

Технологии машинного зрения в спорте: помощь в судействе, анализ игровых ситуаций, фотофиниш и другие применения.

Основные преимущества камер машинного зрения – широкий выбор устройств с большим и малым разрешением, размером сенсора, высокой скоростью захвата изображения, отличной оптикой и современные коммуникационные интерфейсы. Важно отметить возможность синхронизации работы нескольких камер между собой, с часами точного времени, в том числе разнесенных на сотни метров, гибкую систему запуска от внешних событий. Все это предназначено для получения изображения с необходимым для машинной обработки качеством и скоростью, в точно заданный момент времени. В статье рассматриваются новые возможности, которые открываются с применением технологий машинного зрения в спортивной индустрии и связанных с ней областях.

Помимо очевидного применения в наиболее требовательных к качеству изображения (где не всегда справляются IP-камеры видеонаблюдения) комплексах обеспечения безопасности на спортивных аренах (распознавание лиц, опасных предметов и пр.), технологии машинного зрения находят в спорте много других применений:

- тренировочный процесс – высокоскоростная съемка: плавание, гимнастика, легкая атлетика, лыжи и другие виды спорта, где ключевым моментом является техника выполнения движений;
- судейство – видеофиксация положения "офсайд", "аут", "гол" и прочих, фотофиниш в массовых забегах;
- аналитика в игровых видах спорта (перемещение игроков по полю) – футбол, хоккей, баскетбол.

Ниже мы попробуем более подробно остановиться на требованиях к оборудованию в конкретных областях применения, но начнем, пожалуй, с основных технологических особенностей камер машинного зрения.

Два мира – две судьбы

Исторически сложившиеся различия в областях применения IP-камер и камер машинного зрения (МЗ) определяют функциональные возможности, казалось бы, близких, но на поверку весьма разных подходов к разработке отдельных компонентов и систем. Ключевыми критериями IP-видеонаблюдения являются:

- визуальная оценка качества изображения оператором;
- возможность работы при меняющихся условиях освещенности (день/ночь);
- широкое использование автоматических функций: автоэкспозиция, автофокусировка;

- относительно несложные алгоритмы обработки, типа детектора движения, как правило, реализованные в самих камерах;
- сжатие изображений для минимизации сетевого трафика в распределенных системах видеонаблюдения и объемов видеозаписей для архивного хранения;
- удобные средства поиска, просмотра записей и управления большими объемами данных, получаемых с десятков камер (VMS).

На рынке доминируют решения высокой степени готовности и оборудования, и программного обеспечения. Роль интегратора ограничена компоновкой системы из готовых, по сути, стандартных компонентов, согласно требованиям конкретного заказчика, ее привязкой к конкретному объекту, инсталляцией и вводом в эксплуатацию. Работа с программным обеспечением ограничивается настройкой готового пакета для выбранного набора оборудования в соответствии с пожеланиями тех, кто в дальнейшем с системой будет работать.



Незаменимым подспорьем стали камеры машинного зрения в системах помощи судейским бригадам

Таким образом, разрешение и скорость работы камер ограничиваются способностями человека-оператора, а основной упор делается на минимизацию объемов видеозаписей и удобство работы с ними. Функции программного обеспечения также четко определены и реализованы в готовых продуктах для массового применения.

Индустрия машинного зрения имеет несколько иные приоритеты, вытекающие из гораздо более широкого круга решаемых задач, результатом чего стало весьма ограниченное распространение и использование "коробочных" продуктов. В силу природного разнообразия предметов и задач наблюдения требования к системам захвата изображения очень сильно варьируются от задачи к задаче. Изначально

предполагается машинная обработка изображения, что влечет за собой требования к максимальной передаче деталей, разнообразию и единообразию условий съемки для повышения эффективности (детали), скорости и достоверности (условия съемки) работы алгоритмов. Список главных моментов в выборе компонентов машинного зрения видится нам следующим образом:

- качество изображения, степень детализации и скорость (частота кадров) должны соответствовать используемым математическим алгоритмам для решения разных прикладных задач;
- условия освещенности должны быть по возможности стабильными и/или контролируруемыми. В большинстве случаев речь идет об искусственном освещении;
- ограниченное использование или полное отсутствие в камере автоматизированных функций типа автоэкспозиция или автофокусировка. Все контролируется внешним программным обеспечением;
- основная обработка информации выполняется на отдельных вычислителях, так как сложность алгоритмов не позволяет разместить требуемые вычислительные мощности в компактный корпус камеры. В ряде случаев требуется совместная обработка изображений с нескольких камер. Тип и мощность вычислителя определяются требованиями конкретной задачи и используемой математики;
- требуются высокоскоростные интерфейсы для передачи изображений с большим разрешением (детализация) и высокой частотой кадров (фиксация быстро протекающих процессов);
- функциональность программного обеспечения от производителей камер ограничивается набором драйверов для гибкой конфигурации работы оборудования. Прикладные программы разрабатываются под конкретный заказ.

В результате мы получаем гораздо больше вариантов в выборе оборудования (камер, оптики, подсветки, вычислителей) и значительно более глубокое вовлечение интеграторов в процесс разработки программного обеспечения под конкретную задачу, для конкретного заказчика. Ради справедливости нужно сказать, что и в области машинного зрения тоже есть набор стандартных задач, где вполне удачно применяются стандартные решения. Например, контроль маркировки готовой продукции (акцизных марок, штрих- и DMC-кодов), но это лишь малая толика из всего многообразия использования камер машинного зрения в промышленности.

Чтобы в полной мере оценить разницу, давайте рассмотрим ключевые технологические различия между IP-камерами и камерами машинного зрения:

- разрешение, разрядность и скорость захвата изображений;
- высокоскоростные коммуникационные интерфейсы;
- специальные функции.

Когда размер имеет значение

При всем разнообразии сегодня трудно найти камеру для машинного зрения с разрешением меньшим, чем VGA (640x480). Наиболее популярная размерность – 2 Мпкс, FullHD или чуть больше и, наконец, "верхушка" массового диапазона – 12 Мпкс, 4K FullHD. Выпускаются камеры и с большим разрешением – в 20 и даже 50 Мпкс, но их применение ограничено и стоимостью, и сложностью обработки таких гигантских объемов данных. Визуально оценить разницу в масштабах массовых размерностей можно по приведенному рисунку.

Важно обратить внимание на то, что при увеличении разрешения необходимо сохранить достаточную чувствительность, которая при прочих равных определяется физическим размером пикселя. Способность элемента матрицы воспринимать свет прямо пропорциональна его площади – например, квадратный пиксель со стороной в 2 мкм имеет чувствительность в шесть раз ($25/4$) худшую, чем пиксель со стороной в 5 мкм. Менее очевидным следствием использования более крупного пикселя является глубина резкости – она прямо пропорциональна его геометрическому размеру. То есть величина пикселя (и, как следствие, размер матрицы) самым непосредственным образом влияет на количество передаваемых оттенков, возможность работы с малым освещением (или короткой выдержкой) и глубину резкости получаемого изображения. В частности, поэтому компактные Action-камеры никогда не смогут сравниться по качеству картинки с камерами машинного зрения.

Увеличение разрешения само по себе является причиной: во-первых, повышенных требований к емкости каналов передачи данных, во-вторых, при сохранении приемлемой чувствительности – увеличения размеров сенсоров и, соответственно, оптики. Если в IP-камерах большинство объективов предназначено для сенсоров размером 1/3" и 1/2", редко больше, то в машинном зрении "приличная" камера с FullHD/2 Мпкс-разрешением и пикселем, близким к 5 мкм, имеет размер матрицы минимум 2/3" (On Semi Python 2000, CMOSIS CMV2000), а то и близкий к 1" (Sony Pregius IMX174, IMX249). С одной стороны, выбор объективов для машинного зрения в этой размерности нельзя назвать скудным, с другой – привычных в IP-индустрии объективов с управляемой диафрагмой или Zoom для 2/3" совсем немного, а для 1" – единицы. Если же требуется камера с большим разрешением, например 4K HD/12Мп (CMOSIS CMV12000), то размер матрицы неумолимо приближается к профессиональным 35 мм фотоаппаратам, что, с другой стороны, делает возможным использование соответствующей оптики – F-mount (Nikon), EF/EFS (Canon), M42 и других "больших" форматов.

Для того чтобы воспользоваться более чувствительным сенсором, в камерах среднего и выше среднего диапазона устанавливаются не 8- (типичная ситуация для IP-камер), а 10-и даже 12-разрядные аналого-цифровые преобразователи. В комбинации с повышенным динамическим диапазоном сенсора это дает увеличение точности передачи оттенков серого или цвета в несколько раз (8 разрядов – 256 оттенков, 10 – 1024, 12 – 4096). Важно, что картинка при этом не только выглядит лучше визуально,

но и содержит гораздо больше аналитической информации. Несмотря на существенное увеличение стоимости комплекта камера + оптика, при необходимости получения изображения требуемого качества для машинной обработки на эти затраты приходится идти. Относительно высокая цена делает особенно значимым соответствие параметров камеры и оптики. Возможности системы в целом будут определяться ее слабейшим компонентом: камерой или объективом – в такой ситуации очень легко переплатить из желания сэкономить. Как минимум на первых порах целесообразно обращаться за подбором оборудования к профессионалам.

Коммуникационные интерфейсы

Как несложно догадаться, большая размерность и более высокая частота кадров при отсутствии сжатия требуют существенно большей пропускной способности каналов передачи данных. Индустрия машинного зрения, не сегодня столкнувшись с этой проблемой, разработала и продолжает разрабатывать ряд специализированных интерфейсов. Вместе с тем общедоступные компьютерные интерфейсы тоже не стоят на месте, благодаря чему у нас есть достаточно большой выбор стандартных (во всех смыслах) решений для разных задач. Наиболее популярные интерфейсы и оценки их производительности приведены в таблице.

Параметры производительности интерфейсов

Интерфейс	Пропускная способность, Гбит	Длина "медного" кабеля*, м	FullHD/2 Мпкс, кадр/с	4K/12 Мпкс, кадр/с
GigE	1,0	100	50	–
USB 3.0	5,0	8	165	–
CameraLink	10,0	10	340	60
CoaXPress	25,0	25	600	100
PCIe x 4, Gen3	64,0	5	1500	300

* приведена оценочная величина. Для конкретного оборудования длина кабеля может отличаться. В ряде случаев длина может быть существенно увеличена за счет активных и пассивных повторителей и удлинителей.

Наибольшее распространение сегодня получили GigE и USB 3.0. Первый – в силу использования доступной сетевой архитектуры Ethernet, второй – стандартного оснащения им большинства современных компьютеров. В обоих случаях камера может получать питание по тому же проводу: в случае GigE через PoE (Power Over Ethernet) инфраструктуру, в случае USB 3.0 – это делается по умолчанию.

Специальные возможности

Необходимость получения дополнительной, помимо собственно изображения, информации, управления процессом захвата изображения, синхронизации с внешними устройствами и другими камерами, выливается в ряд дополнительных функций, реализованных в камерах машинного зрения. Вот наиболее важные из них:

- цифровой запуск Triggering;

- служебная информация в кадре Chunk Frame;
- предустановленные параметры съемки Sequencer;
- синхронизация по Ethernet – PTP (Precise Time Protocol).

Управляемый цифровой затвор, или Triggering, предоставляет возможность с помощью цифровых линий ввода/вывода гибко управлять как работой самой камеры – открывать затвор или делать серию снимков по цифровому сигналу, так и управлять внешними устройствами, например подсветкой. Интересные возможности в адаптации к сложным условиям съемки открывает оперативное управление длительностью экспозиции посредством длительности сигнала триггера – без перепрограммирования параметров работы самой камеры, то есть очень быстро. В зависимости от используемой вычислительной платформы для принятия решения и перенастройки требуются десятки и даже единицы миллисекунд.

Функция Chunk Frame предназначена для добавления в каждый кадр служебной информации. Это может быть порядковый номер кадра, значение внутреннего счетчика – таймера камеры или состояние цифровых линий. Данная информация используется при последующей обработке последовательности кадров. В новых камерах реализована уникальная возможность задавать несколько наборов настроек, используемых при получении изображения. Вполне очевидными значениями экспозиции и усиления не ограничились: сюда были включены область кадра (ROI), баланс белого и ряд других параметров. В зависимости от установок триггера камера может снять либо последовательность кадров с разными настройками по одному сигналу, либо последовательно выполнять захват изображения кадр за кадром с разными настройками в заданном порядке. Таким образом, появляется возможность без лишних затрат и усилий получить набор кадров, снятых с разными установками, выбрать из них наилучший или использовать несколько для извлечения требуемой информации, которую сложно или вовсе невозможно получить иным способом. Возможность анализа последовательности кадров с получением выраженного цифрой результата – еще одна важная особенность машинной обработки изображений, недоступная оператору.



Одно из применений технологий машинного зрения в спорте - фотофиниш

Как уже упоминалось выше, зачастую для получения требуемого результата при обработке используется информация от нескольких камер. При этом изображения должны быть получены синхронно или через точно заданные промежутки времени. Традиционно схема синхронизации нескольких устройств, в том числе и камер, реализуется на базе выделенных цифровых линий. Каждая камера оснащается как минимум одной цифровой линией для внешнего запуска и одной цифровой линией для выдачи цифрового сигнала на другое устройство – камеру или осветитель. Такая система замечательно работает, если камеры расположены относительно близко друг от друга. В случае если они разнесены на сотни метров, например по стадиону, прокладка дополнительного кабеля и вероятные задержки и искажения по нему делают ее реализацию сложной, если не невозможной. К счастью, эта задача актуальна не только для распределенных систем машинного зрения, но и для многих других промышленных распределенных систем. Несколько лет тому назад был разработан международный стандарт IEEE 1588, который позволяет на основании статистических алгоритмов и PTP-протокола (Precise Time Protocol) синхронизировать внутренние часы отдельных устройств в распределенной системе с точностью до долей миллисекунд, а при использовании специализированных Master-контроллеров точного времени и лучше. Благодаря этому десятки камер могут синхронизировать собственные встроенные часы и выполнять захват изображений синхронно, в заданные моменты времени. Важно отметить, что синхронизация выполняется по обычной сети Ethernet, то есть не требует создания отдельной дополнительной инфраструктуры, с минимальными или даже нулевыми дополнительными затратами.

Зачем все это нужно?

Самым очевидным применением вышеописанных возможностей может быть использование в ходе тренировочного процесса доступной системы высокоскоростной видеорегистрации для изучения техники выполнения движений спортсменами. Требования к такой системе относительно просты: возможность регистрации и последующего воспроизведения движений спортсмена с максимальной детализацией. Камера может быть подключена к обычному ноутбуку или компьютеру по стандартному интерфейсу USB 3.0.

Незаменимым подспорьем стали камеры машинного зрения в системах помощи судейским бригадам. Видеофиксация положений "аут", "офсайд", "гол" разрешает большое количество ненужных споров. В этом случае критически важной будет скорость регистрации, чтобы точно зафиксировать момент или рассчитать траекторию пересечения мячом или шайбой линии. В такой системе в зависимости от дисциплины может использоваться до десятка камер. Часто система устанавливается на поле непосредственно перед матчем, поэтому должна быть достаточно мобильной. Благодаря развитию компьютерных, коммуникационных технологий и технологий машинного зрения стоимость компонентов и сложность интеграции таких систем значительно снизились. Сегодня многие команды, клубы и федерации могут себе позволить подобную "роскошь", казавшуюся недоступной еще несколько лет назад.

Современный спорт предъявляет все более высокие требования не только к самим спортсменам, но и к команде тренеров. В игровых дисциплинах, где ключевым преимуществом команды является не столько наличие "звезд", сколько правильная организации командной игры, оценка действий каждого игрока, их взаимодействия и команды в целом и на тренировке, и в ходе игры является бесценной информацией для выработки тренером эффективной тактики и стратегии. Для получения объективной, оперативной информации уже недостаточно просто видеофиксации игрового поля. Математическая обработка видеопотока позволяет получить в каждый момент времени положение каждого игрока своей команды и команды соперника. Для многих спортивных арен и клубов системы спортивной видеоаналитики стали сегодня весьма прибыльным бизнесом. Создание таких систем требует организации синхронной работы десятков камер, мощных вычислительных возможностей, но с лихвой окупается полученным результатом.

В спорте находят применение системы машинного зрения разной стоимости и масштаба. Стоимость оснащения спортивной арены может исчисляться сотнями тысяч евро, а небольшая локальная система скоростной видеорегистрации уложится в несколько тысяч. При этом львиную долю этой стоимости составляет отнюдь не оборудование, а программное обеспечение и инжиниринг. Отсюда высокая, зачастую запредельная стоимость приобретения, установки и обслуживания готовых систем из-за рубежа, рассчитанных на индустрию спорта с другими запросами и бюджетами. Невысокая стоимость оборудования и понятная архитектура открывают для новых в области машинного зрения игроков относительно легкий путь для продвижения

собственного бизнеса в спортивной индустрии на основании уже имеющегося опыта работы с системами видеонаблюдения. Нужно отметить, что отечественные разработчики, традиционно сильные в интеграции и разработке сложного программного и математического обеспечения, сегодня могут в полной мере помочь воспользоваться этой возможностью.

Автор



Максим Сорока

Генеральный директор ООО "ВиТэк"