

## **Содержание**

### **Глава 1 Компьютерное зрение**

#### **1.1 История развития компьютерного зрения**

#### **1.2 Задачи и функции компьютерного зрения**

##### **1.2.1 Задачи компьютерного зрения**

##### **1.2.2 Функции компьютерного зрения**

### **Глава 2 Области применения компьютерного зрения**

#### **2.1 Видеонаблюдение и безопасность**

#### **2.2 Машинное зрение для роботов**

#### **2.3 Автомобильная отрасль**

#### **2.4 Применение компьютерного зрения в военных целях**

#### **2.5 Потребительский рынок**

#### **2.6 Медицина и здравоохранение**

#### **2.7 Сельское хозяйство**

#### **2.8 Розничная торговля**

#### **2.9 Логистика, доставка товаров**

#### **2.10 Производство**

#### **2.11 Искусственный интеллект общего применения (AGI – Artificial General Intelligence)**

#### **2.12 Другие применения распознавания образов**

### **Глава 3 Рынок**

#### **3.1 Мировой рынок CV**

#### **3.2 Рынок CV в России**

### **Заключение**

## Введение

Систем компьютерного зрения (СКЗ) – молодая область информационных технологий, связанная с распознаванием образов и изображений. История СКЗ берет начало с конца 50-х годов XX века. Тем не менее, современный мир уже невозможно представить без технологий распознавания изображений.

Использование систем для распознавания образов является одним из основополагающих моментов промышленной стратегии развития для большинства предприятий мира. Об этом говорит экспоненциальный рост рынка потребления программных решений, использующих технологии распознавания изображений, наблюдающийся в последние годы. Мировой рынок промышленного внедрения систем с элементами использования компьютерного зрения, составлявший в конце прошлого и начале нынешнего столетия около 8 млрд. долларов, имеет тенденцию увеличиваться в три раза каждые несколько лет. Рост компаний, работающих в области СМЗ, составляют от 15 до 30% ежегодно. Но, имеющиеся коммерческие разработки покрывают менее 10% всех теоретических и потенциальных возможностей систем компьютерного зрения.

Возможности области, в которых могут применяться СКЗ долгое время получали развитие, преимущественно, как машинный аналог человеческих глаз. Основной функцией и назначением систем компьютерного зрения является распознавание внешнего вида и расположение в пространстве наблюдаемых объектов, а также, преобразование оцифровывание изображений, для последующей обработки на персональном компьютере и (или) передачи. Наиболее значимые направления использования систем компьютерного зрения на данный момент являются:

- зрение роботов;
- устройства для дистанционного управления;
- средства обработки и распознавания зрительных образов и символьной информации;
- системы для мониторинга и обеспечения безопасности;
- системы медицинской диагностики и биометрического контроля.

Также, СКЗ имеют достаточно серьезные преимущества перед человеческим зрением. Системы компьютерного зрения, оборудованные сенсорными специфическими устройствами и датчиками, могут увидеть то, что недоступно простому невооруженному глазу. Увеличение и расширение спектрального диапазона восприятия СКЗ за пределы восприятия человеческим глазом расширяет функционал и возможности систем компьютерного зрения значительно. Примером устройств, превосходящих по характеристикам человеческое зрение, являются системы для обеспечения безопасности, телевизоры и приборы ночного видения. Важной особенностью таких систем является и то, что они обеспечивают преобразование "невидимого" представления объектов в доступный для восприятия человека диапазон, тем самым увеличивая возможности человека в оперативном принятии решений.

Реализация возможностей систем компьютерного зрения, которые перечислены выше, стабильно обеспечивается современными устройствами (фотоприёмники, видеокамерами). Они обеспечивают высокую пространственную выборку освещенности наблюдаемых объектов. Спектрально-селективные оптические детекторы и источники

света открывают новый качественный уровень возможностей систем компьютерного зрения, освоение которых находится еще в стадии своего зарождения. Новый уровень связан с использованием систем компьютерного зрения в решении задач для автоматизации мониторинга за наблюдением у контролируемых объектов специфических свойств, которые, например, обусловлены их химическим составом.

Современный уровень развития программно-аппаратных средств и разработка новых принципов компьютерного зрения и анализа изображений раскрывают перспективу создания принципиально нового и уникального поколения систем компьютерного зрения. Имеющиеся исследования указывают, что устройства с использованием алгоритмов распознавания изображений перспективны для многих областей практических разработок программных продуктов: мониторинга качества и безопасности объектов, промышленного и природного происхождения, контроля изменений в экосистемах, экспресс-диагностики, загрязнения объектов несвойственными им, в том числе и неизвестными, соединениями. Так же, современный уровень развития технологий позволяет использовать системы машинного зрения в распознавании лиц, эмоций на фотографиях и в видеорядах, что позволяет открыть новые горизонты в использовании систем безопасности. Тема является малоизученной, так как недостаточно много предприятий и научных лабораторий занимаются вопросами развития компьютерного зрения (вывод сделан из обзора рынка систем компьютерного зрения).

## **Глава 1 Компьютерное зрение**

### **1.1 История развития компьютерного зрения**

Компьютерное зрение (техническое (машинное) зрение, computer vision, зрение роботов) как самостоятельная дисциплина оформилась в середине предыдущего столетия. Данное направление возникло в момент исследования искусственного интеллекта и в период, когда были споры о возможности или невозможности создания "мыслящей" машины и искусственного интеллекта. Из работ по распознаванию образов, выделилась отдельная дисциплина компьютерное зрение именно в тот момент, когда была осознана его математическая специфика, что изображение (фотография) представляет собой двумерную проекцию трехмерного мира, то есть "вырожденное" преобразование. Следовательно, для правильной и полной интерпретации изображения, необходимо иметь дополнительную информацию.

После сравнения компьютерного и человеческого зрения, выделили несколько способов получения вторичной информации о предметах: стали рассматривать активный осмотр, использование стереопары, априорных моделей увиденного, содержательных, семантических, а также геометрических. Рассмотрение было в основном "на бумаге" (а тем более в нашей стране) так как подходящей технической базы еще не изобрели, а получавшиеся математические модели были слишком сложны.

Использование фотоизображений для измерений начали использовать задолго до распространения компьютерной техники (фотограмметрия). Данный подход предлагалось использовать, и он был использован для аэрофотосъемки, а позже – в основном для космической навигации и прочих задач. Но для решения других практических задач использовались и иные способы получения дополнительной информации: в разных видах искусственная организация среды, например, специальная подсветка (щелевая подсветка),

использование разного рода маяков и маятников (маркировка деталей на конвейере или триангуляционные вышки при аэрофотосъемке), а также стандартное расположение предметов.

Первые задачи и результаты были получены в 70-х годах прошлого века (задачи были определены в начале 70-х годов, а первые результаты во второй половине 70-х). Первый этап работ по компьютерному зрению стал самым сложным: приходилось решать несколько задач: техническую и тяжелейшую задачу по вводу изображений в память электронно-вычислительной машины, которая была ограничена (64Кб для БЭСМ-6). Но вывод изображения на просмотр (как-то сформированного в памяти ЭВМ), было невозможно до изобретения системы ГРАФОР. Она позволяла выводить на бумагу при помощи графопостроителя контурный рисунок или график, они очень сильно отличались от изображений, которые мы имеем на экранах своих персональных компьютерах. Немного лучше была ситуация и для ЭВМ, которые могли выводить изображения на дисплей(М-6000), но и на этих машинах выводились только контурные или точечные изображения. Обработку получавшихся изображений (графического препарата) и моделирование трехмерных объектов, предметов и сцен осуществлялась при помощи самописного и нестандартного матобеспечения, поэтому компьютерное моделирование и техническое макетирование на тот момент были очень упрощенными, хотя и требовали усилий.

Качественный рывок в области компьютерного зрения произошел в 1990-х 2000-х годах. Это стало возможным благодаря появлению новых мощных компьютеров (суперкомпьютеров). Но в этом периоде ничего нового в сфере "переворота" устоявшейся дисциплине не было, скачок был обусловлен все время повышающимся интересом к компьютерному зрению и, в следствии, выросшими попытками переноса методов из других дисциплин (например: математическая статистика, искусственный интеллект, фотограмметрия и другие) в компьютерное зрение.

Стоит отметить, что первым результатом, который стал видимым после такого прорыва, явился прогресс в сфере разработки методов описания изображений. Использование свертки изображений при помощи вейвлетов и фильтров Габора – позволяли описывать изображения в компактной форме, уменьшая их размерность. Это позволило открыть целое направление: поиск по базам данных изображений. Более того, так как теперь описания интересных областей изображения стали усовершенствованными (например, дескриптор и его аналоги), стало возможным реализовывать плотную трехмерную конструкцию. И её точность примерно равна точности моделей, которые были получены при активном лазерном сканировании.

Результаты, которые были получены в начале 21-ого века в фотограмметрии, уже позволяли заниматься построением трехмерных моделей в медицине почти в режиме реального времени. Так же трехмерные модели широко применялись и применяются для разработки и построения компьютерных моделей городов. Из программных продуктов существует приложение PhotoSynth. Оно обрабатывает множество фотографий достопримечательностей и пытается создать их трехмерные модели. Получившиеся модели позволяют насладиться видом достопримечательностей, не отходя от экрана компьютера. Математические модели этих объектов получаются не очень точными, но визуально получившиеся трехмерные объекты позволяют передать гораздо больше информации нежели отдельные фотографии и позволяют создавать "эффект присутствия".

Это свидетельствует о том, что синтетический мир в данном контексте ближе, чем может показаться на первый взгляд.

Так же долго не занимались обработкой видео, потому что ощущался недостаток мощности компьютеров в начале 2000-х стала активно развиваться в рамках области компьютерного зрения. Системы видеонаблюдения и биометрии теперь стали использоваться компьютеры, также повсеместно стало внедряться программное обеспечение по распознаванию лиц. На сегодняшний день уже даже ведутся работы по анализу распознавания поведению и действий по видеоинформации.

Слияние таких дисциплин как компьютерное зрение и компьютерная графика позволило создать «дополненную реальность» (augmented reality), состоящую из ощущений реального мира (в частности, изображений) и добавленных к ним "мнимых" объектов вспомогательно-информативного свойства.

Например, моделью дополненной реальностью является трехмерная модель города (например, Помпеи), которая была сгенерирована по изображениям и фотографиям, в которую добавлены графические элементы: виртуальные деревья, мосты и персонажи. Следовательно, дополненная реальность позволяет вносить некоторые искусственные элементы при моделировании реалистичных объектов. Дополненная реальность отличается от виртуальной тем, что не создает новый мир. Она может использоваться как для развлечения, так и в науке: там, где нужен синтез информации разного происхождения, её унификация, представление в компактном виде.

На данный момент крупнейшие поисковики (Яндекс, Google) предоставили возможность навигации по городам мира, но на их сайтах можно найти не только карты городов, но и «прогуляться» по кварталам, увидеть их в таком виде, как увидел бы житель этих городов, который находится непосредственно в том квартале, за которым ведется "наблюдение". Вполне возможно, что с помощью дополненной реальности в скором времени станет возможным получать информацию о строениях, в т.ч. визуализировать их, используя снимки фасадов, а также создавать вид из окна любого этажа.

На сегодняшний момент теория компьютерного зрения как самостоятельный раздел кибернетики сложилась полностью, сейчас он опирается на большую практическую и научную базу знаний. Ежегодно по данной тематике проводятся конференции и симпозиумы, выпускается множество различных программных и программно-аппаратных комплексов, издаются множество книг и пишутся научные статьи по данной тематике. Имеются несколько научных организаций, которые освещают, поддерживают и принимают участие в исследованиях в областях современных информационных технологий, в том числе системах компьютерного зрения. В частности к ним относятся: SPIE (Международное сообщество по оптической инженерии), IEEE Computer Society (Institute of Electrical and Electronics Engineers), РОФДЗ (Общество содействия развитию фотограмметрии и дистанционного зондирования) и ряд других организаций.

## **1.2 Задачи и функции компьютерного зрения**

### **1.2.1 Задачи компьютерного зрения**

Главными задачами, которые ставит перед собой машинное зрение, являются:

- обнаружение объектов;
- их распознавание;

- отслеживание объектов на статичном изображении и/или в видеопотоке при определенных заданных условиях.

#### Распознавание и обнаружение

Стандартное задание в машинном и компьютерном зрении и обработке изображений – это обнаружение характерных объектов, каких-либо характеристик или действий на видеоданных. Зрение человека способно справиться с ним без особой нагрузки и при практически любых условиях. При этом, стоит отметить, что процент распознавания объектов у него будет достаточно высок. Однако в области компьютерного зрения распознавание случайных объектов в случайных ситуациях остается нерешенной проблемой, с которой еще предстоит разобраться.

Безусловно, уже в наличии присутствуют различные методы, которые позволяют находить объекты и справляться с этим успешно. Но они работают только для конкретных предметов: простые геометрические объекты, различного вида символы, человеческие лица и автомобили. При этом им необходимы определенные условия для качественного распознавания. Как правило, это касается освещения, фона и расположения объекта относительно детектора. Подобные проблемы широко освещены в литературе.

Следует разобраться с терминологией:

- распознавание – это обнаружение предварительно заданных объектов или классов, как правило, с их двух или трехмерным положением на изображении. Также используется, когда программа заранее обучена характеристикам объектов;
- идентификация – это распознавание какого-либо единичного объекта и соотнесение его с определенным классом;
- обнаружение – это проверка видеоданных согласно каким-либо условиям, обговоренным заранее. Оно иногда используется для первичного анализа, так как способно работать быстро, основываясь на простых вычислениях. После нахождения нужного участка в видеоматериале или изображении, они передаются для анализа более серьезным методам, которые задействуют большие ресурсы и воссоздающим правильные интерпретации.

Можно выделить несколько особенных задач, которые концентрируются на распознавании конкретно текстов:

- поиск изображений по содержанию: процесс обнаружения всех изображений из базы, в которых заложены определенные характеристики, общие для группы и важные для нахождения;
- оценка положения: идентификация положения конкретного объекта, его направления или ориентации относительно объектива камеры;
- оптическое распознавание знаков: идентификация нужных символов, печатных и рукописных, на изображениях с целью их трансфера в текст, который возможно будет редактировать, индексировать и в целом удобнее работать (например, ASCII);
- восстановление сцены. Когда в наличии есть два или больше изображения сцены, а также видеоданные. Его задача воссоздать трехмерную модель сцены. Для начала моделью может быть набор точек трехмерного пространства. Более сложные методы воспроизводят полную трехмерную модель.
- восстановление 3D формы по 2D изображениям: такое возможно при использовании реконструкции и стереорекострукции поля нормалей и карты

- глубины, с ориентацией на закраску полутонового изображения; также реконструкции карты глубины по текстуре; определения формы по перемещению.
- восстановление изображений. Как правило, здесь решается проблема наличия шума (шум записывающего устройства, размытость т.д.). Самым легким способом решения этой задачи является использование фильтров, например, нижних и средних частот. Этого достаточно для базового улучшения изображения. Для более серьезной работы используется анализ данных на наличие линий или границ, а затем, на основе собранных данных, проводится более глубокое исправление изображения;
  - выделение на изображениях структур определенного вида, сегментация изображений;
  - анализ оптического потока. Он связан с идентификацией перемещения пикселей между двумя изображениями. Также он способен решить задачи, связанные с оценкой движения в целом, где последовательность изображений обрабатывается таким образом, что становится возможным создать оценку скорости каждой точки изображения или 3D сцены. Яркими примерами данного анализа являются: следование камеры за перемещениями животных, людей или машин, а также определение трехмерного движения камеры.

### 1.2.2 Функции компьютерного зрения

Реализация систем компьютерного зрения сильно зависит от области их применения. Некоторые системы являются автономными и решают специфические проблемы детектирования и измерения, тогда как другие системы составляют под-системы более крупных систем, которые, например, могут содержать подсистемы контроля за механическими манипуляторами, планирования, информационные базы данных, интерфейсы человек-машина и т. д. Реализация систем компьютерного зрения также зависит от того, является ли её функциональность заранее определённой или некоторые её части могут быть изучены и модифицированы в процессе работы. Однако, существуют функции, типичные для многих систем компьютерного зрения.

**Получение изображений:** цифровые изображения получают от одного или нескольких датчиков изображения, которые помимо различных типов светочувствительных камер включают датчики расстояния, радары, ультразвуковые камеры и т. д. В зависимости от типа датчика, получающиеся данные могут быть обычным 2D изображением, 3D изображением или последовательностью изображений. Значения пикселей обычно соответствуют интенсивности света в одной или нескольких спектральных полосах (цветные или изображения в оттенках серого), но могут быть связаны с различными физическими измерениями, такими как глубина, поглощение или отражение звуковых или электромагнитных волн, или ядерным магнитным резонансом.

**Предварительная обработка:** перед тем, как методы компьютерного зрения могут быть применены к видеоданным с тем, чтобы извлечь определённую долю информации, необходимо обработать видеоданные, с тем чтобы они удовлетворяли некоторым условиям, в зависимости от используемого метода. Примерами являются:

- Повторная выборка с тем, чтобы убедиться, что координатная система изображения верна
- Удаление шума с тем, чтобы удалить искажения, вносимые датчиком

- Улучшение контрастности для того, чтобы нужная информация могла быть обнаружена
- Масштабирование для лучшего различения структур на изображении

**Выделение деталей:** детали изображения различного уровня сложности выделяются из видеоданных. Типичными примерами таких деталей являются:

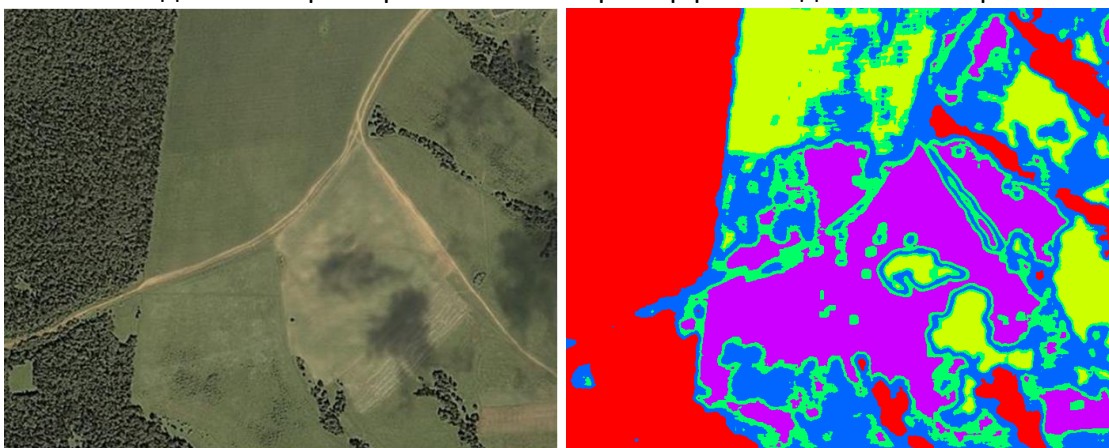
- Линии, границы и кромки
- Локализованные точки интереса, такие как углы, капли или точки: более сложные детали могут относиться к структуре, форме или движению.

**Детектирование/Сегментация:** на определённом этапе обработки принимается решение о том, какие точки или участки изображения являются важными для дальнейшей обработки. Примерами являются:

- Выделение определённого набора интересующих точек
- Сегментация одного или нескольких участков изображения, которые содержат характерный объект

Сегментация — это процедура разделения цифрового изображения на некоторое количество частей (массив пикселей, в свою очередь называемых суперпикселями). Целью сегментации считается облегчение либо изменение представления визуального изображения, для того чтобы его было легче и еще проще рассматривать. Сегментация изображений нередко применяется с целью того, чтобы определить предметы и границы (линии, кривые, и т. д.) в изображениях. Наиболее конкретно, сегментация изображений — это процедура присвоения таких меток каждому пикселю изображения, что пиксели с одинаковыми метками обладают общими визуальными особенностями.

В итоге сегментации изображения возникает множество сегментов, которые совместно покрывают целиком изображение, или множество контуров, выделенных из изображения. Все пиксели в секторе схожи по некоторым чертам или вычисленным свойствам, к примеру, по цвету, яркости либо текстуре. Примыкающие элементы существенно различаются по данным характеристикам. Вот пример работы данного алгоритма.



Пример работы алгоритма сегментации.

В данном случае для поиска сегментов использовались текстурные признаки.

Помимо этого, сегментацию используют:

1. В медицине для обнаружения опухолей или других патологий, диагностике при планировании лечения.
2. В распознавании лиц.
3. Системы управления дорожного движения.



Так как сегментация является принципом, а не алгоритмом для реализации было разработано несколько универсальных алгоритмов:

1. Метод, основанный на использовании кластеризации.
2. Методы с использованием гистограммы.
3. Выделение краёв.
4. Методы разрастания областей.
5. Методы разреза графа.
6. Сегментация с помощью модели.

**Высокоуровневая обработка:** на этом шаге входные данные обычно представляют небольшой набор данных, например набор точек или участок изображения, в котором предположительно находится определённый объект. Примерами являются:

- Проверка того, что данные удовлетворяют условиям, зависящим от метода и применения
- Оценка характерных параметров, таких как положение или размер объекта
- Классификация обнаруженного объекта по различным категориям

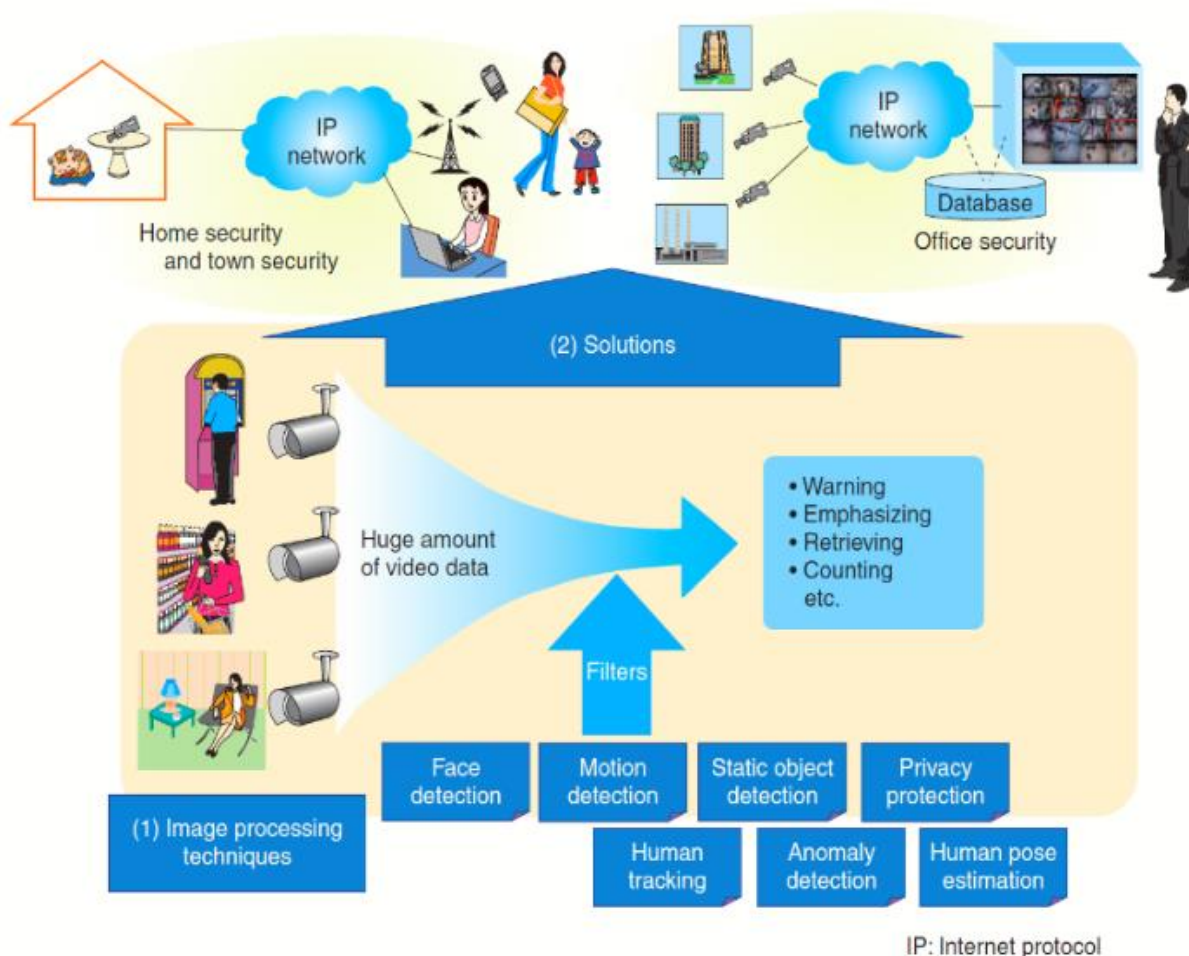
## Глава 2 Области применения компьютерного зрения

Сегодня компьютерное зрение широко применяется во многих отраслях цифровой экономики, таких как «Умный город», автономные автомобили и системы помощи водителю (ADAS), беспилотные летательные аппараты, высокотехнологичное сельское хозяйство, здравоохранение и многих других.

### 2.1 Видеонаблюдение и безопасность

Видеонаблюдение – важная часть физической безопасности. Видеонаблюдение с участием человека, по большей части, сводится к длительным периодам ожидания чего-то необычного на видеомониторе. Это очень важная работа, но очень утомительная. По оценкам психологов, среднее время удержания внимания человека на одном объекте не превышает 14 минут.

Поэтому были созданы т.н. системы интеллектуального видеонаблюдения IVS (intelligent video surveillance) на основе глубокого обучения, задачей которых является распознавание необычных событий или предметов на кадрах видеонаблюдения (см. рисунок ниже).



Интеллектуальное видеонаблюдение (источник: NTT)

Платформа обработки изображений (Image processing) осуществляет распознавание лиц (Face detection), обнаружение движения (Motion detection), обнаружение статических объектов (Static object detection), защиту приватности (Privacy protection), отслеживание маршрута перемещения людей (Human tracking), обнаружение аномалий (Anomaly detection), оценку позы человека (Human pose estimation). При обнаружении каких-либо необычных явлений, выдаются предупреждения, объект выделяется на экране рамкой, и пр. (Warning, Emphasizing, Retrieving, Counting, и т.д.).

В некоторых городах, таких как Лас-Вегас и Дубай, глубокое обучение в системах видеонаблюдения получило практическое применение в системах «Умного города». Например, такие системы могут информировать соответствующие службы о том, когда и где необходимо собирать мусор, обслуживать уличное освещение или управлять сигналами светофоров, например, переключать свет с красного на зелёный, если в поперечном направлении нет машин и пр.

## 2.2 Машинное зрение для роботов

Промышленные роботы-манипуляторы обычно хорошо выполняют повторяющиеся рутинные задачи. Однако практически беспомощны, когда задача меняется, например, когда объект манипуляций будет другого размера или конфигурации. Машинное зрение даёт роботу возможность автоматически адаптироваться к изменениям размеров или неточностям объектов и их произвольному расположению. Таким образом, применение машинного зрения для роботов позволяет производить разные продукты, ничего не меняя в самом роботизированном комплексе и без его полного перепрограммирования.



Адаптивный робот для сварки (источник: журнал Control Engineering Россия #5 (77), 2018)

## 2.3 Автомобильная отрасль

Доля смертей из-за автомобильных аварий составляет 2,2% от общего количества смертей в мире. Это примерно 1,3 млн. в год, или почти 3300 человек в день, не считая того, что от 20 до 50 млн. человек в год получают серьёзные травмы в результате ДТП. Причиной столь высокой смертности чаще всего является «человеческий фактор».

Системы предупреждения о боковом трафике (Cross Traffic Alert) также помогают предотвратить множество аварий, когда водитель не замечает транспорт, движущийся в поперечном направлении. Такие системы обычно строятся на базе радаров, работающих на высокой частоте (20 ГГц и выше). Однако они довольно дороги и могут устанавливаться в автомобилях высокого класса как дополнительная опция.

Компьютерное зрение способно значительно упростить такие системы и сделать их широкодоступными.



Работа системы предупреждения о движении в поперечном направлении (источник: cogentembedded.com)

## 2.4 Применение компьютерного зрения в военных целях

Основные применения CV в военных целях следующие:

- Видеонаблюдение,
- Автономные транспортные средства,
- Средства обезвреживания минных полей,
- Контроль качества при производстве боеприпасов.

## 2.5 Потребительский рынок

*Дрон с компьютерным зрением, распознающий препятствия*

Компания DJI выпустила новейший дрон Phantom 4 который способен распознавать препятствия при помощи встроенной системы CV и машинного обучения. Он способен самостоятельно выбирать маршрут полёта до указанной оператором цели. Процессор GPU дрона был разработан компанией Movidius.



Дрон компании DJI (источник DJI)

Компания Movidius объявила о сотрудничестве с Google в проекте внедрения глубокого обучения в смартфоны, что позволяет выполнять разработку изображений на смартфоне локально, а не опрашивать большое количество графических данных в облако. В дроне DJI используется именно такая технология.

## 2.6 Медицина и здравоохранение



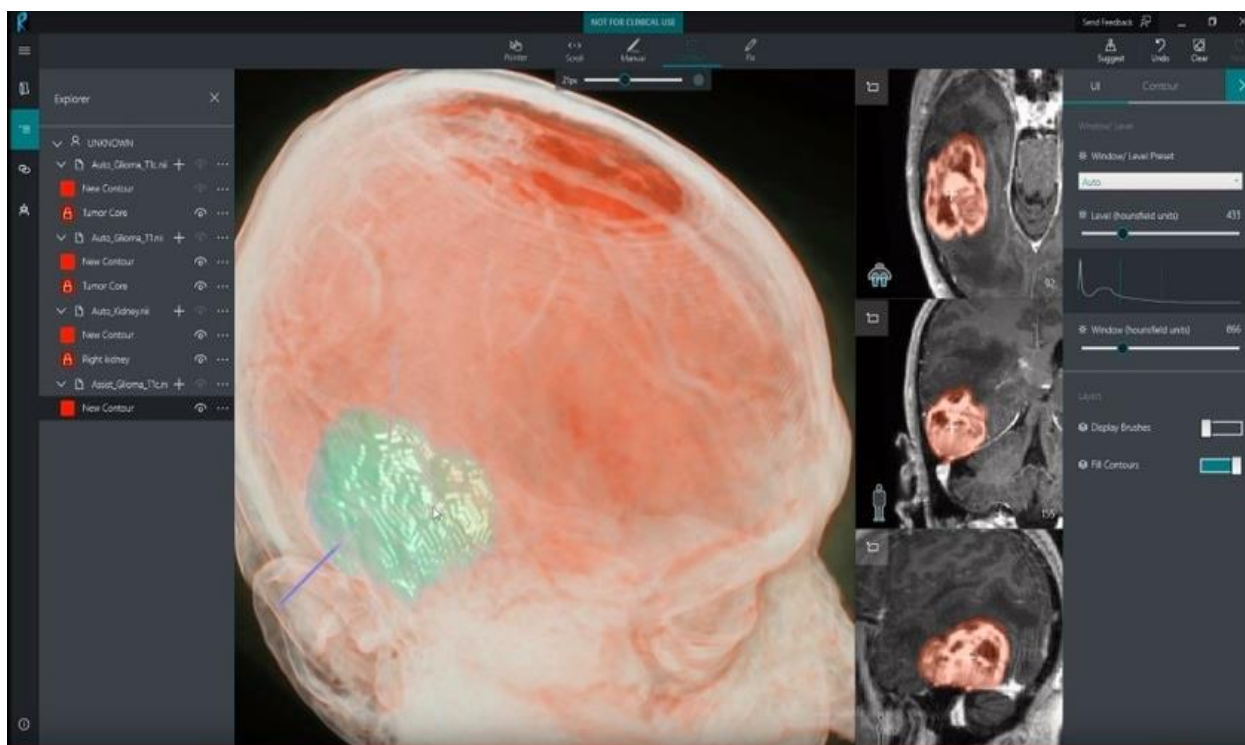
Применение компьютерного зрения для обработки медицинских изображений часто используется в компьютерной диагностике для планирования персональной терапии, медицинского ухода и улучшения принятия решений.

Системы с машинным обучением на базе изображений компьютерного зрения помогают врачу поставить диагноз, поскольку на изображении могут присутствовать мелкие детали, которые врач может не заметить, но такие детали могут быть распознаны системой CV с высокой степенью надёжности.

Кроме того, изображение может быть сравнено с тысячами других подобных изображений в базе данных медицинской системы, и результат сравнения используется для более точной постановки диагноза медицинским специалистом.

#### *Построение 3D-изображения раковой опухоли по данным компьютерной томографии*

Компания Microsoft разработала систему CV InnerEye, которая может визуально идентифицировать и отображать на мониторе врача возможные опухоли и другие аномальные образования по данным компьютерной томографии. Затем лечащий врач может более точно идентифицировать их. Для разработки InnerEye был применён алгоритм глубокого обучения на миллионах сканов компьютерной томографии разных пациентов.



Интерфейс системы CV InnerEye (источник: Emerj Artificial Intelligence Research, 2019)

Несмотря на то, что в здравоохранении существует множество прорывов и технологических достижений, из-за особенностей работы медицины, вероятно, пройдет ещё немало лет, прежде чем технологии CV в здравоохранении получат широкое распространение.

## **2.7 Сельское хозяйство**

Объем продукции сельскохозяйственного производства должен почти удвоиться, чтобы удовлетворить спрос на продукты питания для 9,7 миллиардов людей к 2050 году по данным ООН. Эффективность сельского хозяйства для этого должна вырасти почти на 25%. Применение CV-технологий совместно с системами глобального позиционирования

позволяет вести точное (прецизионное) земледелие (precision agriculture), которое может значительно повысить урожайность и эффективность сельского хозяйства.

Использование беспилотных летательных аппаратов позволяет получать топографические карты местности, а применение технологий обработки изображений позволяет получать 3D-модели участков земной поверхности с возможностью определения любых геометрических размеров. Погрешность геометрических измерений при этом не превышает десятков сантиметров.

#### *Определение зрелости хлопка*

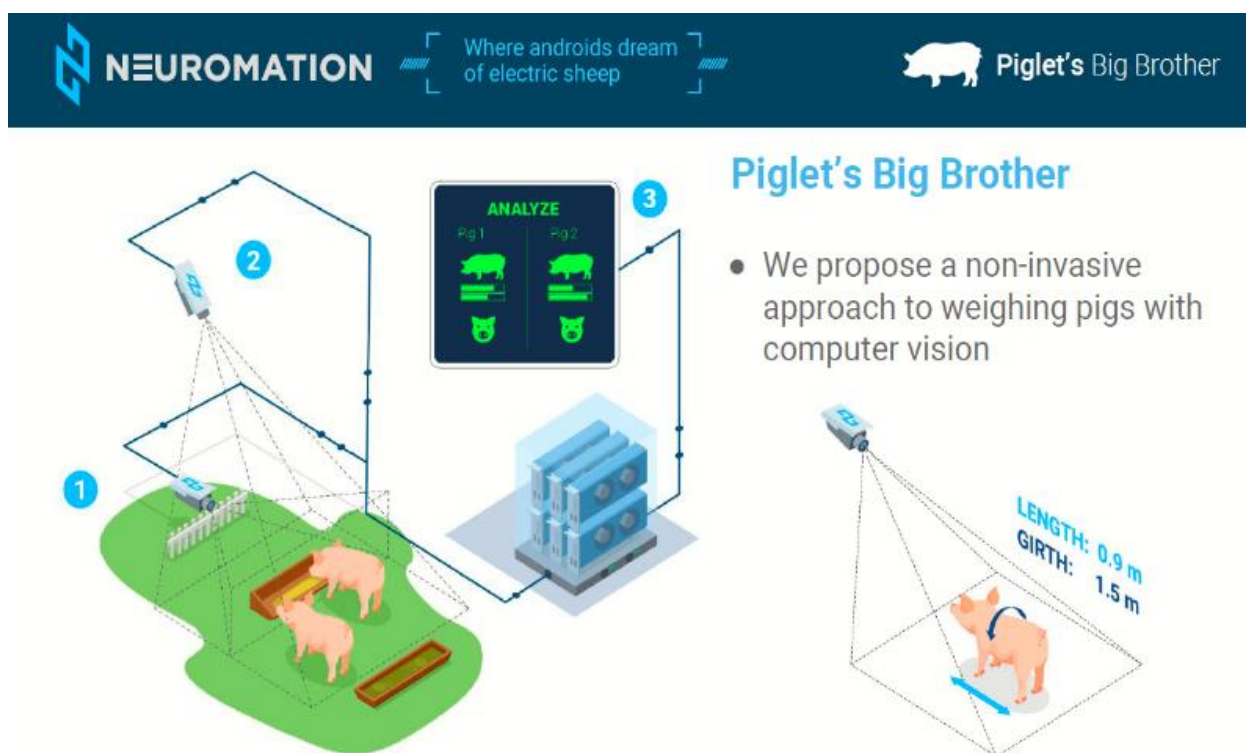
На больших сельскохозяйственных предприятиях, например, хлопковых или кукурузных полях, определение зрелости урожая обычно делается вручную. Такие расчёты, как правило, позволяют получить лишь приблизительную оценку и занимают много времени. Поэтому разработчики из Университета штата Теннесси (США) разработали систему CV с оснащённым камерами квадрокоптером для мониторинга зрелости хлопка.

Полученные с квадрокоптера фотографии обрабатывались с помощью алгоритма распознавания изображений, при этом удалось подсчитать урожай с точностью от 85% до 93% при использовании различных методов и средств анализа.

#### *Определение веса свиней*

Взвешивание свиней обычно делается только два раза за всё время их жизни: в начале и в конце откорма. Загнать животных на весы не очень сложно, но это огромный стресс для животного, а от стресса свиньи теряют вес. Если бы животноводы точнее знали, как идёт процесс откорма каждого поросёнка, то можно было бы составить индивидуальную программу откорма и определить индивидуальный состав пищевых добавок, что существенно улучшило бы общий выход продукции.

Поэтому был разработан новый, неинвазивный метод взвешивания животных на основе системы компьютерного зрения, которая оценивает вес свиней по фото- и видеоданным с использованием машинного обучения. На основании полученных данных корректируется процесс откорма.



### CV для доения коров

Компания «GEA Farm Technologies» разработала систему CV робота для доения коров. Система CV решает задачу точного наведения чашек насосов на соски вымени коровы (объекты), с помощью системы трекинга объектов и системы структурной подсветки, которая необходима для определения дальности от камеры до всех объектов.

Также была разработана система для автоматического обнаружения объектов на видео и их сопровождения в межкадровом пространстве в режиме реального времени: оценка положения, параметров формы и динамики движения в каждый момент времени.

В системе CV важными параметрами слежения за объектами являются время захвата объекта системой CV и его дальнейшее отслеживание. В разработанной CV удалось достичь времени в 3-10 мс на захват всех 4-х объектов, и 0,6 мс – на установку слежения за всеми найденными объектами. Система CV также может определять дальность до объектов в диапазоне 200-700 мм с точностью менее 2,5 мм на всём диапазоне дистанций.



Захват 4-х объектов

Отслеживание найденных объектов

Захват и отслеживание объектов (источник: GEA Farm Technologies)

### Прецизионное земледелие

В настоящее время во всём мире получают широкое распространение решения для прецизионного (точного) земледелия (Precision Agriculture), которые за счёт точного позиционирования сельхозтехники на обрабатываемом поле и, следовательно, более точной обработки пашни, позволяют поднять урожайность на 10% и более.

Компанией «Системы Компьютерного Зрения» разработана система точного земледелия, которая за счёт применения системы CV позволяет достичь точности позиционирования орудия обработки (имплементы) 2 см на дистанции в 6 м (см. рисунок ниже).



Система позиционирования имплементы (источник: compvisionsys.ru)



## 2.8 Розничная торговля

CV в сочетании с алгоритмами искусственного интеллекта позволяет ритейлерам автоматизировать процессы, которые раньше требовалось выполнять вручную. Уведомления о закончившихся товарах или неудовлетворении других требований покупателей можно получать автоматически, создав систему аналитики непосредственно в торговой точке.

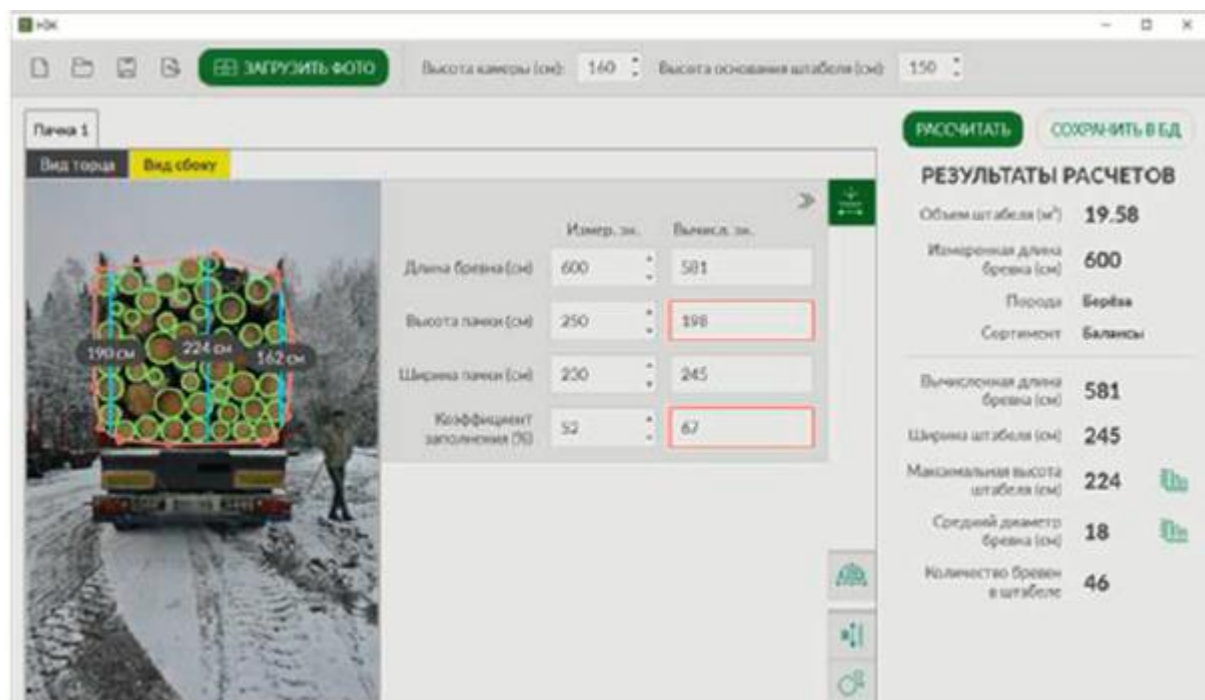
Например, компания X5 планирует внедрить в своих магазинах системы компьютерного зрения, чтобы отслеживать наличие полного ассортимента товаров на полках и вовремя выкладывать раскупленные товары, а также контролировать длину очереди на кассе и, в случае появления большого количества покупателей в залах магазина, сразу же увеличивать количество работающих касс.

В 2018 году компания Amazon открыла магазин Amazon Go, с решением Just Walk Out Shopping, которое позволяет оплачивать товары автоматически при выходе из магазина без подхода к кассе. Камеры CV способны распознавать не только действия покупателя, когда он берёт товар с полки и кладёт в корзину, но и наоборот, когда он кладёт товар обратно на полку. В этом случае, товар удаляется из виртуальной корзины покупателя. Камеры отслеживают покупателя всё время, пока он находится внутри магазина, без распознавания лиц.

## 2.9 Логистика, доставка товаров

### Анализ складских запасов

Компания «Системы компьютерного зрения» разработала новую технологию определения объёмов древесных брёвен с помощью анализа изображений. Для получения точных данных достаточно сфотографировать штабель брёвен с двух сторон. Затем программа обработки изображений самостоятельно определит количество брёвен, плотность укладки и введёт нужные поправки. В качестве дополнительных опций можно определить количество коры, качество древесины (выявление гнили) и некоторые другие параметры.



Работа программы по определению объёма древесины (источник: compvisionsys.ru)



Данная система способна обеспечить погрешность вычисления объёма древесины не более 3%. Точность измерения составляет 97-98%. Для сравнения, при ручном методе точность измерения составляет 85-95%, а при пропуске лесовоза через дорогостоящую лазерную рамку – 90-95%.

## 2.10 Производство

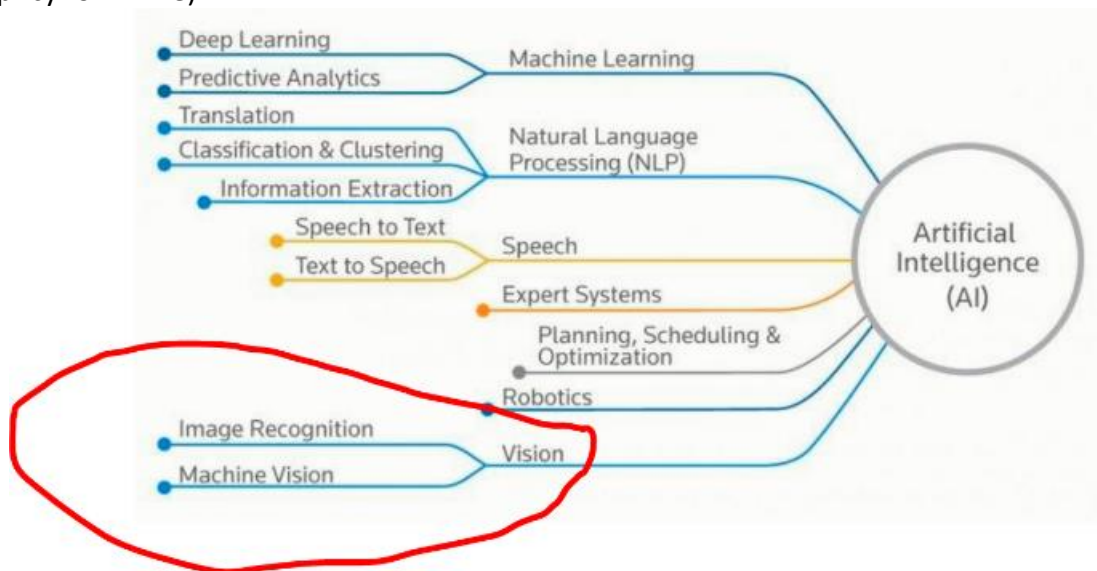
Современное высокотехнологичное производство требует особых подходов к контролю качества выпускаемой продукции. Компьютерное зрение (CV) совершило настоящий технологический прорыв и значительно расширило возможности дефектоскопии в промышленности, перевела ее на новый, более высокий уровень. Теперь технологии позволяют отслеживать качество не только после изготовления изделия или продукта, но и непосредственно во время производственного процесса. Кроме того, системы CV способны значительно упростить и ускорить дефектоскопию производственного оборудования, агрегатов и коммуникаций, находящихся в эксплуатации.

## 2.11 Искусственный интеллект общего применения (AGI – Artificial General Intelligence)

Термин (AGI) Artificial General Intelligence, появившийся относительно недавно, означает способность компьютера совершать абстрагированные умозаключения или, по крайней мере, имитировать этот процесс, тем самым приближая его к мышлению человека. Однако AGI пока находится на самой ранней стадии развития. Абстрактное мышление остаётся пока неразрешимой проблемой для искусственного интеллекта.

Именно по этой причине технологии AGI находятся на кривой Гартнера в самом начале подъёма «триггера инноваций».

CV является одной из важных составляющих технологий для искусственного интеллекта AI (см. рисунок ниже).

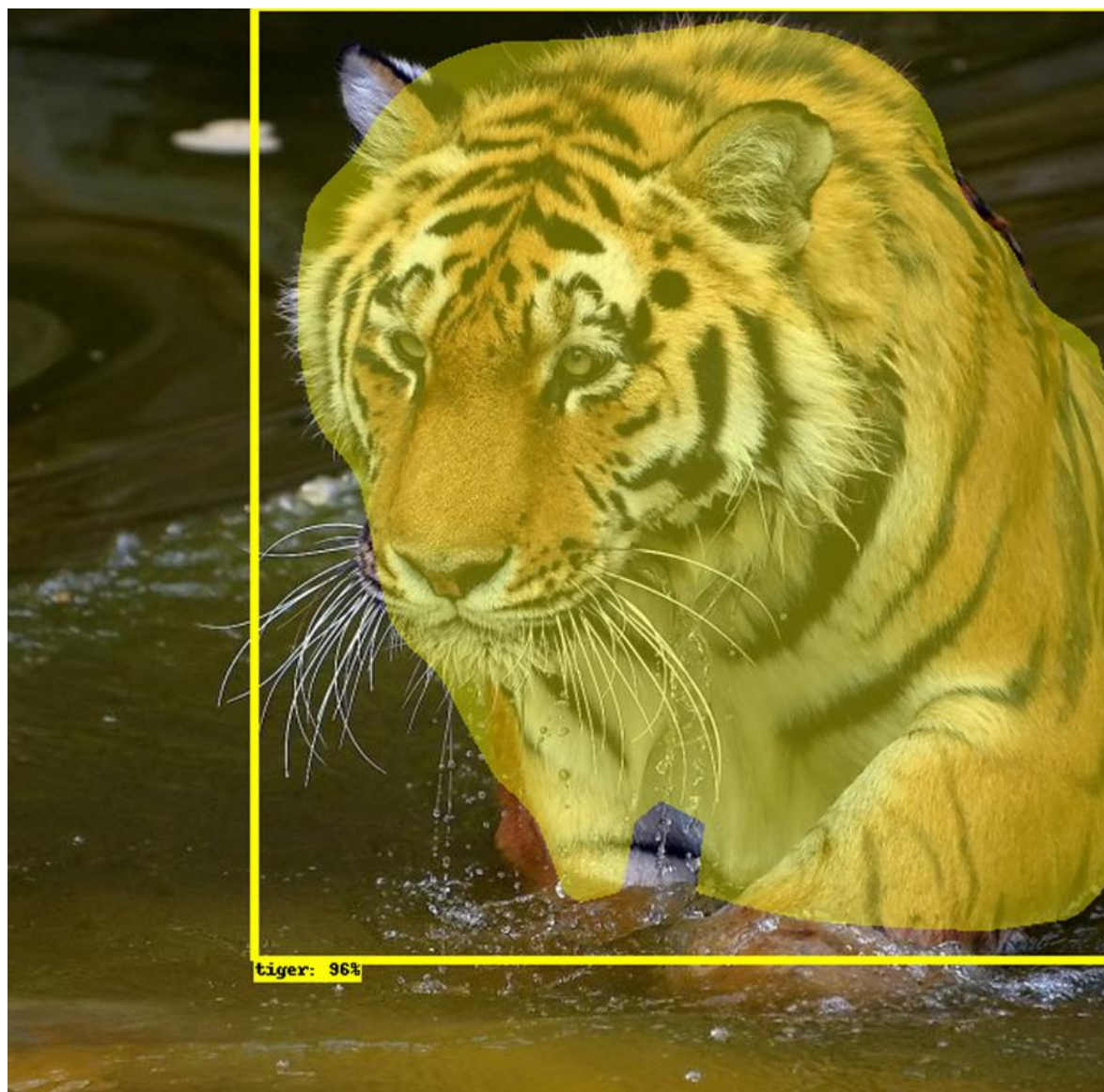


Технологии, лежащие в основе искусственного интеллекта AI (источник: thegalleria.eu)

## 2.12 Другие применения распознавания образов

Компанией «Системы компьютерного зрения» разработана система контроля популяции и передвижения амурских тигров при помощи компьютерного зрения на основе распознавания особей по снимкам с фотоловушек. Система автоматически определяет

принадлежность тигра к конкретному уникальному номеру или имени по изображениям, полученным системой с фотоловушек. Система позволяет вносить информацию о каждом тигре: уникальный идентификатор, имя, пол, возраст, сколько раз был сфотографирован и карту с его фотографиями, иметь ссылки на родственных тигров и возможность проставлять эти родственные связи. Идентификация тигра осуществляется с помощью алгоритмов компьютерного зрения с использованием свёрточных нейросетей.

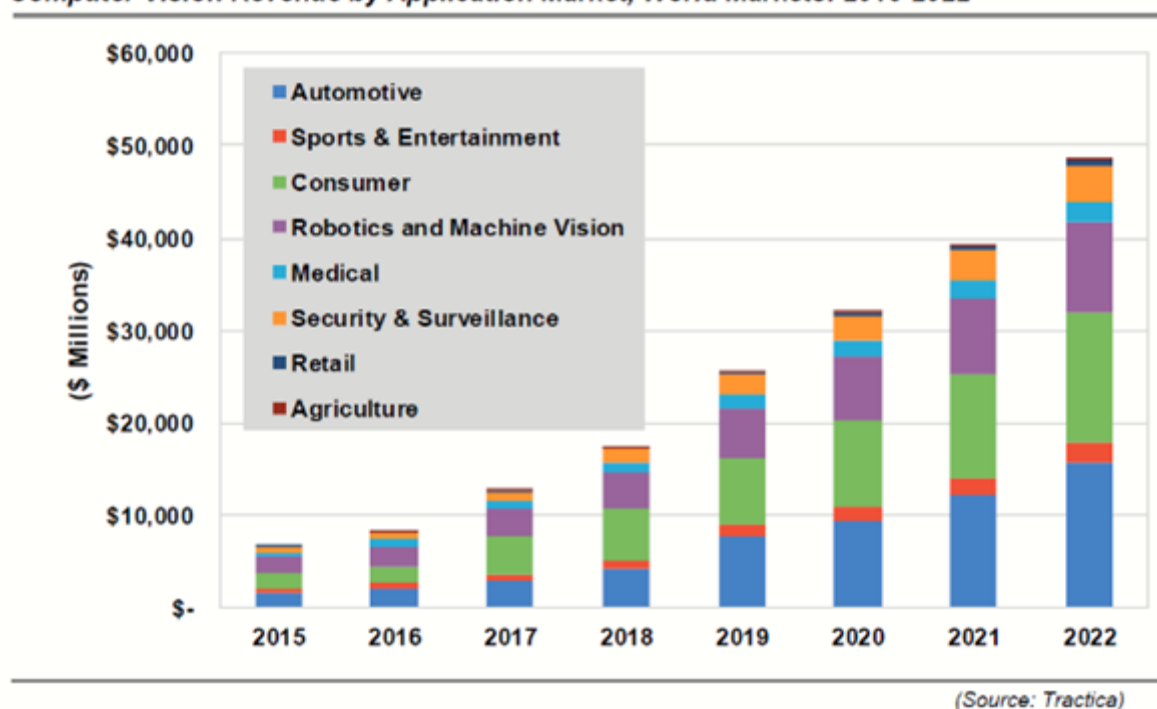


Фотография из системы контроля популяции и передвижения амурских тигров (источник: compvisionsys.ru)

## Глава 3 Рынок

Диапазон применений CV в последние 10-15 лет значительно расширился. Компания Tractica в своём отчёте о рынке компьютерного зрения в 2014 году указывает шесть областей CV. В 2016 году в новой версии отчёта Tractica указывает уже восемь областей применения компьютерного зрения: добавлены Retail (розничная торговля) и Agriculture (сельское хозяйство).

**Computer Vision Revenue by Application Market, World Markets: 2015-2022**



Рост объёма рынка CV 2015-2022 гг. (источник: Tractica, 2016)

### 3.1 Мировой рынок CV

Оценки объёма мирового рынка систем CV у различных аналитических компаний довольно сильно разнятся в зависимости от методологии исследования, таксономии и классификации технологий компьютерного зрения. Разные аналитические компании оценивают рынок по своим методикам, включая или не включая те или иные технологии и области в объём предметной оценки. Например, некоторые компании могут включать в состав технологий CV рентгеновские установки или МРТ, другие считают, что эти технологии к CV не относятся. Некоторые компании выделяют рынок распознавания образов отдельно от рынка CV, причём, по их оценкам, он превосходит рынок CV (или того, что они считают CV).

Одни компании могут оценивать рынок CV вместе с сопутствующей технологией искусственного интеллекта (AI), другие выделяют AI в отдельный рыночный сегмент. С другой стороны, далеко не весь объём AI относится к CV.

Все это затрудняет процесс объективной оценки рынка CV в мире, отдельных регионах и странах.

Например, наиболее авторитетная аналитическая компания McKinsey в своём исследовании об искусственном интеллекте указывает, что границы между многими технологиями CV чётко не очерчены, поэтому объёмы рынка не удаётся точно определить. Так, на 2016 г. McKinsey оценивает рынок Computer Vision с большим «разбросом»: от 2,5 до 3,5 млрд. долл. Причем наибольшую долю инвестиций среди смежных технологий, где может использоваться CV, занимает машинное обучение (Machine Learning) с уровнем инвестиций в 5 – 7 млрд. долл.

### 3.2 Рынок CV в России

Рынок компьютерного зрения весьма сложен для оценки его объёма и прогноза его роста по ряду причин:

- Неопределённость таксономии: что именно следует относить к компьютерному зрению? Некоторые компании оценивают, например, только рынок видеокамер, не включая программное обеспечение. Другие относят к компьютерному зрению только смарт-камеры со встроенным ПО обработки изображений, а отдельные платформы обработки изображений и искусственный интеллект считают другим сегментом и т.д. С другой стороны, очевидно, что далеко не все решения искусственного интеллекта используются в целях CV.
- Несмотря на то, что технологии CV используются достаточно давно, в России рынок систем и решений CV пока нельзя считать окончательно сформировавшимся.
- Компьютерное зрение применяется в широком спектре задач и часто его оценка производится только для каких-то определённых сегментов, например, распознавания лиц, которое, в свою очередь, также может использоваться во многих секторах рынка: охранном видеонаблюдении, ритейле, поисковых мероприятиях и пр.
- Решения с использованием технологий CV в большинстве случаев являются органической частью других, более широких решений, например, «Умного города», и иногда бывает сложно оценить его долю в этих решениях.

В ходе исследования был проведён опрос более чем 50 участников рынка CV. Многие из них затруднились оценить объём рынка CV в России и дать прогноз его развития на ближайшие 3-5 лет. Ответы участников опроса, которые дали оценку объёма рынка CV в России, характеризовались большим «разбросом» - от 1 до 30 млрд. руб. в настоящее время и от 5 до 100 млрд. к 2025 году.

## Заключение

### *Будущее компьютерного зрения*

CV – быстро растущая область цифровых технологий, которая затрагивает многие стороны повседневной жизни.

Компания Apple внедрила функцию распознавания лиц в новые модели iPhone, приобретая такие компании, как PrimeSense, RealFace и Faceshift. Американский портал AngelList, объединяющий стартапы и инвесторов, составил список из 529 новых компаний, которые работают в области компьютерного зрения. Средняя капитализация таких стартапов составляет 5,2 млн. долл. Много стартапов привлекают капитал от 5 до 10 млн. долл. Портал отмечает, что поток инвестиций в компьютерное зрение нарастает. Замена человеческого зрения на компьютерное во многих областях – очень выгодное вложение капитала.

Точность анализа видеоинформации компьютером всё время растёт и применение CV может дать большую экономию средств наряду с улучшением качества.

Можно выделить пять основных тенденций развития CV:

1. Рост промышленных систем компьютерного зрения. CV для медицинских устройств, фармацевтики, производства пищевых продуктов, автомобильной промышленности предоставляет более высокий уровень контроля качества, а CV для промышленности, как ожидается, в 2019 году станет основным трендом в области компьютерного зрения.
2. Облачные системы глубокого обучения. Алгоритмы глубокого обучения и классификаторы нейросетей позволят более быстро и точно проводить классификацию и распознавание изображений от систем CV. В ближайшие годы число таких разработок значительно возрастет.

3. Робототехника. Использование промышленных роботов стремительно увеличивается. Поэтому спрос на системы CV для роботов будет расти.
4. Рост требований к параметрам оптики для CV, который вызван ростом требований к чёткости и разрешающей способности изображений CV. Разрабатываются и производятся сенсоры для CV-камер с большей разрешающей способностью и с большим количеством пикселей, однако, без качественной оптики эти усовершенствования будут малополезны. Поэтому разрабатываются такие инновационные решения, как микролинзы на каждый пиксель и др., которые кардинально могут повысить параметры работы оптических систем, которые уже подошли к своим технологическим пределам в традиционных решениях.
5. Использование термальных изображений при контроле производственных процессов. Обычно термальные камеры использовались в основном для военных целей, в охранном видеонаблюдении. Термальные изображение в комбинации с CV могут обнаруживать такие аномалии в производственном процессе, которые не видны глазу или традиционным системам CV.