

DOI: 10.51790/2712-9942-2024-5-3-03

**АВТОНОМНАЯ МНОГОКАМЕРНАЯ НЕЙРОСЕТЕВАЯ СИСТЕМА КОМПЬЮТЕРНОГО
ЗРЕНИЯ ДЛЯ АВТОМАТИЗАЦИИ КОНТРОЛЯ РУЧНЫХ ОПЕРАЦИЙ****И. М. Клемышев^{1,a}, С. М. Колчин^{2,b}, С. С. Лебедев^{2,6}, С. О. Старков^{3,2}**¹ *Общество с ограниченной ответственностью «Дженерал Компьютер Системс», г. Обнинск,
Российская Федерация*² *Общество с ограниченной ответственностью «КомпВи», г. Обнинск, Российская Федерация*³ *Обнинский институт атомной энергетики — филиал федерального государственного автономного
образовательного учреждения высшего образования «Национальный исследовательский ядерный
университет «МИФИ», г. Обнинск, Российская Федерация*^a *ORCID: <http://orcid.org/0000-0001-7831-9474>, ✉ iklemyshev@gmail.com*^b *skol57@mail.ru*⁶ *ORCID: <http://orcid.org/0000-0002-9626-0223>, compvi19@mail.ru*² *ORCID: <http://orcid.org/0000-0002-0420-7856>, sergeystarkov56@mail.ru*

Аннотация: в данной статье представлен прототип системы для автоматического контроля производственных процессов, выполняемых руками. Обученная система осуществляет мониторинг рабочего места с нескольких ракурсов наблюдения и производит анализ зафиксированных событий на основе искусственного интеллекта в реальном времени. При выявлении аномальных действий сотрудника или некорректности выполнения технологического процесса система оповещает оператора службы безопасности о событии, помогая ему принять решение или осуществить быстрый поиск по архивным записям видеонаблюдения для расследования инцидентов. Предложен механизм автокалибровки многокамерной системы видеонаблюдения, основанный на одновременном наблюдении ключевых точек на руках, определении их расположения в пространстве и математическом решении задачи совмещения с нескольких ракурсов. Аппаратная часть прототипа программно-аппаратного комплекса (ПАК) состоит из двух стереокамер INTEL RealSense435 и вычислительного модуля с графическим процессором Jetson AGX Xavier. Программная часть состоит из подсистем: видеонаблюдения, хранения данных, нейросетевого анализа «цифрового образа».

Ключевые слова: компьютерное зрение, распознавание образов, искусственные нейронные сети, программно-аппаратный комплекс, интеллектуальная видеочамера, нейросетевой детектор.

Благодарности: работы, выполненные ООО «ДженералСиЭс», поддержаны Фондом содействия инновациям (грант № 142ГС1ИИС12-D7/79659).

Для цитирования: Клемышев И. М., Колчин С. М., Лебедев С. С., Старков С. О. Автономная многокамерная нейросетевая система компьютерного зрения для автоматизации контроля ручных операций. *Успехи кибернетики*. 2024;5(3):24–33. DOI: 10.51790/2712-9942-2024-5-3-03.

*Поступила в редакцию: 27.07.2024.**В окончательном варианте: 24.08.2024.***AUTONOMOUS MULTI-CAMERA NEURAL NETWORK COMPUTER VISION SYSTEM FOR
MANUAL OPERATION MONITORING****I. M. Klemyshev^{1,a}, S. M. Kolchin^{2,b}, S. S. Lebedev^{2,c}, S. O. Starkov^{3,d}**¹ *General Computer Systems Limited liability company, Obninsk, Russian Federation*² *COMPVI Limited liability company, Obninsk, Russian Federation*³ *Obninsk Institute for Nuclear Power Engineering, Obninsk, Russian Federation*^a *ORCID <http://orcid.org/0000-0001-7831-9474>, ✉ iklemyshev@gmail.com*^b *skol57@mail.ru*^c *ORCID <http://orcid.org/0000-0002-9626-0223>, compvi19@mail.ru*^d *ORCID <http://orcid.org/0000-0002-0420-7856>, sergeystarkov56@mail.ru*

Abstract: this paper presents a prototype system for the automatic monitoring of manual production processes. The trained system observes the workplace from multiple angles and analyzes recorded events in real time using artificial intelligence. When abnormal employee actions or incorrect execution of technological processes are detected, the system alerts the security service operator, assisting in decision-making or

facilitating a quick search through archived video surveillance records for incident investigation. A mechanism for the auto-calibration of the multi-angle video camera system is proposed, which simultaneously observes key points on the hands, determines their spatial location, and mathematically resolves the problem of combining images from multiple angles. The hardware component of the prototype consists of two Intel RealSense 435 stereo cameras and a computing module equipped with a Jetson AGX Xavier graphics processor. The software component includes several subsystems: video surveillance, data storage, and neural network analysis of the digital images.

Keywords: computer vision, pattern recognition, artificial neural networks, intelligent video camera, neural network detector.

Acknowledgements: this study is supported by the Innovations Promotion Fund (grant No. 142GS1IIS12-D7/79659).

Cite this article: Klemyshev I. M., Kolchin S. M., Lebedev S. S., Starkov S. O. Autonomous Multi-Camera Neural Network Computer Vision System for Manual Operation Monitoring. *Russian Journal of Cybernetics*. 2024;5(3):24–33. DOI: 10.51790/2712-9942-2024-5-3-03.

Original article submitted: 27.07.2024.

Revision submitted: 24.08.2024.

Введение

Цифровизация предприятий идет на разных уровнях и в целом способствует оптимизации технологических процессов. Несмотря на автоматизацию производственных линий, доля ручного труда остается высокой и некоторые технологические циклы, выполняемые вручную, остаются за пределами возможной автоматизации. Тем актуальней является задача контроля правильности выполнения человеком технологического процесса, а также обнаружения событий в процессе работы, которые могут потенциально вести к рискам на производстве, снижению качества продукции, браку, хищению и утрате материальных ценностей.

На сегодняшний день задача безопасности и контроля производственных процессов на предприятиях решается в абсолютном большинстве случаев с помощью систем видеонаблюдения. Классическим решением для мониторинга ситуации на производстве или ином предприятии является система видеонаблюдения (СВН), которая состоит из видеосервера с хранилищем видеoarхива (системы хранения данных), специализированного ПО и подключенных к нему по сети IP-камер, расположенных в различных локациях предприятия. Такие системы являются в настоящее время привычным средством, используемым службой безопасности (СБ) и службой контроля качества технологических процессов предприятия для мониторинга рабочего процесса, соблюдения порядка, обнаружения опасных ситуаций и расследования инцидентов.

Современной СВН можно дать следующие характеристики:

- система дает оператору возможность одновременного наблюдения нескольких видеопотоков;
- камеры стандартно используются для наблюдения помещения, а не рабочего места;
- использование PTZ-камер (камера, которая поддерживает удаленное управление направлением и увеличением), то есть камер, которые дают оператору возможность поворота и приближения изображения;
- система, как правило, включает в себя интеллектуальную аналитику общего назначения:
 - детектирование людей;
 - распознавание лиц;
 - подсчет посетителей;
 - разметка видимой зоны для подсчета пересечения линии и т.п.;
- программная аналитика системы обрабатывает видеопотоки независимо, не связывая результаты их обработки.

Предприятия, оборудованные стандартной СВН, как правило, представляют собой помещения с несколькими рабочими местами/зонами, осматриваемые одной или несколькими камерами, расположенными под потолком. Такая система может помочь службе безопасности в определении действий людей на уровне помещения, связанными с пространственным перемещением, общей жестикულიацией, взаимодействием с окружающими предметами вроде мебели, с проемами и друг с другом.

Когда речь заходит о применении обычной СВН для наблюдения за действиями сотрудников, которые они выполняют на рабочем месте, то можно наблюдать нехватку необходимого функционала для правильного анализа ситуации:

- возможность фокусировки (при использовании PTZ-камеры) только на одном рабочем месте в помещении в один момент времени; без PTZ-камер возможности мониторинга сводятся к пассивному наблюдению с одной точки, как правило, удаленной от каждого рабочего места на расстояние не менее 2 метров;

- наблюдение за рабочими местами только с одной фиксированной позиции, что означает для сотрудника возможность сокрытия действий от наблюдений;

- отсутствие программной аналитики для распознавания действий человека, в том числе выполняемых руками, и системы поддержки принятия решений для СБ на основе данной аналитики.

В свою очередь, предлагаемую нами систему отличает:

- мониторинг с использованием нескольких ракурсов наблюдения, что исключает возможность сокрытия действий и потери системой объекта наблюдения;

- наличие программной подсистемы, набирающей статистику действий сотрудника, за работой которого ведется наблюдение, на основании которой создается «цифровой образ» человека;

- подсистема анализа, способная выявлять аномальное поведение сотрудника в реальном времени, отталкиваясь от собранного «цифрового образа» его паттернов поведения.

В основе функционала нашей системы лежит механизм «захвата движения» по видеоизображению. Сама по себе задача произвести «захват движения» не нова, и на рынке присутствуют датчики, способные делать это.

Работы по теме

Одна из наиболее известных существующих систем, позволяющих производить «захват движения» рук человека, — контроллер Leap Motion, созданный для управления компьютером (или приставкой, ТВ) с помощью жестов и движения рук [1]. Он основан на принципе стереоскопического зрения, что позволяет контроллеру анализировать глубину изображения и, следовательно, получать координаты положения рук в 3D. Встроенные вычислительные мощности позволяют Leap Motion распознавать, анализировать действия человека руками и выполнять соответствующие заданные команды. При всех положительных характеристиках данного устройства оно не может выступать в качестве сенсора предлагаемой многоракурсной системы по причинам, изложенным ниже.

По отзывам пользователей данного контроллера, есть большая доля ошибок в распознавании рук при их близком расположении и сцепке (переплетение и пересечение пальцев двух рук).

Еще одна разработка в данной сфере — проект многоракурсной системы с использованием контроллеров Azur Kinect от Microsoft. Для описания двух рук и объектов в работе использованы 5 камер Azur Kinect [2]. Результаты исследований говорят о том, что система хорошо определяет положение рук в пространстве. За счет того, что система имеет несколько ракурсов, она способна видеть руки в сложных случаях, когда рука видна не полностью, закрыта или невидима с определенной стороны. Однако у этой системы есть следующая особенность: низкое качество работы при сильном освещении и на расстоянии от объекта меньше 80 см. Не считая того, что промышленного варианта такой системы не существует и эта работа является прототипом, означенные особенности не позволяют применить ее в предлагаемой системе.

Для точного распознавания действий человека и анализа жестов рук многоракурсной видеосистемой необходим выбор подхода к калибровке нескольких камер для наиболее точного совмещения пространственной информации, полученной с разных ракурсов наблюдения. В [3], [4] представлены различные методологии калибровки с использованием 2D и 3D подходов. В работе [5] предложено для калибровки системы из нескольких камер использовать ключевые точки скелета человека, которые определяются в 2D с помощью ИНС в реальном времени.

Авторы [6] утверждают, что для точной оценки 3D позы человека с нескольких ракурсов перспективно использовать нейронную сеть СТР, которая напрямую работает в 3D пространстве.

В работе [7] описаны преимущества камер глубины RealSense D400 при создании многоракурсной системы.

Для работы с глубиной изображения одновременно с распознаванием жестов рук, скелета, лица и сегментации человека на основе ИНС в режиме реального времени необходимо аппаратное обеспечение с мощными CPU и графическими процессорами (GPU) и/или тензорными ядрами (TPU) для ускорения обработки данных. При создании автономных систем этим требованиям наилучшим образом соответствуют одноплатные компьютеры от NVIDIA серии Jetson. Связка стереокамер RealSense

серии D400 и одноплатного компьютера Jetson AGX Xavier хорошо себя зарекомендовала при разработке автономных систем компьютерного зрения и робототехники [8, 9].

Математическая модель

Алгоритм для реализации механизма многоракурсного нейросетевого анализа для автоматизации контроля производственных процессов, выполняемых руками, можно разбить на два основных блока операций:

- 1) работа многоракурсной системы видеонаблюдения с автоматической калибровкой;
- 2) работа подсистемы гибридного анализа действий, выполняемых руками.

Совместная работа двух блоков осуществляется путем последовательной передачи данных, полученных в результате работы многоракурсной системы, в подсистему гибридного анализа действий. Результаты работы последней отправляются на хранение в базу данных событий, с возможностью дальнейшего воспроизведения и изучения в графическом интерфейсе пользователя.

В зависимости от настроек системы, результаты работы подсистемы гибридного анализа действий могут непосредственно отображаться в интерфейсе пользователя в реальном времени и давать тревожные сигналы оператору при обнаружении аномальных действий сотрудника или нарушении технологического процесса производства для своевременной реакции на эти события.

Работа многоракурсной системы видеонаблюдения осуществляется одновременно с записью видеопотоков со всех ракурсов в специальное хранилище для дальнейшего воспроизведения.

Внутренний сервер системы предназначен для взаимосвязи основного приложения, запущенного на одноплатном компьютере, с графическим интерфейсом пользователя. С помощью программного интерфейса API осуществляется управление прототипом из графического интерфейса и передача данных от прототипа на панель управления АРМ оператора обратно.

Внутренний программный сервер, в зависимости от конфигурации системы и количества рабочих мест (одноплатных вычислительных модулей), может быть запущен как на одном из одноплатных модулей, так и на отдельной серверной машине, также включенной в сеть предприятия.

Общая схема работы разрабатываемой системы представлена на рис. 1.

Работа многоракурсной системы видеонаблюдения с автоматической калибровкой

Работа этой системы обеспечивает:

- 1) одновременное покадровое чтение видеопотоков с камер программно-аппаратного комплекса, подключенных к вычислительному модулю;
- 2) программную предобработку полученных кадров с целью улучшения качества их дальнейшей обработки;
- 3) трансляцию цветного видеоизображения с каждой камеры через web-сервер по протоколу http;
- 4) программную обработку кадров с целью «захвата движения» человека и его рук, находящихся в зоне наблюдения (включая трансляцию отладочных видеопотоков);
- 5) автоматическую калибровку многоракурсной СВН;
- 6) объединение информации с разных ракурсов наблюдения для получения полных и непрерывных данных о расположении ключевых точек рук человека.

Программная обработка кадров с целью «захвата движения» человека и его рук, находящихся в зоне наблюдения

Кадры видеоизображений с разных ракурсов (камер) наблюдения за рабочим процессом, поступившие в обработку, проходят несколько этапов, результатом которых является «захват движения» человека и его рук, находящихся в зоне наблюдения:

- 1) детекция ключевых точек скелета человека;
- 2) детекция ключевых точек кистей человека;
- 3) детекция (обнаружение) лица человека в кадре;
- 4) сегментация человека в кадре;
- 5) трекинг найденных объектов (ключевых точек, лиц).

Детекция ключевых точек кистей рук и скелета человека на цветном изображении выполняется с помощью специально обученных для этого ИНС: на данный момент это общепринятый и самый действенный способ решения подобного рода задач.

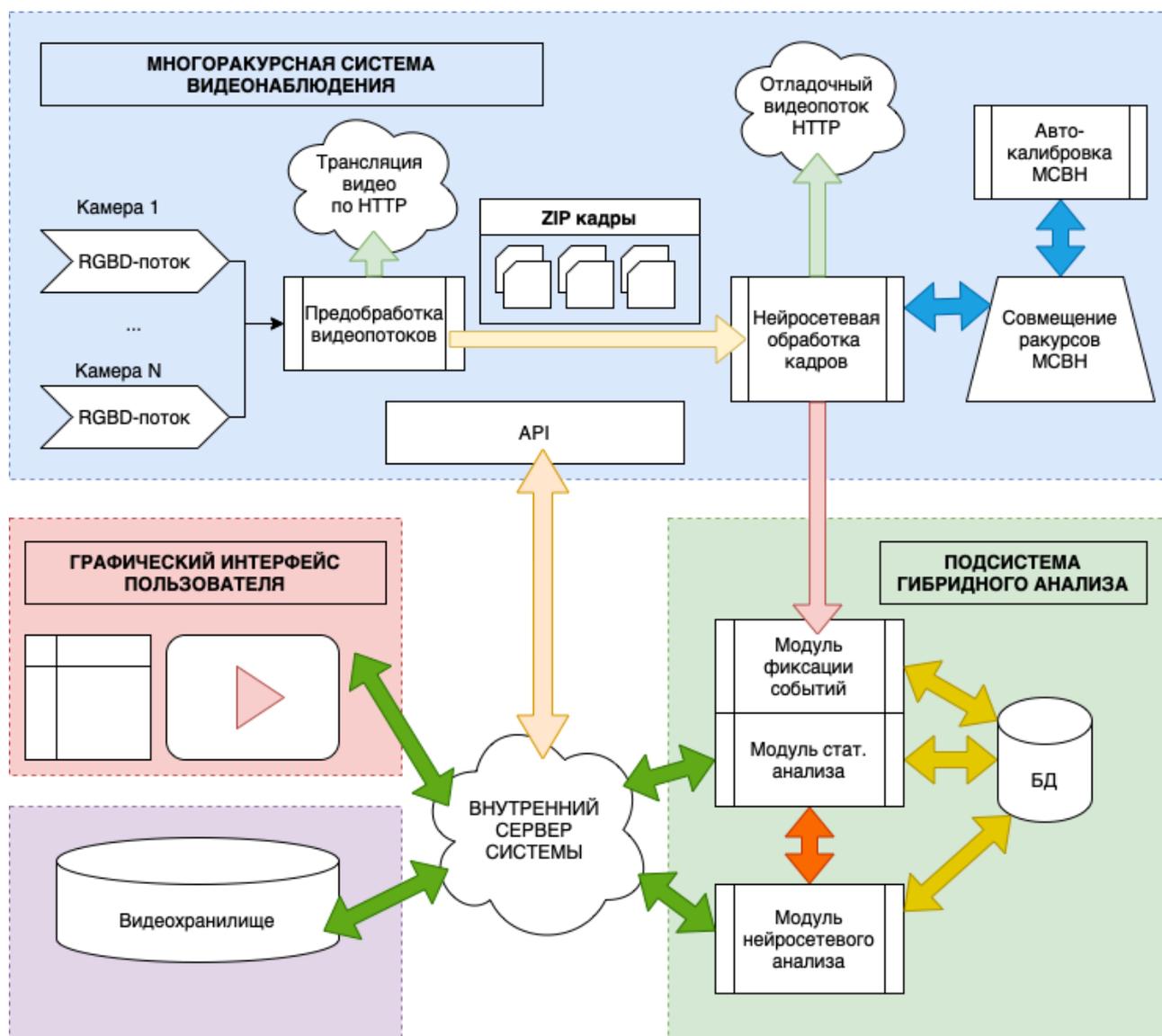


Рис. 1. Общая схема работы автономной системы многокамерного нейросетевого анализа для автоматизации контроля производственных процессов, выполняемых руками

Используемый в данном проекте нейросетевой детектор определяет трехмерное расположение точек кистей рук относительно корневой точки (точки кисти, расположенной в основании запястья) с координатами $(0, 0, 0)$ по осям X, Y, Z . Для того, чтобы определить пространственное положение корневой точки относительно камеры наблюдения, используется информация из «карты глубины» для соответствующего кадра видеоизображения, с которого проводилась детекция ключевых точек кисти руки.

На рис. 2 приведена разметка результата распознавания ключевых точек кисти человеческой руки.

Определение ключевых точек скелета человека (помимо ключевых точек кистей рук) является необходимой частью математической модели системы по следующим причинам:

- 1) является наиболее достоверным признаком наличия человека (или нескольких) в кадре;
- 2) помогает решить проблемы с распознаванием «призрачных» рук, когда нейросетевой детектор «видит» кисть руки человека там, где ее на самом деле нет.

«Захват движения» подразумевает не только детекцию ключевых точек кистей рук и скелета человека, но и трекинг найденных объектов. Трекинг нужен для объединения информации о найденных объектах между кадрами. Это позволяет определять принадлежность найденных ключевых точек к одному и тому же объекту во времени и, соответственно, отслеживать его изменения и траекто-

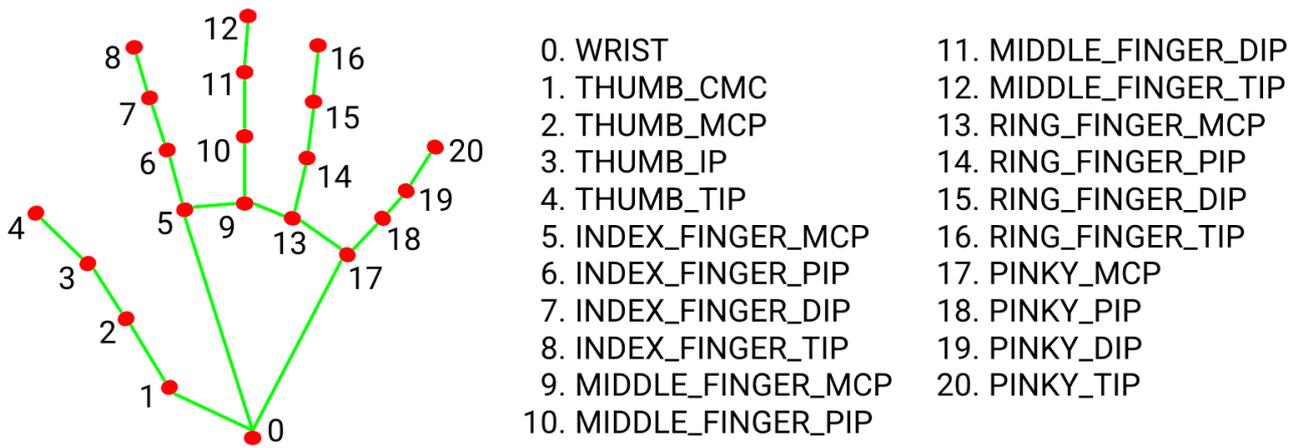


Рис. 2. Ключевые точки кисти руки при распознавании с помощью ИНС

рию движения.

Детекция лиц и сегментация людей необходимы для последующего определения таких событий, как прикосновение руки (кистей и пальцев рук) к телу и к лицу. Фиксация этих событий возможна при известном расположении в кадре лица человека, силуэта его тела, расположении рук и информации о расположении этих элементов в трехмерном пространстве, которую возможно получить из «карты глубины» каждого ракурса за счет использования стереокамер.

Автоматическая калибровка многоракурсной СВН

Под автоматической калибровкой многоракурсной СВН подразумевается определение расположения камер видеонаблюдения относительно друг друга в трехмерном пространстве, производимого без ручного вмешательства. Для выполнения этой операции был разработан авторский метод нахождения их относительного расположения на основании наблюдения за одним и тем же объектом — в данном случае за руками человека.

Рассмотрим выполнение автоматической калибровки на примере двух стереокамер, следящих за руками сотрудника с разных ракурсов наблюдения.

Операция «захвата движения», выполненная для синхронных кадров с обеих камер, дает нам расположение ключевых точек кистей рук.

Пусть одна камера видит одну из двух рук, вторая видит две. Необходимо найти соответствие рук на двух ракурсах. Для этого выполняется поворот трехмерной модели скелета руки (то есть набора ключевых точек кисти руки в трехмерных координатах относительно корневой точки). Когда все найденные модели скелетов кистей рук на всех ракурсах сориентированы одинаковым образом, производится попарное сравнение этих моделей с помощью операции нахождения геометрической дистанции одной модели от другой. Минимальная (в идеальном случае — нулевая, на практике — имеющая наиболее близкое к нулю значение) дистанция говорит о том, что сравниваемые модели описывают одну и ту же руку. При том, что сравниваются модели с разных ракурсов наблюдения, устанавливается соответствие между ними.

При любом количестве рук, видимых с разных ракурсов, производится та же самая операция поиска пар. После нахождения всех пар определяются координаты корневых точек каждой модели по информации о глубине изображения.

Полученная информация об удаленности моделей скелетов кистей рук от каждого ракурса наблюдения совмещается с информацией об углах поворота каждой модели от разных камер (углы получены в процессе ориентирования модели по осям), в результате чего вычисляется расположение и ориентация каждой камеры относительно каждой модели скелета руки.

В зависимости от того, сколько рук являются видимыми на каждом ракурсе и какие пары моделей скелетов кистей рук будут найдены, для каждой камеры будут получены от одного до нескольких наборов координат в трехмерном пространстве. Если наборов несколько, то производится усреднение для получения более точного значения.

В результате описанных выше вычислений получают точные результаты относительного рас-

положения камер в трехмерном пространстве, на основании которых можно выполнять дальнейшие расчеты.

Автоматическая калибровка многоракурсной системы видеонаблюдения выполняется с определенной периодичностью, задаваемой настройками системы.

Объединение информации с разных ракурсов наблюдения для получения полных и непрерывных данных о расположении ключевых точек рук человека

Откалиброванная многоракурсная система видеонаблюдения позволяет корректировать расположение в пространстве обнаруженных ключевых точек кистей рук человека.

С каждого ракурса система определяет расположение ключевых точек кистей рук относительно конкретной стереокамеры, а данные калибровки позволяют объединить эти данные в одно трехмерное пространство и провести усреднение координат каждой точки. В итоге мы получаем более точные значения о расположении «скелетов» рук человека для случаев, когда каждый ракурс видит только часть картины и не видит другой стороны.

Работа подсистемы гибридного анализа действий, выполняемых руками

Работа данной подсистемы обеспечивает:

1) фиксацию событий на основании данных, полученных от многоракурсной системы видеонаблюдения:

- распознавание события «наличие рук нескольких человек на одном рабочем месте»;
- распознавание события «рука вне поля зрения видеокамер»;
- распознавание события «рука сотрудника касается лица»;
- распознавание события «рука сотрудника касается тела»;
- распознавание события «движения глаз по сторонам без поворота головы»;
- определение события, когда рядом с рабочим местом отсутствуют люди (пустое рабочее место);
- определения движения рядом с пустым рабочим местом;
- распознавание действий, создающих риск утраты продукции на основании определения жестов рук сотрудника:

- рука, сжатая в кулак;
- не видно крайних фаланг правой руки;
- не видно крайних фаланг левой руки;
- пальцы правой руки сгруппированы;
- пальцы левой руки сгруппированы;
- плохо (частично) видна кисть правой руки;
- плохо (частично) видна кисть левой руки;

2) статистический анализ накопленных событий, который включает:

- подсчет зафиксированных жестов и событий для каждого работника;
- расчет пороговых значений числа зафиксированных событий для каждого работника на основе накопленной статистики для разных временных масштабов;
- формирование «цифрового образа» сотрудника на основе собранных данных, взвешенных по времени совершения (момента и длительности действия), и статусах зафиксированных действий и их последовательностей;
- формирование данных для отображения оператору в случае превышения пороговых значений по одному или нескольким критериям;

3) нейросетевой анализ «цифрового образа», включающий в себя:

- кодирование данных о входной взвешенной последовательности действий в числовой вектор;
- классификация полученного вектора в соответствии с векторами, полученными из «цифрового образа» паттернов поведения, и определение статуса наблюдаемых действий.

В системе также предусмотрен внутренний сервер, который может работать как на той же аппаратной платформе, что и МСВН, и система гибридного анализа, так и на отдельной платформе, если система сконфигурирована на несколько рабочих мест. Задача сервера — собирать данные со всех подключенных рабочих мест и ретранслировать их в графический интерфейс пользователя. А

также осуществлять обратную связь с вычислительными модулями на рабочих местах для управления запущенными на них системами.

Аппаратная часть прототипа программно-аппаратного комплекса автономной системы многокурсного нейросетевого анализа для автоматизации контроля производственных процессов, выполняемых руками

Прототип системы представляет собой программно-аппаратный комплекс, состоящий из вычислительного блока и двух корпусных стереокамер на штативах.

Вычислительный блок представляет собой одноплатный компьютер в пластмассово-металлическом корпусе. На этом компьютере запускается операционная система и работает вся программная часть прототипа, в том числе нейросетевые детекторы, использующие для быстрой работы графический процессор (GPU) данного компьютера.

Две стереокамеры прототипа являются зрительными сенсорами многокурсной видеоподсистемы, позволяющими получить цветное видеоизображение вместе с информацией о глубине для определения трехмерного расположения точек в наблюдаемой сцене.

Аппаратная часть комплекса состоит из следующих устройств:

- 1) две стереокамеры Intel RealSense 435;
- 2) вычислительный модуль Jetson AGX Xavier (система на модуле с графическим ускорителем и ядрами для ускорения работы нейросетей);
- 3) адаптер питания вычислительного модуля.

При оборудовании рабочего стола сотрудника должны быть соблюдены следующие требования:

- 1) устройства, входящие в состав программно-аппаратного комплекса, должны располагаться таким образом, чтобы не мешать действиям сотрудника при выполнении им рабочих обязанностей и не нарушать условий рабочего места (освещения, свободного пространства для работы);
- 2) камеры, входящие в состав программно-аппаратного комплекса, должны обеспечивать съемку рабочей зоны с разных ракурсов при условии, что руки сотрудника в ходе выполнения рабочих операций находятся преимущественно в пределах кадра каждой из камер.

Программная реализация

Программная часть прототипа состоит из подсистем, общающихся между собой путем передачи сообщений, содержащих различные данные: кадры видеопотоков, результаты «захвата движений» — расположение ключевых точек рук, результаты анализа действий сотрудника, сообщения о тревогах и т.д.

Каждая подсистема строится для выполнения ряда определенных задач и выполняет часть общего функционала всей системы:

- подсистема видеонаблюдения, предназначенная для видеосъемки каждого из рабочих мест с нескольких ракурсов; задача подсистемы — одновременная работа с видеопотоками стереокамер — получение цветных изображений, данных о глубине изображений;
- подсистема хранения данных предназначена для хранения видеозаписей и данных о событиях и объектах;
- подсистема видеоанализа предназначена для определения ключевых точек рук и тела на видеоизображении и осуществления на их основе: автоматической калибровки многокурсной системы, слияния информации с разных ракурсов, поиска и идентификации в видеозаписях различных событий;
- подсистема статистического анализа предназначена для сбора статистических данных по накопленным в системе событиям для каждого сотрудника и формирования «цифрового образа» паттернов поведения;
- подсистема нейросетевого анализа «цифрового образа» паттернов поведения;
- подсистема формирования пользовательского графического интерфейса предназначена для управления и настройки работы прототипа, отображения изображений подсистемы видеонаблюдения, а также для отображения результатов работы подсистем гибридного анализа (статистического и нейросетевого).

Каждая подсистема общается с другими путем передачи информационных сообщений с использованием программных брокеров сообщений и клиент-серверного взаимодействия по протоколу http.

Таким образом, программная архитектура прототипа представляет собой распределенную микросервисную архитектуру. В отличие от монолитной программной архитектуры, микросервисы имеют ряд преимуществ в построении разрабатываемой системы:

- позволяют распределять программную систему различным образом на аппаратных мощностях для эффективного распределения нагрузки;
- позволяют менять конфигурацию как программной, так и аппаратной частей в зависимости от требований клиента к составу и функционалу системы;
- позволяют заменять отдельные сервисы, убирать или добавлять новые функциональные блоки без нарушения работоспособности системы.

Выводы

В статье представлен впервые созданный прототип системы автоматизированного контроля производственных процессов, выполняемых руками. Система представляет собой программно-аппаратный комплекс, который осуществляет мониторинг рабочего места с нескольких ракурсов наблюдения и производит анализ зафиксированных событий на основе искусственного интеллекта, который позволяет выявить аномальные действия сотрудника или проверить корректность выполнения технологического процесса.

Результаты анализа событий представляют собой набор данных, содержащих разметку видеопотока, в котором указано время и тип зафиксированного события. Событие может представлять собой сообщение о нетипичных или подозрительных действиях сотрудника, а также об ошибках в процессе ручного труда.

Разработан уникальный алгоритм, позволяющий системе проводить автоматическую калибровку «на ходу» без участия человека и специальных калибровочных шаблонов. Предлагаемый механизм автокалибровки многоракурсной системы основан на одновременном наблюдении объектов сцены (например, ключевых точек на руках), определении их расположения в пространстве относительно каждой камеры (за счет карт глубины изображения, получаемых благодаря использованию стереокамер) и математическом решении задачи совмещения нескольких ракурсов. В результате решения этой задачи система определяет относительное расположение камер в трехмерном пространстве, в чем и состоит задача калибровки многоракурсной системы.

В этом состоит авторское научно-техническое решение, которое не только позволяет менять количество ракурсов «на ходу», но и сохраняет систему работоспособной при случайном или намеренном сдвиге одной или нескольких камер.

ЛИТЕРАТУРА

1. LeapMotion. Что внутри? *Хабр*. Режим доступа: <https://habr.com/ru/companies/avi/articles/199230/>.
2. Kwon T., Tekin B., Stuhmer J., Bogo F., Pollefeys M. H2O: Two Hands Manipulating Objects for First Person Interaction Recognition. *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021. Режим доступа: <https://taeinkwon.com/projects/h2o/>.
3. Romeo L., Marani R., Perri A.G., D'Orazio T. Microsoft Azure Kinect Calibration for Three-Dimensional Dense Point Clouds and Reliable Skeletons. *Sensors*. 2022;22:4986. Режим доступа: <https://doi.org/10.3390/s22134986>.
4. Bu S., Lee S. Easy to Calibrate: Marker-Less Calibration of Multiview Azure Kinect. *Comput. Model. Eng. Sci.* 2023;136(3):3083-3096. Режим доступа: <https://doi.org/10.32604/cmesci.2023.024460>.
5. Pätzold B., Bultmann S., Behnke S. Online Marker-Free Extrinsic Camera Calibration Using Person Keypoint Detections. *DAGM German Conference on Pattern Recognition (GCPR)*. Konstanz, Germany. 2022. Режим доступа: https://www.researchgate.net/publication/363540896_Online_Marker-free_Extrinsic_Camera_Calibration_using_Person_Keypoint_Detections.
6. Liu H., Wu J., He R. Center Point to Pose: Multiple Views 3D Human Pose Estimation for Multi-Person. *PLoS ONE*. 2022;17(9):e0274450. DOI: 10.1371/journal.pone.0274450. Режим доступа: https://www.researchgate.net/publication/363541752_Center_point_to_pose_Multiple_views_3D_human_pose_estimation_for_multi-person.
7. Yoon H., Jang M., Huh J., Kang J., Lee S. Multiple Sensor Synchronization with the Realsense RGB-D Camera. *Sensors*. 2021;21(18):6276. Режим доступа: <https://doi.org/10.3390/s21186276>.

8. Jeon J., Jung S., Lee E., Choi D., Myung H. Run Your Visual-Inertial Odometry on NVIDIA Jetson: Benchmark Tests on a Micro Aerial Vehicle. *IEEE Robotics and Automation Letters*. 2021;6(3):5332-5339. DOI: 10.1109/LRA.2021.3075141. Режим доступа: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9416140&isnumber=9399748>.
9. Boschi A., Salvetti F., Mazzia V., Chiaberge M. A Cost-Effective Person-Following System for Assistive Unmanned Vehicles with Deep Learning at the Edge. *Machines*. 2020;8(3):49. DOI: 10.3390/machines8030049. Режим доступа: https://www.researchgate.net/publication/343948324_A_Cost-Effective_Person-Following_System_for_Assistive_Unmanned_Vehicles_with_Deep_Learning_at_the_Edge.