

1. ЦИФРОВАЯ ВИДЕОКАМЕРА

1.1. Структурная схема цифровой видеокамеры

В современных цифровых видеокамерах изображение проходит большое число дополнительных стадий предобработки перед тем, как пользователь увидит результат съемки. Как правило, необработанное «сырое» изображение проходит по конвейеру аналоговой или цифровой обработки, где каждый отдельный блок выполняет свой набор функций преобразования и коррекции. Типовая структурная схема видеокамеры представлена на рисунке 1.1.

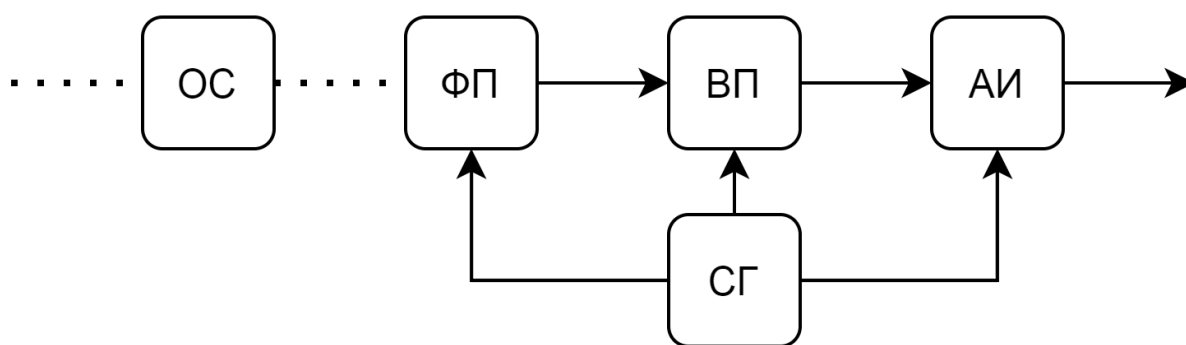


Рисунок 1.1 – Структурная схема цифровой видеокамеры

Для преобразования поля излучения, так называемого «пространства предметов», в поле «пространства изображений», представляющего собой двумерное распределение интенсивности светового потока, проецируемое на поверхность фотоприемника (ФП), необходима оптическая система (ОС).

Выбор оптической системы является важным этапом при разработке любой теле- и видеосистемы, так как во многом определяет возможности всей системы в целом. Объективы имеют множество различных характеристик, которые необходимо учитывать при выборе оптической системы под конкретную задачу: размер кадра и кроп-фактор, фокусное расстояние, угол поля зрения, светосила, разрешающая способность, уровень оптических искажений и многие другие.

Многие из перечисленных характеристик напрямую влияют на выбор фотоприемника видеокамеры. Например, фотоприемник должен располагаться точно в фокусе объектива для максимальной четкости фокуси-

руемого изображения, что напрямую будет отражаться на массогабаритных характеристиках конечного устройства. Кноп-фактор выбираемого ФП должен быть равен или меньше, чем у объектива, в противном случае на изображении появятся черные участки по углам из-за сферической формы фокусирующих линз. Так как оптическая система состоит из набора прозрачных линз, у которых показатель преломления на границе с воздухом отличается от 1 и различен для различных длин волн, то в плоскости фотоприемника произойдет смещение цветовых каналов, так называемые «хроматические аберрации», которые проявляются в виде цветных контуров на изображении. Каждый объектив также характеризуется также своей разрешающей способностью, которая должна быть равна или превышать разрешающую способность ФП. Если это условие не выполняется, то использование дорогого и качественного ФП окажется бессмысленным. Пример объектива фирмы Canon с изменяемым размером диафрагмы при неизменной светосиле на всем диапазоне, который выполнен в пыле- и влагозащищенном корпусе, представлен на рисунке 1.2.

Фотоприемник – полупроводниковый прибор, предназначенный для регистрации и преобразования оптического излучения в электрический сигнал. Фотоматрицы также имеют ряд важных для проектирования характеристик: отношение сигнал/шум, чувствительность, разрешающая способность, физический размер матрицы и формат кадра, и многие другие. Однако помимо различий в перечисленных параметрах, матрицы делятся по применяемой технологии. Наиболее широкие сферы применения нашли матрицы ПЗС и КМОП.

ПЗС–матрица состоит из светочувствительных фотодиодов, в которых накапливается заряд, соответствующий уровню освещенности пиксела, и использует технологию приборов с зарядовой связью для его последующего считывания. На выходе ПЗС-матрицы видеосигнал представлен в аналоговом формате в виде уровня напряжения, соответствующего величине считываемого зарядного пакета.



Рисунок 1.2 – Объектив фирмы Canon EF 24-70mm f/2.8L II USM

КМОП-матрица, в отличие от ПЗС-матриц, позволяет выбирать и считывать уровень накопленного заряда в каждом конкретном пикселе в произвольном порядке, как в микросхемах памяти. Каждый пиксел имеет свой усилитель считывания. При выборе пиксела уровень накопленного заряда через усилитель поступает на видеошину и считывается встроенным АЦП. После чего в виде цифрового кода, соответствующего уровню считанного заряда, передается в выходной интерфейс.

Каждый тип матриц имеет свои преимущества и недостатки, поэтому использование того или иного типа сенсора в большинстве случаев обусловлено требованиями конкретной задачи. Пример цветной КМОП-матрицы фирмы E2V представлен на рисунке 1.3.

На выходе микросхемы фотоприемника изображение представляет собой необработанный «сырой» поток данных. Ввиду неидеального технологического процесса при изготовлении фотоматриц, а также процесса накопления и считывания кадра, необходимо выполнить ряд корректирующих операций по устранению шумов и дефектов сенсора перед тем, как кадр будет передан далее для отображения на экран какого-либо устройства или подвергнут анализу интеллектуальными системами по классификации, распознаванию объектов в кадре и т.д. Эту функцию как правило выполняет видеопроцессор (ВП).

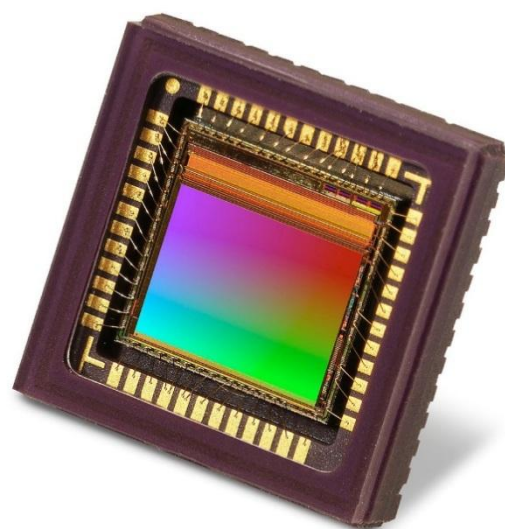


Рисунок 1.3 – CMOS сенсор фирмы E2V EV76C660 с разрешением матрицы 1.3 МП

Для ПЗС-матриц существуют специальные видеопроцессоры, которые позволяют выполнять прием и предварительную обработку видеосигнала в аналоговом виде, формировать управляющие синхросигналы, необходимые для взаимодействия с сенсором, и т.д. Также могут иметь интегрированные горизонтальные и вертикальные драйверы для прямого управления процессом накопления и считывания кадра с сенсора. Пример функциональной схемы такого видеопроцессора для ПЗС-матрицы приведен на рисунке 1.4.

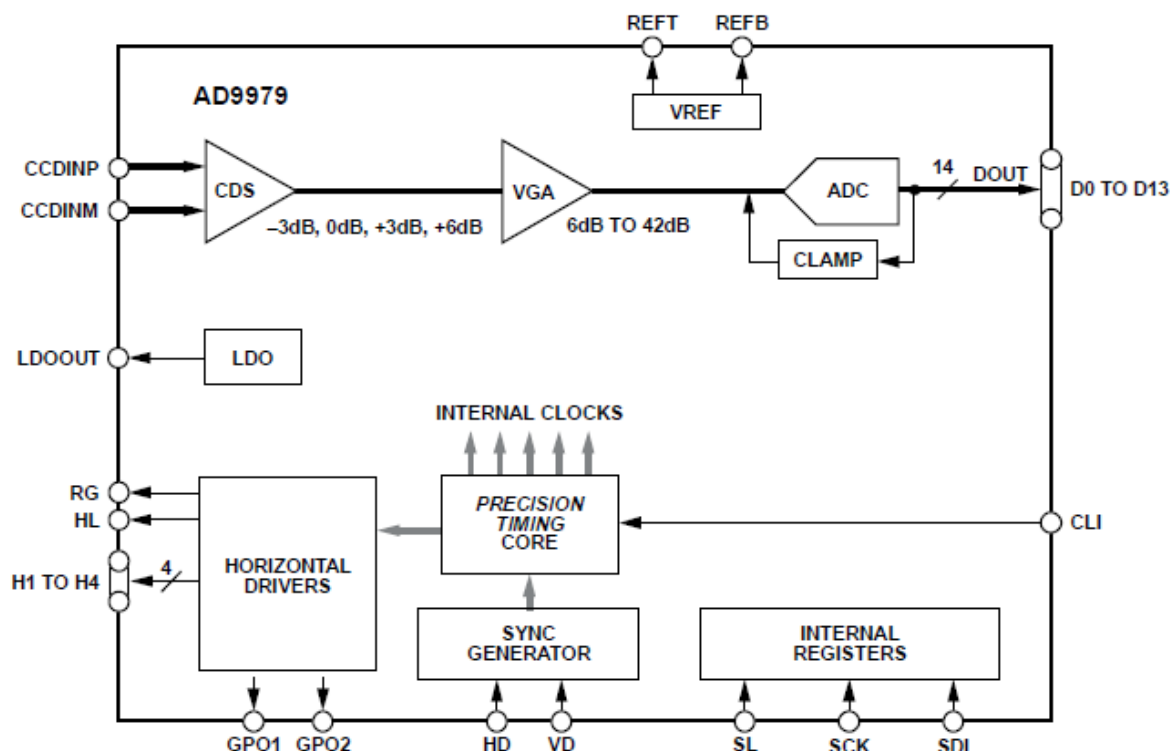


Рисунок 1.4 - Функциональная схема видеопроцессора для ПЗС-матриц фирмы Analog Devices – AD9979

Входной дифференциальный видеосигнал поступает в блок двойной коррелированной выборки (CDS), после чего видеосигнал усиливается в блоке VGA на заданную пользователем величину и преобразуется в цифровой код при помощи интегрированного блока АЦП.

Для управления режимами работы внутренних блоков предусмотрена регистровая модель (Internal registers). А для управления самой ПЗС-матрицей предусмотрен интегрированный горизонтальный драйвер, который формирует управляющие сигналы на основе входных сигналов вертикальной (VD) и горизонтальной (HD) синхронизации.

Для КМОП-матриц обработка может осуществляться на любой аппаратной платформе, способной выполнять обработку данных в цифровом формате, которые будут рассмотрены позднее в подглаве 1.3.

Для синхронизации работы сенсора и видеопроцессора необходим блок тайминг-генератора (ТГ), который формирует сетку частот для работы самого сенсора, взаимодействия его с видеопроцессором, а также для работы внутренней логики видеопроцессора. В некоторых случаях на ТГ может возлагаться задача генерации сигналов синхронизации для управления работой ФП и ВП.

В общем случае необходимо предусмотреть также блок адаптера интерфейса (АИ), задачей которого является передача полученного и обработанного кадра далее по тракту обработки, анализа и визуализации.

Например, если выходом ВП является аналоговый видеосигнал, то в простейшем виде его необходимо преобразовать в цифровой формат с помощью АЦП. В более сложных вариациях, например, для снижения рабочей частоты следующих блоков конвейера блок АИ может выполнять промежуточную буферизацию данных, формируя слова данных большей ширины, и генерировать соответствующие сигналы подтверждения валидности данных. Также для интеграции в сложные встраиваемые системы может потребоваться формирование какого-либо специализированного интерфейса, например: AXI, Avalon, АНВ и т.д.

1.2. Видеопроцессор

1.2.1. Функции видеопроцессора

Видеопроцессор – это аппаратный блок или устройство, осуществляющее обработку видеопотока или цифрового изображения. Кадр, полученный с фотоматрицы, подвержен воздействию шумов различной природы, уровень которых зависит от многих факторов. Поэтому перед выводом его на экран смартфона или монитор компьютера, или дальнейшей обработке алгоритмами компьютерного зрения, необходимо выполнить ряд функций предобработки для увеличения качества получаемого кадра.

На первых этапах необходимо компенсировать все дефекты на изображении, которые связаны с технологическим процессом изготовления самой фотоматрицы или вносятся оптической системой. К таким можно отнести замену или интерполяцию «битых» пикселей, фиксацию уровня черного, устранение виньетирования и т.д.

Если цветное изображение получается методом цветового кодирования с помощью наложения специальных цветофильтров на черно-белую матрицу, то необходимо интерполировать или иным образом заполнить соседние пиксели в каждом цветовом канале.

Для снижения уровня шумов на изображении выполняется его обработка различными алгоритмами шумоподавления.

Помимо этого, необходимо менять «баланс белого» в кадре для правильной цветопередачи, осуществлять гамма-коррекцию, согласовывать его динамический диапазон с динамическим диапазоном визуализирующего устройства и многое другое.

В зависимости от реализации или требований конкретной задачи набор выполняемых операций может варьироваться. Рассмотрим подробнее некоторые из перечисленных функций видеопроцессора.

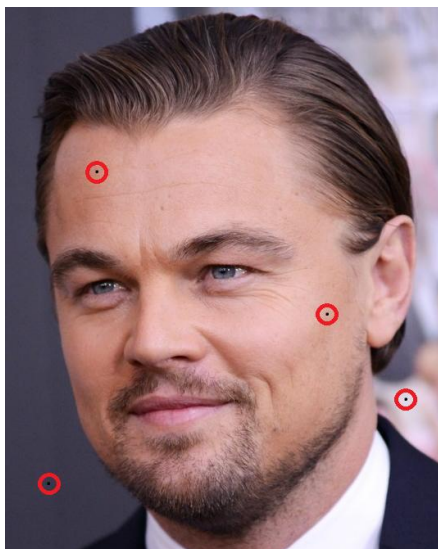
1.2.2. Коррекция битых пикселей

Качество фотоматрицы зависит от большого числа переменных, в том числе от условий и технологии её производства, поэтому могут получаться пиксели, которые не функционируют должным образом и проявляются на изображении в виде черных или белых точек.

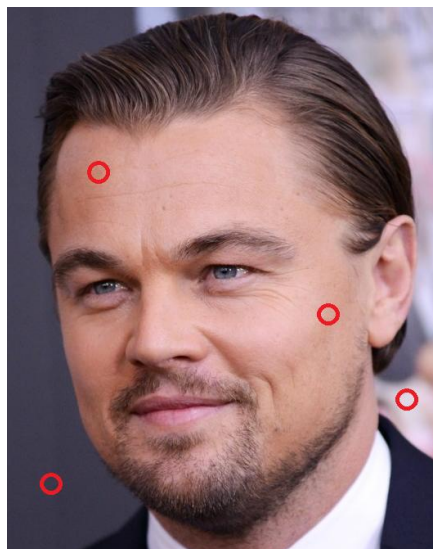
Полностью избежать этого явления не представляется возможным, однако, несколько десятков «битых» пикселей не окажет значительного влияния на качество снимка в целом, где изображение состоит из нескольких миллионов, а в современных матрицах нескольких десятков миллионов пикселей.

Решением этой проблемы является составление специальной «карты битых пикселей» на заводе изготовителя. Впоследствии на каждом снимке необходимо компенсировать дефект согласно этой карте (рисунок 1.5.). Можно выделить несколько основных способов коррекции битых пикселей:

- Заменить на значение соседнего пикселя;
- Заменить на взвешенную сумму соседних пикселей;
- Скорректировать значение, если известна степень искажения, вносимая дефектом;



а)



б)

Рисунок 1.5 – Исходное изображение с «битыми» пикселями матрицы (а) и после их интерполяции (б)

1.2.3. Фиксация уровня черного

Даже при отсутствии освещения на фоточувствительной матрице в её фотодиодах будет накапливаться некоторый заряд, обусловленный процессом термоэлектронной эмиссии. Для того, чтобы скомпенсировать наличие этой составляющей в кадре, из него вычитают среднюю величину уровня черного по изображению. Такая операция возможна из-за конструкции самой матрицы, в которой предусмотрены участки кадра, закрытые от внешнего воздействия, тем самым в них может накапливаться только заряд, обусловленный вышеупомянутым процессом. Структура реального кадра, получаемого с сенсора представлена на рисунке 1.6.

Так как величина накопленного заряда является случайной величиной и нелинейно зависит от температуры, то вычитаемое значение получают усреднением по всем «черным» пикселям.

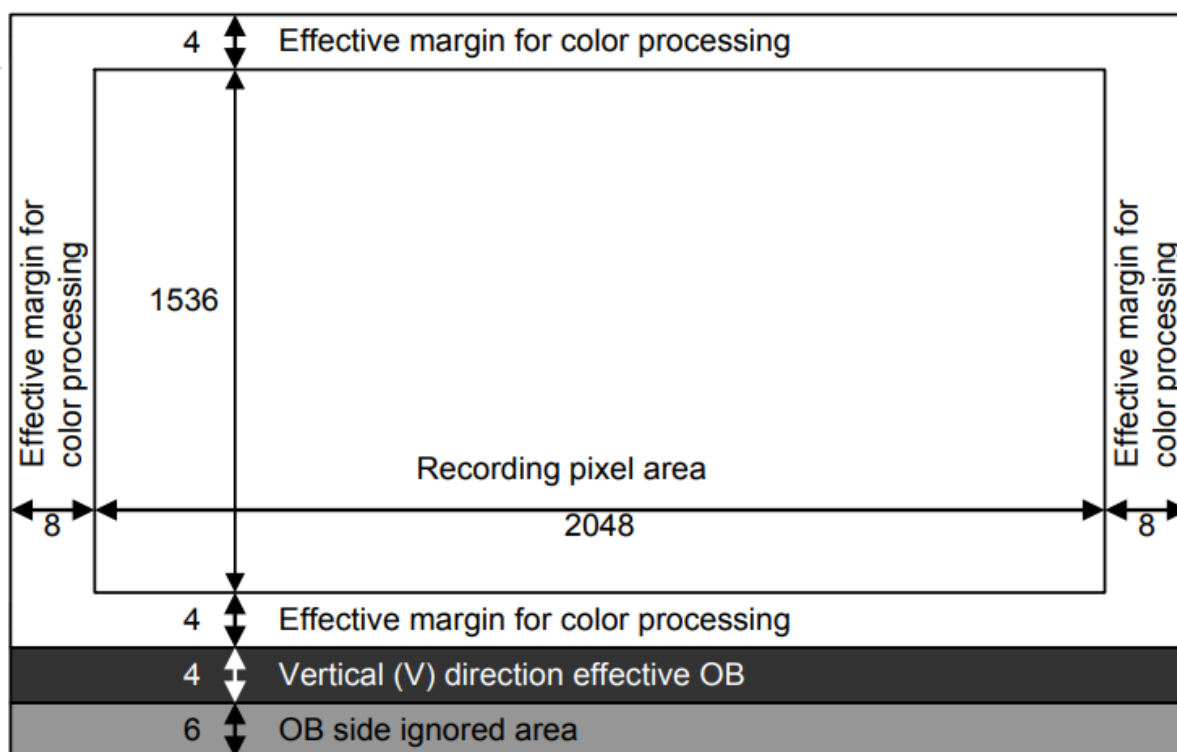


Рисунок 1.6 – Структура информационной части кадра КМОП сенсора фирмы Sony - IMX265

1.2.4. Устранение эффекта виньетирования

Явление «виньетирования» проявляется в виде уменьшения яркости или насыщенности участков изображения на периферии в сравнении с центром. Одной из причин этого дефекта является ослабление оптической системой потока лучей, которые проходят под углом к её оптической. Особенно этот эффект заметен при использовании широкоугольных объективов или объективов с большой светосилой.

Выделяют несколько типов виньетирования:

- Оптическое;
- Натуральное;
- Механическое;
- Пиксельное;

Каждый имеет свою природу и способы борьбы с ним. Однако необходимо помнить, что при компенсации этого дефекта программными способами, побочным эффектом работы алгоритма может оказаться увеличение шумов по краям изображения.

Пример работы алгоритма компенсации эффекта виньетирования представлен на рисунке 1.7.

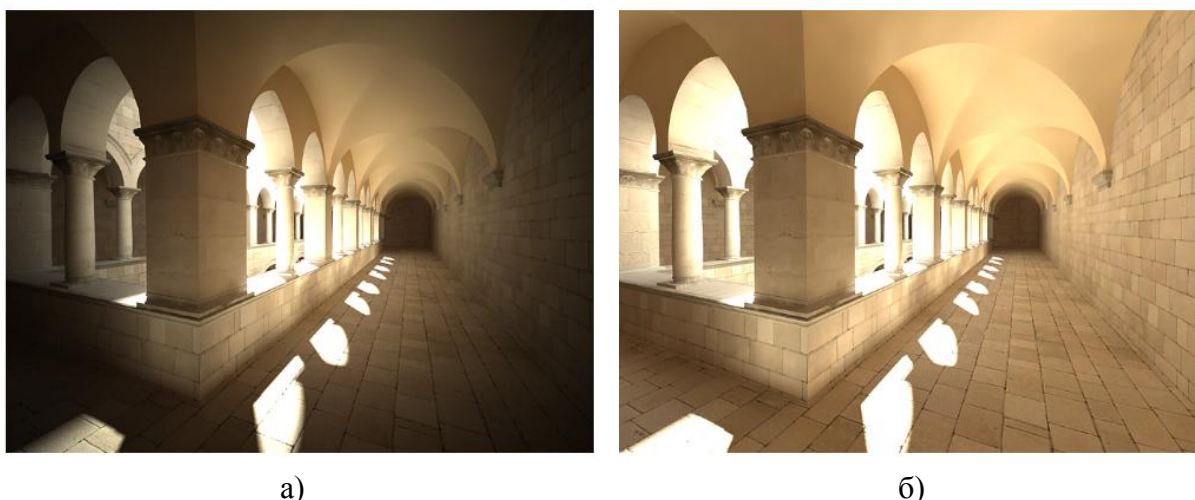


Рисунок 1.7 – Исходное изображение с эффектом виньетирования (а) и после его компенсации специальным алгоритмом (б)

1.2.5. Кодирование цветowych каналов

Получить цветное изображение можно различными способами. Например, падающий на объектив камеры световой поток можно разделить с помощью дихроических зеркал или призм на несколько цветowych каналов. Тогда на каждый канал нужно предусмотреть отдельный фотодатчик (рисунок 1.8).

Однако при таком подходе значительно увеличиваются массогабаритные характеристики конечного устройства из-за тяжелой и громоздкой оптической системы. Поэтому в современных фото- и видеокамерах применяют метод цветовой кодирования с использованием различных цветофильтров, которые накладываются на фотодиоды матрицы.

Одним из часто используемых фильтров является фильтр

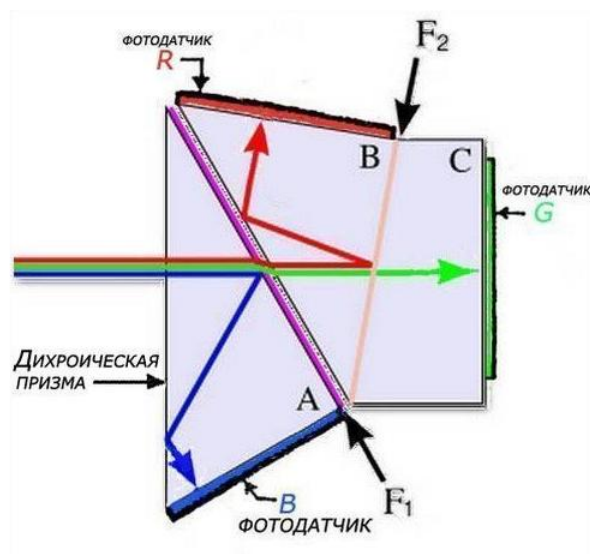


Рисунок 1.8 – Деление светового потока с помощью дихроической призмы

Байера и его различные модификации. Классический фильтр Байера представляет собой двумерный массив цветных фильтров, состоящий из $\frac{1}{4}$ части красных элементов, $\frac{1}{4}$ синих и $\frac{1}{2}$ зелёных элементов, расположенных как показано на рисунке 1.9.

Для получения результирующего цветного изображения необходимо интерполировать полученные значения пикселей в каждом канале на соседние пиксели. Процесс получения цветного изображения таким способом показан на рисунке 1.10.

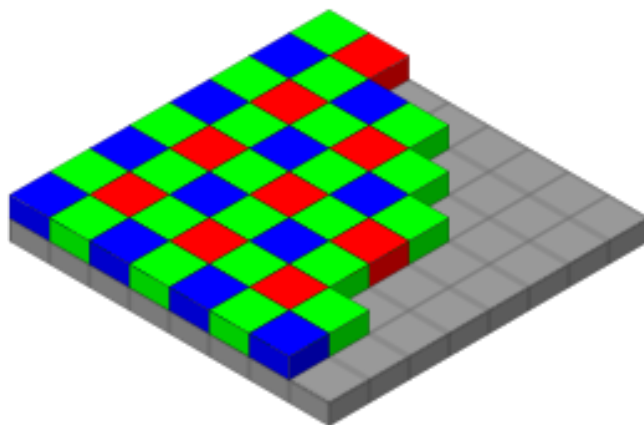
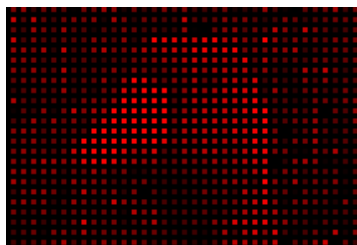
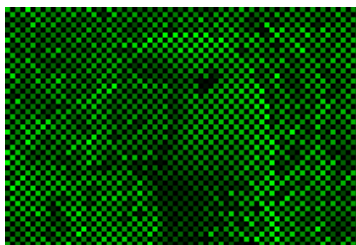


Рисунок 1.9 – Пространственное цветоделение изображения с помощью мозаичного фильтра Байера

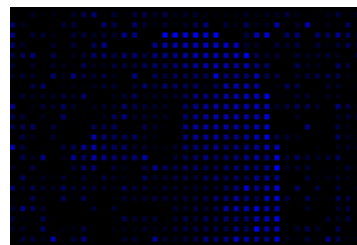
Из-за применения мозаичных цветовых фильтров снижается реальная разрешающая способность системы в цветных деталях, а также появляется эффект цветного муара. Для подавления цветных артефактов необходимо применить фильтр нижних частот, дополнительно размывающий изображение, с последующим применением алгоритмов повышения резкости.



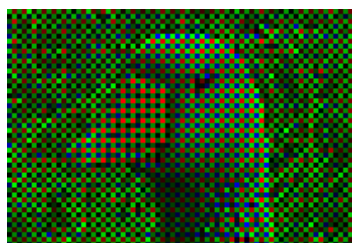
Красный канал (R)



Зеленый канал (G)



Синий канал (B)



Объединение каналов



Интерполяция недостающих пикселей

Рисунок 1.10 – Получение цветного изображения с применением мозаичного фильтра Байера

1.2.6. Шумоподавление

Необходимость в шумоподавлении может быть обусловлена множеством различных причин. Для улучшения визуального восприятия кадра или, например, в медицинских целях для увеличения четкости изображения на рентгеновских снимках и т.д. Пример работы одного из алгоритмов шумоподавления представлен на рисунке 1.11.



Рисунок 1.11 – Исходное зашумленное изображение (а) и после обработки билатеральным фильтром (б)

Также шумоподавление играет важную роль при сжатии видеоряда или изображений. И в видео, и в изображениях сжатие основано на пространственной корреляции значений соседних пикселей, а в случае видео - еще и временной корреляции (на сходстве последовательно идущих кадров между собой). Одной из основных проблем в алгоритмах сжатия является определение локальной зашумленности данной области изображения, поскольку при сжатии сильный шум может быть принят за детали изображения, и это может, во-первых, привести к увеличению сложности с точки зрения сжатия и, во-вторых, отрицательно повлиять на результирующее качество сжатого изображения. При сжатии видеоизображений наличие шума приводит, кроме того, к увеличению межкадровой разницы, пони-

жая, таким образом, степень сжатия, и отрицательно влияет на точность работы алгоритма компенсации движения.

Существует огромное число алгоритмов шумоподавления, отличающихся по сложности, эффективности, объему затрачиваемых ресурсов и прочему. Подробнее данный вопрос рассмотрена в главе 2 настоящей работы.

1.2.7. Регулировка баланса белого

Практически при любом освещении белый цвет будет восприниматься человеческим глазом белым. Такую «автоматическую» цветокоррекцию производит человеческий мозг. Однако цифровая камера в сравнении со сложно организованным зрительным аппаратом человека является достаточно примитивным инструментом.

Поэтому для правильной цветопередачи в кадре пользователю необходимо выставить правильный баланс белого, чтобы относительно этого параметра можно было скорректировать остальные цвета на снимке.

Каким образом баланс белого влияет на визуальное восприятие цветов на изображении можно оценить по рисунку 1.12.



Рисунок 1.12 – Различный баланс белого на фотографии

Как видно на фотографии, в зависимости от баланса белого, снимок приобретает другие цветовые оттенки. Центральная фотография имеет правильный баланс белого, на левой фотографии явно выражен синий

оттенок неба и облаков, в то время как на правой значительно выражены оттенки желтого цвета.

Существуют алгоритмы для автоматической регулировки баланса белого в кадре, которые широко используются в современных фотоаппаратах. Однако для получения профессиональных снимков, где правильная цветопередача является важным аспектом, рекомендуется регулировать баланс белого вручную.

1.2.8. Линейное контрастирование изображения

Задача контрастирования связана с улучшением согласования динамического диапазона изображения и экрана, на котором выполняется визуализация. Изображение может содержать большое число пикселей близких по значению, однако если растянуть гистограмму на весь доступный диапазон, то заметно может улучшиться восприятие снимка. Пример показан на рисунке 1.13.



Рисунок 1.13 – Исходное изображение (а) и после линейного контрастирования (б)

Как можно заметить, изначально снимок сильно высвечен, темные пикселы практически отсутствуют, растянув гистограмму на весь диапазон, контраст был значительно увеличен.

1.2.9. Компенсация хроматических aberrаций

Хроматическими aberrациями называют разновидность aberrации оптической системы, которая обусловлена зависимостью показателя преломления среды от длины волны проходящего через неё излучения. Из-за паразитной дисперсии фокусные расстояния не совпадают для лучей света с разными длинами волн.

Хроматические aberrации приводят к снижению чёткости изображения, а иногда также и к появлению на нём цветных контуров, полос, пятен и т.д. Например, на рисунке 1.14 при детальном рассмотрении вершины горы можно заметить тонкий зеленый контур.

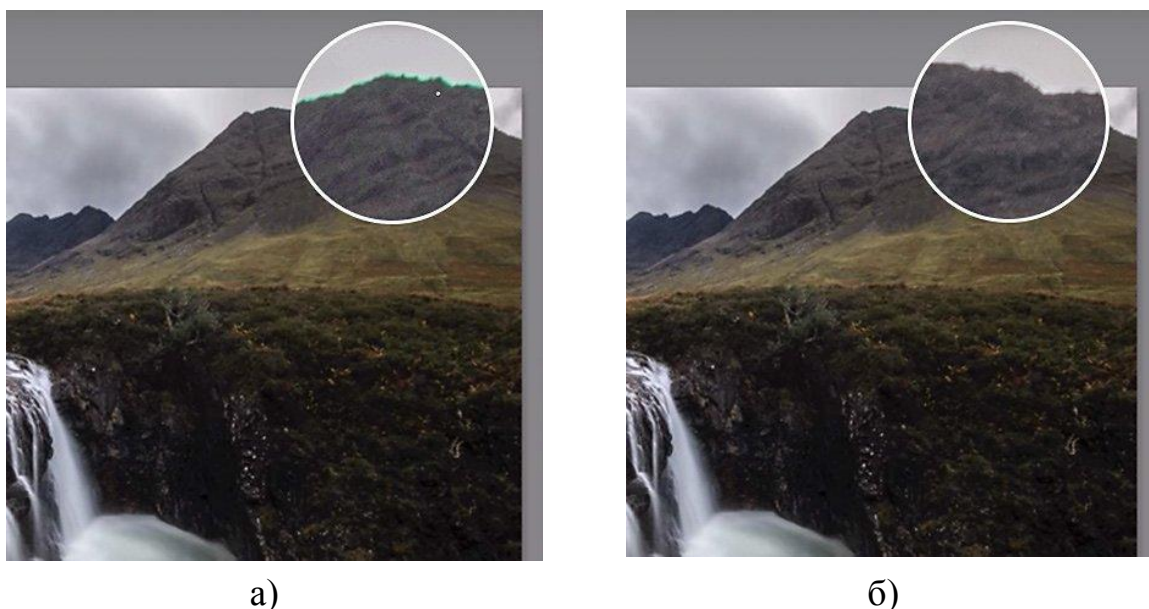


Рисунок 1.15 – Хроматические aberrации на исходном изображении (а) и после их компенсации (б)

Часто коррекцию искажений этого типа выполняют с применением математического аппарата регрессионного анализа. Когда на основе некоторых опорных, эталонных кадров, можно оценить величину искажений, обучить регрессионную модель и затем соответствующим образом, корректировать каждое получаемое изображение.

1.3. Платформы для реализации функций видеопроцессора

1.3.1. Обзор существующих платформ

Видеопроцессор может быть реализован множеством способов в зависимости от требуемого набора алгоритмов и функций.

Основные платформы для реализации ВП:

- Центральный процессор (CPU);
- Графический процессор (GPU);
- Интегральные схемы специализированного назначения (ASIC);
- Программируемые пользователем вентильные матрицы (FPGA);
- Облачные технологии, вычисления на серверах (Cloud);

Рассмотрим более подробно каждый из перечисленных вариантов реализации.

1.3.2. Центральный процессор (CPU)

Современные процессоры обладают большой производительностью, позволяя использовать их в качестве платформы для реализации обработки видеопотока (рисунок 1.16). Однако следует отметить, что зачастую алгоритмы реализованы локальным сканированием изображения скользящим окном. В таком случае гораздо эффективнее выполнять такие операции параллельно, так как они не зависят от результатов обработки на предыдущих этапах. Несмотря на заведомо большую производительность на ядро и более высокую рабочую частоту, порядка нескольких ГГц, такая обработка оказывается малоэффективной в случае захвата и обработки изображения в высоком разрешении с большим числом кадров в секунду.

Так для видеопотока с разрешением 1080p и частотой 60 кадров в секунду для одного цветного канала поток данных будет составлять порядка 3 Гбит/с. С учетом того, что часть кадра необходимо буферизировать для последующей обработки скользящим окном, процессор оказывается не в состоянии справиться с такой нагрузкой, даже при условии наличия нескольких ядер.



Рисунок 1.16 – Центральный процессор компании Intel

Для более эффективной обработки кадра предпочтительнее иметь большое число ядер с сокращенным набором выполняемых инструкций, нежели несколько ядер, поддерживающих всевозможные инструкции для обслуживания операционной системы, различных режимов адресации памяти и т.д.

Процессор, построенный на этой концепции, называется GPU. Он имеет большое количество универсальных простых ядер, число которых может достигать нескольких тысяч.

1.3.3. Графический процессор (GPU)

GPU является универсальным средством для обработки больших объемов данных на нескольких сотнях или тысячах ядер параллельно (рисунок 1.17). За счет этой особенности их использование в задачах обработки видеопотока, рендеринга и т.д. является более предпочтительным, нежели CPU. Однако появляется необходимость в ресурсе, который будет осуществлять загрузку этих маленьких ядер и распределять нагрузку между ними. Таким ресурсом может быть как CPU, рассмотренный выше, так и любое другое устройство, обладающее достаточной производительностью.

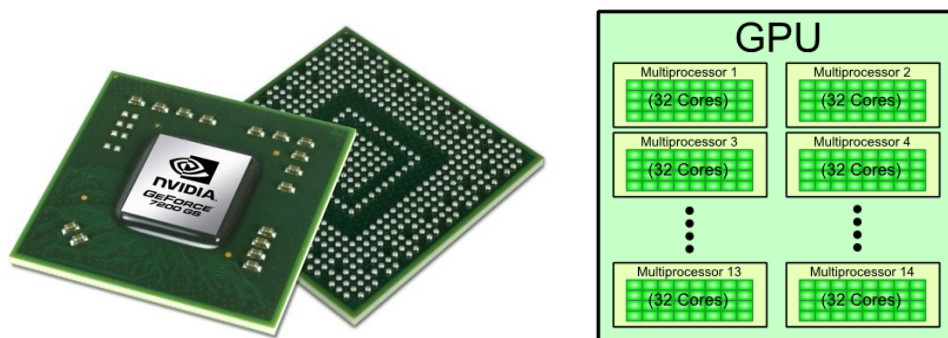


Рисунок 1.17 – GPU от компании NVIDIA – GeForce 7200GS

Современные GPU имеют не только большое количество универсальных ядер, но и высокую производительность каждого ядра, работая на частоте 1-2 ГГц. При таком подходе значительная степень параллелизма позволяет без особого труда производить обработку изображения несколькими алгоритмами последовательно в режиме реального времени, обеспечивая минимальную задержку кадра.

Однако такие процессоры имеют, как правило, значительное энергопотребление, и как следствие, высокий уровень рассеиваемой мощности, что требует дополнительного охлаждения в виде блоков радиаторов, которые охлаждаются вентиляторами, и достаточного мощного блока питания, что не позволяет использовать такое решение в задачах, требующих значительной времени автономной работы, устойчивости к механическим нагрузкам, работу в широком диапазоне температур и т.д. Поэтому их применение ограничено стационарными компьютерами, которые можно подключить к электросети и обеспечить соответствующее охлаждение комплектующих всей системы.

1.3.4. Интегральные схемы специального назначения (ASIC)

Часть функций видеопроцессора могут быть реализованы отдельными блоками и интегрированы в специализированные микросхемы. Например, используемые в смартфонах процессоры на ARM-архитектуре, такие как snapdragon, mediatek и многие другие, имеют встроенные аппаратные блоки по кодированию/декодированию видео, а также блоки приема видеопотока с цифровых камер ультравысокого разрешения 4к с большой частотой кадров в секунду и т.д.

Благодаря интеграции в микросхему, разработчику нет необходимости беспокоиться о вопросах согласования интерфейсов внешних модулей, генерации специальных синхросигналов и схемотехнических тонкостях использования конкретной микросхемы. Основной задачей разработчика в такой ситуации является правильная настройка этих блоков и согласование физического уровня взаимодействия с другими микросхемами.

1.3.5. Программируемые пользователем вентиляные матрицы (FPGA)

Внутренняя структура FPGA (ПЛИС) является конфигурируемой (рисунок 1.18), что позволяет реализовывать необходимые алгоритмы обработки ровно в той степени, в какой они необходимы, а также оптимизировать структуру разрабатываемых блоков с целью обеспечения максимального распараллеливания этапов обработки данных.

Так как обработка выполняется не программными способами, используя итерационный подход, а выполняется в виде аппаратных блоков внутри ПЛИС, то обработку видеопотока можно выполнять в виде конвейера, состоящего из отдельных аппаратных блоков, каждый из которых выполняет свои функции.

Такой подход позволяет добиться не только высокой пропускной способности, но также и значительного снижения энергопотребления в сравнении с CPU, потребление которых может составлять порядка 100 Вт, и GPU – новейшие микросхемы могут потреблять до 400 Вт. Решения на ПЛИС позволяют добиться энергопотребления в единицы Ватт, что дает возможность использовать их в задачах, недоступных для CPU и GPU.

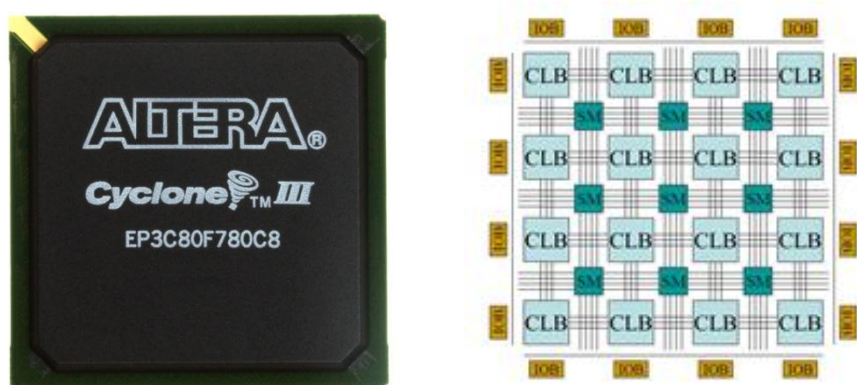


Рисунок 1.18 – ПЛИС Altera Cyclone III и его внутренняя архитектура

1.3.6. Облачные технологии

В современном мире облачные технологии широко используются во многих сферах для хранения данных, удаленной обработки и вычислений пользовательских задач, как средство связи различных модулей системы и т.д.

Как упоминалось выше, обработка видеопотока высокого качества с высокой частотой следования кадров является ресурсоемкой задачей и требует в некоторых случаях достаточно большой вычислительной мощности от используемого видеопроцессора. Поэтому обработку видеопотока можно возложить на мощные серверные системы.

«Сырой» видеопоток упаковывается в кадры и отправляется на сервер, который обрабатывает его и возвращает результат обработки пользователю (рисунок 1.19). Такой подход широко используется, например, при моделировании или технических расчетах сложных алгоритмов в современных САПР.

Безусловно, использование облачных технологий вносит дополнительные задержки на передачу кадра, поэтому использование такого формата возможно в тех случаях, когда задержки по времени не критичны для самой задачи, например: обработка и анализ видеоряда в складских помещениях, парковках, торговых центрах и т.д.

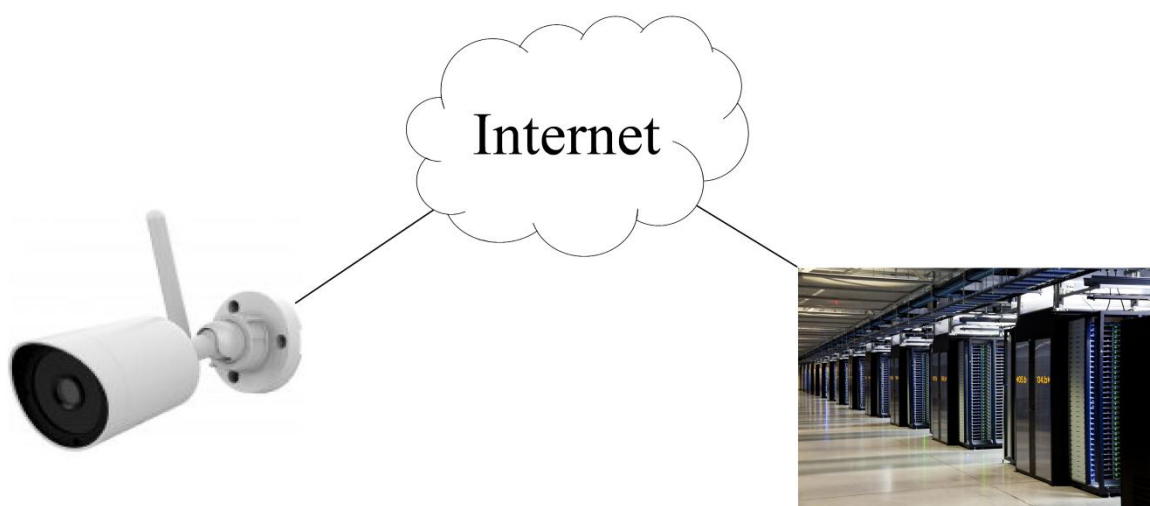


Рисунок 1.19 – Обработка видеопотока на арендуемом серверном оборудовании