

Как ААЕОН развивает технологию искусственного интеллекта в своих решениях

Анна Клекот

С бурным развитием беспроводных технологий и систем на базе ИИ потребность в компактных и производительных вычислительных устройствах увеличилась в разы. Но производителей, сравнимых по качеству с давно известной российскому потребителю ААЕОН, не так много. В статье мы расскажем о направлениях совершенствования продуктов компании.

Введение

ААЕОН производит готовые платформы и компоненты для систем AI, Edge AI и IoT, которые интегрируются в решения концепции Industry 4.0 (например, умные города, умные фабрики, системы видеонаблюдения и видеоаналитики), сотрудничая со всемирно известными хедлайнерами данного направления, включая NVIDIA, Intel, Google и Kneron. Нейросети появились ещё в 70-х годах XX века, но сейчас стало возможным накопление больших объёмов данных для их качественного обучения, при этом вычислительные мощности стали относительно дешёвыми и позволяют проводить глубокое обучение с высокой производительностью.

Технологический союз компании ААЕОН и NVIDIA позволил создать ультракомпактные компьютеры серии BOXER-8000 на базе плат Jetson, которые демонстрируют недостижимую для классических компьютеров производительность в задачах, связанных с искусственным интеллектом. Последующие поколения устройств продолжают показывать всё большую эффективность с увеличением вычислительной мощности. В сегодняшней статье более подробно поговорим о данном направлении. Также, поскольку в 2021 году компания Intel объявила о снятии с производства линейки визуальных процессоров Intel Movidius Myriad X, мы рас-

смотрим конкурентные решения на визуальных процессорах Kneron (рис. 1).

Алгоритмы ИИ имеют иную, нежели обычные программы, структуру и для лучшего результата требуют вычислителей, отличных от привычных компьютеров. Основная разница заключается в необходимости использования многопоточности и большого количества процессорных ядер.

Реализовать это на базе стандартных процессоров не представляется возможным ввиду значительного энергопотребления, которому потребуются мощный теплоотвод, и из-за высокой стоимости подобного решения.

Кроме того, даже если такое решение реализовать физически, оно не поместится ни в одну компактную систему,



Рис. 1. Визуальный процессор Kneron KL530

работающую непосредственно на объекте, например, вблизи камер видеонаблюдения или в системах контроля доступа типа турникетов или проходных.

Устройство должно иметь компактный размер и в нашем случае может быть представлено в виде модуля встраиваемых систем либо уже готового решения в корпусе с готовым набором наиболее востребованных портов ввода-вывода и слотов расширения.

Компания ААЕОН имеет в своём портфолио огромную линейку встраиваемых компьютеров на базе процессоров NVIDIA Jetson разных поколений – это серии BOXER-8600 и BOXER-8200.

NVIDIA Jetson – это гибридный процессор, состоящий из ядер ARM для стандартных вычислений и графиче-



Рис. 2. BOXER-8251AI-KIT

Таблица 1. Технические характеристики серий BOXER-8600 и BOXER-8200

Характеристики	BOXER-8640AI	BOXER-8641AI	BOXER-8256AI	BOXER-8233AI
Визуальный процессор	NVIDIA Ampere	NVIDIA Ampere	NVIDIA Volta	NVIDIA Pascal
Центральный процессор	Arm Cortex-A78AE	Arm Cortex-A78AE	NVIDIA Carmel Arm v8.2	NVIDIA Denver A57
ОС	Linux (NVIDIA Jetpack 5.x и выше)	Linux (NVIDIA Jetpack 5.x и выше)	Linux (NVIDIA Jetpack 4.6.2)	Linux (NVIDIA Jetpack 4.5.1)
ОЗУ	32 Гбайт LPDDR5	32 Гбайт LPDDR5	8 Гбайт LPDDR4	4 Гбайт LPDDR4
Хранение данных	64 Гбайт eMMC, micro SD, 2,5" SATA, M.2 2280	64 Гбайт eMMC, micro SD, 2,5" SATA, M.2 2280	16 Гбайт eMMC, micro SD, 2,5" SATA, M.2 2280	16 Гбайт eMMC, micro SD, 2,5" SATA, 1× mSATA
Видеовыходы	1× HDMI 2.0	1× HDMI 2.0	2× HDMI 2.0	1× HDMI 2.0
Ethernet	4× POE	2× GbE	1× GbE	2× POE, 1× GbE
Ввод-вывод	2× RS-232/422/485, 4× USB 3.2, 1× micro USB, 1× CAN, аудио	2× RS-232/422/485, 3× USB 3.2, 1× USB 2.0, 1× micro USB, 8× DIO	1× RS-232/422/485, 3× USB 3.2, 1× micro USB, 1× CAN, аудио	1× RS-232/485, 4× USB 3.2, 1× micro USB, 1× CAN, 1× DIO, аудио
Слоты расширения	M.2 2230 (E), 40-pin NVIDIA Jetson IO	M.2 2230 (E), M.2 3052 (B), слот для SIM-карты	M.2 3042/52 (B)	1× mini PCIe, M.2 2230 (E), слот для SIM-карты
Питание	12–24 В	12 В	12–24 В	12–24 В
Рабочая температура	–20...+55°C	–20...+55°C	–15...+60°C	–15...+55°C
Габариты	210 × 164,2 × 59 мм; 2,9 кг	180 × 136 × 61,1 мм; 2,5 кг	180 × 136 × 67 мм; 1,3 кг	180 × 136 × 61,1 мм; 1,3 кг

ских процессоров, оптимизированных для работы с параллельными вычислениями и предназначенных для работы с ИИ, т.е. нейронными сетями. Это позволяет достичь высокой производительности при низком энергопотреблении, что важно для компактных встраиваемых систем. А реализованная в графических процессорах NVIDIA архитектура позволяет легко программировать задачи с параллельными вычислениями, используя стандартный язык программирования C++, и создавать приложения для разных ОС.

Также в производственной линейке представлено несколько одноплатных компьютеров различного уровня производительности, включающих несущую плату и систему охлаждения в комплекте. Например, BOXER-8251AI-KIT на базе гибридного процессора NVIDIA Jetson Xavier (рис. 2). Решение обеспечивает до 32 Тфлопс для быстрой работы современных алгоритмов искусственного интеллекта.

Jetson AGX Xavier обладает достаточной производительностью для задач визуальной одометрии, объединения датчиков, локализации и картографии, распознавания объектов и построения маршрутов.

Встраиваемые компьютеры серий BOXER-8600 и BOXER-8200 на базе этих процессоров – это уже готовые решения, подходящие для реализации большого ряда типовых решений.

В таблице выше (табл. 1) приведены технические характеристики серии BOXER-8600 и BOXER-8200 на разных графических процессорах. Производимая линейка встраиваемых компьютеров значительно шире.

Немного о техническом зрении

Вычислительные машины уже давно успешно заменяют человека в решении широкого ряда задач – от элементарных вычислений до задач управления различными исполнительными механизмами, в том числе в задачах, требующих визуального восприятия. На сегодняшний день техническое зрение нашло наибольшее применение в системах управления конвейерных производств (рис. 3), сортировочных и сопроводительных операциях на складах, в горно-обогатительной промышленности и других областях, и решает оно следующие задачи:

- контроль нанесения маркировки и/или этикетки;
- контроль содержимого тары;
- считывание и распознавание штрих-кодов и текста;
- проверка целостности и комплектности упаковки;

- сортировка материала / готового продукта;
- контроль соответствия внешнего вида заданному стандарту;
- обнаружение, измерение, классификация и подсчёт объектов.

Системы технического зрения используются в производстве полупроводниковых и печатных плат, позволяя существенно снизить процент брака и время проверки конечного изделия.

В машиностроении, производстве строительных материалов и металлопрокате техническое зрение решает задачи контроля поверхности изделий, их качества, идентификации изделий и точных измерений.

На конвейерных линиях в производстве поток в зависимости от типа продукции может достигать сотен единиц в минуту, поэтому контроль здесь может осуществляться только автоматическим способом. При правильном проектировании и соблюдении всех необходимых

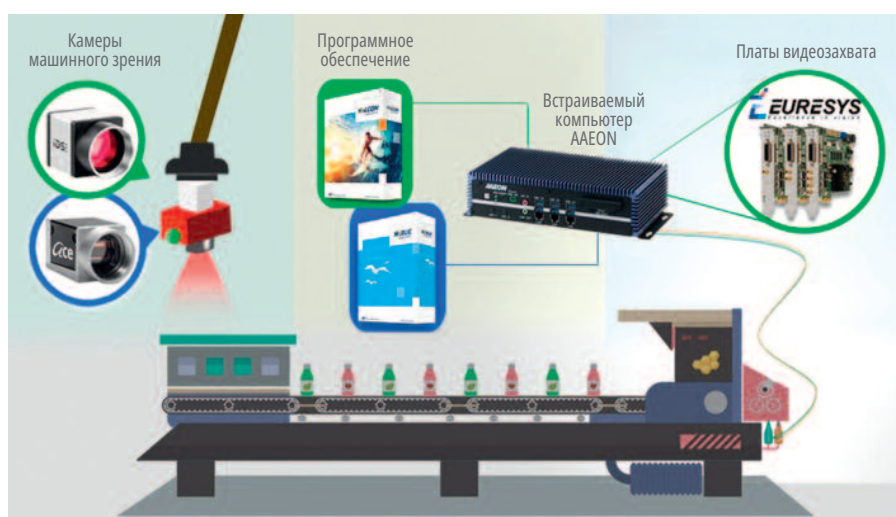


Рис. 3. Пример конвейерного производства с применением продукции AAEON



Рис. 4. BOXER-6842M

факторов построения системы машинного зрения точность инспекции изделий приближается к 100%.

Для комплексных задач применяют классические системы машинного зрения с персональным компьютером в качестве мощной вычислительной платформы, а для решения локальных типовых задач достаточно использования датчиков технического зрения. Такие датчики имеют компактный размер, оснащены стандартными портами ввода-вывода, имеют собственное программное обеспечение и поддерживают большинство распространённых протоколов, таких как PROFIBUS, Profinet, Modbus, OPC, TPC и др.

Для систем технического зрения AAEON предлагает линейку компьютеров BOXER-6800 и BOXER-6600 с возможностью установки полноразмерных плат видеозахвата или ускорителей нейронных сетей формата PCI Express и других малых форматов (рис. 4).

Периферийный искусственный интеллект

Edge computing (периферийные вычисления) вызывают значительный интерес благодаря новым сценариям

использования, особенно после внедрения сетей 5G.

Искусственный интеллект в значительной степени зависит от передачи данных и выполнения сложных алгоритмов машинного обучения, т.е. от того, о чём мы говорили выше.

Периферийные вычисления – это вычислительная парадигма нового поколения, которая переносит ИИ и машинное обучение туда, где происходят генерация данных и вычисления, – на границу сети. Их объединение привело к появлению нового направления – периферийного искусственного интеллекта (Edge AI). Edge AI обеспечивает более быстрые вычисления и извлечение данных, лучшую безопасность данных и эффективный контроль над непрерывностью операций, таким образом повышается производительность приложений и снижаются эксплуатационные затраты. Периферийный искусственный интеллект облегчает машинное обучение, автономное применение моделей глубокого обучения и передовых алгоритмов на самих устройствах Интернета вещей (IoT) уже вне облачных сервисов. То есть создаётся масштабируемый «искусственный разум» для обработки данных разного рода в конкретно поставленной задаче.

Преимуществами машинного обучения на периферии можно отметить следующие пункты.

- **Сокращение времени обработки операции.** В данном случае процесс обработки данных происходит не только на уровне сети, но также и на уровне устройства, таким образом исчезает необходимость пересылки огромных объёмов данных через сеть или на устройства, что в результате улучшает пользовательский опыт.
- **Минимизация нагрузки сетей.** В Edge AI реализована технология клаудлетов (cloudlet), которая представляет собой небольшое облачное хранилище, расположенное на периферии.

Эта технология повышает мобильность и снижает нагрузку на передачу данных. Следовательно, она может снизить стоимость услуг передачи данных и повысить скорость и надёжность потока данных. Минимизируется риск раскрытия конфиденциальной информации во время процесса передачи, так как не нужно передавать данные в облако, проводить аналитику и возвращать данные назад.

- **Конфиденциальность.** Разного рода данные на сегодняшний день – наиболее ценный актив. Прозрачная структура сбора и хранения данных повышает лояльность пользователя к предприятию.
- **Низкая стоимость машинного обучения.** Edge AI сокращает расходы на машинное обучение, выполняемое в облачных дата-центрах, ввиду автономности процессов.

Надо, в свою очередь, понимать, что машинное обучение требует больших вычислительных мощностей на аппаратных платформах периферийных вычислений. В инфраструктуре Edge AI производительность вычислений ограничена производительностью периферийного или IoT-устройства. В большинстве случаев сложные модели Edge AI приходится упрощать перед развёртыванием, чтобы повысить их точность и эффективность, но тем не менее для широкого спектра задач – это панацея.

За последние 10 лет вычислительные платформы для нейронных сетей превратились из полноразмерных серверов в компактные встраиваемые компьютеры, в том числе платформы с пассивным охлаждением, так как сами ускорители из полноразмерных плат PCI express мигрировали в компактные mini PCI express или M.2 устройства. Это более бюджетные, но достаточно производительные решения.

Компания Kneron, основанная в 2015 году, является ведущим поставщи-

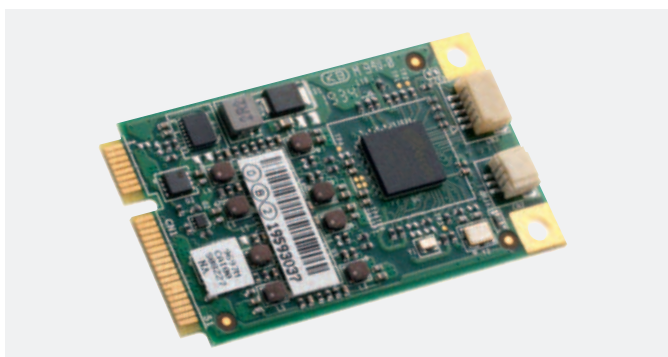


Рис. 5. Mini-AI-520

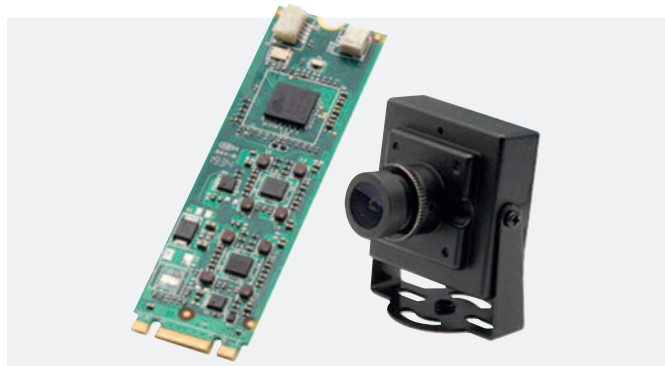


Рис. 6. M2AI-2280-520

Однофазные источники бесперебойного питания



- Мощность от 400 ВА до 10 кВа
- Напольное и стоечное исполнение
- Управляемая группа розеток (для моделей до 3 кВА включительно)
- Карта управления SNMP в комплекте для моделей с индексом NC
- Параллельное резервирование до трех ИБП
- «Горячая» пользовательская замена батарей
- Управляющее ПО в комплекте



Таблица 2. Сравнительные характеристики модулей ААЕОН

Характеристики	Mini-AI-520	Mini-AI-720	M2AI-2280-520	M2AI-2280-720
Чип	Kneron KL520	Kneron KL720	Kneron KL520	2x Kneron KL720
Формат	mini PCIe	mini PCIe	M.2 2280	M.2 2280
Интерфейс подключения	USB	PCIe[x1]	USB	PCIe[x1]
Библиотеки	ONNX, TensorFlow, Keras, Caffe	Pytorch, ONNX, TensorFlow 1.6, Tensorflow lite, Keras, Caffe	ONNX, TensorFlow, Keras, Caffe	Pytorch, ONNX, TensorFlow 1.6, Tensorflow lite, Keras, Caffe
Архитектура нейросетей	Vgg16, Resnet, GoogleNet, YOLO, Tiny YOLO, Lenet, MobileNet, DenseNet	Resnet, GoogleNet, YOLO, Tiny YOLO, MobileNet-SSD, DenseNet, RNN, LSTM	Vgg16, Resnet, GoogleNet, YOLO, Tiny YOLO, Lenet, MobileNet, DenseNet	Resnet, GoogleNet, YOLO, Tiny YOLO, MobileNet-SSD, DenseNet, RNN, LSTM
Производительность	0,35 Терафлопс	1,4 Терафлопс	0,35 Терафлопс	2× 1,4 Терафлопс
Энергопотребление	0,5 Вт	5 Вт	0,5 Вт	5 Вт
Рабочая температура	0...+70°C	0...+50°C (с активным охлаждением)	0...+70°C	0...+50°C (с активным охлаждением)
Примечание	Есть опциональный модуль с камерой		Есть опциональный модуль с камерой	

ком комплексных решений искусственного интеллекта и занимается разработкой интегрированных аппаратных и программных решений для технологий искусственного интеллекта (AI) и периферийного искусственного интеллекта (Edge AI).

Среди инвесторов Kneron – Qualcomm, Horizons Ventures, Alibaba Entrepreneurs Fund, Sequoia Capital, CDIB, Foxconn и Himax Technologies, и даже MTS AI.

Партнёрство ААЕОН с Kneron было естественным шагом, потому что оба партнёра, по сути, собирают разные кусочки головоломки, что позволяет искусственному интеллекту внедряться повсеместно. SoC Kneron AI KL520/KL720 имеет оптимальный баланс производительности, мощности и стоимости, это доступное решение без компромиссов в производительности или энергопотреблении. Edge AI может принести пользу различным отраслям: от улучшения контроля производства на сборочной линии до управления автономными транспортными средствами.

Ускорители нейронных сетей Kneron KL520/KL720 реализованы в модулях искусственного интеллекта ААЕОН Edge. Модули представлены в форматах mini PCIe и M.2: Mini-AI-520 (рис. 5), Mini-AI-720, M2AI-2242-520, M2AI-2280-520 (рис. 6), M2AI-2242-720, M2AI-2280-720.

Сравнительные характеристики модулей ААЕОН Edge см. табл. 2.

Периферийные вычисления с множественным доступом

Mobile Edge computing – MEC (Мобильные Периферийные вычисления, или Периферийные вычисления с множественным доступом) представляют собой совокупность всех мобильных устройств и используемых ими прило-

жений и, соответственно, всех данных, которые они потребляют и воспроизводят. И с внедрением сетей, так же как и Edge computing, данная область стремительно развивается и растёт. Mobile Edge computing – это область пересечения периферии беспроводной сети и периферии инфраструктуры, там, где мобильные сети и Интернет встречаются и осуществляют обмен трафиком. Традиционная цифровая инфраструктура в синергии с мобильными сетями существенно повышает производительность и снижает задержки для новых сценариев использования, таких как Интернет вещей, потоковое видео и мобильность.

Мобильные Периферийные вычисления (MEC) – это не просто концепция, это также система стандартов, разработанная некоммерческой группой ETSI. MEC – это, по сути, взаимосвязь между периферийными вычислениями и мобильностью в самом широком смысле, включая камеры наблюдения, телемедицину, IoT, подключённые к системе транспортные средства и многое другое.

В одном из своих интервью Дэн Флоренс, старший менеджер по управлению сетевой инфраструктурой Micron, отметил, что периферийные вычисления с множественным доступом преследуют те же цели, что и в целом периферийные вычисления (Edge computing): перемещение большого количества вычислительных функций ближе к тому месту, где генерируются данные и где принимаются решения, чтобы обеспечить более быстрое и эффективное реагирование.

В сочетании с архитектурой периферийных вычислений 5G обычно рассматривается как технология, которая решает одну из самых насущных про-

блем мира взаимосвязанных вещей – сетевую задержку – при одновременном обеспечении высокой пропускной способности.

Широкая доступность сетей 5G в ближайшем будущем и огромное количество оборудования и программного обеспечения, которое будет подключаться к этим сетям, не обойдутся без периферийных вычислений с множественным доступом (MEC).

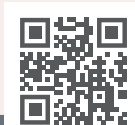
Заключение

При выборе вычислительной платформы на этапе разработки новой системы или расширения существующей стоит предусмотреть возможность применения ускорителей алгоритмов ИИ даже в достаточно простых устройствах. Глобальная перестройка программного обеспечения с учётом новых тенденций потребует обновления вычислительных мощностей. В этой ситуации вам не придётся вкладывать большие ресурсы в мощные сервера, когда ваша система уже имеет в своём составе распределённые компоненты, которые могут взять на себя часть работы.

Выбирая для использования в своих проектах решения от компании ААЕОН, вы сможете подобрать платформы и компоненты с поддержкой различных процессоров и с широкими возможностями интеграции. Высокая надёжность оборудования, ориентированного на промышленное применение, позволит разместить его непосредственно на объектах управления, что также снизит сложности, связанные с внедрением и обслуживанием современных систем на производствах. ●

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

Решения Advantix из реестра Минпромторга



Уже не первый год компания «Авантикс» помогает заказчикам удовлетворять их потребности в промышленном оборудовании в рамках импортозамещения. В связи с повышенным спросом на отечественную продукцию, внесённую в специальный единый реестр российской радиоэлектронной продукции, компания «Авантикс» создала линейку промышленных компьютеров, внесённых в реестр Министерства промышленности и торговли РФ по ПП от 10.06.2019 № 878.

Пока линейка представлена двумя безвентиляторными моделями – IPC-ATOM2-1U и ER-ATOM1.

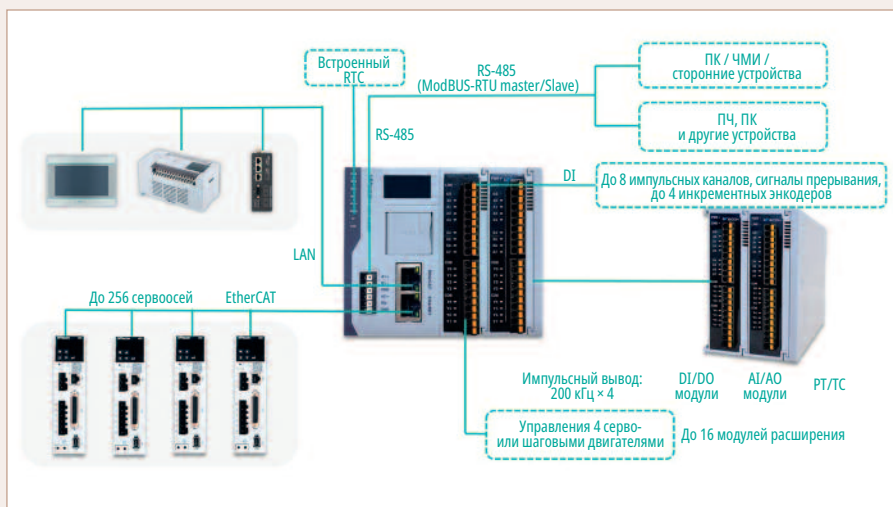
Промышленный компьютер Advantix IPC-ATOM2-1U является решением в корпусе высотой всего 1U для монтажа в 19-дюймовую стойку, работает на базе 2- или 4-ядерного процессора Intel Atom серии E и имеет достаточную производительность для работы с графикой благодаря графическому процессору Intel HD с поддержкой до двух дисплеев.

Несмотря на компактный размер, модель имеет все необходимые порты ввода-вывода, в том числе пять LAN-портов, а также вмещает до 1 Тбайт памяти, что делает её отличным решением для систем информационной безопасности, например, для разворачивания межсетевых экранов или шлюзов.

Advantix ER-ATOM1 – экономичная компактная модель начального уровня для базовых задач автоматизации, которая отличается гибкостью конфигурации благодаря поддержке слота расширения mini PCIe и возможностью увеличения оперативной памяти, а также дисковой подсистемы. Несмотря на компактный размер и лёгкий вес (всего 800 г), модель вмещает в себе три сетевых интерфейса, два USB-порта и один HDMI (или VGA). Компьютер совместим с операционной системой Astra Linux, а также поддерживает Windows 10 IoT Enterprise 64 bit. Обе модели доступны к заказу до 2025+ года.

Для получения цены и сроков поставки присылайте заявки в info@advantix-pc.ru.

Производительный EtherCAT контроллер Wecon – LX6V



Характеристики EtherCAT контроллер Wecon – LX6V

Режим работы	Циклическое сканирование/прерывание
Программирование	Instruction Table / Ladder Diagram / FB / FC
Количество инструкций	Базовые инструкции: 29 / Инструкции приложений: 170
Время выполнения	Базовые инструкции 0,01–0,02 мкс
Системная память EMMC	5 Мбайт
Ввод/вывод, расширение	16 встроенных DIO, до 16 модулей ввода/вывода
Высокоскоростной импульсный выход	Тип транзисторный 4 канала / 200 кГц
Высокоскоростной вход	Одиночный, 4 канала 150 кГц
Высокоскоростной вход	AB фаза, 4 канала 100 кГц, поддерживает 2 или 4 умножения частоты
Фильтр	Для всех входных данных
Последовательный порт	2×COM RS-485
LAN-порт	1×Ethernet, 1×EtherCAT
Рабочая температура	0...55°C

Компания Wecon выпустила новую серию программируемых контроллеров – LX6V. По сравнению с предыдущими сериями в ПЛК серии LX6V увеличен объём программной памяти до 5 Мбайт для хранения большего объёма программ и сохранения высокой производительности, более производительный процессор ARM Cortex с частотой 1,8 ГГц позволяет сократить время выполнения цикла до 0,01–0,02 мкс.

Модульная конструкция, с учётом наличия 16 встроенных дискретных каналов ввода/вывода, позволяет подключить до 16 модулей расширения. Из числа встроенных 4 канала поддерживают высокоскоростной вывод и 8 каналов ввода с частотой 200 кГц.

Основной особенностью новой серии является поддержка высокоскоростной шины EtherCAT, за счёт чего появляется возможность подключения до 256 сервоприводов (прежде всего, серии VD3E), устройств сто-

ронных производителей с поддержкой шины EtherCAT, а также интеграцию в современные системы управления, такие как системы управления движением или системы технического зрения, где требуется скоростная передача данных между компонентами.

