

Проблемы и возможности применения видеоаналитики в сферах образования и развлечений


Анастасия Дмитриевна Окатьева

Студент

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики

Санкт-Петербург, Россия


okataeva@mail.ru

 0000-0002-9482-5673

Поступила в редакцию 15.05.2021

Принята 16.07.2021

Опубликована 15.08.2021

 10.25726/d6189-0390-2258-x

Аннотация

Быстрые изменения и совершенствование технологий произвели революцию в современном мире. Взаимодействие человека и компьютера (HCI) развивалось в течение определенного периода, трансформируя многие аспекты нашей жизни, включая то, как мы учимся. В настоящее время студенты могут извлечь выгоду из быстрого обмена информацией, доступности в Интернете и практической реализации того, что ранее преподавалось только в книгах. Опыт обучения и компетентность зависят от того, насколько хорошо предмет преподается студентам и через какую среду. Книги и текстовые ресурсы со временем зарекомендовали себя как отличный способ доставки и использовались на протяжении веков. Аудио-и видеоматериалы также оказались эффективным способом доставки информации, поскольку они обеспечивают хорошее количество богатого контента за относительно короткий период, что привело к повышению мотивации учащихся в классе и изменению восприятия преподавателей. Однако отсутствие погружения и контроля делает обучение на основе видео менее личным, чем интерактивные классы и моделирование реальной жизни. Видеонаблюдение с помощью видеоаналитики может быть развернуто для мониторинга территорий в определенное время суток. Например, как только школа открывается, не должно быть много активности на парковке или в определенных местах вокруг школы. В таких ситуациях интеллектуальные камеры с видеоаналитикой могут использоваться для обнаружения активности в тех областях, которые представляют интерес, чтобы предупредить службу безопасности школы о том, что что-то может потребовать их внимания. Радиолокационное обнаружение идеально подходит для периметров, где устройство может быть ненавязчиво настроено для оповещения, когда кто-то входит в определенную зону.

Ключевые слова

видеоаналитика, эмоции, распознавание, образование.

Введение

Автоматическое распознавание выражений лица является требовательным и интересным направлением в области компьютерного зрения, а также интерфейса компьютерного мозга. Методы распознавания выражений лица, основанные на компьютерном зрении, имеют определенные ограничения, такие как позиционирование камеры, мобильность пользователя и конфиденциальность. В системе распознавания выражений на основе компьютерного зрения, пользователь должен находиться в диапазоне камеры для записи его / ее выражений лица, например, 60 см и 80 см. «Положение головы и подбородка пользователя для стабилизированного положения головы ограничивает мобильность пользователя». Рассмотрим для начала существующие технологии видеоаналитики в образовании.

Использование развлечений в качестве стратегии социальных изменений-это практика, которая восходит к истокам организованных обществ. Информация, спасающая жизнь, общие ценности и убеждения, а также социально приемлемые практики передаются из поколения в поколение с помощью искусства рассказывания историй, используя сложные навыки для интеграции популярных развлечений с обучением. В высокоопосредованных культурах 21 века практикующие средства массовой информации и искусства активно используют многие формы развлечений для обучения своей аудитории. К концу 20-го века стратегическое использование развлечений для социального влияния и образовательных целей обычно называлось просоциальными развлечениями, развлечениями с добавленной стоимостью, развлечениями для развития, развлекательными, информационно-развлекательными и образовательными драмами, среди других терминов.

В 1999 году несколько ученых, изучающих стратегическое использование развлечений для образования и социальных изменений, остановились на термине коммуникационная стратегия развлечения и образования, который в настоящее время является общим описанием, используемым специалистами в области коммуникации. Развлечения-образование определяется как коммуникационная стратегия, которая намеренно стремится укрепить или изменить отношение, ценности, убеждения или социальные практики путем интеграции образовательного контента в развлекательные программы. В стратегии коммуникации между развлечениями и образованием есть пять основных элементов. Во-первых, это предполагает использование развлечений с целью обучения или убеждения аудитории. Все виды развлечений предоставляют информацию и оказывают влияние. Развлечение-образование, однако, должно быть целенаправленным, чтобы развлечение создавалось с учетом заранее определенного образовательного или убедительного эффекта. Во-вторых, развлечения-образование теоретически обосновано.

Стратегический характер развлекательного образования основан на теориях коммуникации или теориях в тесно связанных дисциплинах, таких как социальная психология, социология и политология. Изучение человеческого общения обычно считается академической областью исследований, которая является междисциплинарной, опирающейся на многие теоретические традиции. Таким образом, коммуникационная стратегия развлечения и образования опирается на множество различных теорий. В-третьих, развлечения-образование стремится создавать высококачественные развлечения, с помощью которых аудитория учится. Это требует сочетания развлекательных ценностей с образовательным контентом, чтобы аудитория училась, пока ее развлекают.

Развлечения-образование осуществляется через все средства массовой информации и искусства, включая кино, театр, музыку, народное искусство, уличную драму, телевизионные программы, радиопрограммы, живые выступления и спортивные мероприятия. Все, что можно считать развлекательным, также может быть образовательным, и, таким образом, коммуникационная стратегия развлечения и образования охватывает все средства массовой информации, искусство и опосредованные события, во время которых происходит общение. В-четвертых, развлекательно-образовательная стратегия - это коммуникационная стратегия, основанная на исследованиях. Она включает в себя формирующее исследование с потенциальной аудиторией до создания продукта или спектакля в области медиа или искусства, а также итоговое исследование для измерения влияния развлекательных и образовательных сообщений на их целевую аудиторию. Наконец, развлекательно-образовательная деятельность требует сотрудничества между специалистами в области средств массовой информации и агентами изменений, такими как специалисты в области здравоохранения и развития и медиаученые, которые создают цели в области образования и социальных изменений. Это сотрудничество может принимать различные формы, но обычно включает в себя переговоры между творческими деятелями и государственными должностными лицами или преподавателями.

Материалы и методы исследования

Высокотехнические области образования, от медицины до инженерии, могут обеспечить более широкий опыт обучения с помощью виртуальной реальности. Например, наиболее важной проблемой для студентов-медиков, изучающих анатомию, является понимание тела в трех измерениях и того, как

различные последствия для здоровья реагируют на различные методы лечения. Возможно, именно поэтому обучение в этой области может быть столь же обширным, как и 10 лет в медицинской школе. VR может решить эту проблему, сократить время, необходимое для обучения, и подготовить высококвалифицированных врачей быстрее, чем традиционные методы обучения (Abadi, 2016).

Например, введение и мониторинг внутривенного введения (внутривенно) является важной частью ухода за больными, поскольку до 70% пациентов имеют ту или иную форму внутривенного устройства, введенного в рамках их лечения. При неправильном выполнении осложнения могут быть такими же серьезными, как инфекции крови и смертность. Приложение VR, EON-XR позволяет получить практический опыт погружения, который позволяет студентам найти вены и выполнить процедуру в безопасной, безрисковой среде. Наряду с процедурными инструкциями студентам предоставляется дополнительная информация, такая как ключевые термины. Этот инструмент является портативным и может быть интегрирован в онлайн-обучение, а также в практические занятия в классе (Domínguez-Jiménez, 2020). VR позволяет использовать обучение, развлечение и исследование в различных отраслях науки и техники полностью безопасно, и эффективно (наглядно), чтобы продемонстрировать различные эффекты и реакции (например, химический опыт с едкими кислотами).

Результаты и обсуждение

1. Технические решения видеоаналитики в образовании и сфере развлечений
- 1.1. Обнаружение эмоций и выстраивание гипотез на основе полученных данных

При обнаружении эмоций в процессе видеоаналитики методом Виолы–Джонса обрабатываются изображения в интегральной форме. Интегральное представление позволяет выполнить быстрое вычисление суммарной яркости произвольного прямоугольника на оригинальном документе с постоянным временем, независимо от размеров этого прямоугольника. Указанное представление изображения – это матрица, совпадающая по размерам с исходным изображением, в каждом элементе которой хранится сумма интенсивностей всех пикселей, находящихся левее и выше данного элемента. Элементы матрицы рассчитываются по формуле:

$$II(x, y) = \sum_{0 \leq i \leq x, 0 \leq j \leq y} I(i, j), \quad (1)$$

где $I(i, j)$ – яркость пикселя исходного изображения.

Каждый элемент матрицы $II(x, y)$ представляет собой сумму пикселей в прямоугольнике от $(0, 0)$ до (x, y) ;

Прямоугольные комбинации, используемые для обнаружения объекта на изображении, не являются настоящими вейвлетами Хаара. Вместо этого они содержат прямоугольные комбинации, которые лучше подходят для задач распознавания. Из-за этой разницы, эти функции называют не вейвлетами, а функциями или примитивами Хаара. На рис. 1 представлены примитивы Хаара.

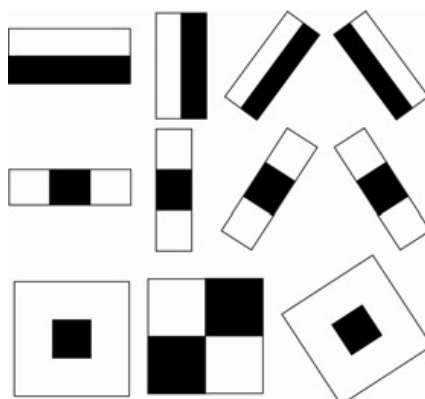
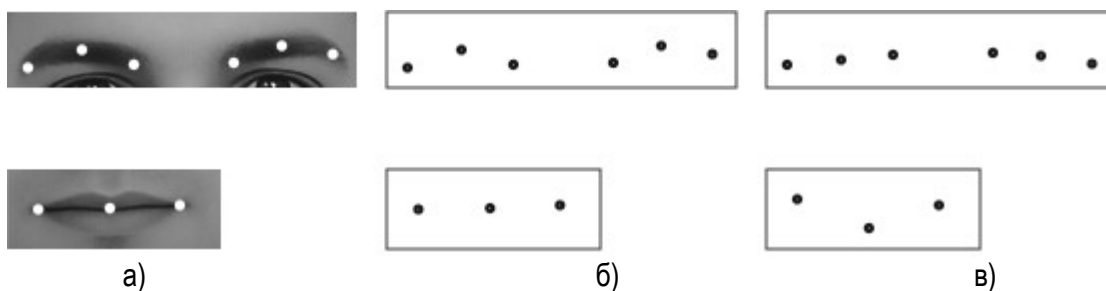


Рисунок 1. Примитивы Хаара

Следующим шагом является нахождение ключевых точек выделенных элементов лица. Определить эмоции можно на основе анализа нескольких ключевых точек. Например, на рис. 2 показаны комбинации точек бровей и рта.



а – изображение бровей и рта; б – комбинация ключевых точек бровей и рта, соответствующая их изображением; в – комбинация ключевых точек, соответствующая иным положениям рассматриваемых элементов на изображениях

Рисунок 2. Ключевые точки элементов лица

Процесс выделения ключевых точек проиллюстрирован на рис. 3.

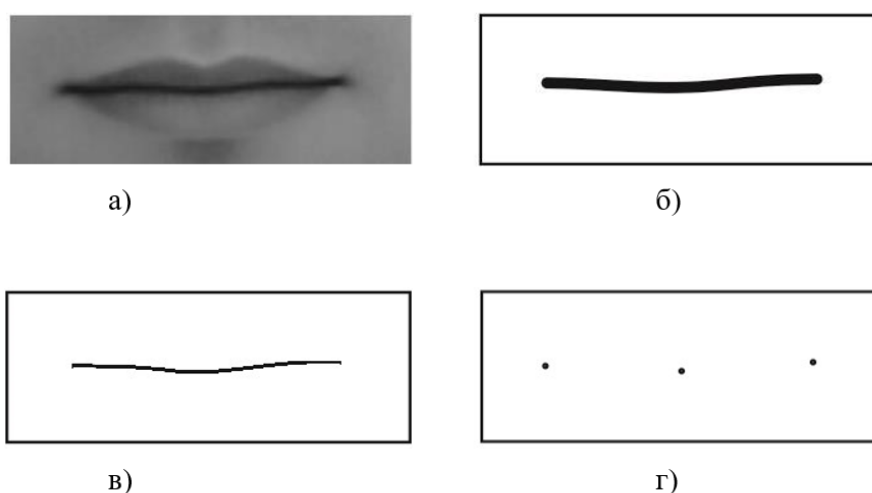


Рисунок 3. Выделение ключевых точек:

- а) полутоновые изображения,
- б) бинарное изображение,
- в) применение градиентной маски,
- г) локализация ключевых точек

После выделения ключевых точек можно выполнить классификацию эмоции. Например, в таблице 1 – представлены характеристики эмоций по сочетанию брови и рта. Проведен подход к автоматическому распознаванию эмоций может быть эффективно применен в разнообразных интеллектуальных человеко–машинных системах [13].

Таблица 1. Характеристики эмоций по сочетанию брови и рта

Эмоция	Бровь	РИТ
Удивление	Поднимается	Открывается
Страх	Поднимается и скривится	Открывается и растягивается
Отвращение	Снижается	Поднимается и концы снижаются
Гнев	Снижается и скривится	Открывается или концы снижаются
Счастье	Поднимается	Концы поднимаются
Грусть	Концы снижаются	Концы снижаются

1.2. Дискретные модели эмоций

Дискретная модель включает в себя шесть основных эмоций (счастье, грусть, страх, удивление, отвращение, гнев), которые являются общепринятыми (Liu, 2019). Все остальные эмоции считаются частью этих основных эмоций. Пространственная модель, как на рис. 4, рисует эмоции на двух шкалах – валентности и возбуждении. Валентность (valence) обозначает полярность эмоций, а возбуждение (arousal) обозначает интенсивность эмоций.

Эта модель определяет эмоции на двумерном пространстве – валентности и возбуждении. Валентность обозначает полярность эмоций (положительные или отрицательные), а возбуждение – интенсивность (высокая или низкая).

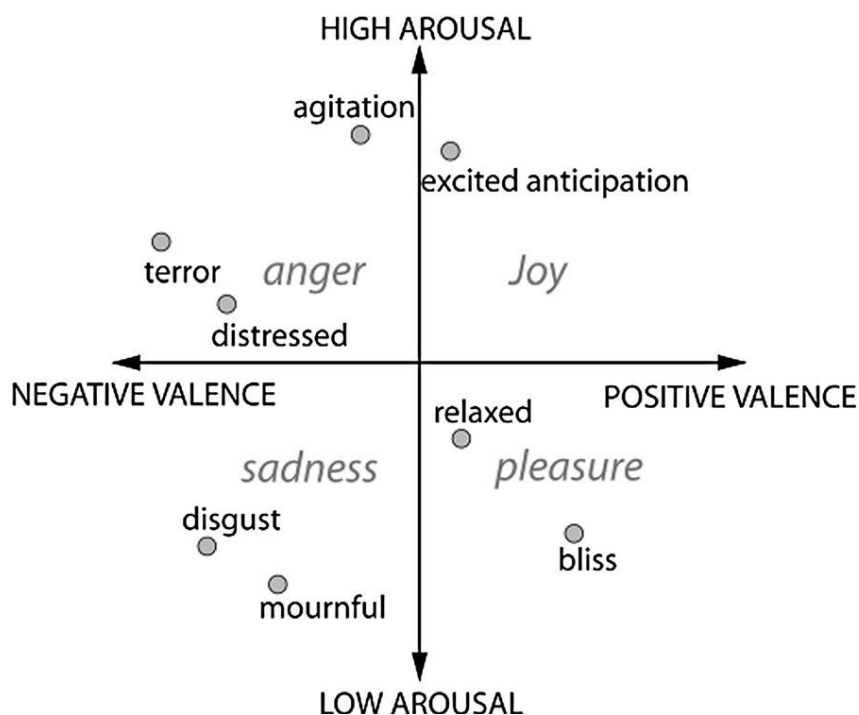


Рисунок 4. Мерная модель эмоций

Нелинейные особенности, такие как приближение энтропии (ApEn, Approximation Entropy), наибольшая экспонента Ляпунова (LE, Largest Lyapunov Exponent), корреляционная размерность (CD, correlation dimension), показатель Херста (H, Hurst) и ошибка нелинейного прогнозирования – передают информацию, связанную с такими свойствами, как сходство, предсказуемость, надежность и чувствительность сигнала (Chi, 2012).

Понятно, что когда люди оценивают качественное свойство объекта, такое как эмоциональное выражение лица, или когда они оценивают общее сходство сложных объектов, они сравнивают объекты по нескольким субъективным шкалам, а не по одной. Комбинируя определенным образом, субъективные выводы, они и осуществляют сложное суждение, похожее на психологическую оценку.

Для решения таких проблем и развивались методы многомерного отражения.

Поэтому сначала мы определяем дистанцию (разность) $\delta_{i,j}$ между кодами ЛБШ (локальные бинарные шаблоны)). Это расстояние отображает сходство шаблонов интенсивности изображения, которые были использованы для создания каждой строке кода ЛБШ. Полная матрица несоответствий, представляющая расстояния между всеми возможными значениями кодов, может быть определена как:

$$\Delta = \begin{bmatrix} \delta_{11} & \cdots & \delta_{1n} \\ \vdots & \ddots & \vdots \\ \delta_{d1} & \cdots & \delta_{dn} \end{bmatrix}, \quad (2)$$

для каждой матрицы Δ , MDS ищет отображение кодов на маломерное метрическое пространство, такое как:

$$\delta_{i,j} \approx \|V_i - V_j\| = \|MDS(C_i) - MDS(C_j)\| \quad (3)$$

Таким образом коды C_i и C_j отражены как V_i, V_j соответственно.

Разница между двумя кодами ЛБШ D может быть оценена как количество значений битов между ними, отличаются (расстояние Хемминга). Тем не менее, это может дать нам неполную информацию о различиях в шаблонах интенсивности, которые образовали данные коды: местонахождение различных значений битов столь же важны об оценке сходства кода, как и сами значения.

В качестве примера, рассмотрим снова три ЛБШ вектора $a = (1, 0, 0, 0, 0, 0, 0, 0)$, $b = (0, 1, 0, 0, 0, 0, 0, 0)$ та $c = (0, 0, 0, 0, 1, 0, 0, 0)$. Количество разных битов (а таким образом и расстояние Хемминга) для пары a и b такое же, как и для a и c - 2. Выходные текстурные шаблоны, которые образовали эти ЛБШ, очень отличаются: пара a и b похожи лишь ротацией одного пикселя в шаблоне вокруг центрального пикселя, в то время как код c представляет собой отражение шаблона интенсивности a , и следовательно менее похожи.

$$a = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad b = \begin{bmatrix} 0 & 1 & 0 \\ 0 & x & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad c = \begin{bmatrix} 0 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

Чтобы принять во внимание разницу в местоположении кодов пикселей, вместо расстояния Хеминга, мы использовали EMD (Earth Mover's Distance). EMD используется для отображения наименьшего усилия для преобразования одного распределения в другое. Здесь она была использована как мера разницы между двумя кодами ЛБШ. EMD между двумя кодами P и Q определяется следующим образом:

$$EMD(P, Q) = \frac{\min_{\{f_{kl}\}} \sum_{kl} f_{kl} d_{kl}}{\sum_{kl} f_{kl}}, s. t., \quad (5)$$

$$0 \leq f_{kl}, \sum_l kl \leq Q_l,$$

$$\sum_{kl} f_{kl} = \min\left(\sum_k P_k, \sum_l Q_l\right)$$

Это отображает усилия необходимые для изменения значения из одного кода в другой, где $f_{kl} \in \{0, 1\}$ это переход значения бита k в P (P_k) в бит l в Q (Q_l), $\{f_{kl}\}$ это полный переход из P в Q и

$$d_{kl} = \|k - l\|_2 \quad (6)$$

Использовано как стандартное расстояние между позициями двух битов и представляет собой усилие необходимое для переноса значения между позициями.

1.3. Системы распознавания лица

Система распознавания эмоций лица, представленная в (Munoz, 2014), состоит из двухступенчатого процесса, то есть обнаружения лица (ограниченное лицо) в изображении с последующим обнаружением эмоций на обнаруженной ограниченной поверхности. Следующие две методики используются для соответствующих упомянутых задач в системе распознавания лиц.

1) характерный каскадный классификатор Хаара: он хорошо обнаруживает фронтальную поверхность изображения. Это в реальном времени и быстрее по сравнению с другими детектором лица. Эта статья использует реализацию с Open-CV.

2) Модель Xception CNN (Convolutional neural network) (Mini_Xception, 2017): Будем готовить архитектуру модели CNN классификации, которая принимает ограниченную грань (48×48 пикселей) в качестве входных данных и прогнозирует вероятности 7 эмоций в выходном слое.

Данные состоят из 48×48 пикселей серых изображений лиц. Грани были автоматически зарегистрированы таким образом, чтобы лицо было более или менее центрированным и занимало примерно одинаковое количество места на каждом изображении. Задача состоит в том, чтобы классифицировать каждое лицо на основе эмоций, проявляющиеся в мимике, к одной из семи категорий (0 = сердит, 1 = мерзость, 2 = страх, 3 = счастлив, 4 = печален, 5 = сюрприз, 6 = нейтральный).

Если охватывать более широко данную тему, в области видеонаблюдения системы распознавания лиц (FR) используются для обнаружения лиц, представляющих интерес, появляющихся в распределенной сети камер. Производительность систем FR still-to-video может значительно снизиться, поскольку лица, снятые в неограниченном рабочем домене (OD) с помощью нескольких видеокамер,

имеют другое базовое распределение данных по сравнению с лицами, снятыми в контролируемых условиях в домене регистрации с помощью неподвижной камеры. В конкретной реализации, основанной на классификации разреженных представлений, синтетические грани, сгенерированные с помощью DSFS, используются для формирования междоменного словаря, учитывающего структурированную разреженность, где блоки словаря объединяют исходные и синтетические грани каждого отдельного человека. Экспериментальные результаты, полученные с помощью видеозаписей из наборов данных Chokerpoint и COX-S2V, показывают, что увеличение набора эталонных галерей систем FR still-to-video с использованием предлагаемого подхода DSFS может обеспечить значительно более высокий уровень точности по сравнению с современными подходами, при этом лишь незначительно увеличив его вычислительную сложность.

2. Проблемы использования систем видеоаналитики в сфере образования

Видеоаналитика не может решить все проблемы безопасности в образовании, но это ценный инструмент и один из наименее навязчивых доступных вариантов. Список проблем безопасности, с которыми может справиться видео, растет с каждым днем. Видеосистемы могут обеспечивать мониторинг школьных помещений в режиме реального времени и способствовать быстрому реагированию на инциденты. Новые достижения, такие как видеоаналитика, в настоящее время недостаточно используются в сфере образования.

Сегодня сеть позволяет обмениваться видеоизображениями по всей школьной системе, путешествовать по существующим сетям, расширяя возможности более централизованной структуры управления безопасностью и делая видео более ценным.

В частности, высшие учебные заведения с большей вероятностью будут просматривать видео в реальном времени, учитывая большие кампусы, большее количество зданий и больше общественных зон, где собираются сотрудники и студенты (Niewiadomski, 2016).

Многочисленные проблемы на рынке образовательных товаров и услуг для обеспечения безопасности (с точки зрения видео) включают в себя широкие открытые пространства, которые затрудняют обеспечение безопасности школ камерами видеонаблюдения, поскольку огромный объем необходимого покрытия может быть непомерно дорогим. Во-вторых, государственные и федеральные правила должны быть приняты во внимание и сбалансированы с необходимостью защиты частной жизни студентов.

Панорамные камеры являются одним из инструментов для решения проблем, поскольку одна 360-градусная камера может заменить от 4 до 5 PTZ-камер

Наконец, школы и колледжи сталкиваются с сокращением бюджетов, что означает, что решения для обеспечения безопасности должны обеспечивать больший охват и функциональность, а также быть экономически эффективными для развертывания. Панорамные камеры являются одним из инструментов для решения этих проблем, поскольку одна 360-градусная камера может заменить от четырех до пяти традиционных камер с панорамированием и наклоном, что приводит к меньшему количеству камер и большему охвату-и все это при меньших затратах на оборудование и лицензирование.

Видеонаблюдение с помощью видеоаналитики может быть развернуто для мониторинга территорий в определенное время суток. Например, как только школа открывается, не должно быть много активности на парковке или в определенных местах вокруг школы (Schwartz, 2013).

В таких ситуациях интеллектуальные камеры с видеоаналитикой могут использоваться для обнаружения активности в тех областях, которые представляют интерес, чтобы предупредить службу безопасности школы о том, что что-то может потребовать их внимания. Радиолокационное обнаружение идеально подходит для периметров, где устройство может быть ненавязчиво настроено для оповещения, когда кто-то входит в определенную зону.

Цель в потенциально опасной ситуации - ускорить время реагирования. Чем быстрее вы сможете обнаружить что-то с помощью технологии, тем быстрее вы сможете отреагировать. Таким образом, возможность идентифицировать что-то, происходящее на парковке, и предупредить сотрудников

школьных ресурсов может обеспечить 30 секунд или минуту форы для реагирования, что может привести к тому, что школа окажется в ситуации блокировки и быстрее получит первых ответчиков на месте.

Было показано, что использование освещения в ночное время может сдерживать преступность. Тем не менее, может быть дорого содержать здание и территорию освещенными всю ночь, каждую ночь. Чтобы смягчить эти проблемы и потенциальные затраты, существуют видеокamеры с чрезвычайно низкой освещенностью, которые позволяют им видеть в почти полной темноте или в некоторых случаях в полной темноте. Это позволяет школе экономить деньги, выключая свет, одновременно достигая уровня эффективности наблюдения, аналогичного дневному развертыванию.

Также, в системах видеоаналитики часто может использоваться VR технологии, для того, чтобы обеспечить полноценное исследование происходящей ситуации с видеоприборов.

Существуют видеокamеры с экстремально низкой освещенностью, позволяющие видеть в почти полной темноте или в полной темноте.

Заключение

Столкнувшись с более высоким уровнем инцидентов с учащимися и дисциплинарными проблемами учащихся в некоторых школах, школьная система в Соединенных Штатах стремилась модернизировать свою систему видеонаблюдения, чтобы обеспечить лучший местный и удаленный мониторинг в важных областях. Камеры высокой четкости Avigilon с самообучающимися решениями для видеоаналитики и контроля доступа были установлены в 101 школе, а программное обеспечение для управления видео ACC 6 с технологией поиска внешнего вида Avigilon обеспечивает расширенный поиск видеоаналитики.

VR технологии в области видеоаналитики могут использоваться с целью улучшения обучения, полного погружения в развлекательных целях, повышения эффективности наглядных экспериментов в полностью безопасных условиях. Благодаря полной наглядности происходящего, VR технологии являются наиболее популярными в данный момент и наиболее недооцененными, с точки зрения количества программного обеспечения к этой технологии.

Поисковая система искусственного интеллекта с глубоким обучением может сортировать многочасовые кадры и позволяет операторам быстро и эффективно нажимать на кнопку и искать все экземпляры человека или транспортного средства по всем камерам на сайте

Список литературы

1. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "Tensorflow: a system for large-scale machine learning," in 12th USENIX Symposium on Operating Systems Design and Implementation (Savannah, GA: USENIX Association), 265–283.
2. Chi, Y. M., Wang, Y.-T., Wang, Y., Maier, C., Jung, T.-P., and Cauwenbe, G. (2012). Dry and noncontact EEG sensors for mobile brain–computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* 20, 228–235. doi: 10.1109/TNSRE.2011.2174652
3. Danelljan, M., Robinson, A., Khan, F. S., and Felsberg, M. (2016). "Beyond correlation filters: learning continuous convolution operators for visual tracking," in European Conference on Computer Vision, eds B. Leibe, J. Matas, N. Sebe, and M. Welling (Amsterdam: Springer), 472–488. doi: 10.1007/978-3-319-46454-1_29
4. Domínguez-Jiménez JA, Campo-Landines KC, Martínez-Santos JC, Delahoz EJ, Contreras-Ortiz SH. A machine learning model for emotion recognition from physiological signals. *Biomedical Signal Processing and Control.* 2020;55:101646.
5. Fu, Q., Luo, Y., Liu, J., Bi, J., Qiu, S., Cao, Y., et al. (2017). "Improving learning algorithm performance for spiking neural networks," in 2017 IEEE 17th International Conference on Communication Technology (ICCT) (Chengdu: IEEE), 1916–1919. doi: 10.1109/ICCT.2017.8359963
6. Gavrilescu, M., (2015) "Recognizing emotions from videos by studying facial expressions, body postures and hand gestures", 23rd Telecommunication forum TELFOR, pp. 720-723.

7. Gu Y, Wang Y, Liu T, Ji Y, Liu Z, Li P, et al. EmoSense: Computational Intelligence Driven Emotion Sensing via Wireless Channel Data. *IEEE Transactions on Emerging Topics in Computational Intelligence*. 2020;4(3):216–226.
8. Heike Brock. (2018) “Deep learning - Accelerating Next Generation Performance Analysis Systems” 12th Conference of the International Sports Engineering Association, Brisbane, Queensland, Australia, pp. 26–29.
9. Hossain MS, Muhammad G. Emotion-Aware Connected Healthcare Big Data Towards 5G. *IEEE Internet of Things Journal*. 2018;5(4):2399–2406.
10. Liu, J., Huang, Y., Luo, Y., Harkin, J., and McDaid, L. (2019). Bio-inspired fault detection circuits based on synapse and spiking neuron models. *Nerocomputing* 331, 473–482. doi: 10.1016/j.neucom.2018.11.078
11. Munoz MO, Foster R, Hao Y. Exploring Physiological Parameters in Dynamic WBAN Channels. *IEEE Transactions on Antennas and Propagation*. 2014;62(10):5268–5281.
12. Niewiadomski Mancini, Varni Volpe, Camurri Automated Laughter detection from full body movements” *IEEE Transactions on Human-Machine Systems*, 46 (1) (2016), pp. 113-213
13. Schwartz G, Tee BCK, Mei J, Appleton AL, Kim DH, Wang H, et al. Flexible polymer transistors with high pressure sensitivity for application in electronic skin and health monitoring. *Nature communications*. 2013;4:1859. pmid:23673644
14. Shirbhate Neha, Talele Kiran, (2016), “Human Body Language Understanding for Action detection using Geometric Features”, 2ndInternational Conference on Contemporary Computing and Informatics, IEEE, pp.603-607.
15. Soroush, M. Z., Maghooli, K., Setarehdan, S. K., and Nasrabadi, A. M. (2019). A novel EEG-based approach to classify emotions through phase space dynamics. *Signal Image Video Process*. 13, 1149–1156. doi: 10.1007/s11760-019-01455-y
16. Wang X, Le D, Cheng H, Xie C. All-IP wireless sensor networks for real-time patient monitoring. *Journal of biomedical informatics*. 2014;52:406–417.
17. Wang, X., Zhang, T., Xu, X., Chen, L., Xing, X., and Chen, C. L. P. (2018). “EEG emotion recognition using dynamical graph convolutional neural networks and broad learning system,” in 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (Madrid: IEEE), 1240–1244. doi: 10.1109/BIBM.2018.8621147

Problems and possibilities of using video analytics in the fields of education and entertainment


Anastasia D. Okatyeva

Student

St. Petersburg National Research University of Information Technologies, Mechanics and Optics

Saint Petersburg, Russia


okataeva@mail.ru

 0000-0002-9482-5673

Received 15.05.2021

Accepted 16.07.2021

Published 15.08.2021

 10.25726/d6189-0390-2258-x

Abstract

Rapid changes and improvements in technology have revolutionized the modern world. Human-computer Interaction (HCI) has evolved over a period of time, transforming many aspects of our lives, including

how we learn. Currently, students can benefit from the rapid exchange of information, accessibility on the Internet, and practical implementation of what was previously taught only in books. Learning experience and competence depend on how well the subject is taught to students and through what medium. Books and text resources have proven to be a great delivery method over time and have been used for centuries. Audio and video materials have also proven to be an effective way to deliver information, as they provide a good amount of rich content in a relatively short period, which has led to increased motivation of students in the classroom and a change in the perception of teachers. However, the lack of immersion and control makes video-based learning less personal than interactive classes and real-life simulations. Video surveillance using video analytics can be deployed to monitor territories at certain times of the day. For example, once a school opens, there shouldn't be a lot of activity in the parking lot or in certain places around the school. In such situations, smart cameras with video analytics can be used to detect activity in areas of interest, to warn the school security service that something may require their attention. Radar detection is ideal for perimeters, where the device can be unobtrusively configured to alert when someone enters a certain area.

Keywords

video analytics, emotions, recognition, education.

References

1. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "Tensorflow: a system for large-scale machine learning," in 12th USENIX Symposium on Operating Systems Design and Implementation (Savannah, GA: USENIX Association), 265–283.
2. Chi, Y. M., Wang, Y.-T., Wang, Y., Maier, C., Jung, T.-P., and Cauwenbe, G. (2012). Dry and noncontact EEG sensors for mobile brain–computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* 20, 228–235. doi: 10.1109/TNSRE.2011.2174652
3. Danelljan, M., Robinson, A., Khan, F. S., and Felsberg, M. (2016). "Beyond correlation filters: learning continuous convolution operators for visual tracking," in European Conference on Computer Vision, eds B. Leibe, J. Matas, N. Sebe, and M. Welling (Amsterdam: Springer), 472–488. doi: 10.1007/978-3-319-46454-1_29
4. Domínguez-Jiménez JA, Campo-Landines KC, Martínez-Santos JC, Delahoz EJ, Contreras-Ortiz SH. A machine learning model for emotion recognition from physiological signals. *Biomedical Signal Processing and Control.* 2020;55:101646.
5. Fu, Q., Luo, Y., Liu, J., Bi, J., Qiu, S., Cao, Y., et al. (2017). "Improving learning algorithm performance for spiking neural networks," in 2017 IEEE 17th International Conference on Communication Technology (ICCT) (Chengdu: IEEE), 1916–1919. doi: 10.1109/ICCT.2017.8359963
6. Gavrilescu, M., (2015) "Recognizing emotions from videos by studying facial expressions, body postures and hand gestures", 23rd Telecommunication forum TELFOR, pp. 720-723.
7. Gu Y, Wang Y, Liu T, Ji Y, Liu Z, Li P, et al. EmoSense: Computational Intelligence Driven Emotion Sensing via Wireless Channel Data. *IEEE Transactions on Emerging Topics in Computational Intelligence.* 2020;4(3):216–226.
8. Heike Brock. (2018) "Deep learning - Accelerating Next Generation Performance Analysis Systems" 12th Conference of the International Sports Engineering Association, Brisbane, Queensland, Australia, pp. 26–29.
9. Hossain MS, Muhammad G. Emotion-Aware Connected Healthcare Big Data Towards 5G. *IEEE Internet of Things Journal.* 2018;5(4):2399–2406.
10. Liu, J., Huang, Y., Luo, Y., Harkin, J., and McDaid, L. (2019). Bio-inspired fault detection circuits based on synapse and spiking neuron models. *Nerocomputing* 331, 473–482. doi: 10.1016/j.neucom.2018.11.078
11. Munoz MO, Foster R, Hao Y. Exploring Physiological Parameters in Dynamic WBAN Channels. *IEEE Transactions on Antennas and Propagation.* 2014;62(10):5268–5281.

12. Niewiadomski Mancini, Varni Volpe, Camurri Automated Laughter detection from full body movements” IEEE Transactions on Human-Machine Systems, 46 (1) (2016), pp. 113-213
13. Schwartz G, Tee BCK, Mei J, Appleton AL, Kim DH, Wang H, et al. Flexible polymer transistors with high pressure sensitivity for application in electronic skin and health monitoring. Nature communications. 2013;4:1859. pmid:23673644
14. Shirbhate Neha, Talele Kiran, (2016), “Human Body Language Understanding for Action detection using Geometric Features”, 2ndInternational Conference on Contemporary Computing and Informatics, IEEE, pp.603-607.
15. Soroush, M. Z., Maghooli, K., Setarehdan, S. K., and Nasrabadi, A. M. (2019). A novel EEG-based approach to classify emotions through phase space dynamics. Signal Image Video Process. 13, 1149–1156. doi: 10.1007/s11760-019-01455-y
16. Wang X, Le D, Cheng H, Xie C. All-IP wireless sensor networks for real-time patient monitoring. Journal of biomedical informatics. 2014;52:406–417.
17. Wang, X., Zhang, T., Xu, X., Chen, L., Xing, X., and Chen, C. L. P. (2018). “EEG emotion recognition using dynamical graph convolutional neural networks and broad learning system,” in 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (Madrid: IEEE), 1240–1244. doi: 10.1109/BIBM.2018.8621147