



Алексей Коченгин

# Камеры для систем компьютерного зрения: сложный выбор

В статье приведена общая информация о технологии компьютерного зрения, рассмотрены основные технические характеристики при выборе камер, на которые следует опираться разработчикам систем компьютерного зрения. Представлен обзор продуктов компании-производителя промышленных камер для систем компьютерного зрения Basler глазами разработчика.

## ВВЕДЕНИЕ

Понятие «компьютерное зрение» относится к междисциплинарной научной области, которая занимается созданием компьютеров для анализа цифровых изображений или видео. С точки зрения инженерии, с помощью компьютерного зрения можно автоматизировать задачи, которые выполняет зрительная система человека.

Компьютерное зрение применяется для решения ряда основных задач:

- распознавание – классическая проблема в компьютерном зрении, заключающаяся в том, чтобы определить наличие искомого действия или объекта на полученном изображении. Задача распознавания часто разбивается на ряд подзадач, таких как идентификация объекта, классификация объекта, обнаружение аномалий и др.;
- восстановление изображения – задача улучшения качества изображения при помощи фильтрации шумов;
- анализ движения – задача расчёта числовых характеристик движущегося объекта, например, определение скорости и пространственного положения;
- реконструкция сцены – определяется возможность построения 3D-модели поверхности или объекта по полученному массиву двух- и трёхмерных изображений.

Также следует отметить, что понятия обработки и анализа изображения машинного зрения не эквивалентны понятию компьютерного зрения и являются лишь областями компьютерного зрения, решающими более узкие спектры задач.

Очевидно, что одним из ключевых компонентов систем компьютерного зрения является камера. Перед разработчиками программно-технического комплекса, решающего задачи компью-

терного зрения, остро встаёт вопрос выбора камеры.

## ПРОБЛЕМЫ ВЫБОРА КАМЕРЫ ДЛЯ СИСТЕМ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Рассмотрим основные технические характеристики камер.

### Тип используемой матрицы

Подавляющее большинство всех матриц строится на двух технологиях –

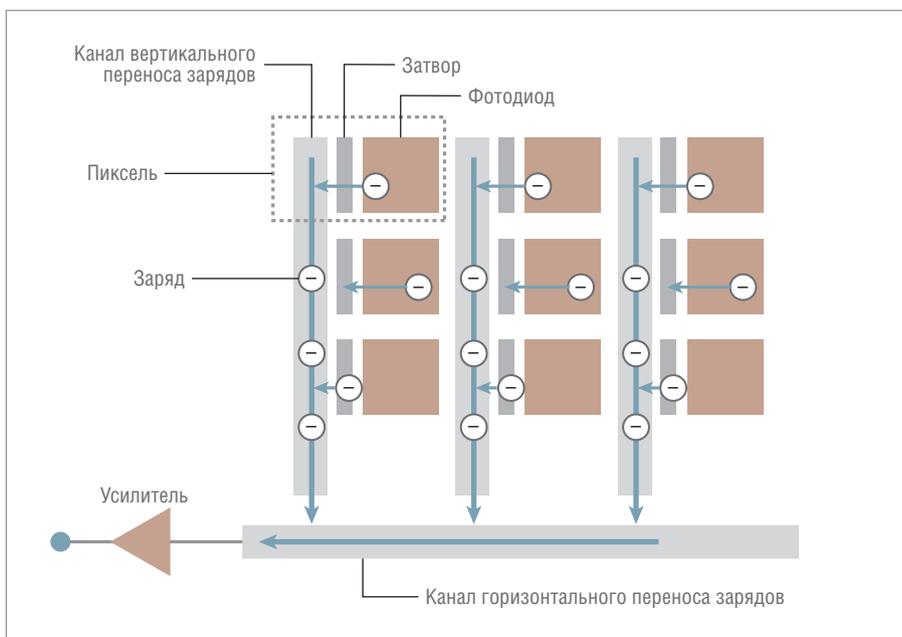


Рис. 1. Структура CCD-матрицы

Иллюстрация с сайта micromed-spb.ru

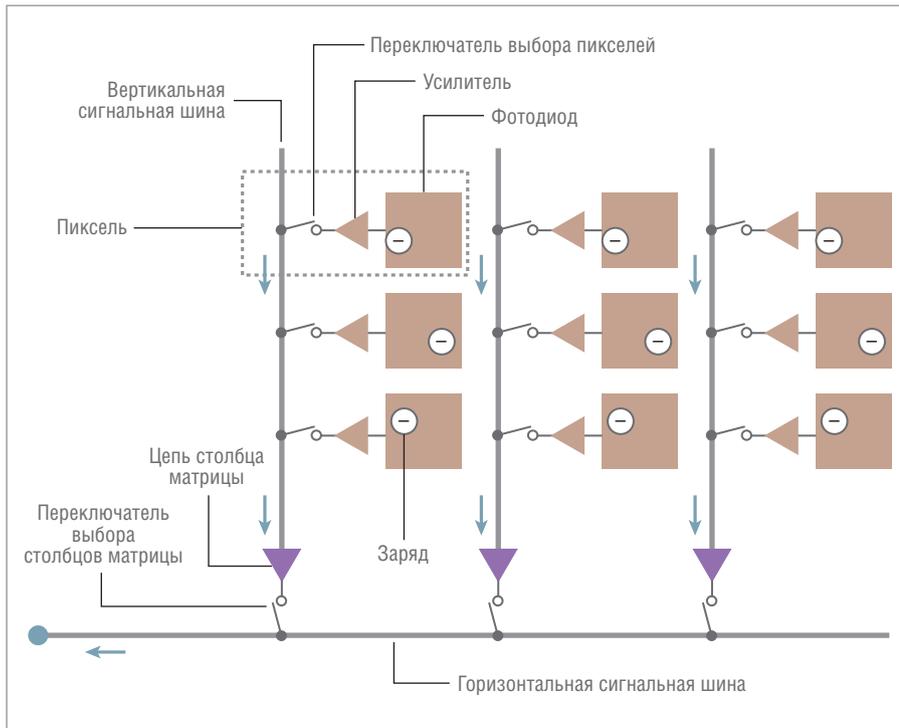


Иллюстрация с сайта micromed-spb.ru

Рис. 2. Структура CMOS-матрицы

CCD, или ПЗС (Charge-Coupled Device – прибор с зарядовой связью) и CMOS, или КМОП (Complementary Metal-Oxide-Semiconductor – комплементарная структура металл-оксид-полупроводник).

В матрице, построенной по технологии ПЗС, сигналы с полупроводников (пикселей) проходят через один усилитель и потом попадают в АЦП, в то время как в КМОП-матрицах пиксели уже имеют собственные усилители и АЦП. За счёт параллельной работы каждого АЦП КМОП-матрицы позволяют достигать гораздо большей скорости чтения.

На рис. 1 и 2 изображены структуры построения матриц по технологиям ПЗС и КМОП соответственно.

Однако стоит отметить, что ПЗС-матрицы менее подвержены «цифровому шуму» и позволяют добиться лучшего качества изображения, жертвуя при этом скоростью обработки. Большинство производителей современных камер используют КМОП-матрицы, из-за более простой схематехники, высокой скорости и низкого энергопотребления. Камеры с ПЗС-матрицей, как правило, используются для съёмки в условиях слабой освещённости.

### Разрешение камеры

Под разрешением матрицы понимается количество пикселей, используемых для формирования изображения на матрице. Оно влияет на детализацию изображения и обозначается в формате

$A \times B$ , где  $A$  – количество пикселей по горизонтали,  $B$  – по вертикали. Для упрощения записи часто пишут либо произведение  $A$  и  $B$ , измеряемое в мегапикселях, либо только количество точек по вертикали, приписывая формат вывода на экран:  $p$  (прогрессивный, от английского progressive scan) или  $i$  (чересстрочный, от английского interlaced).

При выборе камеры по параметру разрешения матрицы разработчику следует учитывать особенности своей задачи. При этом стоит принять во внимание уже готовые решения либо обратиться к стандартам. Например, для задач распознавания лиц в РФ существует ГОСТ Р ИСО/МЭК 19794-5-2013. В данном стандарте можно найти требования к минимальному изображению лица в кадре (240 пикселей по горизонтали и 60 пикселей – расстояние между глазами). Исходя из данного требования, можно напрямую оценить необхо-

димое разрешение камеры. Как правило, для такого рода задач разрешения HD (720p) или Full HD (1080p) будет достаточно. Дальнейшее увеличение разрешения приведёт только к повышению цены камеры, но при этом с большой вероятностью не окажет существенного влияния на точность алгоритмов распознавания.

### Размер матрицы

В зависимости от соотношения сторон (4:3 или 16:9) у матриц с одной и той же диагональю физические размеры различны (табл. 1). Поэтому, например, камера на матрице 1/3" с соотношением сторон 4:3 даёт больший угол обзора по вертикали и меньший по горизонтали, чем камера на матрице с такой же диагональю, но с соотношением 16:9.

В целях облегчения подбора совместимой оптики и углов обзора обычно заявляют ближайшее из стандартных значений для диагонали матрицы: 1", 1/2", 1/2,5", 1/2,7", 1/2,8", 1/3", 1/4". При этом измерять её принято в видеконовых дюймах. Эта единица измерения, равная 2/3 обычного дюйма, была введена со времён зарождения телевидения, когда приёмным элементом в телекамере служила электронная трубка (видикон), а размер обозначал её диаметр (в который должен был вписываться с запасом снимаемый кадр).

Помимо этого, необходимо помнить, что на некоторых режимах работы камеры часть пикселей матрицы не используется. Поэтому при определении угла обзора следует говорить не столько о размере матрицы, сколько о размере активной области матрицы.

### Объектив

Правильный подбор объектива крайне важен при решении задач компьютерного зрения.

Выделяют три типа объективов:

Таблица 1

Зависимость физических размеров матрицы от соотношения сторон

Формат матрицы, видеконовый дюйм	Диагональ матрицы, мм	Соотношение сторон			
		4:3		16:9	
		Ширина, мм	Высота, мм	Ширина, мм	Высота, мм
1/4	4,23	3,39	2,54	3,69	2,08
1/3	5,64	4,52	3,39	4,92	2,77
1/2,8	6,05	4,84	3,63	5,27	2,96
1/2,7	6,27	5,02	3,76	5,47	3,07
1/2,5	6,77	5,42	4,06	5,9	3,32
1/2	8,47	6,77	5,08	7,38	4,15

Таблица с сайта beward.ru

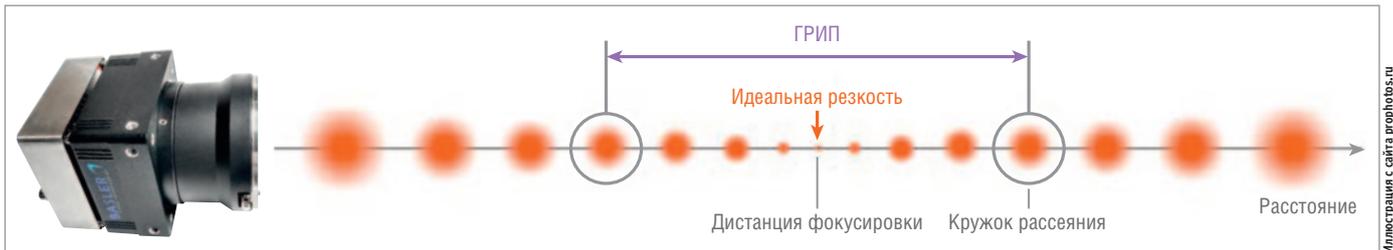


Иллюстрация с сайта prophoto.ru

Рис. 3. Изображение границ глубины резкости

- монофокальные (фиксированные) объективы имеют одну фиксированную величину фокусного расстояния. В таких объективах нельзя регулировать угол обзора видеокамеры. Из-за своей простоты и низкой стоимости они являются наиболее популярными;
- вариофокальные объективы. Среди них различают объективы с автофокусом и с ручной фокусировкой, которые позволяют регулировать фокусное расстояние объектива и, как следствие, менять угол обзора видеокамеры. Объективы этого типа являются универсальными, однако их стоимость выше;
- трансфокальные (зум-объективы) являются самыми дорогими, однако они позволяют реализовать удалённо с пульта управления регулирование углов обзора, масштабирование зоны наблюдения, наводить резкость, фокусировать изображение.

Как можно было заметить из описания типов, основной характеристикой объектива является его фокусное расстояние. Под фокусным расстоянием понимают расстояние от оптического центра объектива (линзы) до точки фокусировки (матрицы). Размер матрицы и фокусное расстояние напрямую влияют на угол обзора камеры. Расчёт угла обзора достаточно тривиален:

$$\alpha = 2 \cdot \arctg\left(\frac{d}{2 \cdot f}\right)$$

Здесь  $\alpha$  – угол обзора видеокамеры в метрических градусах,

$d$  – ширина матрицы в мм,

$f$  – фокусное расстояние объектива в мм.

Также при подборе объектива для задач компьютерного зрения стоит обратить внимание на такой параметр, как глубина резкости, или ГРИП (глубина резко изображаемого пространства). В некоторой степени данный параметр является нечётким, поскольку сложно точно оценить границы области пространства, при нахождении внутри которой объекты получатся резкими. Сложность оценки заключается в опре-

делении диаметра кружка рассеяния (рис. 3).

При расчёте глубины резкости наиболее часто значение диаметра кружка рассеяния выбирается из диапазона от 0,015 до 0,03 мм. Чем меньше диаметр, тем больше сужаются границы резкости, при этом повышаются требования к аппаратной части видеокамеры.

Расчёт глубины резкости производят по следующей формуле:

$$P = R_1 - R_2$$

$$R_1 = \frac{R \cdot f^2}{f^2 + K \cdot (R - f) \cdot Z}$$

$$R_2 = \frac{R \cdot f^2}{f^2 - K \cdot (R - f) \cdot Z}$$

Здесь  $P$  – глубина резкости,  
 $R_1$  – передняя граница ГРИП,  
 $R_2$  – задняя граница ГРИП,  
 $R$  – дистанция фокусировки (дистанция фокусировки отмеряется не от передней линзы объектива, а от места, где находится матрица аппарата, – от фокальной плоскости),

$f$  – фокусное расстояние объектива,  
 $K$  – значение текущей диафрагмы (максимальное значение диафрагмы указывается производителем),  
 $Z$  – диаметр кружка рассеяния.

Ещё один параметр, который может быть обнаружен при эксплуатации уже закупленной камеры, – это дисторсия. Под дисторсией понимают геометрические искажения прямых линий. В большинстве случаев такие искажения изображения не приводят к значительной потере его информативности, в то же время учёт дисторсии довольно сложен, поскольку недостаточно учесть фокусное расстояние объектива и размер сенсора. Потребуется дополнительные данные о технологических тонкостях процесса производства камеры. Для систем видеораспознавания пространственные искажения кадра, такие как дисторсия, могут скрыть ряд признаков, необходимых алгоритму распознавания. Данный факт необходимо учитывать при выборе камеры. Для оценки дисторсии в первом приближении мож-

но учитывать следующее утверждение: при объективе с небольшим фокусным расстоянием и широкими углами обзора будет иметь место бочкообразная дисторсия.

Определившись с данными характеристиками, разработчик должен обратить внимание на коммуникационный интерфейс камеры, её форм-фактор, а также наличие и функциональные возможности SDK (набор средств разработки), необходимого для интеграции камеры в программное обеспечение.

Стоит отметить, что камеры, позиционируемые производителями для решения задач компьютерного зрения, часто становятся не только аппаратной платформой для стримминга видеопотока, но и сложным вычислительным комплексом, имеющим встроенное ПО, представляющее собой вариант решения ряда задач компьютерного зрения от производителя. Одним из таких крупных производителей является компания Basler.

### РЕШЕНИЯ КОМПАНИИ BASLER В ОБЛАСТИ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Компания Basler является ведущим мировым производителем высококачественных камер и аксессуаров для различных областей, включая автоматизацию производства, медицину, контроль дорожного движения и многие другие. Штат группы компаний Basler насчитывает около 800 сотрудников, занятых в главном офисе в Аренсбурге (Германия) и торговых филиалах, расположенных в Европе, Азии и Северной Америке.

При первом обзоре линейки камер Basler бросается в глаза достаточно компактный форм-фактор камер, при этом в ассортименте присутствуют бескорпусные серии, такие как Basler dart (рис. 4). Такие решения снимают головную боль разработчика по внедрению камер в собственные корпуса аппаратного комплекса, имеющие жёсткие требования по габаритным размерам.

В данной статье не рассматривается подробно проблема выбора интерфей-



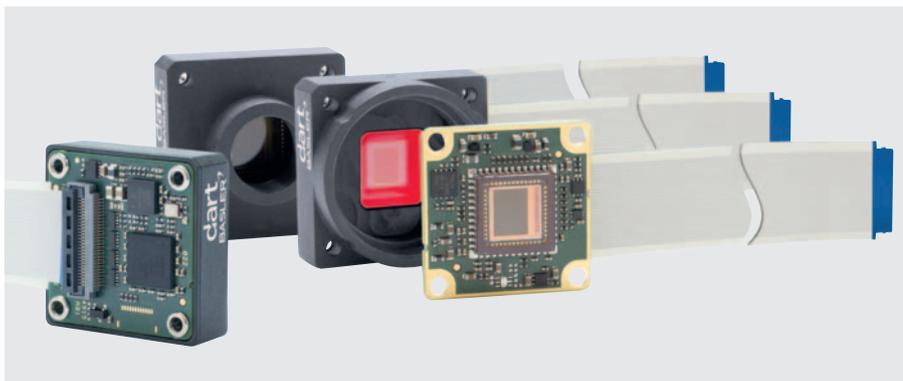


Рис. 4. Камера серии Basler dart

са, однако периодически перед разработчиками возникает задача выбора интерфейса камеры, соответствующего требованиям по удалённости расположения камеры, её пропускной способности, типу системы сбора и обработки информации. Как правило, для этого необходимо провести сравнительный анализ всех типов интерфейсов, составить матрицу характеристик и по ней определиться с интерфейсом. Процесс трудоёмкий и затратный по времени. На своём официальном сайте компания Basler предлагает очень удобный кон-

фигуратор (рис. 5), который позволяет сократить трудозатраты по выбору интерфейса.

Как видно из рис. 5, помимо известных и понятных для разработчиков ПО интерфейсов Ethernet и USB последних поколений Basler предлагает интерфейсы для встраиваемых решений, например, мобильных телефонов, в которых камеры подключаются по интерфейсу MIPI CSI-2 (Mobile Industry Processor Interface Camera Serial Interface – последовательный интерфейс камеры, стандартизированный для мобильных уст-

ройств), либо отдельные матрицы экранов, работающие через интерфейс LVDS (Low-Voltage Differential Signaling – низковольтная дифференциальная передача сигналов).

Для реализации такого решения специалисты Basler разработали интерфейс BCON (Basler Connectivity – интерфейс Basler), с одной стороны подключаемый к камере, а с другой к процессорной плате, выступающей в роли конвертера интерфейсов, а также имеющей ряд портов GPIO и разъёмов питания.

В каталоге продукции компании также можно найти широкий ряд объективов, роль которых при выборе камеры была описана в предыдущем разделе.

Для своих камер Basler предоставляет программное обеспечение Pylon. Данный программный продукт содержит SDK, драйверы и инструментальные средства для работы с любой камерой Basler на ПК под управлением Windows или Linux либо компьютере Mac. SDK включает в себя пример использования API (Application Programming Interface – программный интерфейс приложения) на языках C, C++, C#, VB.Net, а также



Рис. 5. Конфигуратор выбора интерфейсов



документацию с описанием всех классов и методов для каждого языка.

Для обеспечения максимальной скорости и качества изображения в камерах Basler, а также унификации используемого интерфейса, видеопоток обрабатывается и упаковывается по внутренним алгоритмам Basler. Доступ к видеопотоку, отдельным фреймам или любой другой информации с камеры можно получить через интерфейс Basler pylon API, описание которого и представлено в SDK. Такая унификация позволяет без

существенных изменений кода перейти, например, с камеры, подключённой по интерфейсу CameraLink, на Gigabit Ethernet. Также поддерживается последняя версия стандарта GenICam 3.

В области компьютерного зрения огромной популярностью пользуется библиотека OpenCV (Open Source Computer Vision — библиотека машинного зрения с открытым кодом), являющаяся абсолютно бесплатной, с открытым исходным кодом и содержащая в себе различные готовые реализации алго-

ритмов компьютерного зрения, численные алгоритмы общего назначения и удобные программные интерфейсы для работы с камерами. В связи с этим компания Basler выпустила подробное руководство по интеграции своих C++ API с библиотекой OpenCV. Помимо SDK Basler поддерживает Open Source-проект в репозитории на GitHub. В данном репозитории представлена обёртка Basler pylon API для языка Python. В качестве бонуса в python-скриптах имеются примеры взаимодействия с библиотеками OpenCV.

Последнее, что хотелось бы отметить, — это раздел Vision Campus на официальном сайте Basler. В нём собрана база знаний, в которой на доступном языке объяснены понятия, связанные с камерами и системами машинного зрения.

## ЗАКЛЮЧЕНИЕ

Рынок видеонаблюдения оценивается в десятки миллиардов долларов США годового оборота и показывает устойчивые тренды роста порядка 15% в год. В частности, это связано с развитием технологии компьютерного зрения. Данная статья предлагается в помощь разработчикам систем компьютерного зрения, а также знакомит читателя с производителем промышленных камер Basler, имеющих большой опыт успешных внедрений в этой области:

- распознавание номерных знаков Access Gate в лондонском контейнерном терминале;
- система обнаружения дронов Sky Patriot;
- система обнаружения девиантного поведения в людском и транспортном потоке StareServer;
- система видеонализа лиц Pugnatics для оценки эффективности рекламы в ритейле;
- поиск в пространстве и роботизированное перемещение деталей в промышленной среде;
- распознавание лиц в системе СКУД российского колледжа;
- распознавание номерных знаков быстро движущихся автомобилей в Сингапуре.

Можно с уверенностью сказать, что Basler является одним из лидеров среди производителей промышленных камер для систем компьютерного зрения. ●

**Автор – сотрудник  
фирмы ПРОСОФТ  
Телефон: (495) 234-0636  
E-mail: info@prosoft.ru**

