания о типичных размерах объектов. По сравнению со стереозрением методы монокулярной глубины характеризуются меньшей точностью, однако отличаются более низкой вычислительной сложностью и отсутствием необходимости в калибровке камер. С развитием глубокого обучения нейронные сети, обученные на обширных датасетах, стали широко применяться для решения задач оценки монокулярной глубины [8].

Методы машинного обучения основаны на обучении моделей на обширных датасетах, содержащих изображения и соответствующие им значения расстояний. Полученные модели способны с высокой точностью предсказывать расстояние до объектов на новых, ранее не виданных изображениях. Обучение моделей машинного обучения требует значительных вычислительных ресурсов. Для повышения точности перед оценкой рас-

стояния часто применяются алгоритмы предварительной обработки, например сегментация объектов и детектирование границ с использованием сверточных нейронных сетей (CNN) [9].

Из рассмотренных методов наибольший интерес представляет метод машинного обучения, который представляется наиболее перспективным и актуальным для определения расстояния до объекта по фотографии.

Применение метода машинного обучения позволяет достигать высокой точности и надежности в определении расстояния, зачастую опережая традиционные методы стереозрения и монокулярной глубины. Ключевыми преимуществами этих подходов являются их универсальность, достигаемая за счет адаптации к различным сценариям и объектам, а также значительно лучшая устойчивость к сложным условиям, таким как недостаточное освещение, наличие шума и оклюзия объектов.

Однако стереозрение и классические методы монокулярной глубины все еще могут быть полезны в определенных сценариях, методы машинного обучения, особенно глубокое обучение, предлагают наилучший баланс между точностью, надежностью и адаптивностью для решения задачи определения расстояния до объекта по фотографии.

Экспериментальное исследование применения метода машинного обучения для обнаружения автомобилей

Первоочередная задача заключается в нахождении расстояния от объектива видеокамеры до рассматриваемого объекта (Автомобиль). Для этого было выведено следующее соотношение:

$$\frac{W}{D} = \frac{\text{sensor } w}{f}, \quad (1)$$

где w – ширина объекта; D – расстояние до объекта; f – фокусное расстояние; $\text{sensor } w$ – ширина изображения объекта на матрице.

Фокусное расстояние представляет из себя расстояние от центра линзы до точки, где сфокусированные лучи света пересекаются, образуя резкое изображение на сенсоре. Измеряется в миллиметрах.

Взаимосвязь фокусного расстояния и ширины изображения объекта на матрице можно увидеть на рис. 1.

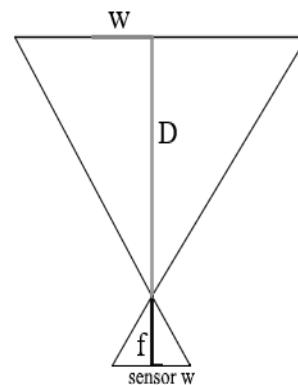


Рис. 1. Пример работы линзы в видеокамере

Fig. 1. Example of lens operation in a video camera

Ширина кадра в пикселях определяет горизонтальное разрешение изображения. Это количество пикселей, составляющих ширину изображения. Чем больше пикселей, тем выше разрешение, тем больше деталей можно захватить на изображении. Ширина кадра, наряду с высотой, определяет общее количество пикселей в изображении.

Изменение фокусного расстояния влияет на угол обзора и, соответственно, на размер объектов в кадре. Однако при неизменном размере сенсора ширина кадра в пикселях остается постоянной, определяясь только разрешением самого сенсора.

Связь ширины изображения в мм и ширины изображения в пикселях:

$$\frac{w}{image\ width} = \frac{sensor\ w}{sensor\ width}, \quad (2)$$

где w – ширина одного пикселя в мм; $image\ width$ – ширина кадра в пикселях; $sensor\ width$ – ширина кадра сенсора в пикселях.

Ширина кадра сенсора – это физический показатель, определяет максимальный размер изображения, которое может быть захвачено, имеет постоянное значение для камеры.

Ширина кадра – это не постоянный показатель, зависит от количества пикселей на фотографии, может быть меньше ширины сенсора, зависит от выбранного разрешения. В данном эксперименте все фотографии имеют разрешение 1280×720 [10].

Из соотношения (2) получим формулу изображения объекта на матрице.

$$sensor = \frac{w \cdot sensor\ width}{image\ width}. \quad (3)$$

Получившуюся формулу подставим в соотношение (1) для того, чтобы вывести формулу получения расстояния до объекта – D .

$$\frac{W}{D} = \frac{\frac{w \cdot sensor\ width}{image\ width}}{f}. \quad (4)$$

Выводим расстояние до объекта из формулы (4).

$$D = \frac{Wf \cdot image\ width}{w \cdot sensor\ width}. \quad (5)$$

Получаем формулу расстояния до объекта, данная формула пригодится нам при написании скрипта на Python по определению рассматриваемого объекта на фотографии и его удаленности от видеокамеры.

$$Distance_mm = \frac{object_width \cdot 1000 \cdot focal_length \cdot image_width_pixels}{bbox_width \cdot sensor_width}. \quad (6)$$

В выражении (6) используется известная ширина объекта ($object_width$), фокусное расстояние ($focal_length$), ширину изображения в пикселях ($image_width_pixels$), ширину $boundingbox$ в пикселях ($bbox_width$) и ширину сенсора. Перемножение на 1000 преобразует результат из метров в миллиметры [12].

Реализация формулы определения расстояния до объекта на примере Python находится на рис. 2.

```
x, y, bbox_width, bbox_height = bbox
sensor_width = 36 # Ширина Full Frame матрицы в мм
distance_mm = (object_width * 1000 * focal_length * image_width_pixels) / (bbox_width * sensor_width)
distance_meters = distance_mm / 1000
print(2 * math.atan(12 / 24) * 57.3)
return distance_meters
def image_to_3d(u, v, width, height, fov_x, fov_y, distance):
    ...
def process_frame(frame, bbox):
    ...
center = (bbox[0] + bbox[2]) / 2, bbox[1] + bbox[3] / 2
```

Рис. 2. Пример определения расстояния до объекта, Python

Fig. 2. Example of Object Distance Determination, Python

Задаем четыре переменные, описывающие параметры рассматриваемого объекта и камеры.

$Object_width = 0.35$ – данная переменная хранит ширину объекта в метрах. Значение 0,35 означает, что объект имеет ширину 35 сантиметров.

$Focal_length = 24$ – фокусное расстояние камеры в миллиметрах. Фокусное расстояние определяет угол зрения камеры (угол, образованный крайними точками рассматриваемого объекта и узловой точки камеры, характеризует взаимосвязь между удаленностью объекта от камеры и его величиной) и увеличение. Чем меньше фокусное расстояние, тем шире угол зрения камеры.

$Image_width_pixels = 1280$ – переменная представляет ширину изображения в пикселях. Данное разрешение было выбрано ввиду удобства обработки изображений скриптом [11].

$Fov_degrees = 84$ – горизонтальное поле зрения (FOV – FieldofView) камеры в градусах. Это угол, который охватывает камера в горизонтальном направлении.

$Distance_to_avto(bbox)$ – функция вычисления расстояния до объекта (Автомобиль) на основе его $boundingbox$.

$X, y, bbox_width, bbox_height = bbox$ – распаковка кортежа $bbox$, содержащего координаты x, y , левого верхнего угла $boundingbox$ и его ширину и высоту.

$Sensor_width = 36$ – ширина сенсора камеры в миллиметрах, ширина матрицы камеры.

Формула для вычисления расстояния, основанная на принципе подобия треугольников:

Пример реализации преобразования 2D-координат в 3D находится на рис. 3.

Само преобразование 2D-координат пикселя на изображении в 3D-координаты, близкие к реальности, выполняется с помощью функции $:image_to_3d(u, v, width, height, fov_x, fov_y, distance)$.

```

def image_to_3d(u, v, width, height, fov_x, fov_y, distance):
    # 1. Нормализация координат изображения
    x_n = (2 * u / width) - 1 # от -1 до 1 по x
    y_n = -(2 * v / height) + 1 # от -1 до 1 по y

    # 2. Вычисление углов
    angle_x = fov_x / 2 * x_n
    angle_y = fov_y / 2 * y_n
    angle_x_rad = angle_x / (180 / math.pi)
    angle_y_rad = angle_y / (180 / math.pi)
    # 3. Вычисление 3D координат
    d = (distance / math.cos(angle_x_rad)) / math.cos(angle_y_rad)

    return (angle_x, angle_y, d)
def process_frame(frame, bbox):

```

Рис. 3. Преобразование 2D-координат пикселя в 3D

Fig. 3. Conversion of 2D Pixel Coordinates to 3D

Нормализация координат изображения осуществляется следующим образом:

$$x_n = \frac{2 \cdot u}{width} - 1, \quad (7)$$

где координата u (горизонтальная) масштабируется и смещается так, чтобы находиться в диапазоне от -1 до 1 . Центр изображения соответствует 0 , левая граница -1 , правая граница 1 .

$$d = \frac{distance}{math.com(angle_x_rad) \cdot math.com(angle_y_rad)}. \quad (10)$$

Используя известное расстояние $distance$, выражение (10) корректирует расстояние d , чтобы учесть наклон относительно центральной оси камеры, предполагая, что расстояние $distance$ измеряется по перпендикуляру к камере. Это упрощение не учитывает искажение объектива и не является точным для больших углов [14].

`Process_frame(frame, bbox)` – функция, принимающая кадр изображения (`frame`) и `boundingbox` (`bbox`) объекта и вычисляющая центр `boundingbox`.

Функция `Center (bbox[0] + bbox[2] / 2, bbox[1] + + bbox[3] / 2)` производит вычисление центра `boundingbox`, используя координаты левого верхнего угла (`bbox[0]`, `bbox[1]`), ширину (`bbox[2]`) и высоту (`bbox[3]`).

В данном коде используется простейшая модель камеры, игнорирующая искажение объектива. В реальных камерах такие искажения существуют и могут значительно влиять на точность. Также стоит учитывать, что линейное приближение как связь между пиксельными координатами и углами является не точным для больших полей зрения.

Преобразуем координаты объекта на изображении в трехмерные 3D-координаты в пространстве.

$U = center[0]$ – представляет собой X-координату центра объекта на изображении, `center` – список или кортеж, где `center[0]` соответствует горизонтальной координате, пиксели от левого края изображения.

$V = center[1]$ – Y-координата центра объекта на изображении, `center[1]` соответствует вертикальной координате, пиксели от верхнего края изображения.

$Width = 1280$ – ширина изображения в пикселях;
 $height = 720$ – высота изображения в пикселях;
 $fov_x = 73$ – угол обзора по горизонтали (fieldofview). Это угол, под которым камера охватывает пространство по горизонтали.

Аналогично, координата v (вертикальная) масштабируется и смещается в диапазон от -1 до 1 . Центр изображения соответствует 0 , верхняя граница 1 , нижняя граница -1 [13].

Вычисление углов осуществляется по формуле

$$\text{angle}_x = \frac{\text{fov}_x}{2 \cdot x_n}, \quad (8)$$

где горизонтальный угол вычисляется путем умножения нормализованной координаты x_n на половину горизонтального поля зрения. Это предполагает линейную зависимость между углом и координатой на изображении, что является упрощением.

Вертикальный угол вычисляется по формуле

$$\text{angle}_y = \frac{\text{fov}_y}{2 \cdot y_n}, \quad (9)$$

Преобразование горизонтального и вертикального углов из градусов в радианы осуществляется функциями `angle_x_rad` и `angle_y_rad`.

Вычисление 3D-расстояния производится по формуле

$$distance = \frac{distance}{math.com(angle_x_rad) \cdot math.com(angle_y_rad)}. \quad (10)$$

$\text{Fov}_y = 53$ – угол обзора по вертикали. Это угол, под которым камера охватывает пространство по вертикали [15].

$Z = \text{distance_to_drone}(bbox)$ – осуществляется вычисление расстояния от дрона до объекта в пространстве, основываясь на аннотации `bbox` (boundingbox, ограничивающая рамка) объекта на изображении. Результат сохраняется в переменной Z .

Для данного скрипта используем сформированную выборку из фотографий автомобилей. Результат работы скрипта находится на рис. 4.

Расстояние, которое было рассмотрено: 1,7 м, 2 м, 3 м, 5 м. Расстояние, которое определила программа на 1-й фотографии, составило $-1,71$ м, угол отклонения по оси X от центра $-27,26^\circ$, угол отклонения по оси Y от центра $-1,14^\circ$, что удовлетворяет изменившемуся расстоянию на дороге.

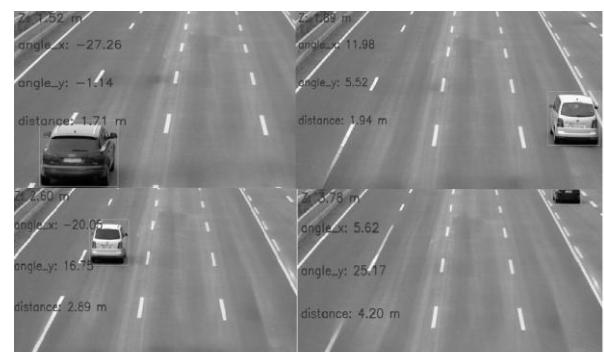


Рис. 4. Пример работы программы

Fig. 4. Example of Program Operation

При увеличении расстояния до 2 метров программа определила 1,94 м. Углы отклонения по осям X и Y составили $11,98^\circ$ и $5,52^\circ$ соответственно.

Абсолютная погрешность оказалась небольшой (0,06 м), а относительная погрешность составила 3 %. Данная процентная погрешность считается допустимой, хотя и не оптимальной, так как отражает отношение ошибки к измеряемой величине.

При дистанции в 3 метра программа показала результат 2,89 м. Углы отклонения составили – 20,05° по X и 16,75° по Y. Абсолютная погрешность возросла до 0,11 м, а процентная – до 3,67 %.

При увеличении расстояния до 5 метров измеренное значение составило 4,20 м. Углы отклонения по X и Y: 5,62° и 25,17° соответственно. Абсолютная погрешность значительно увеличилась до 0,80 м, а процентная погрешность достигла 16 %.

Наблюдается тенденция к увеличению абсолютной и процентной погрешностей с увеличением дистанции. Особенно заметно ухудшение точности при измерении на расстоянии 5 метров.

Заключение

Полученные результаты исследования выявили значительную зависимость между расстоянием до объекта и точностью его определения программой. Анализ продемонстрировал, что с увеличением дистанции наблюдается непропорциональное увеличение погрешности в оценке расстояния. Данный факт свидетельствует о критическом ограничении текущей реализации алгоритмов обработки изображений, используемых для определения расстояния, особенно при работе с объектами, находящимися на значительном удалении.

Снижение точности измерений на больших дистанциях может быть обусловлено рядом факторов. Искажения, возникающие в процессе захвата и обработки изображения, усиливаются с увеличением расстояния, приводя к неточностям в расчетах. Наконец, разработанный алгоритм может быть недостаточно адаптирован к особенностям обработки изображений удаленных объектов, что приводит к систематическим погрешностям.

Учитывая вышеизложенное, крайне важно направить дальнейшие усилия на оптимизацию алгоритма обработки изображений с целью повышения надежности определения расстояния, особенно в условиях больших дистанций. Это может потребовать пересмотра данного подхода и перейти к рассмотрению нового метода, способного более точно учитывать факторы, влияющие на качество и интерпретацию изображений.

Библиографические ссылки

1. Агафонов А. А., Юмаганов А. С. Сравнение методов детектирования трёхмерных объектов в задаче автономного вождения транспортных средств // Информационные технологии и нанотехнологии : сборник трудов по материалам VI Международной конференции и молодежной школы. 2020. С. 277–284.
2. Ковалевский В. А., Волхонский В. В. Методы повышения качества изображения для работы системы видеонаблюдения в условиях недостаточной освещаемости // Известия высших учебных заведений. Приборостроение. 2021. № 10 (64). С. 811–816.
3. Истомин В. И., Привалов А. Н. К вопросу применяемости методов монокулярной оценки глубины для 3D-реконструкции геометрии документов // Известия Тульского государственного университета. 2024. Вып. 9. С. 401–404.

4. Паутова Л. В. Исследование и оптимизация стерео-компьютерного метода формирования изображений и разработка на его основе устройства для диагностики параметров стереоскопического зрения : дисс. ... канд. техн. наук, 2003. 164 с.

5. Использование машинного обучения для управления навигацией робота по показаниям датчиков расстояний / Н. А. Ломовцева, Ю. Е. Кувайская, И. Н. Карпунина, В. Н. Клячкин, Д. М. Ястребов // Информационные технологии и нанотехнологии : сборник трудов по материалам Международной конференции и молодежной школы. 2022. Т. 5. 50252–50581.

6. Бархум М., Пыркин А. А. Использование монокулярной оптики при оценке глубины объектов для двумерного картирования моделируемой среды // Научно-технический вестник информационных технологий, механики и оптики. 2024. № 1. С. 118-122.

7. Зайцев А. К. Автономное управление транспортным средством с использованием Lidar // Вестник науки. 2023. Т. 1, № 6 (63). С. 873–887.

8. Godard, C., Aodha, O. J., & Brostow, G. J. (2019). Diving Into Self-Supervised Monocular Depth Estimation. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), c. 2803-2812.

9. Ranftl, R., Lasinger, K., & Hafner, D. (2022). Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer. IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(3), 1623 - 1637 с.

10. Лупенко Н. Л., Богуш Р. П., Чен Х. Анализ методов определения абсолютного расстояния до объекта по изображению с одной видеокамеры с использованием нейронных сетей // Вестник полоцкого государственного университета. 2024. № 2 (43). С. 24–32.

11. Khan, F., Salahuddin, S., & Javidnia, H. (2020). Deep Learning-Based Monocular Depth Estimation Methods—A State-of-the-Art Review. Sensors, 20(8), 1- 16 с.

12. Тихонов М. К. Интеграция федеративного обучения и yolov11 для обнаружения объектов в автономных транспортных средствах // Научный результат. Информационные технологии. 2024. № 9. С. 58–64.

13. Skibicki, J., Golijanek-Jędrzejczyk, A., & Dzwonkowski, A. (2020). The Influence of Camera and Optical System Parameters on the Uncertainty of Object Location Measurement in Vision Systems. Sensors, 20(108), 80-233 с.

14. Devo, A., Mao, J., Costante, G., Loianno, G., & Authors, A. (2022). Autonomous Single-Image Drone Exploration With Deep Reinforcement Learning and Mixed Reality. IEEE Robotics and Automation Letters, 7(2), 5031-5038 с.

15. Zhu, Z., Wu, F., Sun, W., & Wu, Q. (2025). Depth Estimation Based on MMwave Radar and Camera Fusion with Attention Mechanisms and Multi-Scale Features for Autonomous Driving Vehicles. Electronics, 14(2), 1-17 с.

References

1. Agafonov A.A., Yumaganov A.S. Sravnenie metodov detektirovaniya tryohmernykh ob'ektov v zadache avtonomnogo vozhdeniya transportnykh sredstv [Comparison of 3D object detection methods for autonomous driving applications]. Informacionnye tekhnologii i nanotekhnologii : sbornik trudov po materialam VI Mezhdunarodnoj konferencii i molodezhnoj shkoly [Information technology and nanotechnology: a collection of papers based on the materials of the VI International Conference and Youth School]. 2020. Pp. 277-284 (in Russ.).

2. Kovalevskij V.A., Volhonskij V.V. [Methods for improving image quality for video surveillance systems operating in low light conditions]. Izvestiya vysshih uchebnyh zavedenij. Priborostroenie. 2021. No. 10. Pp. 811-816 (in Russ.).

3. Istomin V.I., Privalov A.N. [On the applicability of monocular depth estimation methods for 3D reconstruction of document geometry]. *Izvestiya Tul'skogo gosudarstvennogo universiteta*. 2024. Vol. 9. Pp. 401-404 (in Russ.).
4. Pautova L.V. *Issledovanie i optimizaciya stereokomp'yuternogo metoda formirovaniya izobrazhenij i razrabotka na ego osnove ustrojstva dlya diagnostiki parametrov stereoskopicheskogo zreniya* [Research and optimization of a stereocomputer method for image formation and development of a device based on it for diagnosing stereoscopic vision parameters]: PhD thesis, 2003. 164 p. (in Russ.).
5. Lomovceva N.A., Kuvajskova Yu.E., Karpunina I.N., Klyachkin V.N., Yastrebov D.M. *Ispol'zovanie mashinnogo obucheniya dlya upravleniya navigacij robot po pokazaniyam datchikov rasstoyanij* [Using machine learning to control robot navigation based on distance sensor readings]. *Informacionnye tekhnologii i nanotekhnologii: sbornik trudov po materialam Mezhdunarodnoj konferencii i molodezhnoj shkoly* [Information technology and nanotechnology: a collection of papers based on the materials of the International Conference and Youth School]. 2022. Vol. 5. 50252–50581 (in Russ.).
6. Barhum M., Pyrkin A.A. [Using monocular optics to estimate the depth of objects for two-dimensional mapping of a simulated environment]. *Nauchno-tehnicheskiy vestnik informacionnyh tekhnologij, mehaniki i optiki*. 2024. No. 1. Pp. 118-122 (in Russ.).
7. Zajcev A.K. [Autonomous vehicle control using Lidar]. *Vestnik nauki* Вестник науки. 2023. Vol. 1, no. 6. Pp. 873-887 (in Russ.).
8. Godard, C., Aodha, O. J., & Brostow, G. J. (2019). Digging Into Self-Supervised Monocular Depth Estimation. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), c. 2803-2812.
9. Ranftl, R., Lasinger, K., & Hafner, D. (2022). Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(3), 1623 - 1637 c.
10. Lupenko N.L., Bogush R.P., Chen H. [Analysis of methods for determining the absolute distance to an object from an image from a single video camera using neural networks]. *Vestnik polockogo gosudarstvennogo universitetaa*. 2024. № 2 (43). C. 24–32.
11. Khan, F., Salahuddin, S., & Javidnia, H. (2020). Deep Learning-Based Monocular Depth Estimation Methods-A State-of-the-Art Review. *Sensors*, 20(8), 1- 16 c.
12. Tihonov M.K. [Integrating Federated Learning and yolov11 for Object Detection in Autonomous Vehicles]. *Nauchnyj rezul'tat. Informacionnye tekhnologii*. 2024. No. 9. Pp. 58-64 (in Russ.).
13. Skibicki, J., Golijanek-Jędrzejczyk, A., & Dzwonkowski, A. (2020). The Influence of Camera and Optical System Parameters on the Uncertainty of Object Location Measurement in Vision Systems. *Sensors*, 20(108), 80-233 c.
14. Devo, A., Mao, J., Costante, G., Loianno, G., & Authors, A. (2022). Autonomous Single-Image Drone Exploration With Deep Reinforcement Learning and Mixed Reality. *IEEE Robotics and Automation Letters*, 7(2), 5031-5038 c.
15. Zhu, Z., Wu, F., Sun, W., & Wu, Q. (2025). Depth Estimation Based on MMwave Radar and Camera Fusion with Attention Mechanisms and Multi-Scale Features for Autonomous Driving Vehicles. *Electronics*, 14(2), 1-17 c..

* * *

Object Distance Determination by Means of Technical Vision

A. V. Koluzov, Head of Laboratory, Military Innovation Technopolis «ERA», Anapa, Russia
 K. S. Korneev, Junior Researcher, Military Innovation Technopolis «ERA», Anapa, Russia
 A. S. Goy, Senior Operator, Military Innovation Technopolis «ERA», Anapa, Russia

The purpose of this paper is to evaluate the effectiveness and accuracy of an integrated approach for object distance estimation based on a camera model and machine learning libraries in the context of advanced driver assistance systems (ADAS) and autonomous driving.

This paper examines how camera model parameters (sensor resolution, focal length, field of view, and aperture characteristics) influence the quality of input data for algorithms implemented using machine learning libraries and, consequently, the overall distance estimation accuracy. Particular attention is paid to the practical application of modern machine learning algorithms, including deep neural networks, to solve the problem of monocular depth estimation from a single image.

This study aims to identify the potential dependence of distance estimation accuracy on object distance. In particular, it will analyze how distance increasing can lead to a disproportionate increase in estimation error, which will help identify potential limitations of current implementations of machine learning algorithms when working with distant objects. Factors, contributing measurement accuracy reduction at long distances will be considered, such as increased distortion during image capture and processing, as well as issues of adapting algorithms to the characteristics of data from remote objects.

Keywords: camera parameters, image processing algorithms, distance determination accuracy, computer vision, stereo vision, focal length, monocular depth, aperture, exposure time, image segmentation, algorithm testing.

Получено: 30.07.25

Образец цитирования

Колузов А. В., Корнеев К. С., Гой А. С. Определения расстояния до объекта с использованием технического зрения // Интеллектуальные системы в производстве. 2025. Т. 23, № 4. С. 33–39. DOI: 10.22213/2410-9304-2025-4-33-39.

For Citation

Koluzov A.V., Korneev K.S., Goy A.S. [Object distance determination by means of technical vision]. *Intellektual'nye sistemy v proizvodstve*. 2025, vol. 23, no. 4, pp. 33-39 (in Russ.). DOI: 10.22213/2410-9304-2025-4-33-39.