

РЕАЛИЗАЦИЯ КОМПЬЮТЕРНОГО ЗРЕНИЯ СРЕДСТВАМИ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ НА БЮДЖЕТНЫХ МИКРОКОНТРОЛЛЕРАХ

Приймак А.А.

Пенза

Ключевые слова: компьютерное зрение, распределенные вычисления, микроконтроллер, нейронные сети, распознавание объектов.

Аннотация. Цель статьи состоит в выработке принципов, решений и схем, позволяющих реализовать компьютерное зрение с наименьшими экономическими и временными затратами. В частности, рассматривается построение системы распознавания объектов. Показаны, популярные, на сегодняшний день, решения для реализации компьютерного зрения, отмечены их преимущества и недостатки. В качестве альтернативы, предложено использовать несколько, параллельно работающих, микроконтроллеров, соединенных с одной камерой. Недорогие, но сравнительно «слабые», микроконтроллеры предложено объединить в логические блоки, между которыми распределить кадры анализируемого видеопотока. Также, высказано предложение, о том, чтобы делить каждый кадр видеопотока на несколько частей и анализировать эти части отдельно – каждую на отдельном микроконтроллере, входящем в состав блока.

IMPLEMENTATION OF COMPUTER VISION BY MEANS OF DISTRIBUTED COMPUTING ON LOW-COST MICROCONTROLLERS

Priymak A.A.

Penza

Keywords: computer vision, distributed computing, microcontroller, neural networks, pattern recognition.

Abstract. The purpose of the article is to develop principles, solutions and schemes that allow implementing computer vision with the least economic and time costs. In particular, the construction of an object recognition system is considered. Shown, popular today, solutions for the implementation of computer vision, showing their advantages and disadvantages. As an alternative, it is proposed to use several microcontrollers operating in parallel, which are connected to one camera. Inexpensive, but relatively "weak", microcontrollers are proposed to be combined into logical blocks, between which the frames of the analyzed video stream are distributed. Also, a proposal was made to divide each frame of the video stream into several parts and analyze these parts separately - each on a separate microcontroller, which is part of the block.

Введение. В настоящее время разрабатывается много дронов и роботов, гражданского и военного назначения, утрата которых предполагается, в ходе их эксплуатации. Для них могут быть актуальны технические решения, которые предполагают низкую стоимость и высокую скорость разработки этих устройств.

Одной из проблем, актуальной для робототехники, является задача распознавания объектов. Для её решения используются различные алгоритмы – перебор изображений объекта под разными углами, поиск и исследование контура, искусственные нейронные сети [1-2]. Последний набирает сегодня всё большую популярность.

Использование нейронных сетей для распознавания изображений, обычно, предполагает применение метода глубокого обучения. Реализация этого метода может потребовать значительных вычислительных ресурсов. Поэтому, в качестве компонента, отвечающего за обнаружение объекта, используют одноплатные компьютеры, которые обладают нужными вычислительными ресурсами. Однако, цена таких компьютеров высока, а их доступность на рынке не очевидна.

В связи с этим предлагается альтернатива – использование нескольких, параллельно работающих, микроконтроллеров [3-4]. Микроконтроллеры, которые подразумевается использовать, уступают одноплатным компьютерам по вычислительным ресурсам, но зато они дешевле и всегда в наличии на российском рынке электронных компонентов. При этом, задача распознавания объекта, такими микроконтроллерами, имеет реальные перспективы.

Сравнительный анализ популярных решений. Популярным решением для реализации компьютерного зрения, сегодня, является использование одноплатного компьютера Raspberry Pi. На него может быть инсталлирована операционная система Linux или Windows. Под этими ОС возможен запуск программ, реализующих управление дроном или роботом. Обычно, программы, реализующие компьютерное зрение, используют библиотеку OpenCV [1], которая доступна для большинства популярных языков программирования.

Вычислительных ресурсов компьютера Raspberry Pi достаточно для реализации компьютерного зрения и решения задач распознавания, в частности. Модель Raspberry Pi 4 имеет четырёхядерный процессор Broadcom BCM2711 с частотой 1500 МГц и объём оперативной памяти 4096 Мб.

Однако, стоимость этого одноплатного компьютера велика, а его приобретение может быть затруднительным. При этом, возможности Raspberry Pi, могут быть даже избыточными для некоторых проектов, где требуется, например, только определить наличие объекта на изображении, попавшем в объектив камеры.

Альтернативой использованию Raspberry Pi, для задачи распознавания объекта, может служить несколько микроконтроллеров, реализующие параллельные вычисления. Нагрузка, при этом, будет равномерно распределена между несколькими микроконтроллерами.

Предлагаемое решение. Некоторое время назад такое решение казалось не реализуемым, потому что частота тактирования, производимых микроконтроллеров (16 МГц у Arduino NANO), была меньше, чем минимальная частота тактирования модуля камеры (24 МГц у OV7670). Но, сегодня доступны микроконтроллеры с более высокой частотой, например, ESP32, у которого она составляет 160 МГц. Также, у данного микроконтроллера есть Flash – память, объемом в 4 Мб. Такого объема памяти, вполне достаточно, для хранения в ней небольшого изображения, являющегося фрагментом целого кадра. Таким образом, отпадает необходимость в приобретении и подключении внешней памяти к микроконтроллеру. Стоимость же самого микроконтроллера ESP32 гораздо меньше (даже при условии приобретения нескольких штук), чем у Raspberry Pi, а с его доступностью на рынке не возникает никаких проблем.

Предлагаемая схема (рис. 1) состоит из модуля камеры (например, OV7670 с FIFO буфером) и четырех микроконтроллеров ESP32, а также кварцевого резонатора. Камера соединена с микроконтроллерами по интерфейсу I²C (совместим с SCCB). Данный интерфейс выбран, по причине того, что его поддерживает вышеназванная камера, однако может быть использован и интерфейс SPI.

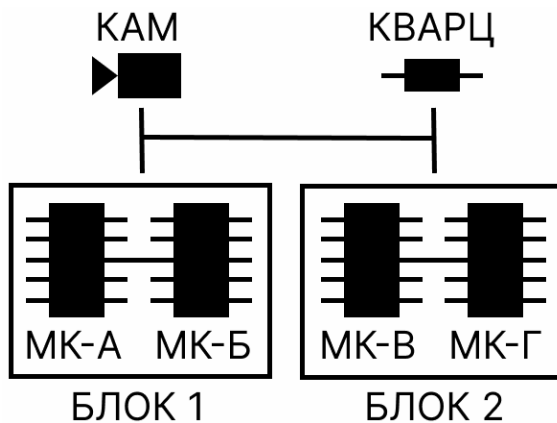


Рис. 1. Предлагаемая схема

Микроконтроллеры логически объединены в блоки, которые функционируют параллельного, с одной тактовой частотой, задаваемой резонатором. Каждый блок осуществляет анализ отдельного кадра, а каждый микроконтроллер выполняет анализ фрагмента этого кадра. В рамках предлагаемой схемы, может быть минимум два блока, внутри каждого из которых может находиться минимум два микроконтроллера.

Работа схемы (при наличии двух блоков, внутри каждого из которых размещено два микроконтроллера) начинается с того, что на первом такте (рис. 2) происходит фотоснимок (кадр). Микроконтроллер *A* получает одну половину этого кадра, а микроконтроллер *B* его другую половину. Таким образом, первый блок начинает анализировать первый кадр. На такте $1 + n$ происходит второй фотоснимок и микроконтроллер *B* получает одну его половину, а микроконтроллер *Г* другую его половину. Таким образом, второй блок начинает анализировать второй кадр. На такте $1 + n + n$ происходит третий фотоснимок, половинки которого, анализируются в первом блоке, а на такте $1 + n + n + n$ делается четвертый фотоснимок, две части которого достаются микроконтроллерам второго блока. То есть, происходит непрерывное получение фотоснимков с камеры. Каждый из снимков анализируется либо микроконтроллерами первого блока, либо микроконтроллерами второго блока, которые работают параллельно.

Предложенное решение позволяет совместить во времени процесс получения фотоснимка с камеры и его дальнейший анализ. Под понятием «анализ», в данной статье, понимается поиск объектов на фотоснимке и, возможно, выработка управляющего сигнала для какого – либо исполнительного механизма (например, сервопривода). Здесь может быть реализована выработка различного сигнала, в зависимости от того, в какой именно части кадра (каким микроконтроллером) будет обнаружен объект.

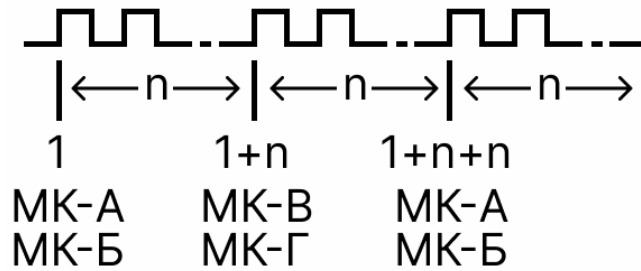


Рис. 2. Работа схемы по тактам с указанием задействованных МК

Предлагается реализовать поиск объекта на фотоснимке с помощью алгоритмов, построенных на нейронных сетях. Реализация работы нейронных сетей на микроконтроллерах, типа ESP32, возможна, и, хотя их вычислительные способности ограничены, распараллеливание вычислений должно обеспечить работоспособность решения и обеспечить приемлемую частоту кадров.

В качестве типа, используемых нейронных сетей, предлагается выбрать свёрточные нейронные сети (CNN) [2], которые наиболее хорошо подходят для решения задач распознавания изображений. В качестве технологии – метод глубокого обучения [2].

Важным, в предлагаемой схеме, является случай попадания искомого объекта «на стык» двух половинок кадра, каждая из которых достаётся отдельному микроконтроллеру. Предлагается учесть такой случай, при создании алгоритмов анализа фрагментов фотоснимка. Для этого, нейронную сеть следует «обучать» не только нахождению полного изображения детектируемого объекта, но и тренировать её искать его части. Также, для повышения надёжности, предлагается суммировать вероятности нахождения объекта на каждой из частей фотоснимка, приняв за максимум 1 или 100%. То есть, например, если микроконтроллер А, находящийся в первом блоке, «выдал» вероятность нахождения объекта в 20%, а микроконтроллер В – 40%, то следует суммировать эти две вероятности для подсчёта общей – 60%. Таким образом, предлагается вычислять суммарную вероятность нахождения объекта для всего логического блока, содержащего несколько микроконтроллеров, анализирующих части фотоснимка. Другими словами, вероятность нахождения объекта должна рассчитываться для всего кадра, а не для только для его фрагментов.

Выводы. Автор пришел к следующим выводам, анализируя предлагаемое решение.

1. На момент написания работы моделирование, с использованием реальных микроконтроллеров и других компонентов схемы, ещё не проводилось. Исходя из этого обстоятельства, пока не представляется возможным спрогнозировать время, которое понадобится каждому микроконтроллеру для анализа фрагмента фотоснимка, определенного размера. Поэтому, трудно сделать выводы об оптимальном количестве блоков и количестве микроконтроллеров, внутри каждого из них.

2. Преимуществом схемы, помимо низкой стоимости её компонентов (относительно Raspberry Pi), является её масштабируемость и гибкость. На время анализа и минимальную частоту кадров можно влиять как изменением количества блоков, так и изменением количества микроконтроллеров в каждом из них.

3. Также, есть предположение о том, что переход от анализа видеосигнала к анализу отдельных кадров, на физически разных компонентах, приведет к повышению надежности обнаружения объекта.

4. Система, использующая предлагаемую схему, может применяться и как предварительная. То есть, при обнаружении объекта системой, построенной на нескольких маломощных микроконтроллерах, изображение может передаваться, для последующего анализа, в другую систему, построенную на высокопроизводительных и дорогих компонентах.

Список литературы

1. Постоит А.В. Основы искусственного интеллекта в примерах на Python. – СПб.: БХВ, 2021. – 448 с.
2. learn.microsoft.com [Электронный ресурс]. – URL: <https://learn.microsoft.com/ru-ru/azure/machine-learning/concept-deep-learning-vs-machine-learning>.
3. Косяков М.С. Введение в распределённые вычисления. – СПб.: НИУ ИТМО, 2014. – 155 с.
4. Шипов И.А. Реализация распределенных вычислений на отечественных микропроцессорных устройствах // Известия ЮФУ. Технические науки. – 2022. – №1. – С. 218-225.

Сведения об авторе:

Приймак Антон Александрович.