

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА БИНОКУЛЯРНОГО ФАСЕТОЧНОГО ЗРЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ АЗИМУТА И РАССТОЯНИЯ ДО ОБЪЕКТА НА ПЛОСКОСТИ

© 2025 К.Н. Белов¹, Э.А. Бибикина^{1,2}, И.В. Булдашев¹, Н.Д. Кундикова^{1,2},
Ю.В. Мухин^{1,2}, А.Н. Николаев¹, А.В. Портнов¹, Я.М. Ридный¹,
Л.Б. Соколинский¹, А.Е. Старков¹, А.А. Шульгинов¹

¹Южно-Уральский государственный университет
(454080 Челябинск, пр. им. В.И. Ленина, д. 76),

²Институт электрофизики УрО РАН
(620016 Екатеринбург, ул. Амундсена, д. 106)

E-mail: belovkn@susu.ru, bibikovaea@susu.ru, buldashviv@susu.ru, kundikovand@susu.ru,
mukhinyv@susu.ru, nikolaevan@susu.ru, portnovav@susu.ru, imridnyi@susu.ru,
leonid.sokolinsky@susu.ru, starkovae@susu.ru, shulginovaa@susu.ru

Поступила в редакцию: 10.03.2025

Статья посвящена прототипу системы искусственного бинокулярного зрения для определения азимута и расстояния до объекта на плоскости с использованием искусственной нейронной сети. Дается критический обзор современных систем определения расстояния и азимута на основе активных и пассивных сенсоров. Предлагается интеллектуальная система бинокулярного зрения, представляющая собой пассивный оптический датчик, позволяющий определять азимут и расстояние до круглого объекта произвольного размера, излучающего в видимом или инфракрасном диапазонах электромагнитного спектра. Рассматривается общая архитектура системы фасеточного зрения. Основными структурными элементами системы являются: оптический модуль, аппаратно-программный контроллер и нейросетевой модуль. Оптический модуль с помощью пары объективов преобразует световой сигнал от объекта в два пиксельных Фурье-изображения, которые поступают на вход аппаратно-программного контроллера. Контроллер выполняет первичную обработку пиксельных Фурье-изображений и преобразует их в две битовые маски, элементы которых соответствуют отдельным фасеткам (каждая фасетка интегрирует четыре смежные колонки пиксельного изображения). Полученные битовые маски поступают в нейросетевой модуль, который на основе их анализа определяет координаты объекта в виде расстояния и азимута.

Ключевые слова: фасеточное зрение, оптическая модель, определение расстояния и азимута, ПЗС-матрица, нейросетевая модель, прототип.

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Белов К.Н., Бибикина Э.А., Булдашев И.В. и др. Интеллектуальная система бинокулярного фасеточного зрения для определения азимута и расстояния до объекта на плоскости // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2025. Т. 14, № 1. С. 5–29. DOI: 10.14529/cmse250101.

Введение

Одной из важных функций роботизированных систем, используемых в различных областях техники, является определение положения объекта относительно наблюдателя. Для этой цели в зависимости от поставленных задач может быть достаточно датчика определения расстояния и азимута. Сенсорные системы для определения расстояния можно разделить на активные или пассивные [1].

Активный датчик излучает акустические или электромагнитные волны, которые отражаясь от объекта, попадают на приемник данного датчика. Анализ параметров волн, приобретенных при отражении, позволяет определить расстояние до объекта, скорость сближения, относительные координаты объекта. Существуют три основных вида таких датчиков.

Первым способом определения расстояния активным датчиком является, так называемый «времяпролетный» метод или метод dToF (direct time of flight), основанный на прямом измерении времени, которое проходит между отправкой сигнала и его приемом [2]. Данный метод может быть использован для зондирующих волн разной природы, генерируемых в импульсном режиме. Однако при использовании электромагнитных волн датчику требуется высокоскоростные, высокочувствительные приемники, быстродействующая электроника и источники коротких мощных импульсов излучения. Преимуществом метода является малое время излучения датчика, ввиду испускания редких и коротких импульсов, что при необходимости делает его хоть и не до конца, но достаточно скрытным, к тому же согласование времени работы приемника и источника приводит к возможности использования в одном месте большого количество устройств, не мешающих друг другу, и к исключению внешнего влияния, вносящее помехи.

Активные датчики расстояния, основанные на втором широко известном методе, называемом фазовым или iToF (indirect time of flight), испускают синусоидальный модулированный непрерывный электромагнитный сигнал и определяют сдвиг фаз между излученной и отраженной от объекта волнами [3]. Диапазон и точность измерения расстояния зависит от частоты модуляции сигнала. Для расширения этого диапазона может применяться волны с модуляцией на нескольких частотах. Недостатком этого метода является высокие требования на измерительную часть датчика, которая позволяла бы измерять сдвиг фаз волны модуляции с высокой точностью. К тому же датчик дает усредненную информацию о расстоянии до всех объектов, расположенных на пути волны, эта проблема может быть решена использованием узкого луча и сканированием пространства по углу. Датчик функционирует в режиме синхронного детектирования, то есть его приемная часть работает на частоте модуляции источника, и тем самым обеспечивается высокая помехозащищенность. Использование непрерывного излучения делает его заметным стороннему наблюдателю, затрудняет использование множества таких датчиков в одном месте и порождает высокое электромагнитное загрязнение пространства, что ограничивает сферы его применения. Следует отметить, что на обоих вышеперечисленных принципах (dToF и iToF) основана также работа 3D-приемников изображения [4, 5].

Третьим способом определения расстояния до объекта с помощью активного датчика является метод триангуляции [6, 7]. Детектор датчика, устанавливается на некотором расстоянии от его источника излучения, в результате вместе с объектом они образуют вершины треугольника (базисные точки). Детектор представляет собой ПЗС-матрицу, на которую с помощью линзы фокусируется отраженный от объекта свет. Место на матрице, где зафиксирована фокальная перетяжка, зависит от направления падающего на нее света. Расстояние до объекта и его расположение рассчитываются исходя из принципов геометрии. Излучает датчик как в импульсном, так и в непрерывном режимах, и имеет все соответствующие вышеперечисленные недостатки. К преимуществам метода относится высокая скорость обнаружения объекта и точность измерения.

Таким образом активные датчики измерения расстояния имеют два основных недостатка: во-первых, они потребляют дополнительную энергию для генерации излучения, а

во-вторых, их работа может быть обнаружена внешним наблюдателем, что не всегда допустимо. К тому же, если объект изготовлен из материалов, не позволяющих отражать излучение, испускаемое датчиком, он им не различим. Пассивный датчик лишен этих недостатков. Главное преимущество пассивного датчика измерения расстояния до объекта — это то, что он не излучает сигнал, а использует излучение, испущенное самим объектом, или излучение естественного происхождения (фоновое), отраженное или рассеянное от объекта. Разрешение приборов, работающих для оптического диапазона излучения выше, чем для радио диапазона, благодаря более короткой длине волны. Пассивные датчики для оптического диапазона излучения более предпочтительны в использовании, так как для оптической локации сложнее создать преднамеренные помехи, чем для радиолокации. Основным элементом пассивных оптических датчиков является цифровая камера регистрации изображений, ПЗС-матрица, которая используется в совокупности с оптической линзовой системой.

Один из простейших принципов оценки расстояния до приближающегося или отдаляющегося от камеры объекта является анализ информации, полученной на последовательно сделанных кадрах, например, оценка изменения размера изображения объекта. Этот подход требует дополнительной информации: начальное расстояние до объекта R_0 , его начальный размер A_0 . Текущее расстояние определяется через актуальный на данный момент размер объекта A по формуле: $R_A = R_0 \sqrt{A/A_0}$. Этот метод применим в случае достаточно крупного объекта (его изображение должно быть представлено большим количеством пикселей ПЗС-матрицы). Также на кадрах может оцениваться изменение интенсивности I излучения ближнего инфракрасного диапазона, исходящего от движущегося объекта и поглощенного в атмосфере по закону Бугера—Ламберта—Бера. Расстояние до объекта рассчитывается по формуле:

$$R_I = R_0 + \frac{1}{\sigma} \ln \frac{I_0}{I},$$

где R_0 и I_0 — начальное расстояние до объекта и интенсивность излучения на начальном кадре, σ — коэффициент поглощения атмосферы. Этот метод в отличие от предыдущего хорошо подходит для оценки дальности именно точечной цели. Еще один метод предполагает оценку изменения контраста изображения объекта, испускающего инфракрасное излучение. Изменение контраста на кадрах также зависит от поглощения ИК-излучения в атмосфере, поэтому для расчета расстояния до объекта используется схожее выражение:

$$R_C = R_0 + \frac{1}{\sigma} \ln \frac{C_0}{C}.$$

Здесь R_0 и C_0 — начальное расстояние до объекта и контраст изображения на начальном кадре, σ — коэффициент поглощения атмосферы, C — контраст изображения на текущем кадре. Контраст можно рассчитать по формуле: $C = (L_t - L_b)/L_b$, где L_t — значение яркости изображения объекта, а L_b — яркость фона. Случайные изменения яркости фона создают дополнительные погрешности при измерении расстояния этим методом, также когда объект близок к камере возможен эффект ее насыщения, как и в случае метода измерения расстояния от объекта по интенсивности его излучения. Эти три пассивных монокулярных метода ограничены в условиях применения, также требуют знания начальных условий, что не всегда возможно [8].

В исследованиях [9] предложена модель монокулярного зрения для определения трехмерного положения овальных и сферических объектов с известными размерами. Подобная

модель использует двумерное изображение, являющееся перспективной проекцией объекта, и эффективное фокусное расстояние камеры для определения местоположения объекта относительно камеры. Описанный метод может быть обобщен для трехмерных квадратичных объектов, таких как эллипсоид, параболоид, гиперболоид и цилиндроид, но не для объектов произвольной формы. Кроме того, этот метод обладает относительно высокой вычислительной сложностью и не может обеспечить достаточную точность измерения расстояния до наблюдаемого объекта, к тому же нужно заранее знать размеры объекта.

Также одним из простейших методов измерения расстояния до объекта и его азимута на основе пассивных оптических датчиков является измерение расстояния по пеленгу [10]. Этот метод позволяет вычислять расстояние d_0 до объекта от камеры, расположенной на известной высоте h_c от земли, по формуле $d_0 = \operatorname{tg}\varphi \cdot (h_c - h_0)$, где φ — угол наклона камеры и h_0 — высота объекта от земли. Такая модель достаточно проста, но применима только в том случае, если $h_c > h_0$ и $\varphi < 90^\circ$. Кроме того, необходимо знать высоту объекта, что не всегда выполнимо на практике. Таким образом, этот метод неприменим, когда высота расположения камеры сопоставима с высотой объекта или объект выше камеры.

Еще один метод пассивной локации, который также можно назвать монокулярным, основан на законах геометрической оптики, принцип действия заключается в том, что изображение объекта строится с помощью фокусирующей линзы с фокусным расстоянием F на ПЗС-матрице, которая передвигается на небольшое расстояние вдоль оптической оси линзы [11]. Задача заключается в поиске расположения плоскости, где изображение обладает наибольшей четкостью, то есть в определении расстояния d между линзой и матрицей, расположенной в этой плоскости. Тогда расстояние от линзы до объекта D можно рассчитать по формуле тонкой линзы: $1/F = 1/D + 1/d$, откуда $D = (d \cdot F)/(d - F)$. В качестве оптической системы может использоваться и более сложная система линз. Подобным образом можно также рассчитать расстояние при условии смещения объекта относительно оптической оси системы. Оптическая система должна обладать малой глубиной фокуса. Анализ изображений, полученных на разных расстояниях от линзы, с целью определения наиболее четкого из них занимает время и вычислительные мощности [11, 12]. К тому же датчик содержит движущиеся части, что также влияет на быстродействие, его надежность в процессе эксплуатации и делает его использование малоэффективным.

Схожий принцип определения расстояния до объекта, но обладающий большей точностью и быстродействием, описан в работе [13], в нем используется линзовая система с переменным фокусным расстоянием с известными граничными фокусными расстояниями f_1 и f_2 . Получаются два изображения объекта, обладающих размерами y_1 и y_2 , соответствующими этим фокусным расстояниям. Расстояние от линзовой системы до объекта находится по формуле:

$$D = f_2 \cdot \left(1 - \frac{y_1}{y_2}\right) / \left(\frac{y_1}{y_2} \cdot \frac{f_1}{f_2} - 1\right).$$

Другой пассивный метод измерения расстояния до объекта заключается в использовании пленоптической камеры. Пленоптическая камера или камера светового поля представляет собой массив микрокамер и регистрирует не распределение интенсивности света в плоскости ПЗС-матрицы, а векторное световое поле [14]. Между объективом камеры и ПЗС-матрицей располагается растр микролинз, каждая из которых строит на поверхности матрицы изображение, даваемое основным объективом под своим ракурсом. При анализе комплекса полученных изображений возможно создание полной картины векторного светового поля, содержащей информацию о направлении и интенсивности световых пучков в

пространстве изображений объектива. Недостатком является значительное снижение эффективного разрешения изображения. В работе [15] предложена оптико-геометрическая модель для измерения расстояния до объекта с использованием пленочной камеры. Расстояние d_{out} между основной линзой и объектом можно рассчитать по следующей формуле:

$$d_{\text{out}} = \frac{D}{2 \tan \varphi} + \sqrt{R^2 + D^2/4} - R + T/2,$$

где R , T и D — параметры основной линзы: радиус кривизны, толщина и диаметр зрачка соответственно, угол φ рассчитывается по формуле: $\varphi = \arcsin(n_1 \sin \psi) - \arcsin(D/2R)$, где n_1 — показатель преломления основной линзы, ψ — угол преломления лучей в основной линзе. Описанный метод также требует сложного процесса калибровки камеры [16], не обеспечивает высокую точность измерений больших расстояний [17] и плохо подходит для автономных мобильных роботов.

Распространенным методом определения расстояния до объекта и его азимута является пассивная оптическая триангуляция [18]. Расстояние до объекта может быть рассчитано с помощью принципов геометрии на основе визуальной информации, полученной из стереоизображения: пары изображений, полученных при различных ракурсах объекта. Для этого может быть использована одна ПЗС-матрица ориентированная по-разному в процессе съемки или одновременно две матрицы. Управление такой стереокамерой требует точной настройки таких параметров, как панорамирование, наклон, масштабирование, фокусировка, а также отслеживания перемещений движущегося объекта [19]. Количество используемых матриц может быть увеличено для повышения точности измерения и расширения видимой области поиска объекта.

В последнее время стало популярным использование нейросетей при работе датчиков. Это позволяет автоматизировать процесс распознавания объекта, который можно совместить с последующим автоматическим решением задач по использованию этой информации, что особенно актуально для роботизированных систем.

В настоящей работе реализована интеллектуальная система бинокулярного зрения, представляющая собой пассивный оптический датчик, позволяющий определять расстояние до круглого объекта произвольного размера, излучающего в видимом или инфракрасном диапазонах электромагнитного спектра, а также его азимут. Основным элементом этой системы является устройство, представляющее собой виртуальный фасеточный глаз, то есть его оптический аналог.

1. Архитектура системы фасеточного зрения

Обобщенная структурная схема интеллектуальной системы бинокулярного фасеточного зрения приведена на рис. 1. Основными структурными элементами системы являются: оптический модуль, аппаратно-программный контроллер и нейросетевой модуль. Оптический модуль с помощью пары объективов преобразует световой сигнал от объекта в два пиксельных Фурье-изображения, которые поступают на вход аппаратно-программного контроллера. Контроллер выполняет первичную обработку пиксельных Фурье-изображений и преобразует их в две битовые маски, элементы которых соответствуют отдельным фасеткам (каждая фасетка интегрирует четыре смежные колонки пиксельного изображения). Полученные битовые маски поступают в нейросетевой модуль, который на основе их анализа определяет координаты объекта в виде расстояния и азимута.



Рис. 1. Архитектура системы фасеточного зрения

2. Оптический модуль

В данном разделе дается обоснование возможности использования в качестве фасеточного глаза одиночной линзы или объектива.

Общепринято считать, что фасеточные (составные) глаза могут обеспечить лучшее временное разрешение в системах, которые регистрируют движение объекта. Хотя в живой природе такие глаза состоят из набора отдельных линзовых элементов — фасеток, функции такого глаза можно реализовать с помощью одной-единственной линзы, воспользовавшись тем фактом, что линза осуществляет Фурье-преобразование из координатной области в область направлений. Рассмотрим выражения, необходимые для расчета основных параметров оптического модуля, в котором в качестве фасеточного глаза используется одна линза и одна ПЗС-матрица. На рис. 2 приведена схема такого виртуального фасеточного глаза.

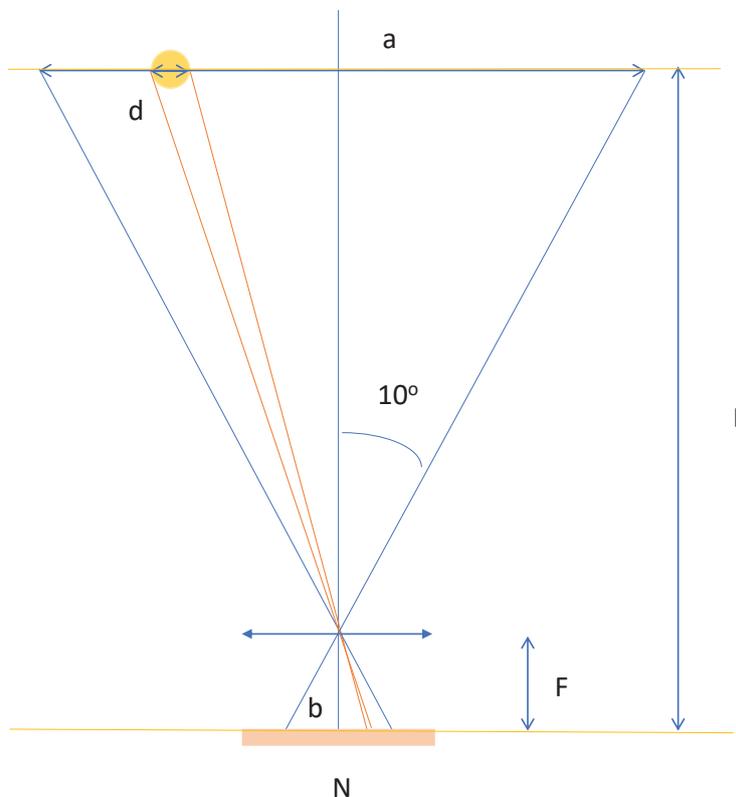


Рис. 2. Оптическая схема виртуального фасеточного глаза. L — расстояние от плоскости движения объекта до плоскости ПЗС-матрицы, F — фокусное расстояние линзы, b — размер рабочей области матрицы, N — количество пикселей матрицы по горизонтали, d — размер объекта

Рассмотрим основные соотношения между геометрическими размерами виртуального фасеточного глаза. Половина угла обзора α определяется половиной рабочей области матрицы $b/2$ и фокусным расстоянием линзы F , а именно, $\operatorname{tg}\alpha = b/(2F)$. Размер рабочей области матрицы b зависит от числа пикселей N и их размера l , а именно, $b = l \cdot N$, таким образом, половина угла обзора однозначно определяется параметрами приемника и фокусным расстоянием линзы:

$$\operatorname{tg}\alpha = \frac{l \cdot N}{2F}. \quad (1)$$

Рассмотрим связь между размером объекта и его размером, регистрируемым в фокальной плоскости линзы. Размер объекта на матрице, то есть в фокальной плоскости линзы, M определяется углами видимости краев объекта β и γ , а именно, $M = F(\operatorname{tg}\beta - \operatorname{tg}\gamma)$, учитывая, что

$$\operatorname{tg}\beta = \frac{D + d/2}{L - F} \quad (2)$$

и

$$\operatorname{tg}\gamma = \frac{D - d/2}{L - F}, \quad (3)$$

где d — размер объекта, D — положение центра объекта, L — расстояние от плоскости движения объекта до плоскости ПЗС-матрицы, окончательно получаем:

$$M = \frac{d \cdot F}{L - F}. \quad (4)$$

Размер изображения в пикселях m равен $m = M/l$. Если считать, что минимальное количество омматидиев, приходящихся на объект, равно трем, то количество пикселей на один омматидий равно $K = m/3$, тогда

$$K = \frac{d \cdot F}{3l \cdot (L - F)}. \quad (5)$$

Оценим основные параметры для двух виртуальных фасеточных глаз (далее — «глаз»), параметры которых приведены в табл. 1, размера объекта $d = 1.0$ см, половины угла обзора $\alpha = 10^\circ$ и количества омматидиев $NN = N/K = 82$.

Таблица 1. Параметры двух «глаз»

	Первый «глаз»	Второй «глаз»
N , пикс.	1024	1600
l , мкм	4.65	2.2
b , см	0.48	0.35
F , см	1.20 ± 0.01	0.80 ± 0.01

Если задан половинный угол обзора для первого «глаза», то для второго «глаза» активными будут не все пиксели матрицы, а только их часть, определяемая выражением (1):

$$N_2 = 2F \frac{\operatorname{tg}\alpha}{l}. \quad (6)$$

Если количество омматидиев $NN = 82$, то в среднем на один омматидий приходится $K = N/NN$ пикселей, или

$$K = 2F \frac{\operatorname{tg}\alpha}{NN \cdot l}. \quad (7)$$

Размер объекта в пикселях в соответствии с выражением (4) равен

$$m = \frac{M}{l} = \frac{d \cdot F}{l(L - F)}. \quad (8)$$

В табл. 2 приведены рассчитанные значения размера объекта в пикселях и количество пикселей на один омматидий для двух «глаз».

Таблица 2. Рассчитанные значения m и K

	Первый «глаз»	Второй «глаз»
m , пикс.	36	50
K , пикс.	11	16

Использование двух одинаковых однолинзовых фасеточных «глаз» позволяет определить положение объекта в декартовых координатах, а также расстояние и азимут. На рис. 3 представлена оптическая схема, необходимая для расчета положения объекта y_0 и x_0 , если известно расстояние от центра каждой матрицы до изображения объекта в фокальной плоскости на левой (a_L) и правой (a_R) ПЗС-матрицах, а также расстояние между центрами матриц $2b$ и фокусное расстояние используемой линзы. Простейшие тригонометрические

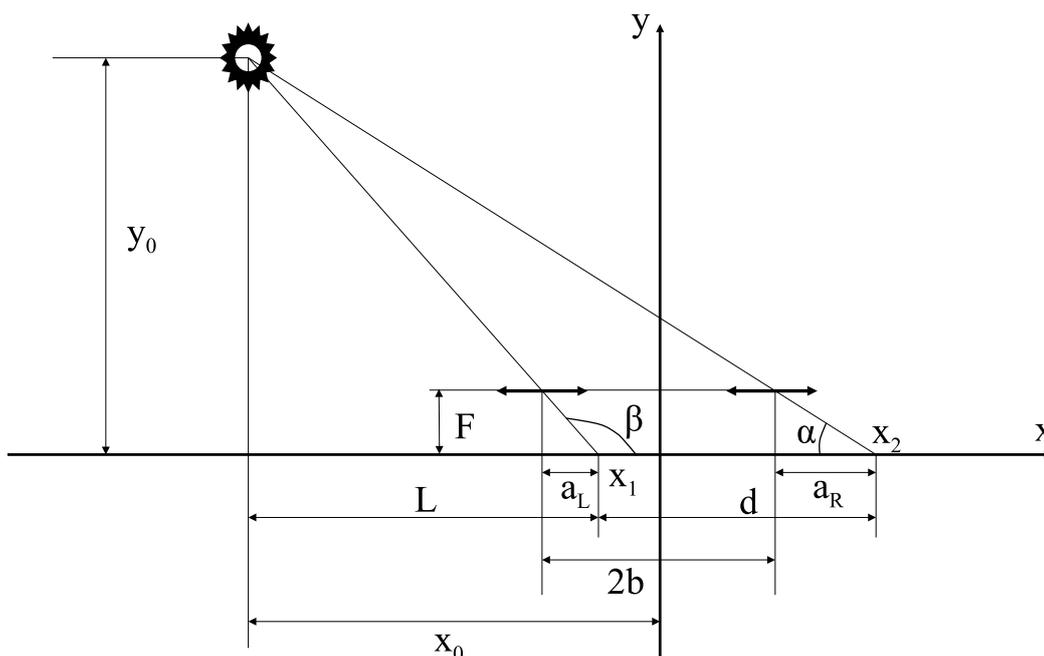


Рис. 3. Оптическая схема бинокулярного зрения на основе виртуальных фасеточных «глаз»

преобразования приводят к следующим значениям

$$x_0 = b \frac{a_R + a_L}{a_R - a_L}, \quad y_0 = F \frac{2b + a_R - a_L}{a_R - a_L}. \quad (9)$$

Выражения (9) позволяет определить расстояние до объекта и азимутальный угол по известным из тригонометрии соотношений.

3. Аппаратно-программный блок

Аппаратно-программный блок решает задачи первичной обработки информации в макете и обеспечивает интерфейс между оптическим и нейросетевым модулем. Основой аппаратно-программного блока является специализированный контроллер на базе отладочной платы с ПЛИС (программируемая логическая интегральная схема) Intel (Altera) CycloneV (модель 5CEFA5F23I7). Данная ПЛИС обладает логической емкостью 77000 элементов, имеет блочную память объемом 4460 Кб, 6 PLL (Phase-Locked Loop — внутренние генераторы со схемой подстройки частоты), а также 240 пользовательских входов/выходов. Тактовая частота ПЛИС составляет 50 МГц. Логика обработки реализована в виде программного обеспечения ПЛИС, функциональная схема которого представлена на рис. 4.

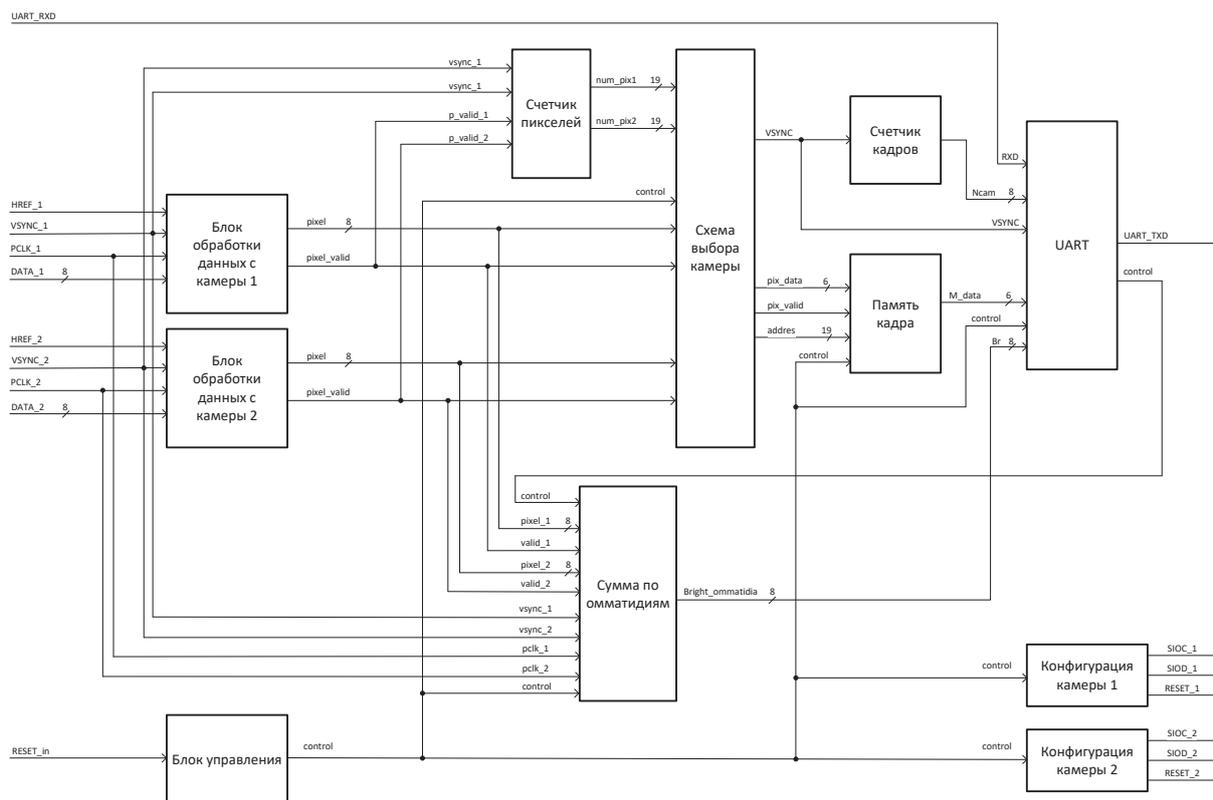


Рис. 4. Функциональная схема контроллера

Контроллер обеспечивает интерфейс информационного обмена с камерами типа OV2640 (или совместимыми) с протоколом обмена SCCB (Serial Camera Control Bus). Используемые камеры имеют формат информационного потока RGB565/555 (также возможны форматы YUV(422/420)/YCbCr422). Максимальное разрешение камер составляет 1600×1200 пикселей. Частота кадров при максимальном разрешении составляет 15 кадров в секунду. Максимальная частота синхронизации пикселей составляет 36 МГц. Контроллер имеет возможность принимать сигналы от оптических модулей, реализует первичную обработку принятых данных, а также производит кодирование выделенной из сигналов информации и передачу информационных посылок в нейросетевой модуль. Для этого реализованы соответствующие блоки управления и обработки. Для настройки камер служат блоки «Конфигурация камеры 1» и «Конфигурация камеры 2». Настройка подразумевает под собой запись необходимых параметров в соответствующие регистры.

Для приема данных с камер служат «Блок обработки данных с камеры 1» и «Блок обработки данных с камеры 2». На выходе данных блоков устанавливаются значения каждого пикселя, а также сигнал подтверждения правильности данных. «Счетчик пикселей» служит для формирования адреса, с помощью которого данные по каждому пикселю записываются в память (блок «Память кадра»). Для выбора камеры, с которой производится запись в память, служит блок «Схема выбора камеры», в котором происходит коммутация входных сигналов. Для контроля передачи кадров служит блок «Счетчик кадров», информация о номере текущего кадра позволяет оценивать количество пропущенных кадров при передаче в нейросетевой модуль. «Блок управления» формирует все сигналы синхронизации. Связь с нейросетевым модулем осуществляется по протоколу UART (Universal Asynchronous Receiver-Transmitter) с помощью блока «UART». Связь с контроллером возможна в обе стороны (настройка необходимых параметров и режима работы контроллера и передача данных для дальнейшей обработки).

Используемая отладочная плата также содержит разъем JTAG (Joint Test Action Group) для программирования ПЛИС и внешнюю память типа SDRAM (Synchronous Dynamic Random Access Memory) общим объемом 64 Мб. В качестве адаптера UART используется внешний модуль на базе преобразователя интерфейсов FT232RL. Органы управления и индикации представлены в виде двух пользовательских тактовых кнопок и светодиодов.

4. Нейросетевой модуль

В данном разделе приведено описание процессов проектирования, разработки и обучения нейросетевого модуля, включающего в себя две искусственные нейронные сети прямого распространения для определения расстояния и азимута до целевого объекта на плоскости, отдельно рассмотрены метрики и способы оценки обученных моделей. Раздел содержит алгоритм генерации обучающих множеств, результаты поиска лучших топологий ИНС и оценку точности полученных моделей.

Бинокулярная модель фасеточного зрения [20] включает в себя два круглых сложных глаза, расположенных симметрично относительно оси y , центры A и B , расположенных на оси x . Расстояние от начала координат до центра каждого глаза равно h . Сложный глаз состоит из m омматидиев, являющихся равнобедренными треугольниками равного размера. Стороны данных треугольников имеют длину l , а основания имеют длину s . Угол между сторонами равен γ .

Алгоритм 1 Generating uniform training data set.

```

1: input  $h, l, m, n, r_{min}, r_{max}, g_{max}$ ;
2:  $\mathcal{M}, DS := \emptyset$ 
3:  $k := 1$ ;
4: for  $i = r_{min} \dots r_{max}$  with step 5 do
5:    $DS := generate(h = h, l = l, m = m, r_{min} = i, r_{max} = i + g_{max},$ 
6:      $g_{max} = g_{max}, n = 1000 + 200 * k)$ 
7:    $\mathcal{M} := \mathcal{M} \cup DS$ 
8:    $k := k + 1$ 
9: end for
10: return  $\mathcal{M}$ .

```

На основе рассмотренной модели фасеточного зрения был разработан алгоритм 1 для параметризуемой генерации аннотированных синтетических обучающих наборов данных. Набор данных $\mathcal{M} \subset \mathbb{R}^2 \times \{0, 1\}^{m/2} \times \{0, 1\}^{m/2}$ состоит из записей вида $(r, \varphi, \Omega_A, \Omega_B)$. Каждая запись соответствует положению наблюдаемого целевого объекта. Значения (r, φ) определяют полярные координаты центра наблюдаемого объекта. Параметр Ω_A представляет собой битовую маску длиной $m/2$, полученную для левого глаза: $\Omega_A[i] = 1$ тогда и только тогда, когда i -й омматидий левого глаза наблюдает целевой объект ($i = 0, \dots, m/2 - 1$). Параметр Ω_B представляет собой битовую маску аналогичной длины ($m/2$), полученную для правого глаза: $\Omega_B[j] = 1$ тогда и только тогда, когда j -й омматидий правого глаза наблюдает целевой объект ($j = 0, \dots, m/2 - 1$). Итоговые характеристики подготовленных наборов данных приведены в табл. 3.

Кратко прокомментируем шаги разработанного алгоритма 1. Шаг 1 осуществляет ввод параметров алгоритма:

- h : расстояние от центра координат до центра сложного глаза;
- l : радиус сложного глаза;
- m : количество омматидиев в глазе;
- n : количество записей в обучающем наборе;
- r_{min} : минимальное расстояние от центра системы до центра целевого объекта;
- r_{max} : максимальное расстояние от центра системы до центра целевого объекта;
- g_{max} : максимальный радиус целевого объекта.

Шаги 2, 3 инициализируют переменные \mathcal{M} , DS и счетчик k , необходимые для равномерной генерации набора данных. Шаги 4–8 реализуют цикл **for**, в котором итерационно генерируется обучающий набор данных \mathcal{M} . На 5 шаге алгоритма формируется промежуточный набор данных DS с помощью метода `generate`, подробно описанного в работе [20]. По завершении цикла **for**, алгоритм возвращает итоговый обучающий набор данных на шаге 9.

Для простоты было принято решение спроектировать две глубокие ИНС прямого пространства: первая для определения расстояния, а вторая для определения азимута целевого объекта. Для поиска оптимального набора гиперпараметров нейронной сети была построена общая гипермодель для обеих сетей. Гипермодель включает входной слой, три скрытых слоя и выходной слой. Все слои являются полносвязными. Для входного слоя и всех скрытых слоев доступны следующие функции активации: $[ReLU, Tanh, Sigmoid]$. Входной слой имеет 1476 нейронов, получающих внешние данные: 738-битное изображение от левого фасеточного глаза и 738-битовое изображение от правого фасеточного глаза. Выходной слой имеет единственный нейрон, дающий конечный результат: расстояние для первой ИНС и азимут для второй ИНС. Для первого скрытого слоя количество нейронов на выбор: $[2048, 3072, 4096, 5040, 5760]$. Для второго и третьего скрытых слоев количество нейронов на выбор: $[3072, 4096, 5120, 5760, 6480, 7200]$ и $[2160, 2880, 3600, 4096, 5120, 6144]$ соответственно.

На основе представленной гипермодели был выполнен ограниченный случайный поиск оптимальных наборов гиперпараметров нейронных сетей с использованием платформы W&B [21]. В качестве алгоритма оптимизации были протестированы SGD (*stochastic gradient descent*) и $RMSProp$ [22]. Размер батча варьировался в пределах от 4 до 64. В качестве функции потерь была использована метрика MAE (*mean absolute error*) [23], рас-

считываемая следующим образом:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - x_i|,$$

где n — число записей в обучающем наборе данных, y_i — результат работы ИНС, x_i — истинное значение (расстояния/азимута).

Таблица 3. Наборы данных и результаты обучения нейросетевых моделей

Набор данных	Параметр	Определение азимута	Определение расстояния
ДС-1 ($g = 5$)	Оптимизатор	RMSProp	SGD
	Batch Size	32	32
	Learning rate	0.00015	0.00073
	MAE	0.0059	9.09
	MAPE	0.403	1.448
	R^2	0.9998	0.9969
	Train size	83000 (80%)	
Test size	25150 (20%)		
ДС-2 ($g = 10$)	Оптимизатор	RMSProp	SGD
	Batch Size	32	32
	Learning rate	0.00023	0.00037
	MAE	0.0079	12.36
	MAPE	0.506	1.951
	R^2	0.99975	0.995
	Train size	20750 (80%)	
Test size	6300 (20%)		

Для обучения и тестирования моделей нейронных сетей использовалась облачная платформа Google Colab [24], предоставляющая вычислительные ресурсы видеокарты $nVidia$ Tesla P4. Для оценки качества полученных нейросетевых моделей были использованы две дополнительные метрики: $MAPE$ (*mean absolute percentage error*) [25], определяемая уравнением

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - x_i}{x_i} \right|,$$

и коэффициент детерминации R^2 [26], представляющий собой

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - x_i)^2}{\sum_{i=1}^n (y_i - y_{mean})^2},$$

где

$$y_{mean} = \frac{1}{n} \sum_{i=1}^n y_i.$$

$MAPE$ часто используется на практике по причине интуитивной интерпретации ошибки с точки зрения относительной погрешности. Коэффициент детерминации дает некоторую информацию о степени соответствия моделей нейронных сетей набору обучающих данных.

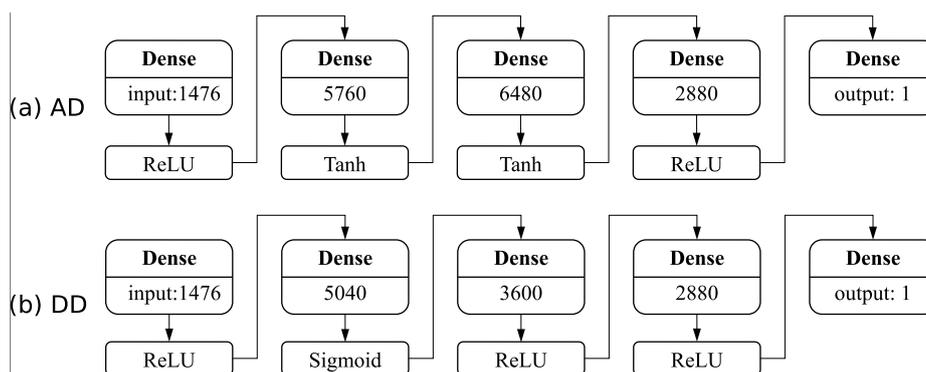


Рис. 5. Топология нейронных сетей для определения а) азимута; б) расстояния до целевого объекта

Для решения задачи определения положения целевого объекта на плоскости разработаны и обучены две нейронные сети прямого распространения для определения азимута и расстояния. Топологии разработанных нейронных сетей приведены на рис. 5. Лучшие результаты обучения разработанных нейронных сетей для различных наборов данных приведены в табл. 3.

5. Экспериментальный прототип системы фасеточного зрения

Фотоприемник системы фасеточного зрения выполнен в виде однолинзового фасеточного глаза. Фасеточный глаз состоит из задаваемого количества омматидиев, являющихся треугольниками имеющими одинаковое основание, т.е. одинаковым количеством пикселей приходящимися на один омматидий. Один омматидий — программно выбираемое количество пикселей ПЗС-матрицы. Для исследования возможности использования однолинзового фасеточного глаза для определения положения объекта была создана экспериментальная площадка, оптическая схема которой была приведена на рис. 2. На начальном этапе работы использовался светодиодный модуль, установленный в держателе на двух взаимно перемещаемых рельсах. В качестве однолинзового фасеточного глаза использовалась коммерчески доступная линза с фокусным расстоянием 1.2 см и ПЗС-матрица ThorLabs DCU223C. Изображения объекта, получаемые на ПЗС-матрице, расположенной в фокальной плоскости линзы, приведены на рис. 6: под буквами а), б), в) фотографии светящегося объекта с разницей в расстоянии 4 см.

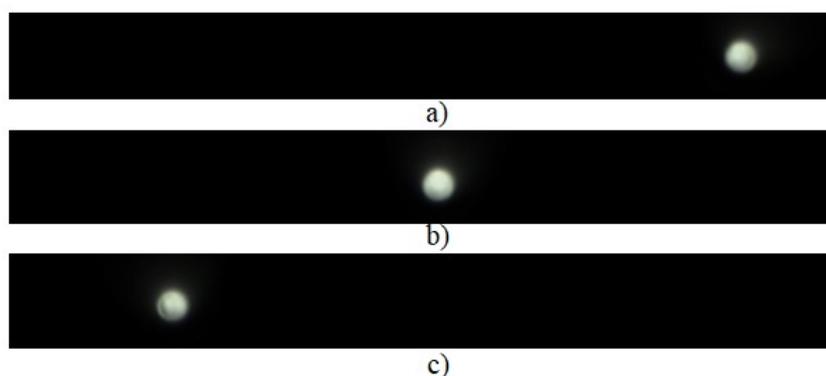


Рис. 6. Изображения объекта, получаемые на ПЗС-матрице, расположенной в фокальной плоскости линзы

Однолинзовый фасеточный глаз непригоден для измерений в реальных условиях. Если бы размер объекта был фиксирован, тогда его можно использовать, но в реальных условиях размеры объекта могут очень сильно отличаться. Для проведения исследований в реальном динамическом режиме был изготовлен прототип движущегося объекта. Эскиз конструкции прототипа показан на рис. 7.

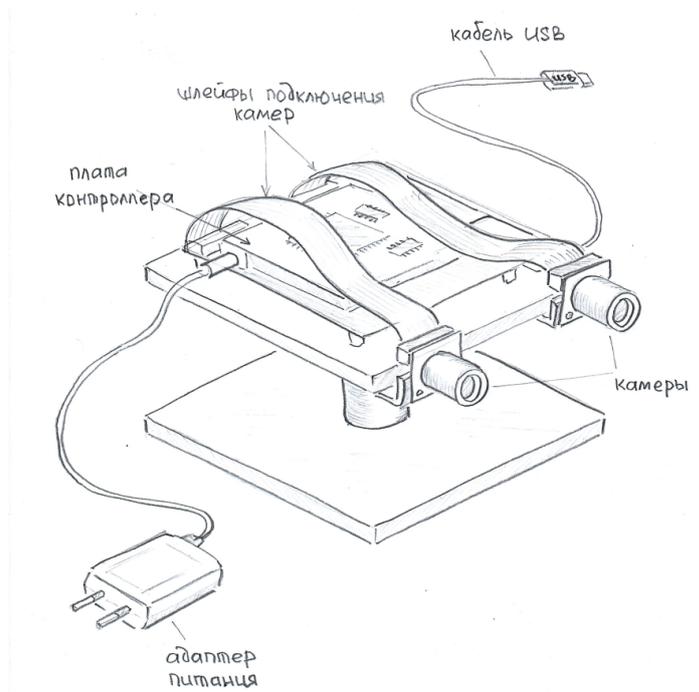


Рис. 7. Эскиз конструкции прототипа системы фасеточного зрения

Конструкция прототипа включает верхнюю металлическую платформу, на которую крепится модуль аппаратно-программного контроллера и оптические модули, нижнюю опорную платформу, опорный штوك с возможностью регулировки высоты установки верхней платформы. В конструкции предусмотрена возможность регулировки положения оптических модулей. Внешний вид изготовленного прототипа приведен на рис. 8.

Объект устанавливается на тележку, которая приводится в движение электрическим двигателем. На рис. 9 приведена оптическая схема установки: а) ноутбук; б) плата «Плис» соединенная с двумя камерами; в) передвижная тележка со светодиодом. Скорость тележки можно регулировать. Движение тележки осуществляется по двутавру длиной 119 см. Свободный ход ~ 100 см. Источник напряжения — батарейка типа «Крона» + DC-DC-преобразователь напряжения.

В качестве объекта использовался светодиод, установленный на стойке. Питание от регулируемого источника напряжения (батарейка типа «Крона» + DC-DC-преобразователь напряжения) позволял регулировать яркость источника излучения. Исследования можно проводить с различными светодиодами, параметры которых представлены в табл. 4. Конструкция экспериментальной демонстрационной площадки позволяла вместо однолинзового фасеточного глаза установить макет системы фасеточного бинокулярного зрения.

Для определения ориентации матриц и юстировки фасеточных глаз макета была разработана специальная оптическая методика. Фасеточные глаза должны быть ориентированы на макете под определенным углом к друг другу. Это означает также, что должны быть

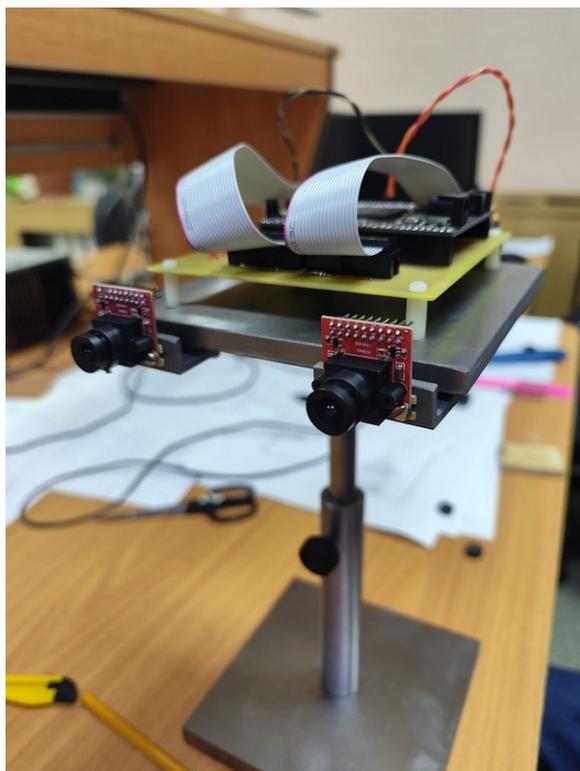


Рис. 8. Внешний вид прототипа системы фасеточного зрения

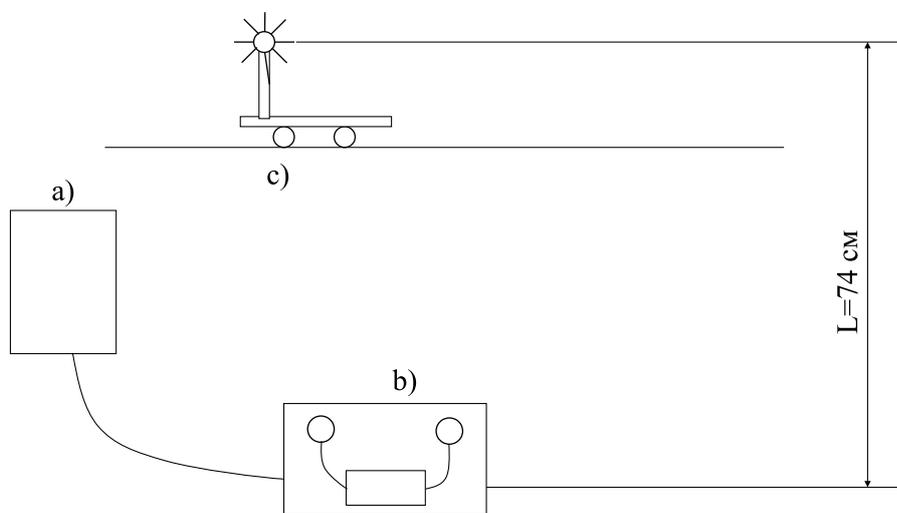


Рис. 9. Оптическая схема установки

определены оптические оси «глаз». Для двухмерного «глаза» необходимо определить плоскость перпендикулярную матрице, которая содержит оптическую ось объектива для каждого «глаза». Экспериментально это означает определение оптического центра ПЗС-матрицы, то есть определение номеров омматидиев «глаза» ближайших к оптическому центру. Плоскости матриц будут при этом перпендикулярны оптическим осям объективов. Проблематика создания такой методики заключается в двух факторах: а) необходимость осуществления ориентировки матриц одновременно для двух «глаз» с последующим выставлением

Таблица 4. Характеристики светодиодов

Цвет, диапазон длин волн	Мощность, Вт
Зеленый (520–540 нм)	1
Красный (600–625 нм)	3
Красный (620–625 нм)	1
ИК (850 нм)	3
Синий (460–465 нм)	1
Полный спектр (400–840 нм)	3

углов наклона «глаз» к друг другу; б) оптическая недоступность матриц мимо объективов для «глаз» в собранном виде. Первый фактор требует, таким образом, использование одного и того же опорного направления (опорного луча) для юстировки одновременно обоих «глаз». Второй фактор требует последовательной работы с каждым разобранным (без объектива) и собранным (с объективом) «глазом» с последующим выставлением фокальной плоскости объектива на поверхность матрицы для обоих «глаз». Такая методика была реализована на опытно-демонстрационном стенде в нашей оптической лаборатории. Процесс установки ориентации матриц продемонстрирован на рис. 10.

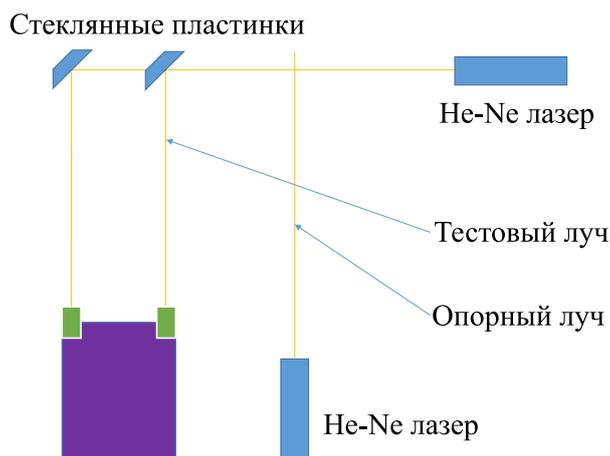


Рис. 10. Определение ориентации поверхностей матриц фасеточных глаз

Всегда использовался один и тот же опорный лазерный луч перпендикулярный установочному рельсу на опытно-демонстрационном стенде. На соседнем со стендом оптическом столе, на расстоянии примерно 330 см от рельса, выставлялся источник тестового луча (отражение лазерного луча от стеклянной пластинки). Тестовый луч всегда выставлялся параллельно опорному лучу, и его интенсивность минимизировалась серыми стеклянными фильтрами. Тестовые лучи направлялись (параллельно опорному лучу) на матрицы со снятыми объективами. Матрицы поворачивались так, чтобы отраженные лучи шли точно назад (параллельно опорному лучу). Положение матриц на макете жестко фиксировалось винтами. После этого объективы аккуратно наворачивались на фрамугу матриц, чтобы матрицы при этом не сдвигались, для чего макет жестко фиксировался на оптическом столе стенда в положении на боку. Потом проводилась юстировка объективов, чтобы их фокальные плоскости совпадали с поверхностью матриц. Для этого мы добивались получения резкого изображения объектов в дальней зоне (на расстоянии примерно 12 метров); при

этом, изображения объектов ближней зоны (около 1 метра) оставались размытыми. Оптические центры матриц определялись экспериментально следующим образом. После установки и юстировки объективов, макет устанавливался в исходное положение с матрицами перпендикулярно опорному лучу. Тестовые лучи направлялись на центры «своих» объективов параллельно опорному лучу. Изображения каустик на матрицах, таким образом, обозначали оптические центры матриц. Результаты определения оптических центров «глаз» оказались следующими. Оптический центр левого «глаза» расположен между 92-ой и 93-ей омматидиями слева направо. Оптический же центр правого «глаза» расположен между 112-ой и 113-ой омматидиями слева направо. Точность определения центров проверялась на вдвое большем числе омматидиев на матрицу, и оценивается меньше одной пятой доли рабочего омматидия. Заданный угловой размер омматидиев соответствует 41 омматидиям на 10 градусов. Он также может быть выбран пользователем соответственно потребностям пользователя или измененным параметрам нейронной сети. Угловые размеры омматидиев определялись экспериментально следующим образом. Тестовые лучи (рис. 11) направлялись на центры «чужих» объективов, и измерялись смещения изображений каустик на каждой матрице. Они оказались одинаковыми по величине и разными по направлению. Угловой размер омматидиев оказался совпадающим с заданным размером с точностью около двух процентов. Практически, точное совпадение реальных угловых размеров омматидиев с заданным размером достигается при общем числе омматидиев в 204 на каждую матрицу.

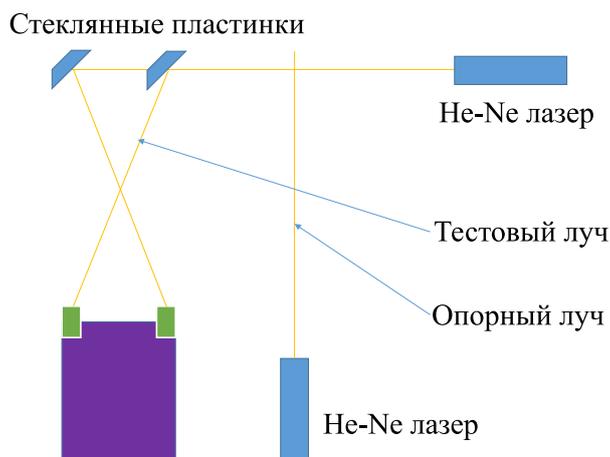


Рис. 11. Определение углового размера омматидиев

Заданное «косоглазие» макета в соответствии с оптико-геометрическими параметрами и требованиями разработанной нейронной сети было определено в 5.4 градуса. То есть каждый «глаз» должен быть повернут внутрь на 2.7 градуса. Подобный поворот «глаз» невозможно реализовать обычными инженерными способами с требуемой точностью. Только с применением оптических методов удастся этого добиться. Требуемая методика была разработана и реализована для нашего демонстрационного макета. Точный поворот «глаз» осуществлялся следующим образом. Стеклянные пластинки (рис. 12) смещались в противоположных направлениях перпендикулярно опорному лучу. Они смещались так, чтобы тестовые лучи, направленные на центры «своих» объективов, составляли одинаковые углы по отношению к опорному лучу в 2.7 градуса. Потом каждый «глаз» поворачивался так,

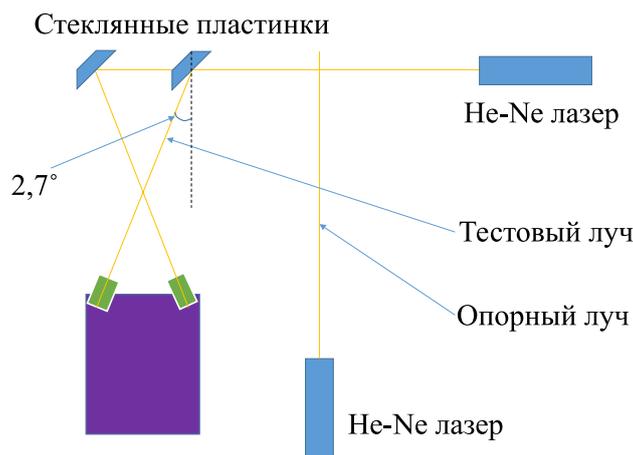


Рис. 12. Установка угловых поворотов фасеточных глаз

чтобы изображения каустик на матрицах совпадали с оптическими центрами матриц. После выставления углов фасеточных глаз крепеж «глаз» жестко фиксировался с помощью винтов. Данная методика выставления углов оказалась очень точной, обеспечивая суммарную погрешность в угле «косоглазия» не более 0.005 градуса.

Проведя настройку установки вычислено максимальное значение интенсивности света для различных положений светодиода. Расстояние от светодиода до матрицы камеры составило $L = 74$ см. Получена зависимость интенсивности света от номера омматидия для левой и правой камеры (рис. 13: а) левая камера, б) правая камера). Всего использовалось 200 омматидиев, линий по вертикали 600. Чтобы ускорить расчеты и ограничить ненужные детали проводилось суммирование не всех линий по горизонтали при вычислении интенсивности света, а только выбранных. В данной работе это названо «Прищур». На данном этапе, прищур равнялся 70. Использовался светофильтр «НС10». Анализ экспериментальных результатов проведен аппроксимацией полиномами второго порядка с коэффициентами $y = a * x^2 + b * x + c$. Значения коэффициентов полинома представлены в табл. 5. В экспериментах были задеты не все омматидии, а только по левой камере с 37 по 155, а по правой с 48 по 170.

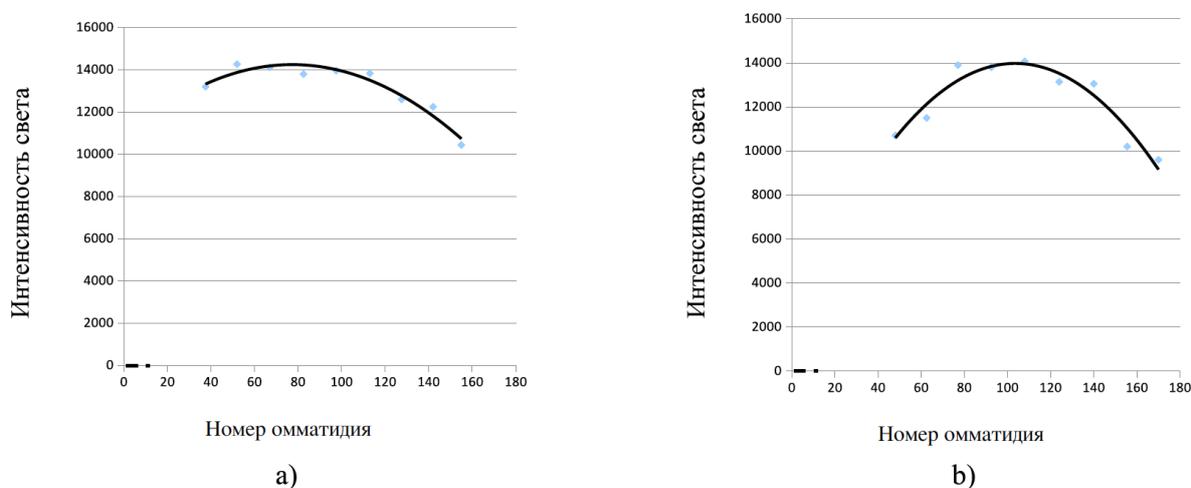


Рис. 13. График зависимости интенсивности света от номера омматидия

Таблица 5. Значения коэффициентов полинома второго порядка

Коэффициенты	<i>a</i>	<i>b</i>	<i>c</i>
Левая камера	-0.58	89.9	10772
Правая камера	-0.54	208.4	6868

Из анализа рис. 13 видно, что получились две параболы ветками направленными вниз. Однако, они не совсем симметричны. Поэтому был проведен следующий вариант обрезки.

Используя нейронные сети проведено сравнение результатов полученных из эксперимента с данными, выдаваемыми интеллектуальной системой бинокулярного фасеточного зрения (табл. 6). Расстояние до объекта измерялось в миллиметрах, азимут измерялся в градусах. Точность в таблице определяется в процентах по формуле $100 - MARE$.

Таблица 6. Сравнение реальных данных по расстоянию и азимуту с работой нейронных сетей

Координаты объекта	Расст-е (мм)	Предсказ. расст-е (мм)	Точность (%)	Азимут (градусы)	Предсказ. азимут (градусы)	Точность (%)
33 см, 70°	330	354.9	93.0	70	69.8	99.8
30 см, 90°	300	294.3	98.1	90	90.8	99.1
33 см, 110°	330	328.7	99.6	110	109.3	99.3
65 см, 70°	650	694.0	93.2	70	70.2	99.7
60 см, 90°	600	608.0	98.7	90	89.5	99.4
65 см, 110°	650	671.0	96.7	110	110.0	99.9
95 см, 70°	950	932.4	98.1	70	71.4	98.1
90 см, 90°	900	909.3	98.9	90	90.7	99.2
95 см, 110°	950	914.3	96.2	110	109.1	99.2

Таблица 6 показывает, что использование предложенных нейронных сетей позволяет получить координаты объекта с точностью 93% для расстояния и 98% для азимута.

Заключение

В статье представлен прототип системы бинокулярного фасеточного зрения для определения азимута и расстояния до объекта на плоскости. Искусственный фасеточный глаз моделируется на основе однолинзовой оптической схемы с использованием ПЗС-матрицы. Пиксели ПЗС-матрицы разбиваются на регулярные непересекающиеся области, каждая из которых соответствует отдельному омматидию. Основой аппаратно-программного блока является специализированный контроллер на базе отладочной платы с ПЛИС Intel Altera CycloneV, модель 5CEFA5F23I7. Контроллер обеспечивает интерфейс информационного обмена с камерами типа OV2640 (или совместимыми) с протоколом обмена SCCB. Нейросетевой модуль включает в себя две глубокие нейронные сети прямого распространения: первая для определения расстояния, а вторая для определения азимута целевого объекта. Конструкция прототипа включает верхнюю металлическую платформу, на которую крепится модуль аппаратно-программного контроллера и оптические модули, нижнюю опорную платформу, опорный штوك с возможностью регулировки высоты установки верхней плат-

формы. В конструкции предусмотрена возможность регулировки положения оптических модулей. Для проведения натуральных экспериментов был изготовлен испытательный стенд, включающий в себя ноутбук, прототип системы фасеточного зрения (плата «Плис», соединенная с двумя камерами), передвижная тележка со светодиодом. Скорость тележки можно регулировать. Движение тележки осуществляется по двутавру длиной 119 см. Свободный ход составляет 100 см. Расстояние от светодиода до камер равно 74 см. С использованием разработанного стенда были проведены натурные эксперименты, которые показали, что созданный интеллектуальный прототип позволяет получить координаты объекта с точностью 93% для расстояния и 98% для азимута.

Литература

1. Ben-Ari M., Mondada F. Elements of robotics. Springer Nature, 2017. 308 p. DOI: 10.1007/978-3-319-62533-1.
2. Mahajan A., Walworth M. 3D position sensing using the differences in the time-of-flights from a wave source to various receivers // IEEE Transactions on Robotics and Automation. 2001. Vol. 17, no. 1. P. 91–94. DOI: 10.1109/70.917087.
3. Journet B.A., Poujouly S. High-resolution laser rangefinder based on a phase-shift measurement method // Three-Dimensional Imaging, Optical Metrology, and Inspection IV. Vol. 3520 / ed. by K.G. Harding, D.J. Svetkoff, K. Creath, J.S. Harris. International Society for Optics, Photonics. SPIE, 1998. P. 123–132. DOI: 10.1117/12.334326.
4. Zhang C., Lindner S., Antolovic I.M., *et al.* A 30-frames/s, 252 × 144 SPAD Flash LiDAR With 1728 Dual-Clock 48.8-ps TDCs, and Pixel-Wise Integrated Histogramming // IEEE Journal of Solid-State Circuits. 2018. Vol. 54, no. 4. P. 1137–1151. DOI: 10.1109/JSSC.2018.2883720.
5. Bamji C., Godbaz J., Oh M., *et al.* A Review of Indirect Time-of-Flight Technologies // IEEE Transactions on Electron Devices. 2022. Vol. 69, no. 6. P. 2779–2793. DOI: 10.1109/TED.2022.3145762.
6. Dorsch R.G., Häusler G., Herrmann J.M. Laser triangulation: fundamental uncertainty in distance measurement // Appl. Opt. 1994. Mar. Vol. 33, no. 7. P. 1306–1314. DOI: 10.1364/AO.33.001306.
7. Liebe C.C., Coste K. Distance Measurement Utilizing Image-Based Triangulation // IEEE Sensors Journal. 2013. Vol. 13, no. 1. P. 234–244. DOI: 10.1109/JSEN.2012.2212428.
8. Bondžulić B.P., Mitrović S.T., Barbarić Ž.P., Andrić M.S. A comparative analysis of three monocular passive ranging methods on real infrared sequences // Journal of Electrical Engineering. 2013. Vol. 64, no. 5. P. 305–310. DOI: 10.2478/jee-2013-0044.
9. Li L. Building an accurate 3D model of a circular feature for robot vision // Opto-Electronics Review. 2012. Vol. 20. P. 120–125. DOI: 10.2478/s11772-012-0017-y.
10. Jüngel M., Mellmann H., Spranger M. Improving vision-based distance measurements using reference objects // RoboCup 2007: Robot Soccer World Cup XI 11. Vol. 5001. Springer, 2008. P. 89–100. DOI: 10.1007/978-3-540-68847-1_8.
11. Гейдаров П.Ш.О. Алгоритм определения расположения и размеров объектов на основе анализа изображений объектов // Компьютерная оптика. 2011. Т. 35, № 2. С. 275–280.

12. Witus G., Hunt S. Monocular visual ranging // *Unmanned Systems Technology X*. Vol. 6962. SPIE, 2008. P. 38–44. DOI: 10.1117/12.778686.
13. Зуев С.В. Моностатический способ определения расстояния до объекта, его направления и скорости движения. 2008.
14. Lumsdaine A., Georgiev T. The focused plenoptic camera // *2009 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2009. P. 1–8. DOI: 10.1109/ICCPHOT.2009.5559008.
15. Chen Y., Jin X., Dai Q. Distance measurement based on light field geometry and ray tracing // *Optics Express*. 2017. Vol. 25, no. 1. P. 59–76. DOI: 10.1364/OE.25.000059.
16. Heinze C., Spyropoulos S., Hussmann S., Perwass C. Automated robust metric calibration algorithm for multifocus plenoptic cameras // *IEEE Transactions on Instrumentation and Measurement*. 2016. Vol. 65, no. 5. P. 1197–1205. DOI: 10.1109/TIM.2015.2507412.
17. Sardemann H., Maas H.-G. On the accuracy potential of focused plenoptic camera range determination in long distance operation // *ISPRS Journal of Photogrammetry and Remote Sensing*. 2016. Vol. 114. P. 1–9. DOI: 10.1016/j.isprsjprs.2016.01.012.
18. Mahammed M.A., Melhum A.I., Kochery F.A. Object distance measurement by stereo vision // *International Journal of Science and Applied Information Technology (IJSAIT)*. 2013. Vol. 2, no. 2. P. 05–08.
19. Jeong C.-J., Park G.-M. Real-time Auto Tracking System using PTZ Camera with DSP // *International Journal of Advanced Smart Convergence*. 2013. Vol. 2, no. 1. P. 32–35. DOI: 10.7236/IJASC.2013.2.1.032.
20. Starkov A.E., Sokolinsky L.B. Building 2D Model of Compound Eye Vision for Machine Learning // *Mathematics*. 2022. Vol. 10, no. 2. DOI: 10.3390/math10020181.
21. Biewald L. Experiment Tracking with Weights and Biases. 2020. URL: <https://docs.wandb.ai/>.
22. Goodfellow I., Bengio Y., Courville A. *Deep Learning (Adaptive Computation and Machine Learning)*. MIT Press, 2016. 800 p. URL: <http://www.deeplearningbook.org>.
23. Willmott C.J., Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance // *Climate Research*. 2005. Vol. 30, no. 1. P. 79–82. DOI: 10.3354/CR030079.
24. Bisong E. *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Apress, 2019. 709, XXIX p. DOI: 10.1007/978-1-4842-4470-8.
25. Myttenaere A., Golden B., Rossi F. Mean Absolute Percentage Error for regression models // *Neurocomputing*. 2016. Vol. 192. P. 38–48. DOI: 10.1016/J.NEUCOM.2015.12.114.
26. Barten A.P. The coefficient of determination for regression without a constant term // *The Practice of Econometrics. International Studies in Economics and Econometrics*. Vol. 15 / ed. by R. Heijmans, H. Neudecker. Springer, 1987. P. 181–189. DOI: 10.1007/978-94-009-3591-4_12.

Белов Кирилл Николаевич, ассистент кафедры оптоинформатики, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Бибикова Эвелина Анатольевна, к.ф.-м.н., доцент кафедры оптоинформатики, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация); научный сотрудник Лаборатории нелинейной оптики, Институт электрофизики УрО РАН (Екатеринбург, Российская Федерация)

Булдашев Иван Владимирович, старший преподаватель кафедры оптоинформатики, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Кундикова Наталия Дмитриевна, д.ф.-м.н., профессор, заведующий кафедрой оптоинформатики, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация); главный научный сотрудник Лаборатории нелинейной оптики, Институт электрофизики УрО РАН (Екатеринбург, Российская Федерация)

Мухин Юрий Викторович, к.ф.-м.н., доцент кафедры оптоинформатики, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация); Институт электрофизики УрО РАН (Екатеринбург, Российская Федерация)

Николаев Андрей Николаевич, доцент кафедры радиоэлектроники и систем связи, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Портнов Андрей Владимирович, аспирант кафедры системного программирования, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Ридный Ярослав Максимович, к.ф.-м.н., доцент кафедры оптоинформатики, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Соколинский Леонид Борисович, д.ф.-м.н., профессор, заведующий кафедрой системного программирования, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Старков Артем Евгеньевич, аспирант кафедры системного программирования, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

Шульгинов Александр Анатольевич, к.ф.-м.н., доцент кафедры оптоинформатики, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Российская Федерация)

INTELLIGENT BINOCULAR COMPOUND EYE VISION SYSTEM FOR DETECTING AZIMUTH AND DISTANCE TO OBJECT ON PLANE

© 2025 K.N. Belov¹, E.A. Bibikova^{1,2}, I.V. Buldashev¹, N.D. Kundikova^{1,2}, Y.V. Mukhin^{1,2}, A.N. Nikolaev¹, A.V. Portnov¹, Y.M. Ridnyi¹, L.B. Sokolinsky¹, A.E. Starkov¹, A.A. Shulginov¹

¹South Ural State University (pr. Lenina 76, Chelyabinsk, 454080 Russia),

²Institute of Electrophysics of the Ural Branch of the Russian Academy of Sciences (st. Amundsen 106, Yekaterinburg, 620016 Russia)

E-mail: belovkn@susu.ru, bibikovaea@susu.ru, buldasheviv@susu.ru, kundikovand@susu.ru, mukhinyv@susu.ru, nikolaevan@susu.ru, portnovav@susu.ru, imridnyi@susu.ru, leonid.sokolinsky@susu.ru, starkovae@susu.ru, shulginovaa@susu.ru

Received: 10.03.2025

The article is devoted to the prototype of an artificial binocular compound eye vision system for detecting the azimuth and distance to an object on a plane using an artificial neural network. A analytical review of modern distance and azimuth detection systems based on active and passive sensors is given. An intelligent binocular vision system is proposed, which is a passive optical sensor that allows you to determine the azimuth and distance to a round object of arbitrary size, emitting in the visible or infrared ranges of the electromagnetic spectrum. The general architecture of the compound eye vision system is considered. The main structural elements of the system are: an optical module, a hardware and software controller and a neural network module. The optical module uses a pair of lenses to convert the light signal from the object into two pixel Fourier images, which are fed to the input of the hardware and software controller. The controller performs primary processing of pixel Fourier images and converts them into two bit masks, the elements of which correspond to separate facets (each facet integrates four adjacent columns of the pixel image). The resulting bit masks are fed into a neural network module, which, based on their analysis, determines the coordinates of the object in the form of distance and azimuth.

Keywords: compound eye vision, optical model, distance and azimuth detection, CCD image sensor, neural network model, prototype.

FOR CITATION

Belov K.N., Bibikova E.A., Buldashev I.V., *et al.* Intelligent Binocular Compound Eye Vision System for Detecting Azimuth and Distance to Object on Plane. Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering. 2025. Vol. 14, no. 1. P. 5–29. (in Russian) DOI: 10.14529/cmse250101.

This paper is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 License which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is properly cited.

References

1. Ben-Ari M., Mondada F. Elements of robotics. Springer Nature, 2017. 308 p. DOI: 10.1007/978-3-319-62533-1.
2. Mahajan A., Walworth M. 3D position sensing using the differences in the time-of-flights from a wave source to various receivers. IEEE Transactions on Robotics and Automation. 2001. Vol. 17, no. 1. P. 91–94. DOI: 10.1109/70.917087.

3. Journet B.A., Poujouly S. High-resolution laser rangefinder based on a phase-shift measurement method. *Three-Dimensional Imaging, Optical Metrology, and Inspection IV*. Vol. 3520 / ed. by K.G. Harding, D.J. Svetkoff, K. Creath, J.S. Harris. International Society for Optics, Photonics. SPIE, 1998. P. 123–132. DOI: 10.1117/12.334326.
4. Zhang C., Lindner S., Antolovic I.M., *et al.* A 30-frames/s, 252 × 144 SPAD Flash LiDAR With 1728 Dual-Clock 48.8-ps TDCs, and Pixel-Wise Integrated Histogramming. *IEEE Journal of Solid-State Circuits*. 2018. Vol. 54, no. 4. P. 1137–1151. DOI: 10.1109/JSSC.2018.2883720.
5. Bamji C., Godbaz J., Oh M., *et al.* A Review of Indirect Time-of-Flight Technologies. *IEEE Transactions on Electron Devices*. 2022. Vol. 69, no. 6. P. 2779–2793. DOI: 10.1109/TED.2022.3145762.
6. Dorsch R.G., Häusler G., Herrmann J.M. Laser triangulation: fundamental uncertainty in distance measurement. *Appl. Opt.* 1994. Mar. Vol. 33, no. 7. P. 1306–1314. DOI: 10.1364/AO.33.001306.
7. Liebe C.C., Coste K. Distance Measurement Utilizing Image-Based Triangulation. *IEEE Sensors Journal*. 2013. Vol. 13, no. 1. P. 234–244. DOI: 10.1109/JSEN.2012.2212428.
8. Bondžulić B.P., Mitrović S.T., Barbarić Ž.P., Andrić M.S. A comparative analysis of three monocular passive ranging methods on real infrared sequences. *Journal of Electrical Engineering*. 2013. Vol. 64, no. 5. P. 305–310. DOI: 10.2478/jee-2013-0044.
9. Li L. Building an accurate 3D model of a circular feature for robot vision. *Opto-Electronics Review*. 2012. Vol. 20. P. 120–125. DOI: 10.2478/s11772-012-0017-y.
10. Jüngel M., Mellmann H., Spranger M. Improving vision-based distance measurements using reference objects. *RoboCup 2007: Robot Soccer World Cup XI 11*. Vol. 5001. Springer, 2008. P. 89–100. DOI: 10.1007/978-3-540-68847-1_8.
11. Geidarov P. Algorithm for determining the position and size of objects based on image analysis. *Computer Optics*. 2011. Vol. 35, no. 2. P. 275–280. (in Russian).
12. Witus G., Hunt S. Monocular visual ranging. *Unmanned Systems Technology X*. Vol. 6962. SPIE, 2008. P. 38–44. DOI: 10.1117/12.778686.
13. Zuev S.V. Monostatic method for measuring the distance to object, its direction and movement speed. 2008. (in Russian).
14. Lumsdaine A., Georgiev T. The focused plenoptic camera. 2009 IEEE International Conference on Computational Photography (ICCP). IEEE, 2009. P. 1–8. DOI: 10.1109/ICCPHOT.2009.5559008.
15. Chen Y., Jin X., Dai Q. Distance measurement based on light field geometry and ray tracing. *Optics Express*. 2017. Vol. 25, no. 1. P. 59–76. DOI: 10.1364/OE.25.000059.
16. Heinze C., Spyropoulos S., Hussmann S., Perwass C. Automated robust metric calibration algorithm for multifocus plenoptic cameras. *IEEE Transactions on Instrumentation and Measurement*. 2016. Vol. 65, no. 5. P. 1197–1205. DOI: 10.1109/TIM.2015.2507412.
17. Sardemann H., Maas H.-G. On the accuracy potential of focused plenoptic camera range determination in long distance operation. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2016. Vol. 114. P. 1–9. DOI: 10.1016/j.isprsjprs.2016.01.012.

18. Mahammed M.A., Melhum A.I., Kochery F.A. Object distance measurement by stereo vision. *International Journal of Science and Applied Information Technology (IJSAIT)*. 2013. Vol. 2, no. 2. P. 05–08.
19. Jeong C.-J., Park G.-M. Real-time Auto Tracking System using PTZ Camera with DSP. *International Journal of Advanced Smart Convergence*. 2013. Vol. 2, no. 1. P. 32–35. DOI: 10.7236/IJASC.2013.2.1.032.
20. Starkov A.E., Sokolinsky L.B. Building 2D Model of Compound Eye Vision for Machine Learning. *Mathematics*. 2022. Vol. 10, no. 2. DOI: 10.3390/math10020181.
21. Biewald L. Experiment Tracking with Weights and Biases. 2020. URL: <https://docs.wandb.ai/>.
22. Goodfellow I., Bengio Y., Courville A. *Deep Learning (Adaptive Computation and Machine Learning)*. MIT Press, 2016. 800 p. URL: <http://www.deeplearningbook.org>.
23. Willmott C.J., Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*. 2005. Vol. 30, no. 1. P. 79–82. DOI: 10.3354/CR030079.
24. Bisong E. *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Apress, 2019. 709, XXIX p. DOI: 10.1007/978-1-4842-4470-8.
25. Myttenaere A., Golden B., Rossi F. Mean Absolute Percentage Error for regression models. *Neurocomputing*. 2016. Vol. 192. P. 38–48. DOI: 10.1016/J.NEUCOM.2015.12.114.
26. Barten A.P. The coefficient of determination for regression without a constant term. *The Practice of Econometrics. International Studies in Economics and Econometrics*. Vol. 15 / ed. by R. Heijmans, H. Neudecker. Springer, 1987. P. 181–189. DOI: 10.1007/978-94-009-3591-4_12.