

С. В. Сай, А. В. Каменский, М. И. Курячий

СОВРЕМЕННЫЕ МЕТОДЫ АНАЛИЗА И ПОВЫШЕНИЯ КАЧЕСТВА ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

ISBN 978-5-7389-3162-8



9 785738 931628

Хабаровск 2020

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«ТИХООКЕАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

С. В. САЙ, А. В. КАМЕНСКИЙ, М. И. КУРЯЧИЙ

Современные методы анализа и повышения качества цифровых изображений

Монография

Хабаровск
Издательство ТОГУ
2020

УДК 004.932
ББК 3973.235
С149

Р е ц е н з е н т ы:

Доктор технических наук *В. С. Титов*

Доктор технических наук *О. С. Амосов*

Сай, С. В.

С149 Современные методы анализа и повышения качества цифровых изображений : монография / С. В. Сай, А. В. Каменский, М. И. Курячий ; Министерство науки и высшего образования Российской Федерации, Тихоокеанский государственный университет. – Хабаровск : Изд-во Тихоокеан. гос. ун-та, 2020. – 173, [1] с.

ISBN 978-5-7389-3162-8

В монографии приводятся результаты исследования качества передачи цифровых изображений в системах видеокompрессии на основе стандартов JPEG, JPEG 2000 и цифровых видеокодеков (MPEG-2, MPEG-4, H.264, H.265/HEVC). Анализируются современные методы и метрики оценки качества цифровых изображений. Предлагаются способы повышения качества цифровых изображений в стандартах JPEG и JPEG2000 на основе адаптивного квантования коэффициентов трансформации. Приводятся результаты исследований методов повышения качества изображений на основе рекурсивно-сепарабельных фильтров повышения четкости изображений.

Предназначена для специалистов в области цифровой обработки сигналов и изображений, а также научных работников, аспирантов и студентов технических специальностей вузов.

УДК 004.392
ББК 3973.235

ISBN 978-5-7389-3162-8

© Сай С. В., Каменский А. В., Курячий М. И., 2020
© Тихоокеанский государственный университет, 2020

Введение

Большие массивы передаваемых видеоданных, ограничения каналов связи и хранилищ цифровой информации требуют применения эффективных методов компрессии изображений. Для уменьшения объема передаваемых видеоданных применяются алгоритмы сжатия с потерями, что приводит к искажениям и, следовательно, к ухудшению качества оригинальных изображений.

Контроль качества компрессированного изображения достаточно сложен и неоднозначен. В настоящее время существует множество методов и метрик анализа качества изображений. Поиском компромиссных решений между качеством и степенью компрессии заняты исследователи многих стран.

Четкость относится к основным показателям качества изображения и зависит от разрешающей способности видеосистемы, которая традиционно оценивается количеством различимых глазом штриховых линий тестового изображения (таблицы). При этом используются различные типы таблиц, как статические, так и динамические. Увеличение количества пикселей формата видеокadra до стандартов высокой четкости (HD) приводит к повышению четкости изображения, но при этом для обеспечения необходимой скорости передачи требуется применение более эффективных методов компрессии. В процессе компрессии формат кадра не меняется, но при этом (при высоких коэффициентах сжатия) «размываются» или «теряются» мелкие детали, что в сильной степени ухудшает визуальное качество изображения. Ухудшение четкости в свою очередь приводит к снижению резкости границ крупных деталей изображения.

При решении проблемы сохранения четкости изображения возникают два основных вопроса. Какой критерий будет использован для оценки допустимых искажений, при которых ухудшение четкости и резкости будет незаметно или малозаметно для глаза? Какой алгоритм компрессии будет наиболее эффективным при сохранении четкости с допустимыми искажениями по выбранным критериям?

В изданной в 2003 г. монографии С.В. Сая [1] приводится описание объективных методов и алгоритмов оценок качества передачи и воспроизведения мелких деталей цветных телевизионных изображений.

В книге предложены методы анализа четкости, резкости и точности цветопередачи мелких деталей в равноконтрастном цветовом пространстве, что по

мнению автора позволяет приблизить результаты объективных оценок к субъективным зрительным оценкам.

В настоящей работе приводятся результаты авторских исследований за последние годы в области современных методов анализа и повышения качества цифровых изображений.

В первой главе проводится обзор современных цифровых систем и методов передачи изображений. Описываются особенности сжатия фотореалистичных изображений на основе стандартов JPEG и JPEG 2000, приводится краткое описание современных цифровых видеокодеков (MPEG-2, MPEG-4, H.264, H.265/HEVC) и характеристик визуальных искажений, возникающих при компрессии.

Во второй главе приводится описание современных методов оценки качества цифровых изображений: на основе испытательных таблиц; с помощью объективных метрик оценок фотореалистичных изображений; с помощью программно-аппаратных средств оценки качества видеоданных. В п. 2.5 приводится описание метрик оценки качества реалистичных изображений, предложенных одним из авторов (С.В. Сай) в качестве альтернативных существующим метрикам.

В третьей главе приводится описание способов повышения качества цифровых изображений в стандартах JPEG и JPEG 2000 на основе адаптивного квантования коэффициентов трансформации. Рассматриваются особенности корректирующих фильтров для улучшения субъективного качества изображений на выходе видеокодеков. Предлагаются результаты сравнения качества и эффективности сжатия видеофайлов с помощью анализатора VQMT. Приводится описание способа повышения точности фокусировки цифровой камеры.

В четвертой главе приводятся результаты авторских (А.В. Каменский и М.И. Курячий) исследований методов повышения качества четкости изображений на основе цифровых фильтров. Приводится описание: особенностей измерения разрешения по полю изображения; методов и алгоритмов повышения качества изображений на основе цифровой фильтрации; рекурсивно-сепарабельных фильтров повышения четкости изображений; способов повышения качества изображений при их позонной обработке.

1. Системы и методы передачи изображений

1.1. Общие принципы передачи изображений

Новые цифровые технологии в телевидении позволили устранить многие недостатки присущие аналоговым телевизионным системам и значительно повысить качество передачи и воспроизведения изображений. На их основе разработаны эффективные алгоритмы сжатия потоков цифровых сигналов, которые успешно применяются как в системах цифрового ТВ-вещания, так и в системах прикладного телевидения.

Системы прикладного и измерительного телевидения [2] используются в различных областях жизнедеятельности современного общества. С помощью таких систем решаются задачи видеосвязи, телемедицины, дистанционного образования, мультимедиа, обнаружения и распознавания объектов, подводного телевидения, технического зрения и многие другие.

Развитие цифрового телевидения [3] привело к разработке новых стандартов и мультимедийных приложений в области обработки видеоданных. Современные видеокодеки способны конвертировать исходное видео в различные форматы, включая H.264, H.265/HEVC и др. Разрешающая способность видеокадра в новых форматах увеличена вплоть до 4K (4096x2160). Для хранения видеофайлов используются различные контейнеры (*.mp4, *.avi, *.mkv и др.). Очевидно, что увеличение разрешающей способности требует разработки более эффективных методов и алгоритмов компрессии.

В современных методах видеокомпрессии наибольшее применение имеют алгоритмы, основанные на дискретно-косинусном (ДКП) или вейвлет-преобразовании (ДВП). К таким методам относятся алгоритмы сжатия неподвижных изображений, основанные на стандартах JPEG и JPEG 2000. В системах цифрового телевидения H.264 (H.265), реализованных на основе стандарта MPEG-4, эти алгоритмы используются для компрессии опорных (I) кадров видеоряда. Регулировка соотношения качество-компрессия в кодерах обычно выполняется путем изменения параметров квантования трансформированных сигналов изображения в коэффициенты ДКП или ДВП. Поиском компромиссных решений между качеством изображений и степенью компрессии и/или битовой скоростью передачи цифровых видеоданных заняты исследователи многих стран.

Обобщенная структурная схема передачи цифрового изображения в ТВ системе показана на рис. 1.1.

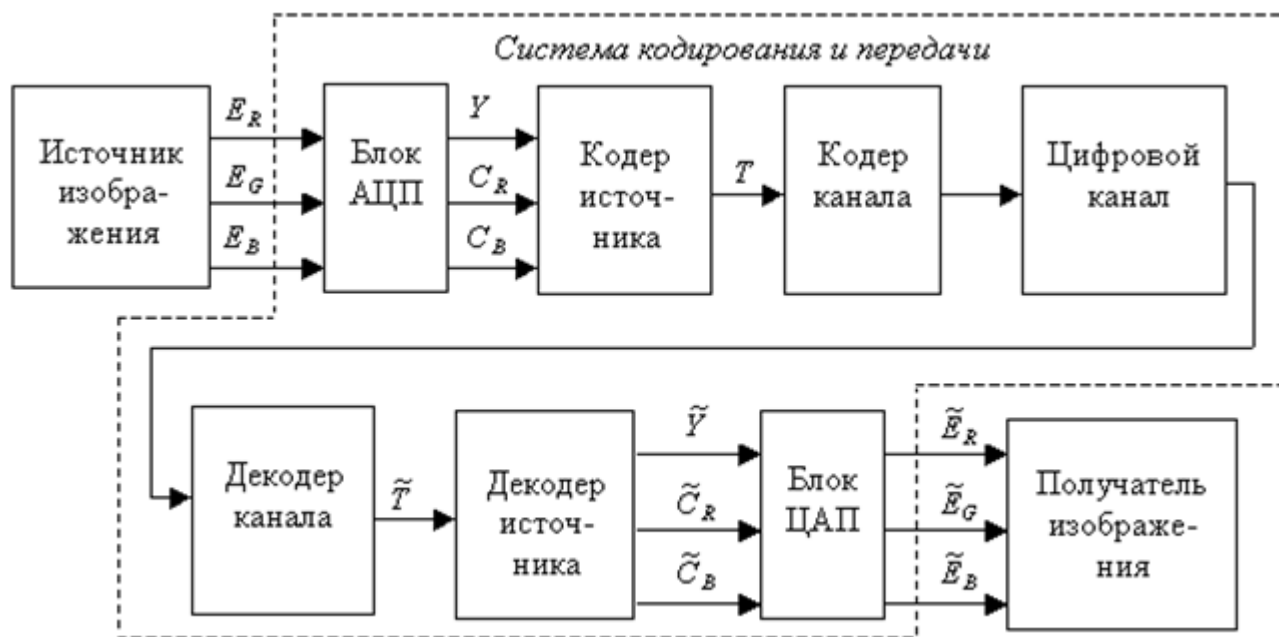


Рис. 1.1. Схема передачи цифрового изображения

В качестве источника изображений используются оптические сенсоры цифровых фото или видеокамер, на выходе которых исходное изображение представлено аналоговыми сигналами красного E_R , зеленого E_G и синего E_B цвета. На выходе блока аналого-цифровых преобразователей (АЦП) формируются цифровые сигналы яркости (Y) и два цветоразностных сигнала (C_R , C_B). В кодере источника выполняется основной алгоритм сжатия видеоданных и формируется транспортный поток T . Кодер канала формирует данные с защитой от ошибок, обеспечивает протокол передачи и реализует метод модуляции. На приемной стороне выполняются обратные преобразования и восстановленные сигналы $\tilde{E}_R, \tilde{E}_G, \tilde{E}_B$ поступают на вход получателя, в качестве которого используется экран ЖК дисплея.

В общем случае, качество декодированного изображения зависит от искажающих факторов всех блоков системы, суммарное влияние которых обычно оценивается по результатам субъективных экспертиз или по выбранному объективному критерию.

Не рассматривая методологию измерений и оценки качественных показателей изображения, определим основные типы и характеристики искажений для каждого блока цифровой видеосистемы.

На выходе источника, изображение уже отличается от его оптического представления (оригинала), что обусловлено неидеальностью апертурно-частотных, переходных, шумовых и др. характеристик преобразователя свет-сигнал. Неидеальность характеристик преобразователя сигнал-свет (ЖК дисплея) вносит дополнительные искажения. Очевидно, что выбор технических характеристик оптического сенсора изображения и дисплея зависит от назначения системы. Например, [4] в цифровых камерах HD используются высокочувствительные матричные ПЗС с числом элементов 1920×1080 , с высокой эффективностью переноса и отношением сигнал/шум не менее 50 дБ. Следовательно, для обеспечения высокой четкости требуется дисплей с таким же разрешением с высокой точностью преобразования сигнал-свет.

Так как, основные искажения вносятся блоками системы кодирования и передачи, полагаем, что искажения, вносимые оконечными устройствами, не превышают допустимых значений и, при этом, датчик и дисплей обеспечивают высокие (эталонные) показатели качества изображения.

Рассмотрим особенности формирования цифровых сигналов. В любой цифровой системе, на первой стадии обработки сигналов изображения используется импульсно-кодовая модуляция (ИКМ), где выполняются следующие операции:

- ограничение спектра;
- дискретизация во времени (или пространстве);
- квантование дискретных отсчетов;
- представление отсчетов в виде бинарных слов.

При обработке изображений оцифровке могут подвергаться непосредственно аналоговые сигналы основных цветов E_R , E_G и E_B , что позволяет сохранять изображения в исходном цифровом формате RAW [5], содержащий необработанные данные об электрических сигналах с фотоматрицы цифрового фотоаппарата, цифровой видеокамеры, а также сканеров неподвижных изображений.

Для повышения эффективности сжатия обычно кодируются не сигналы основных цветов, а их матричные преобразования, которые представлены цифровыми сигналами яркости Y и сигналами цветности U и V [6]:

$$\begin{aligned} Y &= 0,299R + 0,587G + 0,114B; \\ U &= -0,16875R - 0,33125G + 0,5B; \\ V &= 0,5R - 0,41869G - 0,08131B, \end{aligned} \tag{1.1}$$

где R , G и B – цифровые сигналы основных цветов, полученные из аналоговых сигналов E_R , E_G и E_B .

Отметим, что на практике могут использоваться разные варианты преобразования (1.1) по сигналам цветности [7]. Преобразование RGB сигналов в цветное пространство YUV широко применяется в системах цифровой передачи изображений на первых этапах обработки в соответствии с MPEG или JPEG стандартами.

Известно, что глаз человека легче различает градации яркости и оттенки зеленого [8], поэтому сигналы U и V можно отображать с меньшей точностью, что способно существенно уменьшить количество информации, сохраняемой при сжатии. Таким образом, используя особенности человеческого зрения, еще до того, как будет применяться компрессия данных, мы можем получить выигрыш только за счет перехода к другому цветовому пространству.

Для сокращения избыточности по компонентам цветности применяется дискретизация в соответствии с форматами 4:2:2, 4:1:1 или 4:2:0 (рис. 1.2).

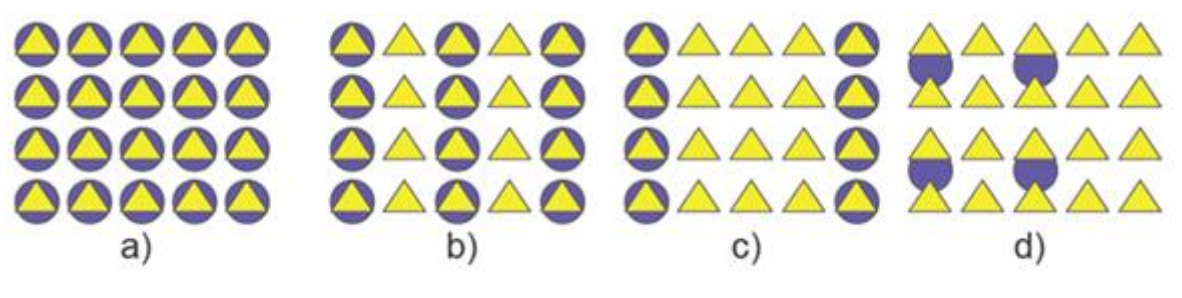


Рис. 1.2. Форматы дискретизация сигналов U и V : а) без дискретизации 4:4:4; б) 4:2:2; в) 4:1:1; г) 4:2:0.

Формат 4:4:4 (рис. 1.2, а) означает, что сигналы яркости и цветности имеют одинаковое количество отсчетов в кадре изображения и, следовательно, обеспечивают одинаковую четкость, как по яркости, так и по цветности. Данный формат практически не используется в вещательном телевидении, т.к. не учитывает свойства зрительного восприятия цветности мелких деталей. Однако может использоваться на первой стадии цифровой обработки для получения форматов с пониженной цветовой четкостью, методом прореживания матриц цветности.

Формат 4:2:2 (рис. 1.2, б) – матрицы цветности в два раза меньше матрицы яркости по горизонтали и равны по вертикали. Каждой паре отсчетов Y по гори-

горизонтальной соответствует по одному отсчету U и V . В данном случае, уже на входе кодера, объем видеоданных сокращается на 50 %.

Формат 4:1:1 (рис. 1.2, в) – матрицы цветности в четыре раза меньше матрицы яркости по горизонтали и равны по вертикали. Данный формат соответствует соотношениям полос частот яркостного и цветоразностных сигналов стандарта SECAM (6.0/1.5/1.5) и практически не используется в цифровом телевидении из-за низкой цветовой четкости и резкости изображения.

Наиболее популярным форматом считается 4:2:0 (рис. 1.2, г), где на каждые четыре пикселя яркости приходится по два (U и V) пикселя цветности. Нетрудно подсчитать, что общий объем кадра изображения уменьшается в два раза.

На выходе системы выполняются обратные преобразования и в блоке цифро-аналоговых преобразователей (ЦАП) восстанавливаются аналоговые значения сигналов основных цветов ($\tilde{E}_R, \tilde{E}_G, \tilde{E}_B$).

Точность преобразований зависит от выбора числа n уровней квантования и частоты дискретизации, а также от формы пространственно-частотных характеристик (ПЧХ) анализирующих и восстанавливающих фильтров.

Представление сигналов конечным числом уровней квантования приводит к искажениям восстановленного изображения в виде шума квантования. При недостаточном числе уровней квантования, шум квантования проявляется в форме случайного шумового компонента яркости и цветности мелких деталей изображения, а также в виде ложных контуров в низкочастотной области.

Для уменьшения заметности шумов квантования, кроме увеличения числа n , используют следующие способы квантования: неравномерное, векторное, блочное и другие. Эти алгоритмы, также как подходы к оценке шума квантования, достаточно широко представлены в литературе [9].

Средний квадрат шума квантования и отношение сигнал/шум для линейного квантователя можно определить по известным формулам

$$\sigma_{кв}^2 = h^2 / 12, \quad \Psi_{кв} \approx n^2,$$

где h шаг квантования.

В цифровом телевидении используются линейные квантователи с 256 уровнями для каждого сигнала, что обеспечивает достаточно высокое качество воспроизводимых изображений и, при этом, шумы квантования не оказывают существенного влияния на точность воспроизведения мелких деталей и резких границ.

Выбор частоты дискретизации сигналов определяется теоремой отсчетов (теоремой Котельникова), согласно которой частота дискретизации должна превышать максимальную частоту спектра исходного сигнала более чем в 2 раза. Согласно теореме отсчетов, точное восстановление аналоговой формы сигналов возможно при идеальных (прямоугольных) частотных характеристиках анализирующего и восстанавливающего фильтров. В реальных системах данное условие выполнить не удастся. Низкочастотная фильтрация приводит к спаду амплитудно-частотных спектров сигналов в области верхних частот, что приводит к снижению резкости и точности передачи мелких деталей. Неидеальность восстанавливающих фильтров добавляет в спектр сигналов паразитные высокочастотные составляющие, которые могут привести к возникновению ложных пространственных гармоник и ложных узоров на изображении.

Наиболее существенное влияние на ухудшение качества изображения оказывают искажения, возникающие в процессе цифрового кодирования и сжатия сигналов источника.

В настоящее время к наиболее передовым алгоритмам цифровой обработки сигналов изображения относятся стандарты JPEG (JPEG 2000), MPEG-2, MPEG-4, используемые для кодирования статических и динамических изображений с различным разрешением. Имеющиеся стандарты определяют лишь поток кодируемых данных и не регламентируют принципы его формирования, следовательно, имеется возможность более эффективной обработки сигналов в рамках этих стандартов с целью обеспечения высокой четкости и резкости изображений.

Рассмотрим особенности цифровой передачи закодированных видеоданных [10]. Для повышения помехоустойчивости в кодере канала выполняются следующие основные операции:

- внешнее кодирование канала с прямым исправлением ошибок, например, с помощью кода Рида-Соломона;
- рандомизация цифрового сигнала;
- введение сигнала вспышки для обеспечения стабильного восстановления несущей на приеме при низком отношении сигнал/шум; внутреннее кодирование канала с прямым исправлением ошибок, например, с помощью решетчатого или сверточного кода.

Для сокращения полосы частот применяют многоуровневую или многопозиционную модуляцию, например, типа QAM (квадратурная АМ) либо QPSK (квадратурная фазовая манипуляция).

В общем случае, качество приема цифрового потока оценивается вероятностью ошибок $P_{ош}$, на значение которой влияют:

- шумы и помехи канала связи;
- эффективность методов помехоустойчивого кодирования;
- погрешности работы приемо-передающих устройств.

В цифровой системе качество доставки закодированных видеоданных высокое, если помехи и искажения не превышают некоторый порог. Но, если порог превышен, и вероятность ошибок превысила допустимое значение, наблюдается резкое ухудшение качества декодированного изображения. При этом тип и характер искажений определяются особенностями алгоритма работы кодера источника.

При анализе искажений и решения метрологических задач целесообразно разделить цифровые системы кодирования на следующие классы:

1. Системы кодирования без потерь (малая степень сжатия).
2. Системы кодирования с потерями (высокая степень сжатия).
3. Системы кодирования статических изображений.
4. Системы кодирования динамических изображений.

В системах кодирования без потерь используются алгоритмы Хаффмана, Лемпела-Зива-Уолша (LZW), арифметическое сжатие, RLE – кодирование и др. Само название «без потерь» говорит о том, что данные методы не вносят искажений в восстановленное изображение при высокой помехоустойчивости канала передачи. Основой таких методов является положение о том, что изображение обладает статистической избыточностью, которая может быть устранена без потери информации. Отметим, что эти алгоритмы могут использоваться в системах кодирования с потерями для сжатия трансформированных сигналов изображения в пространственно-частотную область.

Алгоритмы цифрового сжатия с потерями основаны на том, что цветное изображение обладает визуальной (субъективной) избыточностью, которую можно устранить с частичной потерей данных, без заметного для глаза ухудшения качества воспроизводимого изображения.

К основным методам цифрового сжатия с потерями относятся [11]:

- кодирование на основе преобразований (JPEG и JPEG 2000),
- межкадровое кодирование с компенсацией движения (MPEG).

Кроме основных методов, в литературных источниках [12] приводится описание других алгоритмов, таких как: гибридное кодирование; фрактальное сжатие и др.

Каждый из алгоритмов цифрового сжатия с потерями вносит специфические искажения в декодированное изображение. Особенности выполнения основных алгоритмов и характеристики искажений приведены в следующих разделах первой главы.

1.2. Особенности сжатия изображений на основе JPEG

Кодирование на основе преобразования представляет собой косвенный метод. Исходное изображение подвергается унитарному математическому преобразованию, далее полученные коэффициенты преобразования (трансформанты) кодируются для передачи по каналу связи или для записи на промежуточный носитель. Сжатие визуальных данных основано на том, что для большинства изображений естественного происхождения значения многих коэффициентов преобразования сравнительно малы. Такие коэффициенты отбрасываются, или на их кодирование отводится малое число двоичных разрядов.

Рассмотрим основные процедуры кодирования [13].

Кадр изображения обычно разбивается на блоки (например, размером 8×8 элементов) и подвергается унитарному преобразованию поблочно. В общем случае для блока яркости изображения коэффициенты двумерного разделимого унитарного преобразования определяются как

$$Y^T(k, l) = \sum_{m=1}^N \sum_{n=1}^N Y(m, n) A_c(m, k) A_r(n, l), \quad (1.2)$$

где A_r и A_c – ядра преобразования вдоль строк и столбцов; $N \times N$ – число элементов в блоке. Цветоразностные сигналы U и V преобразуются аналогично выражению (1.2) поблочно и последовательно, после обработки четырех или двух (значение зависит от коэффициента прореживания) блоков яркости. Сформированный таким образом “макроблок” трансформант, состоящий из четырех блоков коэффициентов преобразования яркости и двух блоков коэффициентов цветности, далее подвергается кодированию и нормализации.

На этапе кодирования производится отбор трансформант зональным или пороговым методами. При зональном отборе выделяются совокупности коэффициентов, занимающие некоторые заранее фиксированные области спектра трансформант, обычно соответствующие низкочастотным составляющим. Число уровней квантования берется пропорционально ожидаемой дисперсии соответствующего коэффициента. Область передаваемых трансформант может иметь различную конфигурацию, например, прямоугольную, эллиптическую

или треугольную. Если следовать критерию среднеквадратичной ошибки воспроизведения, то, согласно теоретическим и экспериментальным исследованиям, в передаваемую область должны попадать те коэффициенты, которые имеют наибольшую дисперсию.

Метод порогового отбора сохраняет лишь те коэффициенты, величина которых превышает заданный порог. При этом для отобранных трансформант устанавливается единая шкала квантования и остальным коэффициентам (значения которых меньше порогового) присваивается ноль. В этом случае, при передаче значащих коэффициентов, необходимо указывать их местоположение.

Следует отметить, что в настоящее время большое внимание уделяется адаптивным методам кодирования, которые позволяют повысить эффективность цифрового сжатия с сохранением высокого качества изображения. В адаптивных методах перестраивают шкалу квантования в соответствии с локальными характеристиками различных частей передаваемого изображения. Поскольку преобразование изображения производится поблочно, квантование также реализуется поблочно посредством перехода с одной матрицы на другую.

После квантования, трансформанты каждого блока кодируются модифицированным методом длин серий (RLE), при этом часто используется сравнение элементов блока по “зигзагообразной” траектории.

В процессе сжатия обычно используется алгоритм Хаффмана, где наиболее вероятным значениям трансформант присваиваются короткие кодовые слова и наоборот. При выполнении алгоритма формируется кодовая таблица распределения указывающая число бит присвоенных трансформантам блока, которая хранится как на передающей, так и на приемной сторонах.

В настоящее время существует большое число преобразований, которые можно использовать в системе кодирования, например, преобразования Фурье, Адамара, Карунена–Лоэва, Хаара и др. Преобразование Карунена–Лоэва обеспечивает минимальную среднеквадратичную ошибку, но требует знания статистических характеристик передаваемого изображения, также для данного преобразования не существует быстрого вычислительного алгоритма.

В аппаратуре цифрового сжатия наиболее часто используется дискретное косинусное преобразование (ДКП), имеющее быстрый вычислительный алгоритм и приближающееся по эффективности к преобразованию Карунена–Лоэва применительно к изображениям, для которых подходит статистическая модель Гаусса–Маркова. Следует отметить, что в основу стандарта JPEG [14] заложен именно данный тип преобразования.

После разбиения изображения на блоки выполняется дискретно-косинусное преобразование. Прямое и обратное DCT для блоков $N \times N$ имеет следующий вид:

$$\begin{aligned} Y_{k,l}^T &= \frac{C_{k,l}}{\sqrt{2N}} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} Y_{m,n} \cos \frac{(2m+1)k\pi}{2N} \cos \frac{(2n+1)l\pi}{2N}; \\ Y_{m,n} &= \frac{1}{\sqrt{2N}} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} C_{k,l} Y_{k,l}^T \cos \frac{(2m+1)k\pi}{2N} \cos \frac{(2n+1)l\pi}{2N} \end{aligned} \quad (1.3)$$

где $C_{k,l} = \begin{cases} \frac{1}{\sqrt{2}} \text{ if } k=0, l=0 \\ 1.0 \text{ if } k \neq 0, l \neq 0 \end{cases}$,

$Y_{k,l}^T$ – коэффициент трансформации сигнала яркости. Аналогично выполняются преобразования по сигналам цветности V и U . Коэффициенты трансформации представляют собой коэффициенты гармоник пространственно-частотного спектра каждого блока исходного изображения.

Важное свойство ДКП – для большинства изображений основная часть энергии представляется малым числом коэффициентов трансформации.

Вычисления (1.3) удобнее выполнять в матричном виде, что ускоряет процесс вычислений алгоритмов кодирования с преобразованием.

Таким образом, кодирование цветного изображения на основе преобразования можно представить в виде следующих этапов [15]:

1. Преобразование RGB сигналов в YUV .
2. Преобразование сигналов цветности в формат 4:2:0 или др.
3. Разбиение кадра изображения на блоки яркости и цветности – N на N элементов и выполнение преобразования над каждым блоком яркости и цветности.
4. Квантование и пороговый отбор коэффициентов трансформации.
5. Зигзаг сканирование.
5. Сжатие коэффициентов преобразования для каждого блока с помощью метода RLE + Хаффмана.

Этапы выполнения кодирования и сжатия изображения на основе стандарта JPEG показаны на рис. 1.3.

Алгоритмы выполнения первого и второго этапов рассмотрены выше.

Прямое вычисление ДКП (1.3) требует больших вычислительных затрат, поэтому на практике используется алгоритм быстрого преобразования Фурье.



Рис. 1.3. Этапы выполнения кодирования изображения на основе стандарта JPEG

Также выполнение преобразований часто реализуется с помощью последовательности матричных перемножений. При этом прямое и обратное дискретное косинусное преобразование ДКП, описываются следующими выражениями:

$$\begin{aligned} Y^T &= M_{\text{ДКП}} * Y * M_{\text{ДКП}}^T; \\ Y &= M_{\text{ДКП}}^T * Y^T * M_{\text{ДКП}}, \end{aligned} \quad (1.4)$$

где Y – блок 8×8 элементов яркости или цветности исходного изображения; Y^T – блок коэффициентов трансформации; M – прямая матрица преобразования; M^T – транспонированная матрица.

На рис. 1.4. показана прямая матрица преобразования, используемая в вычислениях (1.4). Преобразование (1.3) является обратимым, т.е. при восстановлении изображения искажения в блоках яркости будут отсутствовать.

$$\begin{pmatrix} .353553 & .353553 & .353553 & .353553 & .353553 & .353553 & .353553 & .353553, \\ .490393 & .415735 & .277785 & .097545 & -.097545 & -.277785 & -.415735 & -.490393, \\ .461940 & .191342 & -.191342 & -.461940 & -.461940 & -.191342 & .191342 & .461940, \\ .415735 & -.097545 & -.490393 & -.277785 & .277785 & .490393 & .097545 & -.415735, \\ .353553 & -.353553 & -.353553 & .353553 & .353553 & -.353553 & -.353553 & .353553, \\ .277785 & -.490393 & .097545 & .415735 & -.415735 & -.097545 & .490393 & -.277785, \\ .191342 & -.461940 & .461940 & -.191342 & -.191342 & .461940 & -.461940 & .191342, \\ .097545 & -.277785 & .415735 & -.490393 & .490393 & -.415735 & .277785 & -.097545 \end{pmatrix}$$

Рис. 1.4. Матрица косинус-преобразования

Однако это условие выполняется при представлении результатов преобразования вещественными числами с плавающей запятой или с фиксированной точкой. Для дальнейшей обработки коэффициенты трансформации Y^T округляются до целых значений с фиксацией определенного количества разрядов после запятой.

По стандарту JPEG считается, что восстановление выполняется без потерь при 12-ти битном представлении значений Y^T .

Таким образом после выполнения ДКП размер блока яркости 8×8 байт увеличится в 1,5 раза. Однако дискретно-косинусное преобразование позволяет выполнить декорреляцию блока, что приводит к появлению в блоке коэффициентов трансформации малозначащих или нулевых коэффициентов. Данное обстоятельство является предпосылкой для использования эффективного сжатия.

Например, если все значения яркости исходного блока равны, что соответствует фоновым участкам изображения, то после ДКП получим всего один ненулевой коэффициент $Y_{0,0}^T$. Остальные коэффициенты в соответствии с выражением (1.3) будут равны нулю. Если использовать сжатие на основе модифицированного RLE, то такой блок сжимается в 25 раз.

На следующем этапе выполняется квантование коэффициентов ДКП и обнуление малозначащих коэффициентов. Квантование коэффициентов можно выполнить с помощью матрицы квантования M :

$$\tilde{Y}_{k,l}^T = \text{round} \left[\frac{Y_{k,l}^T}{M_{k,l}} \right] Q, \quad (1.5)$$

где Q – параметр качества, который равен единице для индексов $k = 0$ и $l = 0$; round – операция округления до целого значения. Матрицы квантования рекомендованы стандартом отдельно для блоков яркости и блоков цветности с размерами 8×8 пикселей.

Типовые матрицы квантования [16], используемые на практике, показаны на рис. 1.5. С увеличением номера k и l значения чисел в матрицах квантования увеличиваются, что обусловлено спадом контрастно-частотной характеристики зрения в области верхних пространственных частот. Для компонент цветности используется более грубое квантование, что обусловлено меньшей чувствительностью зрения к искажениям по цветности.

Значения коэффициентов матрицы квантования по яркости и по цветности можно получить по упрощенным формулам:

$$M_{k,l}^Y = \begin{cases} 16 & \text{if } k=0, l=0 \\ 1+(1+k+l) & \text{if } k \neq 0, l \neq 0 \end{cases}, \quad (1.6)$$

$$M_{k,l}^{UV} = \begin{cases} 16 & \text{if } k=0, l=0 \\ 1+C_N(1+k+l) & \text{if } k \neq 0, l \neq 0 \end{cases}$$

где коэффициент $C_N > 1$ учитывает менее высокую чувствительность зрения к искажениям коэффициентов трансформации блоков цветности в области высоких пространственных частот.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

а)

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

б)

Рис. 1.5. Матрицы квантования для блоков яркости (а) и цветности (б)

В процессе квантования происходит обнуление малозначащих коэффициентов, что позволяет выполнить эффективное сжатие на следующем этапе. Однако этот процесс вносит основные искажения в декодированное изображение, обусловленные потерей малозначащих коэффициентов и округлением до целых значений остальных коэффициентов.

Параметр качества Q позволяет регулировать эти потери и, следовательно, задавать требуемое качество изображения.

На последних этапах (см. рис. 1.3) выполняется сжатие коэффициентов DST на основе энтропийного кодирования. Для повышения эффективности сжатия обычно выполняется преобразование двумерного массива блока в одномерный с применением операции зигзаг-сканирования.

1.3. Особенности сжатия изображений на основе JPEG 2000

Рассмотрим особенности алгоритма сжатия изображений в стандарте JPEG 2000. В кодере выполняются следующие основные процедуры: а) преобразование цифровых RGB сигналов исходного файла изображения в сигналы YVU ; б) выполнение вейвлет-преобразования для каждой из YVU компонент изображения; в) квантование коэффициентов ДВП и обнуление малозначащих коэффициентов; г) кодирование битовых плоскостей и арифметическое сжатие.

Преобразование сигналов основных цветов в цифровые сигналы яркости (Y) и цветности (U , V) для вещественного режима выполняется аналогично (1.1).

Дискретное вейвлет-преобразование одномерного сигнала S_n длиной N можно выполнить по следующим формулам [17]:

$$S_{j,k}^H = \sum_{n=0}^{N-1} S_{j-1,n} h_{n-2k}; S_{j,k}^G = \sum_{n=0}^{N-1} S_{j-1,n} g_{n-2k}, \quad (1.7)$$

где n – номера отсчетов исходного дискретного сигнала; k – номера коэффициентов ДВП преобразованного сигнала в диапазоне от 0 до $(N-1)/2$; j – номер итерации, изменяющийся от 1 до заданного значения J ; h_n – низкочастотные коэффициенты и g_n – высокочастотные коэффициенты импульсной характеристики симметричного биортогонального фильтра. В стандарте JPEG 2000 рекомендовано [18] использование симметричного биортогонального фильтра Козна-Добеши-Фово (CDF 9/7).

Из преобразования (1.7) следует, что коэффициенты ДВП S^H представляют собой низкочастотные составляющие, а S^G высокочастотные составляющие исходного дискретного сигнала. При этом за счет сдвига импульсной характеристики на $2k$, в результате преобразования получаем две составляющие длиной $N/2$. Таким образом, в результате первой итерации ($j = 1$) имеем масштабированную (уменьшенную) в два раза копию (S^H) исходного сигнала и его дополнение (S^G) в виде высокочастотной составляющей. На следующей итерации ($j =$

2) процесс повторяется, где исходным сигналом будет низкочастотная составляющая, полученная после первой итерации. После выполнения всех итераций получим масштабированную (уменьшенную) копию исходного сигнала в 2^J раз и J дополнительных высокочастотных составляющих.

Обратное преобразование и восстановление исходного сигнала выполняется итерационно по следующей формуле:

$$S_{j-1,n} = \sum_{k=0}^{(N/2j)-1} S_{j,k}^H h_{n+2k} + \sum_{k=0}^{(N/2j)-1} S_{j,k}^G g_{n+2k}, \quad (1.8)$$

где j изменяется от максимального значения (J) до единицы.

Вычисление ДВП на очередном шаге можно наглядно представить в виде схемы двухполосного банка фильтров [19], изображенного на рис. 1.6.

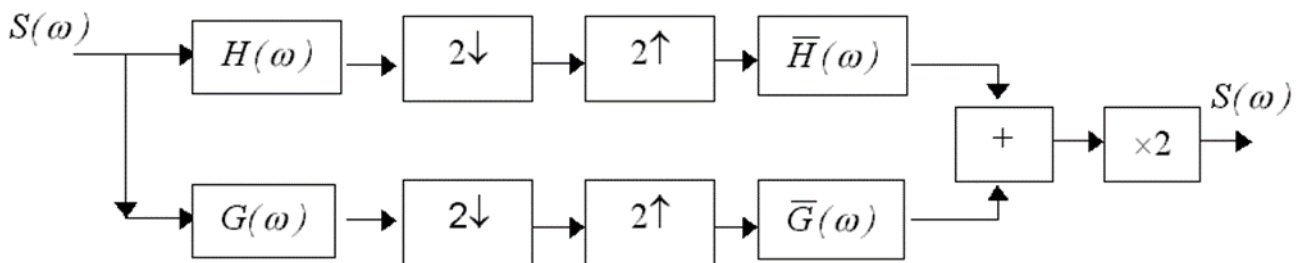


Рис. 1.6. Эквивалентная схема ДВП

Частотные характеристики $G(\omega)$ и $H(\omega)$ соответствуют фильтрующим последовательностям g_k и h_k , а $\bar{G}(\omega)$ и $\bar{H}(\omega)$ соответствуют g_{-n} и h_{-n} . Между фильтрами прямого преобразования (анализа) и обратного (синтеза) вставлены дециматор и интерполятор. Дециматор удаляет избыточные компоненты, возникшие после фильтрации, а интерполятор восстанавливает первоначальную битовую скорость.

Преимуществом представления вейвлет-преобразования в виде блока фильтров – схемы анализа-синтеза – является возможность наглядного анализа иерархически построенных преобразований. Если подобная система обладает свойством полного восстановления, то промежуточный сигнал может подаваться на вход другой аналогичной системы, в результате чего получается каскадно-соединенная система.

Если дальнейшему разложению подвергается каждый промежуточный сигнал, субдиапазоны получают равной ширины, и мы имеем равномерную

систему. В противном случае система – неравномерная. Октавополосное разбиение частотного плана выделяется среди других способов тем, что позволяет более адекватно отразить особенности человеческого зрения – увеличения чувствительности к изменениям с уменьшением частоты.

Если фильтры анализа и синтеза различаются при сохранении условия полного восстановления, то мы имеем биортогональное преобразование. Биортогональное преобразование позволяет использовать симметричные последовательности h_k и g_k длины большей двух, что особенно важно для обработки реальных изображений, так как возможно симметричное продолжение сигнала для вычисления по формулам (1.7) – (1.8) и, соответственно, гладкая склейка на границах.

Для двумерного сигнала изображения вейвлет-преобразование выполняется следующим образом (рис. 1.7). На первой итерации, преобразование (1.7) выполняется последовательно по строкам и столбцам, после чего формируются четыре матрицы (НН, GH, HG, GG) коэффициентов ДВП с размерами в два раза меньшими ширины и высоты изображения. Принятые обозначения матриц следующие: НН – низкочастотная фильтрация по строкам и столбцам; GH – низкочастотная фильтрация по строкам и высокочастотная по столбцам; HG – высокочастотная фильтрация по строкам и низкочастотная по столбцам; GG – высокочастотная фильтрация по строкам и столбцам. На второй итерации выполняется аналогичное преобразование для НН матрицы и т.д. Обычно достаточно 4-6 итераций. Такой процесс преобразований эквивалентен операции фильтрации сигналов изображения по частотным субдиапазонам для каждой итерации j .

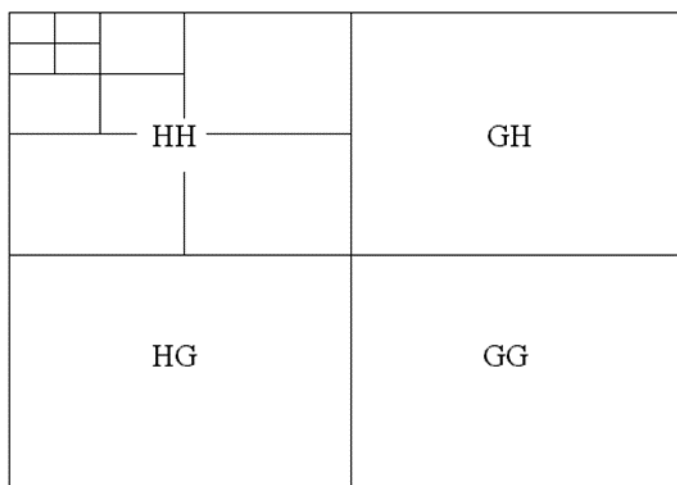


Рис. 1.7. Расположение коэффициентов ДВП двумерного сигнала

После выполнения преобразования коэффициенты ДВП квантуются и преобразуются в целочисленные значения, где для каждого субдиапазона итерации j выбирается свой постоянный коэффициент Δ (шаг квантования), на который делятся все элементы матрицы.

Процесс квантования коэффициентов ДВП для каждой итерации (j) можно описать следующими формулами:

$$\widehat{S}_j^{GG} = \left\lfloor \frac{S_j^{GG}}{\Delta_j^{GG}} \right\rfloor; \quad \widehat{S}_j^{GH} = \left\lfloor \frac{S_j^{GH}}{\Delta_j^{GH}} \right\rfloor; \quad \widehat{S}_j^{HG} = \left\lfloor \frac{S_j^{HG}}{\Delta_j^{HG}} \right\rfloor, \quad (1.9)$$

где S_j – исходное значение коэффициента; $\widehat{S}_j$ – квантованное значение. Операция $\lfloor \dots \rfloor$ означает округление вниз до ближайшего целого. Исключение составляет отрезок «мертвой зоны» около нуля, в которой квантованные вещественные коэффициенты обнуляются с помощью условия:

$$\text{if } |S/\Delta| < 1; S=0; \text{ else } S=S.$$

Отметим, что для самого низкочастотного субдиапазона шаг квантования выбирается равным единице ($\Delta_j^{LL} = 1$). Значения коэффициентов Δ_j^{HH} , Δ_j^{LH} и Δ_j^{HL} образуют шкалу квантования.

Коэффициенты квантования рекомендованы стандартом и имеют более высокие значения для матриц НН, что обусловлено спадом пространственно-частотной характеристики зрения в области верхних частот.

В стандарте рекомендовано оптимальное соотношение параметров квантования для разных субдиапазонов:

$$\frac{\Delta_b}{\Delta_0} = \sqrt{\gamma_b}, \quad (1.10)$$

где величина γ_b оценивается среднеквадратичной величиной искажения, которое вносится в изображение при единичном искажении коэффициента из субдиапазона b . Эта величина может быть определена на основании анализа вейвлет-преобразования «CDF 9/7».

Из таблицы значений соотношения шагов квантования (5) представленной в работе [20] следует, что соотношение шагов квантования для соседних субдиапазонов приблизительно равно двум. В табл. 1.1 в качестве примера приведена шкала квантования для пяти итераций при заданном шаге для высокочастотного диапазона: $\Delta_1^{HH} = 32$.

Процесс квантования и обнуления коэффициентов ДВП позволяет эффективно сжимать изображения с помощью алгоритма арифметического сжатия или на основе других алгоритмов. В тоже время, этот процесс вносит основные потери качества в декодированное изображение.

Таблица 1.1. Шкала квантования коэффициентов ДВП

j	Δ_j^{HH}	Δ_j^{LH}	Δ_j^{HL}
1	32	16	16
2	16	8	8
3	8	4	4
4	4	2	2
5	2	1	1

Изменение коэффициентов квантования с помощью параметра качества Q позволяет регулировать качество изображения и, следовательно, скорость цифрового потока сжатых визуальных данных.

На последнем этапе выполняется арифметическое сжатие квантованных коэффициентов ДВП. С этой целью матрица коэффициентов разбивается на одинаковые блоки (тайлы), по умолчанию размером 64×64 . Далее каждый блок кодируется независимо от других.

В процессе кодирования коэффициенты в блоке представляются в виде битовых плоскостей. Одну из таких плоскостей составляют знаки коэффициентов. Остальные плоскости соответствуют различным разрядам величин коэффициентов (положение бита в плоскости соответствует положению коэффициента в блоке). При этом сначала кодируется плоскость, соответствующая старшему разряду коэффициентов, затем следующая по убыванию, и т.д. Кодирование одной битовой плоскости одного блока осуществляется в три этапа: кодирование старших бит; уточняющий проход; очищающий проход. Для каждого прохода используется бинарное адаптивное арифметическое кодирование и контекстное моделирование.

Основные отличия JPEG от JPEG 2000 заключаются: в выборе преобразования ДВП вместо ДКП, отсутствия блочного разбиения кадра изображения и в стратегии квантования коэффициентов трансформации. Эти отличия приводят к различным характеристикам искажений цифровых изображений. Алгоритмы сжатия без потерь стандартами жестко не регламентированы, поэтому у разработчиков имеется свобода выбора.

1.4. Цифровые видеокодеки

1.4.1. Стандарт MPEG-2

Формат видеокодека, обычно, не определяет все алгоритмы, которые будет использовать видео кодек. К примеру, часто в видео кодировании для сжатия выходного сигнала используется алгоритм нахождения различий и совпадений между предыдущим и следующим блоком изображения (block-matching). После этого создаются ссылки на уже закодированные большие фрагменты (macroblocks) в новом фрагменте с добавлением различий между предыдущим макроблоком и текущим блоком. Нахождение оптимальной комбинации главных блоков изображения и блоков-различия – это NP-сложная задача, означающая, что невозможно найти оптимально-быстрое решение данной проблемы. Однако, это вовсе не означает, что результат не может быть получен. Хотя кодек обязан поддерживать компрессию, указанную в спецификации для выходного файла, но в выборе алгоритма поиска и других алгоритмов кодирования кодек свободен от следования каким-либо спецификацией. Более того, в спецификациях чаще всего нет указаний в выборе алгоритма или такой выбор помечен как рекомендуемое поведение.

Единственное, что обычно требуется от кодека это достигнуть определенного формата файла и определенного уровня компрессии. Для примера можно взять спецификацию популярного кодека H.264 в котором указано, что кодек не обременен какими-либо ограничениями на реализации алгоритмов кодирования и алгоритм кодирования не является частью спецификации [21].

Это позволяет более гибко выбирать размер выходного файла или потока с учетом времени потраченного на кодирование/декодирование. Например, для потокового видео с минимальной задержкой необходимо выбирать алгоритмы быстрого декодирования использующий более грубые оценки, при этом стоит ожидать увеличение выходного размера файла.

Спецификация формата видео кодирования может дополнительно определять уровни и профили формата видео кодирования. Уровни профиля налагают ограничения на разные параметры выходного файла/потока.

Так, уровень накладывает на формат ограничения в максимальном разрешении и максимальном и, иногда, минимальном размере потока, выдаваемом за секунду. В то же время профиль ограничивает кодер в техниках оптимизации. Вместе эти параметры позволяют создавать оборудование и программы в раз-

ном ценовом и качественном диапазоне. А также, косвенно, регулировать потребление ресурсов декодером и кодером.

Рассмотрим особенности стандарта MPEG-2, который является основой для других стандартов семейства MPEG и им подобным [21].

Поток видеоданных, определяемый спецификацией 13818-2, представляет собой иерархическую структуру, элементы которой строятся и объединяются друг с другом в соответствии с определенными синтаксическими и семантическими правилами. Существует 6 типов элементов этой иерархической структуры:

- Видеопоследовательность;
- Группа изображений;
- Изображение;
- Срез;
- Макроблок;
- Блок.

Видеопоследовательность – элемент потока видеоданных высшего уровня. Она представляет собой серию последовательных кадров телевизионного изображения. MPEG-2 допускает как построчные, так и чересстрочные последовательности. Чересстрочная последовательность – это серия телевизионных полей. В процессе компрессии поля могут кодироваться отдельно. Это дает изображения типа «поле». Два поля, кодируемые как телевизионный кадр, образуют изображение типа «кадр». В одной чересстрочной последовательности могут использоваться и изображения-поля, и изображения-кадры. В последовательностях с построчным разложением каждое изображение представляет собой кадр. В соответствии с используемыми методами дифференциального кодирования различают три типа изображений: I, P и B. Изображение типа I кодируется с использованием только той информации, которая содержится в нем самом (I - Intra-coded picture). В нем устраняется только пространственная избыточность. При кодировании P и B изображений используется межкадровое кодирование. При кодировании изображения типа P формируется разность между исходным изображением и предсказанием, полученным на основе предшествующего или последующего изображения типа I (P – Predictive-coded picture). Изображение типа B – это изображение, при кодировании которого используется предсказание, сформированное на основе предшествующего и последующего изображений типа I или P (B – Bidirectionally-predicted-coded picture). В изображениях типа P и B устраняется и пространственная, и временная избыточность. Серия

изображений, содержащих одно I-изображение, называется группой изображений. Чем больше группа изображений, тем большая степень компрессии может быть достигнута.

С информационной точки зрения каждое изображение представляет собой три прямоугольных матрицы отсчетов изображения: яркостную Y и две матрицы цветности U и V . Стандарт MPEG-2 допускает различные структуры матриц. Соотношение между количеством отсчетов яркости и цветности определяется форматом дискретизации (4:2:0, 4:2:2 или 4:4:4).

Каждое изображение делится на срезы, которые состоят из макроблоков. Макроблок складывается из блоков размером 8×8 пикселей. Каждый макроблок содержит группу из 4 блоков с отсчетами яркости (из области изображения с размерами 16×16 пикселей) и группу блоков с отсчетами цветности, взятых из той же области изображения, что и отсчеты блоков яркости. Число блоков с отсчетами цветности зависит от формата дискретизации: по одному блоку U и V в формате 4:2:0, по два – в формате 4:2:2, по 4 – в формате 4:4:4 (см. рис 1.2).

Все структурные элементы потока видеоданных, полученного в результате внутрикадрового и межкадрового кодирования (кроме макроблока и блока), дополняются специальными и уникальными стартовыми кодами. Каждый элемент содержит заголовок, за которым следуют данные элементов более низкого уровня. В заголовке видеопоследовательности (как элемента высшего уровня) приводится разнообразная дополнительная информация, например, размеры и соотношение сторон изображения, частота кадров, скорость потока данных, матрица квантования, формат дискретизации цветности изображения, координаты основных цветов и белого цвета, параметры матрицы для формирования яркостного и цветоразностных сигналов, параметры передаточной характеристики (гамма).

Принципы видеокомпрессии.

Сокращение пространственной избыточности выполняется в изображениях типа I и достигается на уровне блока. Набор операций такого кодирования аналогичен стандарту JPEG – дискретное косинусное преобразование; взвешенное квантование, определяемое элементами матрицы квантования; энтропийное кодирование серии коэффициентов косинусного преобразования, полученной в результате зигзаг-сканирования матрицы коэффициентов (см. рис. 1.3).

Метод сокращения временной избыточности – дифференциальная импульсно-кодовая модуляция с компенсацией движения при кодировании изображений типа P и B. При кодировании формируется разность между исходным

изображением и предсказанием, полученным на основе предшествующего и/или последующего изображения. Полученная разность подвергается дискретному косинусному преобразованию, взвешенному квантованию и энтропийному кодированию. Для повышения точности предсказания и, тем самым, сокращения объема необходимых для представления изображения данных, используется компенсация движения. Оценивается скорость перемещения движущихся объектов от кадра к кадру и при определении предсказания производится соответствующая коррекция в положении опорного изображения, по отношению к которому находится ошибка предсказания. Определение величины и направления смещения движущихся объектов от кадра к кадру, называемого вектором движения, производится на уровне макроблока. На рис. 1.8 показана упрощенная структура кодера MPEG.

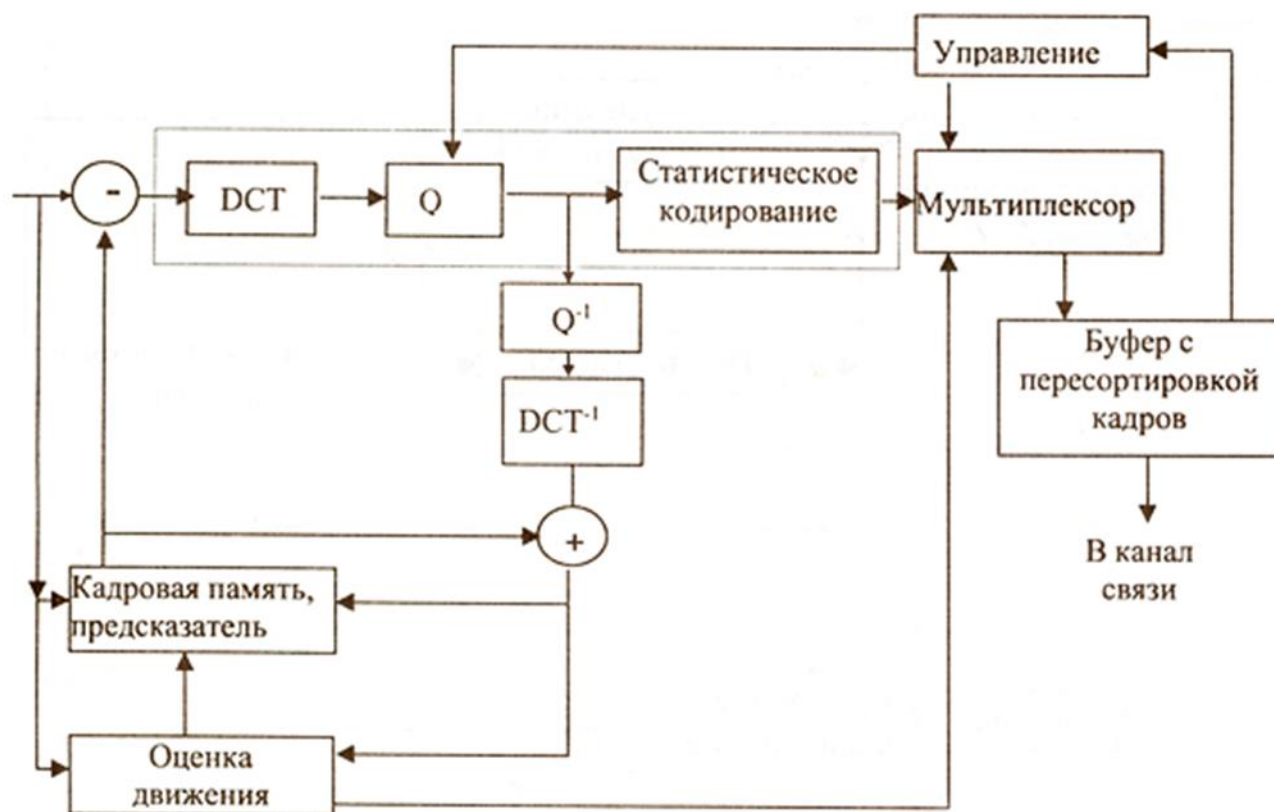


Рис. 1.8. Структурная схема кодера MPEG

Кодер работает следующим образом. Видеокадр разбивается на макроблоки размером 16×16 пикселей, каждый из которых содержит по 4 блока яркости размером 8×8 пикселей и по одному блоку отсчетов сигналов U и V . При кодировании блоков в составе I-кадра вычисляются коэффициенты ДКП, затем они квантуются с использованием таблицы квантования, имеющейся в памяти и ко-

дера, и декодера и применяемой «по умолчанию» (рис. 1.9). Кодер может изменить элементы таблицы, при этом он сообщает об изменении декодеру и посылает новую таблицу.

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Рис. 1.9. Таблица квантования блоков ДКП для I – кадров

Если кодируется макроблок из состава предсказанного кадра, ищется сопряженный блок из опорного кадра, его значения поэлементно вычитаются из элементов кодируемого блока, и уже для этих разностей вычисляются коэффициенты ДКП и производится их квантование. Таблица квантования для макроблоков из Р- и В-кадров, применяемая по умолчанию, содержит число 16 во всех позициях, но также может быть изменена кодером. Отметим, что при поиске сопряженных блоков используются только отсчеты яркости. Полученные при этом векторы перемещения применяются и при кодировании блоков отсчетов цветоразностных сигналов.

Важным звеном алгоритма кодирования остается обеспечение постоянства выходной скорости цифрового потока. Отсчеты сигнала от источника поступают с постоянной скоростью, однако на выходе квантователя скорость поступления битов может варьироваться в широких пределах. Она зависит от типа видеокadra (I-кадр требует значительно больше битов, чем Р- и В-кадры), от его содержания («спокойный» кадр с однородными участками требует меньше битов, чем кадр с выраженной мелкозернистой структурой или/и насыщенный динамическими объектами). Попытка кодировать все кадры одинаковым числом битов приведет к изменению качества изображения от кадра к кадру.

Единственным средством выравнивания скорости потока является применение буферной памяти в кодере и декодере. Буфер может заполняться со сто-

роны квантователя неравными порциями, а опустошаться в сторону канала равномерно, обеспечивая постоянную скорость цифрового потока. Но и здесь не все так просто. Быстрая смена сюжетов и связанное с этим увеличение доли I-кадров вскоре приведет к переполнению буфера, а длинный сюжет без движения — к его опустошению. Простое увеличение размера буфера вызовет возрастание задержки сигнала в кодере, что в некоторых случаях мешает телезрителям (беседа двух теле ведущих в разных студиях).

Для сохранения приемлемого качества изображения при постоянной скорости потока необходимо, во-первых, регулировать шкалу квантования и, во-вторых, адаптивно распределять имеющийся ресурс битов между разными типами кадров с учетом их сложности. Эти задачи решаются в кодере специальным устройством — контроллером битов. Адаптивное управление процессом квантования осуществляется умножением всех элементов матрицы на масштабирующий множитель величиной от 8 до $1/4$, общий для всех отсчетов макроблока, который также сообщается декодеру. Величина множителя больше 1 означает увеличение доли битов, выделяемой данному макроблоку или кадру, меньше 1 — соответствующее уменьшение.

Перераспределение ресурса между I-, P- и B-кадрами производится кодером на основе оценки соотношения битов в предыдущих кадрах. Некоторые кодеры осуществляют кодирование «в два прохода», оценивая на первом проходе сложность видеокadra, выделяя ему на основе этой оценки определенный ресурс битов и уже на втором проходе кодируя отсчеты с учетом выделенного ресурса.

Основные принципы межкадрового кодирования сигналов изображения на базе стандарта MPEG-2 достаточно подробно представлены в литературных источниках.

Оценка движения основана на методе согласования блоков. Самым точным считается корреляционный метод. Из текущего m -го кадра выбирается макроблок яркости размером 16×16 элементов. Затем макроблок перемещают по предыдущему $(m-1)$ -му опорному кадру, вычисляя корреляцию между блоком и покрываемой им площадью предыдущего кадра. Максимум корреляции указывает на правильно найденное смещение, и положение выбранного блока в m -ом кадре определяет вектор движения.

Существует ряд различных подходов к определению вектора движения, в которых основное внимание уделяется проблеме выбора эффективного алгоритма анализа движения, обладающего малым числом операций на пиксел

и достаточной степенью исключения межкадровой избыточности, необходимой для эффективного сжатия динамических видеопоследовательностей.

Наиболее простым и универсальным является алгоритм поиска векторов движения макроблоков на основе полного перебора. Данный метод может использоваться в качестве эталонного для оценки качества других методов. Недостатком метода является низкое быстродействие. С целью упрощения вычислений вместо оценки межкадровой корреляции используется мера среднего абсолютного различия сигналов яркости макроблока в текущем кадре и участка опорного кадра со сдвигом на вектор движения $\vec{V} = (V_i, V_j)$:

$$H(V_i, V_j) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} [Y_m(i, j) - Y_{m-1}(i + V_i, j + V_j)], \quad (1.11)$$

где $N \times N$ – размер блока; (i, j) – пространственные координаты пикселей.

Значение $\vec{V}$, для которых сумма абсолютных разностей $H \leq H_{\min}$ имеет наименьшее значение, принимается за искомый вектор. $H_{\min}$ малая величина, характеризующая максимально допустимую точность компенсации движения. Если $H > H_{\min}$, то принимается, что эквивалентный участок изображения в опорном кадре отсутствует, и данный макроблок кодируется без компенсации движения. Вектора движения ищутся методом полного перебора в заданной ограниченной окрестности: $-N \leq V_i, V_j \leq N$.

К ускоренным алгоритмам поиска векторов движения относятся: логарифмический, комбинированный по двум направлениям, трехшаговый, иерархический и др.

1.4.2. Стандарт MPEG-4

MPEG-4 является первым реальным мультимедийным стандартом для представления данных [22], позволяющим интерактивно работать с комбинациями натурального и синтетического материала, закодированного в виде объектов (он моделирует аудиовизуальные данные, как комбинацию таких объектов). MPEG-4 предоставляет стандартизованные технологические элементы, допускающие интеграцию производства, распределения и доступа к мультимедийному материалу. Это относится к интерактивному и мобильному мультимедиа, интерактивной графике и улучшенному цифровому телевидению.

Стандарт MPEG-4 Видео допускает гибридное кодирование естественных (пиксельных) изображений и видео вместе с синтезированными сценами

(генерированными на ЭВМ). Это, например, допускает виртуальное присутствие участников видеоконференций. Видео стандарт содержит в себе средства и алгоритмы, поддерживающие кодирование естественных (пиксельных) статических изображений и видео последовательностей, а также средства поддержки сжатия искусственных 2-D и 3-D графических геометрических параметров. MPEG-4 поддерживает следующие форматы и скорости передачи:

1. Скорости передачи: обычно между 5 кбит/с и 10 Мбит/с.
2. Форматы: progressive а также interlaced видео.
3. Разрешение: обычно от sub-QCIF вплоть до HDTV.

Эффективное сжатие видео поддерживается для всех скоростей обмена. Сюда входит компактное кодирование текстур с качеством, регулируемым от уровня “приемлемо” (для высоких сжатий данных) вплоть до «практически без потерь».

Ниже приводятся некоторые функциональные возможности стандарта MPEG-4 (Видео) в процессе обработки и кодировании видеоданных, которые отсутствуют в стандарте MPEG-2.

1. Кодирование, учитывающее содержимое изображения и видео позволяет разделить кодовое преобразование и реконструкцию видеообъектов произвольной формы.

2. Произвольный доступ к содержимому видео последовательности открывает возможность реализации функций пауза, быстрый переход вперед или назад для записанных видео-объектов.

3. Расширенное манипулирование видео последовательностями позволяет наложения естественный или синтетический текст, текстуры, изображения и видео. Примером может служить наложение текста на движущийся видео объект, когда текст движется синфазно с объектом.

4. Масштабируемость сложности в кодировщике позволяет кодировщикам различной сложности генерировать корректный и осмысленный поток данных для данной текстуры, изображения или видео.

5. Масштабируемость сложности в декодере позволяет декодировать потоки текстур, изображений или виде декодерами различного уровня сложности. Достигаемое качество, вообще говоря, зависит от сложности используемого декодера. Это может подразумевать, что простые декодеры обрабатывают лишь часть информационного потока.

6. Пространственная масштабируемость позволяет декодерам обрабатывать некоторую часть общего потока, сформированного кодировщиком, при реконструкции и отображении текстур, изображений или видео-объектов при пониженном пространственном разрешении. Для текстур и статических изображений поддерживается не более 11 уровней масштабируемости. Для видео последовательностей поддерживается не более трех уровней.

7. Временная масштабируемость позволяет декодерам обрабатывать некоторую часть общего потока, сформированного кодировщиком, при реконструкции и отображении видео при пониженном временном разрешении. Поддерживается не более трех уровней.

8. Масштабируемость качества позволяет разбить поток данных на несколько составляющих различной мощности так, чтобы комбинация этих составляющих могла при декодировании давать осмысленный сигнал. Разложение потока данных на составляющие может происходить при передаче или в декодере. Полученное качество, вообще говоря, зависит от числа компонент, используемых при реконструкции.

9. Кодирование формы помогает описанию и композиции изображений и видео, а также видео-объектов произвольной формы. Приложения, которые используют двоичные побитовые карты изображения, служат для презентаций баз данных изображений, интерактивных игр, наблюдения, и анимации.

10. Устойчивость к ошибкам обеспечивает доступ к изображениям и видео через широкий спектр систем памяти и передающих сред. Это включает в себя операции алгоритмов сжатия данных в среде, подверженной сбоям при низких скоростях передачи (т.е., меньше чем 64 Кбит/с).

Также в стандарте предусмотрены функции «анимация лица» и «кодирование 2-D сеток с нечетко выраженной структурой».

Некоторые другие особенности работы системы MPEG-4:

- Общая модель событий или отслеживание действий пользователя; общая обработка событий и отслеживание взаимодействий объектов на сцене пользователем или с помощью событий, генерируемых на сцене.
- Средство объединения большого числа потоков в один общий поток, включая временную информацию (мультиплексор FlexMux).
- Средство для запоминания данных MPEG-4 в файле (файловый формат MPEG-4, *.mp4)

- Интерфейсы для различных терминалов и сетей в виде Java API (MPEG-J)
- Независимость транспортного уровня.

Из анализа обзорного материала следует, что стандарты семейства MPEG не определяют схему и конструкцию кодера и декодера, а лишь описывают средства, используемые для обработки аудиовизуальной информации и мультимедийных приложений, определяют синтаксис (правила построения последовательности символов) совместимого цифрового потока и дают примеры реализации декодера. Состав и построение кодера оставлены на усмотрение разработчика. Это может быть аппаратное или программное устройство любой сложности, дающее на выходе синтаксически правильный цифровой поток и позволяющее получить на выходе эталонного декодера желаемый эффект. Таким образом, не накладывается никаких ограничений на технологию, алгоритмы работы, сложность построения кодера и его будущие усовершенствования. В отношении декодера существует одно жесткое ограничение: он должен декодировать любой цифровой поток, совместимый со стандартом MPEG.

В связи с этим, в процессе проектирования цифровой видеосистемы, важное значение приобретают меры оценки качества изображений, по результатам которых можно делать выводы об эффективности используемых алгоритмах сжатия и о работоспособности системы в целом.

1.4.3. Стандарт H.264

H.264/AVC — это международный стандарт видеокодека (кодера и декодера), который предназначен для сжатия и распаковки цифрового видео с целью уменьшения полосы пропускания, требуемой для передачи и хранения видео. Предложен группой экспертов JVT (Joint Video Team) в мае 2003 г.

Как известно, стандартный цифровой видеопоток DVD качества CCIR601 в несжатом виде при разрешении 720×480 пикселей 4:2:2 видео и 30 кадр/с составляет 158 Мбит/с. Это значительно превышает возможности Интернет-сетей передачи данных и приводит к тому, что один час такого видео занимает объем около 79 Гб.

Внедрение стандарта H.264/AVC позволило начать новый этап в развитии Интернет-телевидения, мобильного телевидения и сделало широкодоступным телевидение высокой четкости HDTV.

В стандарте H.264/AVC используется усовершенствованная технология кодирования видеосигнала по методам, схожим с технологией кодирования в предыдущих стандартах MPEG. При этом для обеспечения меньшего битрейта

при сохранении высокого качества изображения используются, новые инструментальные средства.

Оптимизация параметров кодирования.

Недостатком классического кодирования является принятие локально оптимальных решений на каждом этапе видеообработки. В итоге суммарное решение не будет оптимальным. В H.264/AVC используется алгоритм оптимизации параметров кодирования PDO (Rate distortion optimization). При этом в процессе кодирования выбираются такие его параметры, которые обеспечивают получение наилучшего результата.

Улучшенная оценка движения.

Для поисков субмакроблоков различных размеров, от 16×16 до 4×4 пикселей, используется оценка движения. При этом точность определения вектора движения увеличена, как для сигналов цветности, так и для сигнала яркости. Для улучшения кодирования векторов движения используется их предсказание.

Улучшение кодирования при плавных движениях.

Одним из важных отличий стандарта H.264/AVC от предшествующих является введение ряда новых условий для кодирования макроблоков в режиме «skip». В итоге макроблок не кодируется, а вместо него используется другой макроблок в той же позиции, но с другого кадра, что позволяет значительно снизить битрейт. При этом особенно большой выигрыш в уменьшении видеопотока достигается при малых битрейтах или когда вся картинка движется одинаково, например, при плавных движениях видеокамеры.

Фильтрация границ блоков.

Отличием стандарта H.264/AVC от предшествующих является использование деблокирующего фильтра. Этот фильтр сглаживает блочные артефакты на границах макроблоков в изображении. Это приводит к улучшению визуального восприятия каждого кадра и всего видео в целом.

Пространственное предсказание.

В зависимости от направления движения в H.264/AVC используется до 9 разных способов пространственного внутрикадрового предсказания для I кодированных блоков.

Модифицированное дискретное косинусное преобразование.

Для предотвращения ошибок округления при преобразовании остаточной информации используется модифицированное целое дискретное косинусное преобразование (МДКП). При этом в отличие от других стандартов сжатия ви-

деоинформации, размеры блока для МДКП существенно уменьшены, до размеров 8×8 или 4×4 пикселей.

В стандарте H.264/AVC используются более производительные процессы энтропийного кодирования:

1. Context-adaptive binary arithmetic coding (CABAC) — контекстно-адаптированное двоичное арифметическое кодирование, которое представляет собой арифметический кодер. Это кодирование позволяет добиться практически максимально возможной эффективности сжатия.

2. Context-adaptive variable length coding (CAVLC) — контекстно-адаптированное кодирование с различной длиной кодового слова. Этот кодер, основанный на алгоритме сжатия Хаффмана, позволяет ускорить процесс сжатия информации, но в отличие от CABAC не обеспечивает максимально возможной эффективности её сжатия.

Всё это привело к тому, что величина битрейта при использовании стандарта H.264 заметно меньше, чем у MPEG-4 part 2 и других его предшественников. Это сильно проявляется, когда в кадре отсутствует движение.

Стандарт H.264 предоставляет более широкий выбор структуры видеопотока и инструментов, чем MPEG-2, и дает возможность создать лучший видеокодер. Т.е. этот видеокодер может создавать более качественное видео при той же скорости передачи данных и то же качество видео, что и MPEG-2 при меньшей скорости передачи данных. По сравнению с MPEG-4 part 2 стандарт H.264 может обеспечить экономию трафика цифрового потока от 20% до 25% в обычных условиях и свыше 50% в период, когда в кадре нет движения. Это не только снижает общие требования к полосе пропускания для видеосистемы, но что более важно, может значительно сократить объем хранилища, требуемого для записи видео, которое часто является одним из наиболее долгих элементов системы.

Структурная схема кодера H.264 показана на рис. 1.10, где использованы следующие обозначения: F_n — текущий кадр (frame); P — предсказание; uF'_n — восстановленный, но без фильтра, текущий кадр; F'_{n-1} — один или два ранее закодированных кадра; D_n — разница между блоком и его предсказанием; ME — поиск наиболее подходящего предсказания (motion estimate); MC — компенсация движения (motion compensation); T — преобразование; Q — квантование.

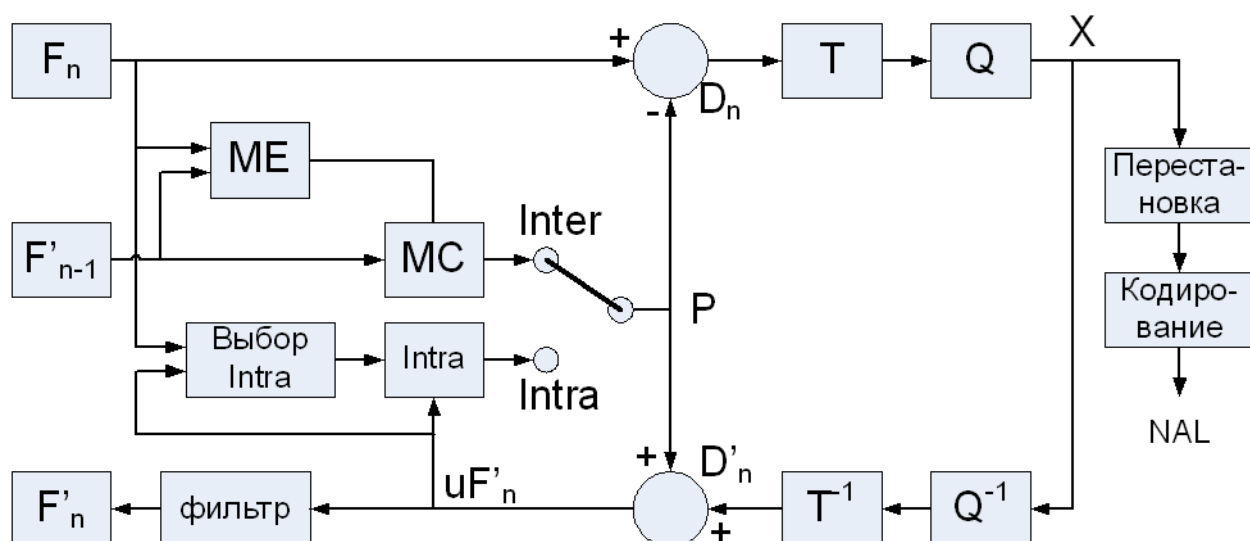


Рис. 1.10. Структурная схема кодера H.264

1.4.4. Стандарт H.265/HEVC

В 2013 г. Международный союз электросвязи (ITU) утвердил стандарт сжатия H.265/HEVC (High Efficiency Video Coding, т.е. высокоэффективное кодирование видеосигнала), который позволит передавать видео с вдвое меньшим битрейтом, чем H.264, при сохранении качества «картинки» [23]. Этот кодек поддерживает видео в разрешении до 7680×4320 пикселей, при этом для просмотра 4k-видео, закодированного в стандарте H.265, достаточно Интернет-соединения со скоростью около 25 МБ/с.

Основные цели, которые преследуются при разработке новых видеокодексов, в частности H.265, остаются неизменны: снижение битрейта, т.е. повышение степени компрессии оцифрованного видеопотока, при сохранении качества его декодированной версии (за счет более сложной обработки сигнала); уменьшение задержки для приложений, которые работают в реальном времени (системы безопасности, видеочаты, видеоконференции и др.); уменьшение аппаратных ресурсов, требуемых при обработке видеоданных; увеличение устойчивости к потерям сигнала при передаче мультимедийных данных.

Помимо попытки достижения этих целей, в стандарте H.265/HEVC используется работа с высоким разрешением HD, а именно в форматах 4K (3840×2160) и 8K (7680×4320) пикселей где применена архитектура с параллельной обработкой данных. H.265/HEVC основан на подходе, который предусматривает сокращение избыточности изображения за счёт [24]:

1. Уменьшения избыточности в соседних кадрах (inter-кодирования), благодаря оценке движения в кадре и его компенсации.

2. Уменьшения пространственной избыточности в ссылочном или разностном кадрах благодаря intra-кодированию. Такой же подход используется в кодеке H.264, однако в стандарте H.265/HEVC предусмотрена более широкая адаптация параметров кодека и расширен диапазон их изменения.

При этом используются: переменные размеры макроблоков изображения; макроблоки изображения, которые имеют древовидную структуру; деблокирующий фильтр, работающий в режиме параллельной обработки; фильтр в цепи обратной связи. Для улучшения качества цветопередачи используется субдискретизация цветовых (хроматических) компонент 4:2:2 и 4:4:4 (помимо 4:2:0).

В стандарт заложена иерархия, которая включает профайлы, слои и уровни. В будущем их можно будет наращивать в зависимости от особенностей использования видеокodeка, например, при предполагаемом развитии его масштабируемой версии SVC.

В стандарте также уделено внимание возможностям гибкого представления видео различных разрешений. Для этого используются: блок кодирования (CU); блок предсказания (PU); блок преобразования (TU). CU – это базовый блок сжатия видеоинформации. Его структура такая же, как у макроблока в стандарте MPEG-2, но она выполнена гораздо более гибкой. Для эффективного сжатия HD и 4k видео в стандарте H.265/HEVC поддерживаются размеры CU большие, чем 16×16 (как в стандарте H.264), например, 32×32 и 64×64 . PU – это блок предсказания, причём один блок CU может содержать несколько PU. Асимметричное разбиение на участки движения (AMP) используется для эффективного кодирования нерегулярных узоров видеокадра. TU – это блок, как блок CU, может включать один или несколько PU. Кроме стандартных преобразований 4×4 и 8×8 , для TU также поддерживаются быстрые преобразования 16×16 и 32×32 .

В стандарте H.265/HEVC также улучшено предсказание векторов движения и модифицированы $1/4$ –пиксельные интерполяционные фильтры.

Особое внимание уделено уменьшению искажений между исходным и восстановленным видеокадрами, для чего объединены несколько фильтров декодера H.265. Для снижения средних искажений локальных областей в дополнение к модифицированному деблочному фильтру введен адаптивный SAO фильтр (Sample Adaptive Offset).

Новшеством, по сравнению с H.264/AVC, в H.265/HEVC стало введение нового понятия тайл (tile). Тайлы представляют собой разбиение кадра на прямоугольные области, которые могут кодироваться и декодироваться независимо друг от друга. В отличие от слайсов (slices), тайлы не являются отдельными синтаксическими единицами кодирования, за счет чего достигается дополнительное увеличение степени сжатия. Использование тайлов позволяет увеличить производительность кодирования и декодирования видеоданных благодаря одновременному проведению этих процессов на современных многоядерных настольных или мобильных платформах.

1.5. Характеристики визуальных искажений

1.5.1. Искажения JPEG

Из анализа этапов выполнения алгоритма кодирования (см. рис. 1.3) следует, что основные потери качества возникают в процессе квантования (1.5) коэффициентов ДКП и обнуления малозначащих коэффициентов. Искажения изображения в основном возникают за счет отбрасывания коэффициентов трансформации в высокочастотной области, значения которых не превышают заданный порог или равны нулю после выполнения операции нелинейного квантования.

Основные типы искажений (артефактов) и их визуальная характеристика, после выполнения алгоритма JPEG с низким параметром качества Q, достаточно полно описаны в [25], поэтому приведем их краткое описание.

1. «Блокинг-эффект» - возникает из-за ненулевых ошибок на границах блоков, которые идентифицируются глазом как скачки яркости от одного блока к другому.

2. «Мозаичный эффект» - выглядит подобно блокинг-эффекту, но обуславливается не резкими переходами между различными блоками, а заметным глазу различием яркости в соседних блоках.

3. «Искажения типа ступеньки» - возникают как результат неправильного восстановления или передачи краев изображений внутри блоков. При переходе к следующему блоку граница объекта терпит скачок и, в целом, выглядит как «ступенька» с элементами, параллельными границам блоков.

4. «Искажения, имеющие вид базисных функций ДКП» - проявляются в виде структур, очень похожих по форме на базисные функции ДКП.

5. «Размытие изображения» - обусловлено значительным ограничением либо полным занулением высокочастотной части спектра ДКП. При этом мелкие детали становятся либо размытыми, либо полностью пропадают в изображении.

6. «Размытие цветовых границ» - проявляется на участках изображения с резкими переходами в сигналах яркости и цветности, обусловленных отсутствием высокочастотных компонент спектра.

7. «Окантовки на границах» - возникновение характерных окантовок на резких переходах яркости и цветности изображения.

Искажения типа 5-6 характеризуют ухудшение визуальной четкости и резкости изображения, которое существенно зависит от выбора характеристик квантователя коэффициентов ДКП.

На рис. 1.11 показан пример тестового изображения (таблица ISO-12233) и на рис. 1.12 показано декодированное изображение после выполнения алгоритма JPEG с низким параметром качества и с высоким коэффициентом сжатия.

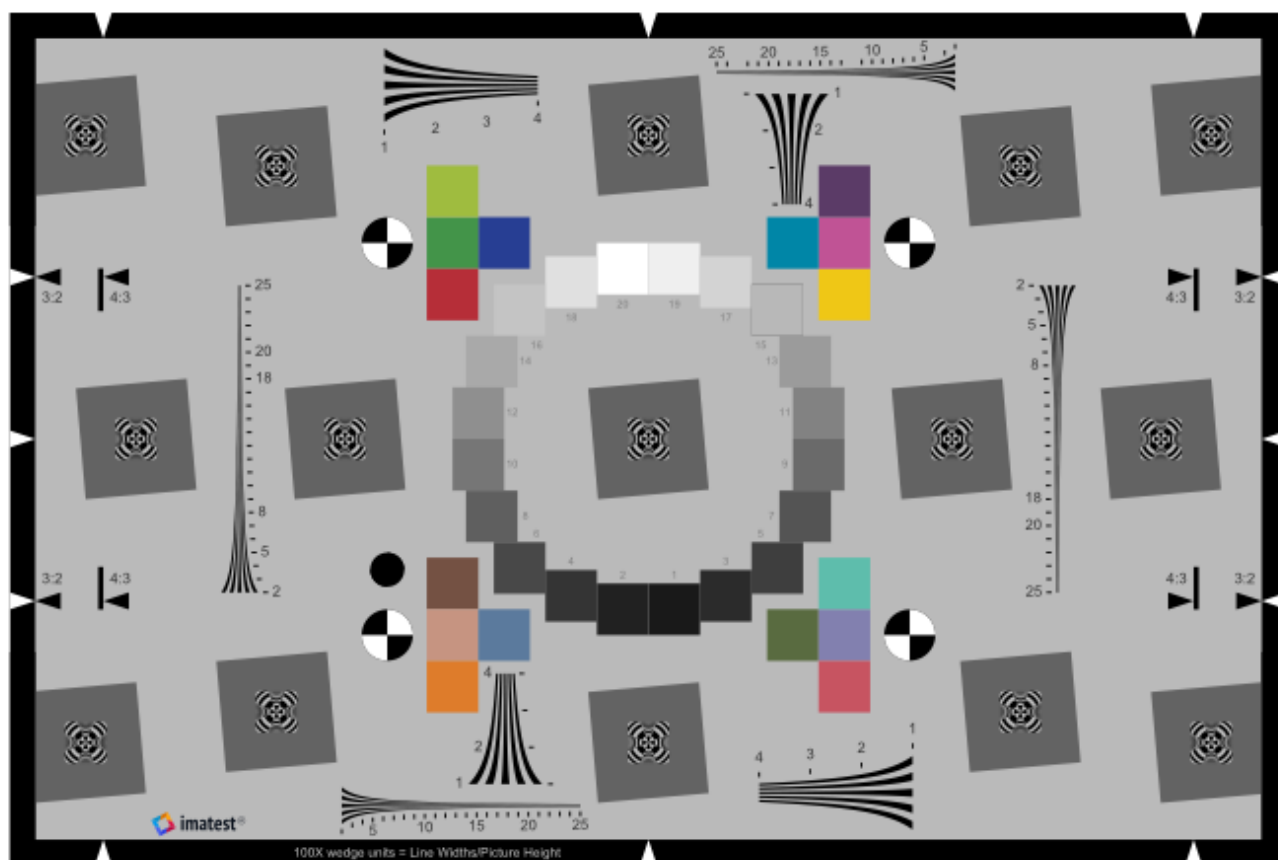


Рис. 1.11. Тестовое изображение ISO-12233

Сжатие выполнено с помощью фоторедактора Adobe Photoshop CS6 Extended с регулируемой шкалой качества со следующими параметрами:

Low – Q (0...4);

Medium – Q (5...7);

High – Q (8...9);

Maximum – Q (10...12).

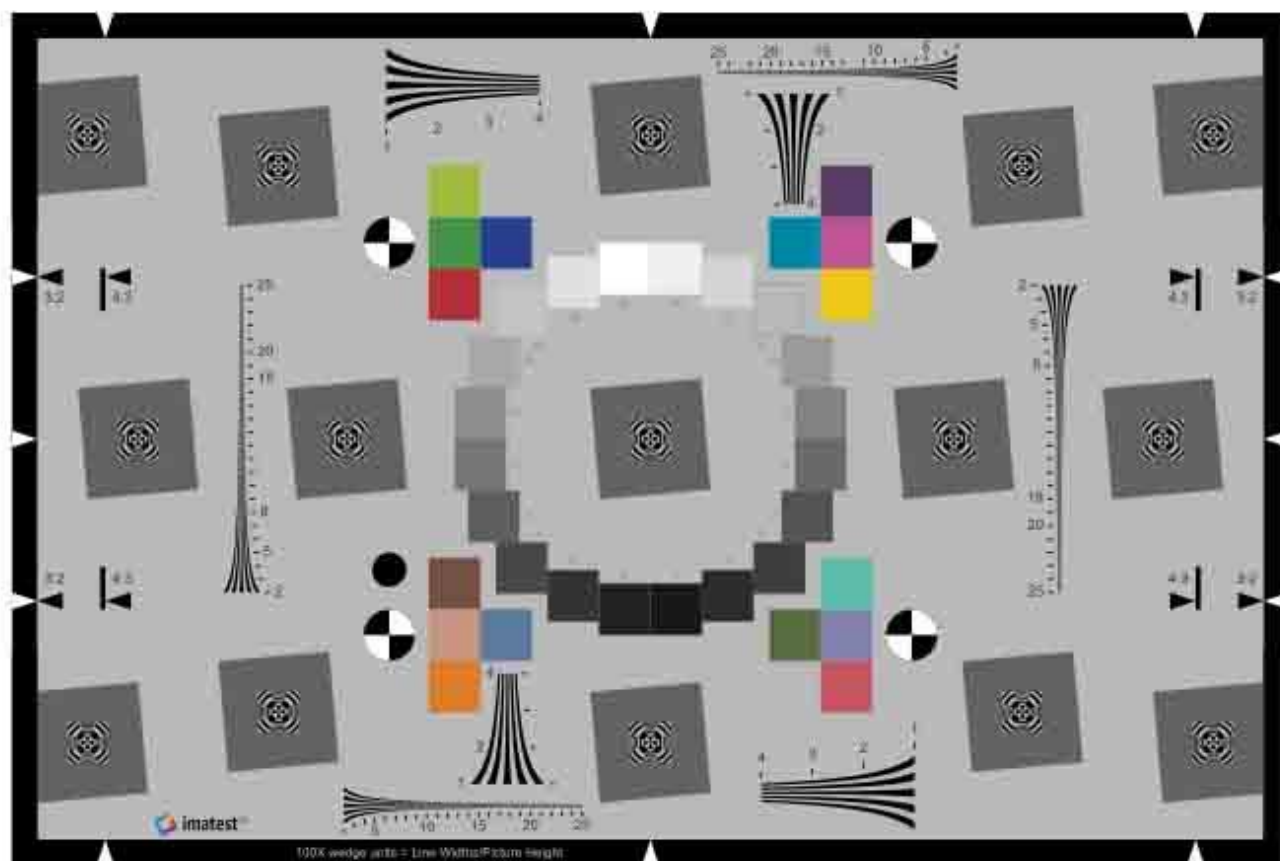


Рис. 1.12. Компрессированное изображение JPEG (Q = 1, $K_{сж} = 17$)

Визуальное сравнение этих изображений показывает, что все описанные артефакты заметны глазом. Отметим, что оценки качества по синтезированным тестовым таблицам не всегда соответствуют реальным визуальным оценкам качества фотореалистичных изображений. Объективные метрики оценки качества реалистичных изображений будут рассмотрены в следующей главе.

1.5.2. Искажения JPEG 2000

Анализ этапов алгоритма JPEG 2000 показывает, что основные потери качества обусловлены квантованием коэффициентов ДВП и обнулением малозначащих коэффициентов. Процесс квантования отличается от алгоритма JPEG и описан в п. 1.3. формулой (1.9). Основные искажения изображения возникают в

процессе квантования, за счет грубого квантования ВЧ составляющих и отбрасывания (обнуления) малозначащих коэффициентов ДВП в каждом из субдиапазонов. Характер и специфика искажений, также зависят от выбора базисной функции материнского вейвлета и количества итераций вейвлет-преобразования.

Отличия от JPEG в том, что преобразование ДВП выполняется итерационно по всей матрице изображения, т.е. отсутствует разбиение на блоки. Поэтому все артефакты, обусловленные блочными эффектами, отсутствуют.

На рис. 1.13 показан пример декодированного тестового изображения JPEG 2000 в Adobe Photoshop CS6 Extended с низким параметром качества, где значение Q регулируется в пределах 1...100.

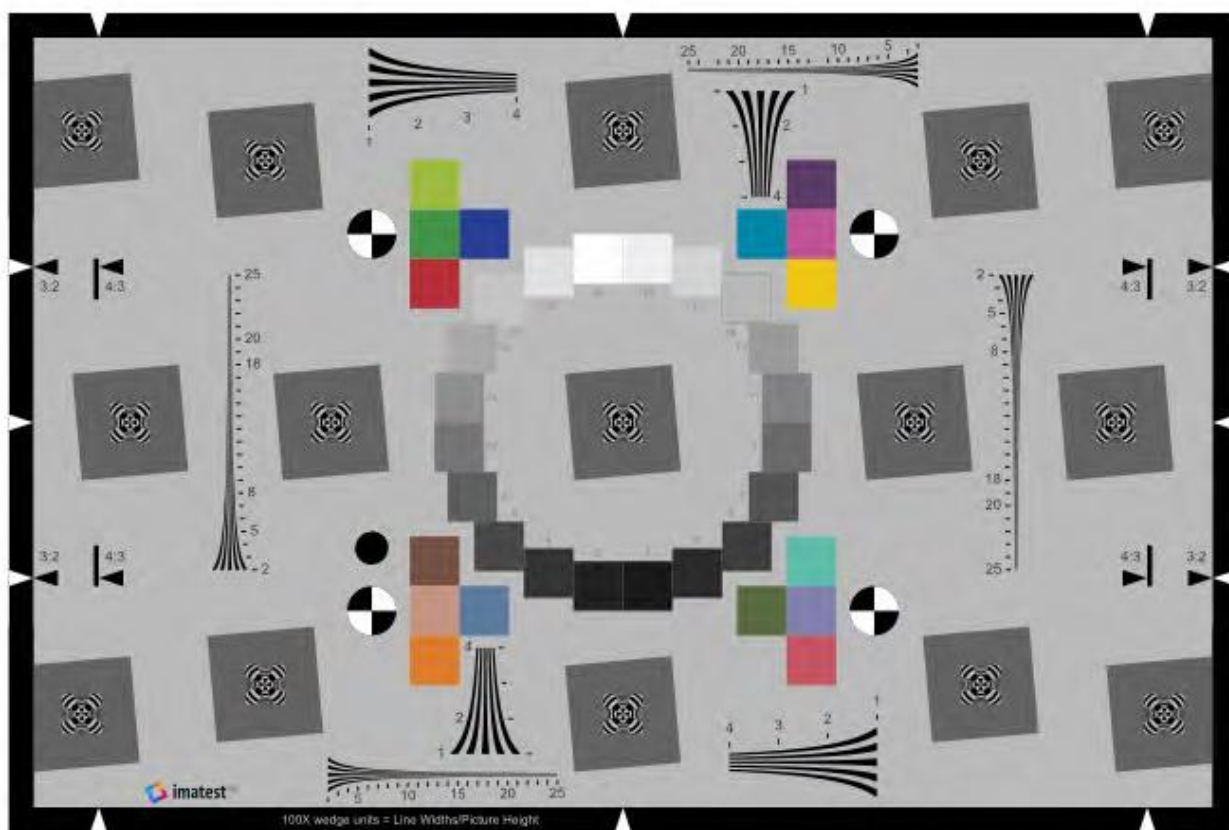


Рис. 1.13. Компрессированное изображение JPEG 2000 ($Q = 30$, $K_{сж} = 17$)

Основные типы искажений (артефактов) и их визуальную характеристику, после выполнения алгоритма JPEG 2000 с низким параметром качества Q можно описать следующими параметрами:

1. «Размытие изображения» - обусловлено значительным ограничением части спектра и обнулением коэффициентов ДВП в высокочастотных субдиапазонах.

2. «Размытие цветовых границ» - проявляется на участках изображения с резкими переходами в сигналах яркости и цветности, обусловленных отсутствием высокочастотных компонент спектра.

3. «Окантовки на границах» - возникновение характерных окантовок на резких переходах яркости и цветности изображения.

1.5.3. Искажения видеокодеков

Рассмотренные выше искажения JPEG будут присутствовать при декодировании видеоданных (MPEG-2, MPEG-4, H.264, H.265) во всех кадрах, что обусловлено артефактами компрессии опорных I кадров в группе изображений. При этом обеспечение постоянной скорости битового потока достигается двумя способами, во-первых, регулировкой шкалы квантования и, во-вторых, адаптивным распределением ресурса битов между разными типами кадров с учетом их сложности. Следовательно, основным источником искажений будет ограничение пропускной способности канала передачи. При низких скоростях передачи битового потока используется более грубое квантование и артефакты компрессии (см. рис. 1.12) будут заметны глазом.

Основные типы и визуальная характеристика искажений, обусловленные алгоритмом межкадрового кодирования, подробно описаны в [26], поэтому ограничимся их кратким обзором.

1. «Ложные границы» – возникают за счет переноса блокинг и мозаичного эффектов из опорного в новый кадр. При больших значениях коэффициентов квантования количество ложных границ увеличивается от кадра к кадру, что приводит к существенной деградации изображения.

2. «Эффект комаров» - проявляется как флуктуации яркости или цветности в блоке на границе между движущимся объектом и фоном. Возникает вследствие различной степени квантования ошибки предсказания от кадра к кадру. Часть блока, который подвергается компенсации движения, может передаваться правильно, но другая часть, соответствующая фону, и ошибка предсказания которой не мала, может передаваться с искажениями вследствие грубого квантования коэффициентов ДКП.

3. «Зернистый шум в стационарной области» - проявляется как медленно движущиеся дрожащие шумы низкой интенсивности в областях, в которых присутствует лишь малое движение, либо движение отсутствует полностью.

4. «Неправильные цвета» - проявляется, в основном, как неправильный цвет всего макроблока как по отношению к собственному цвету, так и по отно-

шению к цвету окружающей области. Наиболее вероятной причиной неправильного цвета является тот факт, что для анализа вектора движения используется только сигнал яркости.

5. «Эффект привидения» – проявляется в том, что за движущимися объектами возможно образование следов («привидений»), которые могут сохраняться в течение сравнительно большого промежутка времени. Зависит от скорости объектов, участвующих в движении.

Задача видеокodeка – это достигнуть определенного формата файла и определенного уровня компрессии с приемлемым качеством изображений. Для примера можно взять спецификацию codeка H.264 в котором указано, что codeк не обременен какими-либо ограничениями на реализации алгоритмов кодирования и алгоритм кодирования не является частью спецификации.

Это позволяет более гибко выбирать размер выходного файла или потока с учетом времени потраченного на кодирование/декодирование. Например, для потокового видео с минимальной задержкой необходимо выбирать алгоритмы быстрого декодирования использующий более грубые оценки, при этом стоит ожидать увеличение выходного размера файла.

Спецификация формата видео кодирования может дополнительно определять уровни и профили формата видео кодирования. Уровни профиля налагают ограничения на разные параметры выходного файла/потока.

Так, уровень накладывает на формат ограничения в максимальном разрешении и максимальном и, иногда, минимальном размере потока, выдаваемом за секунду. В то же время профиль ограничивает codeк в техниках оптимизации. Вместе эти параметры позволяют создавать оборудование и программы в разном ценовом и качественном диапазоне. А также, косвенно, регулировать потребление ресурсов декодером и codeром.

В заключение отметим, что в данном разделе рассмотрены источники и характеристики искажений возникающие в процессе кодирования изображений. В целом на качество передаваемых изображений будут оказывать влияние и другие факторы, такие как шумы и искажения цифровых камер и их характеристики. Также визуальное восприятие будет зависеть от характеристик монитора.

Шумы канала связи при цифровой передаче могут приводить к ошибкам декодирования и к полной деградации изображений. Особенности помехоустойчивого кодирования в данной работе не рассматриваются.

2. Метрики качества цифровых изображений

2.1. Общие характеристики качества изображений

Контроль качества компрессированного изображения достаточно сложен и неоднозначен. В настоящее время существуют множество методов и метрик анализа качества изображений. К традиционным метрикам относятся PSNR и MSE, к более сложным – метрики структурного подобия (SSIM) и мульти-масштабная MS-SSIM [27]. Метод SSIM выполняет сравнение двух изображений и оценивает их изменения по яркости, контрасту и структуре [28]. Сравнение пикселей изображений выполняется с помощью окна, в котором оценивается качество изображения, количественным значением близким к единице. Этот метод более популярен, чем PSNR и MSE, т.к. он более информативен и ближе к зрительному восприятию различий между изображениями [29].

На сайте [30] лаборатории LIVE (Laboratory for Image & Video Engineering) Техасского университета в Остине приводятся современные публикации с описанием и сравнительным анализом наиболее популярных метрик (SSIM, MS-SSIM, VSNR, IFC, VIF, MSVD и др.) по оценке качества изображений с использованием различных баз данных оригинальных и искаженных изображений.

Результаты обзорных исследований показывают, что во всех известных объективных метриках оценки качества компрессированных изображений практически мало учитываются свойства человеческого зрения в области восприятия мелких деталей и структур изображения. До сих пор самыми надежными считаются субъективные экспертные оценки (MOS). Также, пока что не существует единой методики количественной оценки эффективности алгоритмов анализа качества изображений [31].

В настоящей главе приводится описание современных методов оценки качества цифровых изображений: на основе испытательных таблиц; с помощью объективных метрик оценок фотореалистичных изображений; с помощью программно-аппаратных средств оценки качества видеопоследовательностей.

В п. 2.5 приводится описание метрик оценки качества реалистичных изображений, предложенных автором (Сай С.В.) в качестве альтернативных существующим метрикам.

Под термином качество изображения понимают совокупность потребительских свойств изображения, проявляющихся при его восприятии и оценке наблюдателем на основе форм зрительных образов, что предполагает наличие в структуре качества не только объективных (физических), но и субъективных

свойств. Данное соотношение между физическими и субъективными характеристиками изображения выглядит следующим образом.

Физические характеристики [32]:

- градиентные (на основе свойств пограничной кривой);
- частотно-контрастные (спектральные и амплитудные);
- интенсивность (яркость);
- стабильность;
- пространственные (физические размеры, линейные и нелинейные искажения, мерность пространства).

Субъективные характеристики:

- фокусировка, резкость, отчетливость;
- контрастность, различимость градаций, яркости, плотность, тоновые оттенки, цвет;
- минимальная и максимальная видимая яркость, четкость, детальность;
- мерцание, дрожание;
- адекватность формы и величины, объемность.

Существующие методы оценки качества изображений можно разделить на два больших класса [33]:

- субъективные оценки (экспертные);
- математические оценки (метрики).

Для метода субъективных оценок необходима группа(ы) экспертов, чтобы получить информацию о качестве обработанного или восстановленного изображения. Существуют два вида экспертных оценок: абсолютные и сравнительные. В первом случае наблюдатель должен оценить качество изображения по одной из выбранных шкале оценок. Наблюдатель принимает решение и выставляет оценку качества предъявляемому изображению на основании собственного опыта. Следует отметить, что в данном случае разброс результатов оценок может быть достаточно высоким, т. к. у каждого наблюдателя имеется свой личный опыт и мнение о том, какое изображение лучше. Например, кому-то более предпочтительны изображения с высоким контрастом и цветовой насыщенностью, а другой наблюдатель предпочитает более “мягкие” тона.

При сравнительных оценках наблюдатель сравнивает изображения друг с другом из конкретного набора и по определенной шкале производит оценку каждого изображения.

Наиболее распространенным видом шкалы оценок является шкала “общего качества”, где изображениям приписываются баллы, соответствующие града-

циям от “неудовлетворительно” до “отлично”, приведенным в табл. 2.1. На практике наблюдатель выставляет соответствующий балл той категории, которая, по его мнению, наиболее соответствует качеству предъявляемого изображения. Вариантом данной шкалы является “относительная” шкала (табл. 2.2), когда наблюдатель оценивает качество изображения, сравнивая его с другими изображениями, входящими в рассматриваемую группу. Еще одной распространенной шкалой оценок является шкала погрешностей (табл. 2.3), согласно которой наблюдатель должен оценить в баллах степень искажений, изменяющихся от “незаметных” до “крайне нежелательных”.

Таблица 2.1. Шкала общего качества

Балл	Оценка качества изображения
1	Неудовлетворительно
2	Плохо
3	Посредственно
4	Хорошо
5	Отлично

Таблица 2.2. Относительная шкала

Балл	Оценка качества изображения
1	Самое плохое в группе
2	Существенно хуже среднего
3	Несколько хуже среднего
4	Среднее в группе
5	Несколько лучше среднего
6	Заметно лучше среднего
7	Лучшее в группе

Таблица 2.3. Шкала погрешностей

Балл	Оценка качества изображения
1	Незаметные
2	Еле заметные
3	Вполне заметные, но слабо ухудшающие изображение
4	Ухудшающие изображение, но допустимые
5	Несколько нежелательные
6	Определенно нежелательные
7	Крайне нежелательные

Шкала общего качества (5-балльная) позволяет привлекать к экспертным оценкам любых наблюдателей практически без подготовки. Относительная шкала и шкала погрешностей (7-балльные) требуют специальной подготовки наблюдателей или привлечения к услугам наблюдателей-специалистов. Одной из трудностей, связанных с балльными оценками, является возможная нелинейность шкалы.

В стандарте определения качества телевизионных передач ITU-TB T.500-11 описан субъективный пятибалльный метод DSIS (Double Stimulus Impairment Scale). Изображения показываются последовательно парами: сначала демонстрируется исходное или оригинальное (эксперты знают об этом), затем

обработанное или восстановленное изображение. По окончании просмотра эксперты оценивают качество изображения по 5-балльной шкале:

- 5 – различия не воспринимаются;
- 4 – различия воспринимаются, но не раздражают;
- 3 – различия немного раздражают;
- 2 – различия раздражают;
- 1 – различия очень раздражают.

Следует отметить, что на результаты субъективных экспертных оценок влияют характер рассматриваемых изображений и условия наблюдения. Если наблюдатель видел подобные изображения, то он более строго оценит погрешности, поскольку имеет сложившийся образ о структуре изображения. С другой стороны, в незнакомом изображении некоторые искажения могут оставаться незамеченными из-за отсутствия сравнительных образов. Условия наблюдения (контраст, яркость, расстояние до наблюдаемого изображения) должны быть максимально приближены к реальной обстановке.

Результаты экспертных оценок обычно представлены с помощью среднего арифметического значения и среднеквадратического отклонения. Считается, что для надежной оценки качества изображения необходимо опросить не менее двадцати наблюдателей.

Методика субъективной оценки качества телевизионных изображений подробно представлена в Рекомендации МСЭ-R BT.500-13 (01/2012) [34].

2.2. Оценка качества изображений по испытательным таблицам

Испытательные (тестовые) таблицы представляют собой синтезированные изображения заданного формата, которые содержат набор элементов для субъективного контроля качества кодирования и передачи, а также для измерения характеристик и основных показателей качества [35]. Пример такой таблицы показан на рис. 1.11. Дополнительно, для оценки качества видео компрессоров могут использоваться динамические таблицы.

Разработчики видеосистем используют специальные испытательные таблицы (ИТ), которые более критичны к искажениям в канале передачи, чем обычные (средние) сюжеты. ИТ, выполняемые в виде нормализованного рисунка (фотографии, слайда), позволяют обеспечить контроль тракта изображения «от света до света». В настоящее время широко используются электронные ИТ (ЭИТ), создаваемые специальными генераторами электрических сигналов или с

помощью синтезированных цифровых изображений. С помощью ЭИТ можно контролировать работу всех блоков тракта изображения, за исключением преобразователя свет-сигнал. Достоинством ЭИТ является высокая стабильность исходных параметров таблицы. В ТВ применяют как оптические, так и электронные ИТ (ТИТ 0249, ИТ 72, УЭИТ и др.). Методика использования ИТ и ЭИТ для контроля параметров отдельных звеньев тракта передачи описана в специальной литературе.

Качество изображения оценивается рядом показателей, среди которых основными являются: 1) контрастность изображения; 2) количество воспроизводимых градаций яркости; 3) максимальная яркость изображения; 4) четкость в горизонтальном и вертикальном направлениях; 5) искажения границ между деталями с разной яркостью (резкость); 6) правильность цветопередачи; 7) допустимый уровень помех различного рода, искажающих изображение; 8) геометрическое подобие изображения оригиналу; 9) размер кадра и его формат; 10) стабильность положения изображения (устойчивость синхронизации).

В зависимости от способов измерения перечисленные выше параметры можно объединить в две группы: световые показатели и растровые показатели. К световым показателям изображения относят перечисленные в пп. 1 - 7. К растровым показателям изображения относят перечисленные в пп. 8 - 10.

Измерение тех или иных показателей изображения обычно производится с помощью какой-либо части испытательной таблицы (испытательного сигнала): сетчатого или шахматного поля; вертикальных и горизонтальных черно-белых полосок разных размеров; вертикальных или горизонтальных полосок со ступенчато изменяющейся яркостью; вертикальных и горизонтальных рядов точек и т. д.

Разработчики цифровых видеосистем предлагают разные таблицы [36], позволяющие оценивать определенные показатели качества (четкость, резкость, точность цветопередачи и др.). На рис. 2.1 показаны примеры таких таблиц.

Для быстрой проверки основных показателей качества изображения, которые может обеспечить видеосистема, а также для оценки основных характеристик кодера и декодера используют универсальные телевизионные испытательные таблицы, которые содержат набор всех необходимых элементов.

В цифровых видеосистемах, в отличие от аналоговых, практически отсутствуют геометрические искажения. При этом основные искажения возникают в процессе «грубого» кодирования сигналов источника изображения. Визуальные оценки искажений элементов испытательной таблицы позволяют оценить каче-

ство передачи по степени заметности артефактов кодирования (см. п. 1.4) с помощью, выбранной шкалы оценок (п. 2.1).

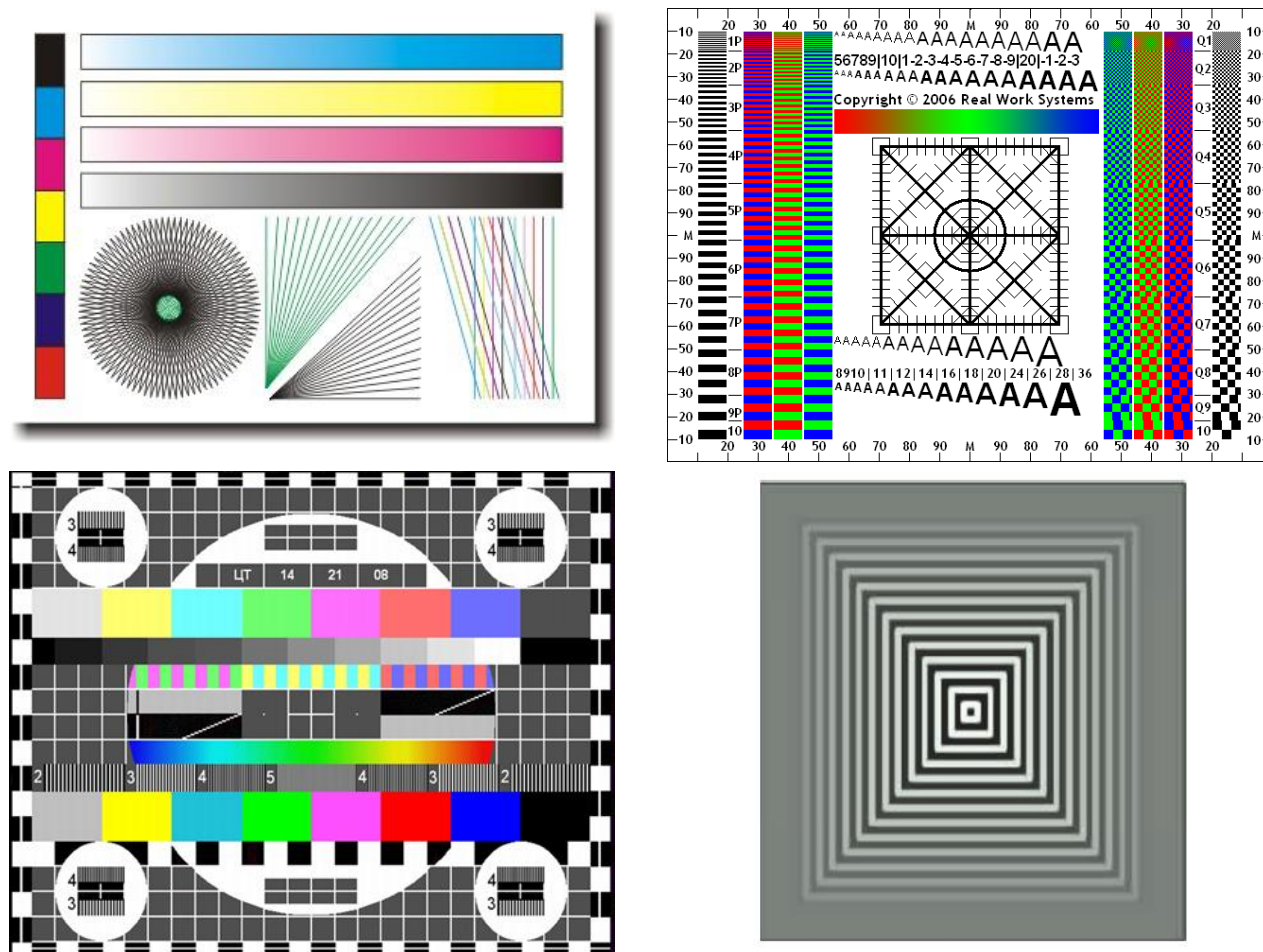


Рис. 2.1. Примеры испытательных таблиц

Отдельные искажения изображения таблицы можно оценивать с помощью измерений параметров цифровых сигналов. Для примера рассмотрим способы оценки разрешающей способности и резкости.

Разрешающая способность. Традиционным способом оценки разрешающей способности системы считается метод измерения максимальной пространственной частоты, при которой обеспечивается пороговый контраст воспроизводимого изображения предельно контрастной штриховой миры. В аналоговых системах цветного телевидения измерения проводятся отдельно по каналам яркости и цветности. В этом случае, характеристики изображения представлены двумя параметрами: разрешающей способностью по яркости и по цветности. В цифровых системах соотношение разрешающей способности по яркости

и цветности определяется выбранным форматом дискретизация (4:2:2, 4:1:1 или 4:2:0).

Метод измерения заключается в следующем. На вход системы подаются сигналы испытательной таблицы в виде изображения штриховой миры синусоидальной или прямоугольной формы с заданной частотой ω_1 и амплитудой A_m первой гармоники. На выходах яркостного и цветоразностных каналов системы измеряется амплитуда соответствующих сигналов и сравнивается с пороговым значением. Линии штриховой миры считаются различимыми, если амплитуда первой гармоники превышает пороговое значение: $\tilde{A}_m(\omega_1) > \Delta_{nor}$.

Оценка разрешающей способности во многом зависит от выбора порогового значения. Очевидно, что уменьшение значения Δ_{nor} приводит к более высокой оценке разрешающей способности системы и наоборот. Минимальное значение Δ_{nor} ограничивается суммарным уровнем шума, а максимальное значение определяется пороговой величиной получателя изображения, т.е. зависит от контрастно-частотной характеристики (КЧХ) зрения.

В конечном итоге, разрешающая способность линейной ТВ системы зависит от: уровня спада сквозных пространственно-частотных характеристик (ПЧХ) каналов яркости и цветности в области верхних частот; величины выбранного порога Δ_{nor} , а также от уровня шума датчика изображения.

В цифровой видеосистеме со сжатием видеоданных, такой способ может привести к неоднозначным результатам, что объясняется нелинейными преобразованиями в процессе кодирования.

В качестве примера рассмотрим, каким образом изменяется амплитуда яркостного сигнала наклонной штриховой миры на выходе системы для двух вариантов обработки: а) двумерная линейная цифровая фильтрация; б) кодирование методом ДКП.

Из рис. 2.2. следует, что линейная фильтрация приводит к снижению амплитуды сигнала яркости в соответствии со спадом ПЧХ низкочастотного фильтра в области верхних частот и, при этом, значение амплитуды не изменяется вдоль оси x . Восстановленное изображение методом ДКП имеет неравномерное распределение амплитуд вдоль оси x , что обусловлено нелинейным процессом квантования коэффициентов ДКП.

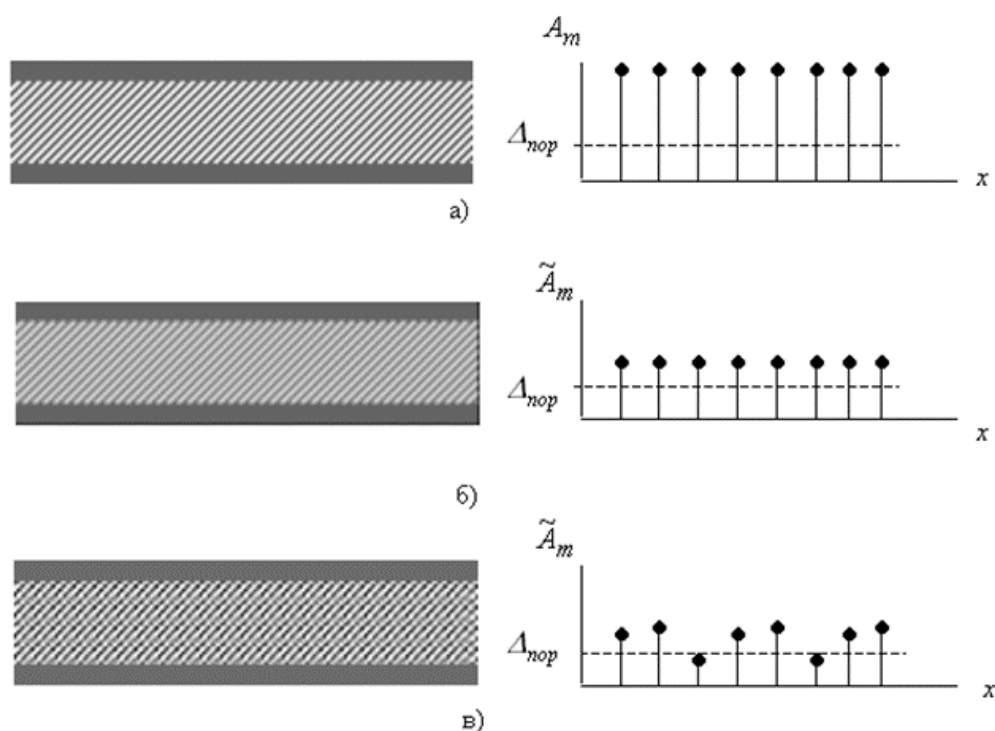


Рис. 2.2. Фрагменты изображения штриховой миры: а) исходный фрагмент; б) после линейной фильтрации; в) после декодирования методом ДКП

Следовательно, рассмотренный выше, метод оценки разрешающей способности дает однозначную оценку при линейной фильтрации и не гарантирует точность результатов при цифровом кодировании методом ДКП.

Отметим другие недостатки стандартного метода.

1. Зависимость порогового контраста Δ_{nop} от чувствительности зрения к восприятию цвета мелких деталей приводит к неоднозначности оценок, что затрудняет анализ точности передачи мелких деталей по сигналам изображения цветной штриховой миры.

2. Испытательное изображение штриховой миры с предельным контрастом, которое наиболее часто используется на практике, не позволяет оценить точность воспроизведения мелких малоконтрастных деталей изображения.

Резкость. В аналоговых системах цветного телевидения традиционным способом оценки резкости границ деталей изображения является определение длительности переходных характеристик (ПХ) между перепадами сигналов яркости или цветности. При этом резкость изображения обратно пропорциональна длительности ПХ. Типичная форма переходной характеристики, например, канала яркости, показана на рис. 2.3.

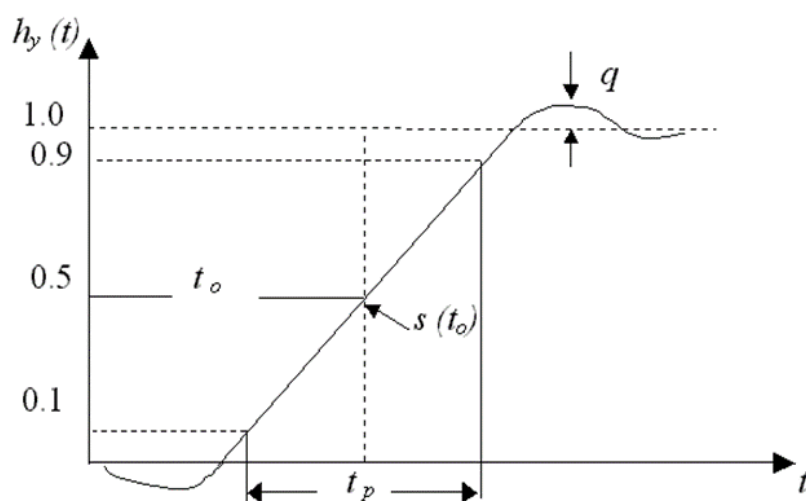


Рис. 2.3. Переходная характеристика канала яркости

Основными параметрами ПХ, влияющих на качество воспроизведения границ крупных деталей изображения, являются: t_p — время нарастания или длительность; t_o — время задержки; $s(t_o)$ — крутизна; q — величина выбросов.

Длительность t_p переходной характеристики обычно оценивается временным интервалом или числом элементов раstra, в течение которого ПХ изменяется от 10 до 90 % ее установившегося значения. В некоторых случаях более простым определением длительности может служить отношение установившегося значения ПХ к крутизне $s(t_o)$ этой характеристики в некоторой точке наклонного участка (например, на уровне 50 %). Увеличение длительности t_p приводит к "размытию" границ и потери резкости изображения в горизонтальном направлении

Время t_o задержки определяется фаза-частотной характеристикой канала и для сигналов цветного телевидения имеет важное значение при совмещении центров ПХ каналов. Расхождение во времени сигналов яркости и цветности приводит к возникновению цветовых окантовок границ деталей, что приводит к дополнительному ухудшению качества изображения. Однако данный тип искажений может быть легко скомпенсирован с помощью аппаратных средств в декодирующем устройстве.

Величина q выбросов зависит от формы частотной характеристики канала в области верхних частот и приводит к возникновению повторов границ деталей в виде узких дополнительных полосок (окантовок). Считается [37], что выбросы не заметны глазом при $q < 3...5\%$ и, при этом, обеспечивается требуемая «чистота переходов». Способ определения длительности переходных характе-

ристик каналов яркости и цветности заключается в следующем. На входы каналов подается единичный скачок напряжения $E(t)$ и по форме выходных сигналов анализируется интервал t_p . При ограничении полос пропускания каналов цветности, например, в соотношении 4:1:1, длительность ПХ каналов цветности в четыре раза больше, чем канала яркости.

Такой способ позволяет оценить резкость вертикальных границ крупных деталей для двух предельных случаев: 1) передаваемое изображение содержит только яркостную информацию - максимальная резкость; 2) передаваемое изображение содержит только информацию по цветности - минимальная резкость. Для цветовых границ с разными уровнями перепадов сигналов яркости и цветности, резкость изображения будет принимать промежуточное значение.

В цифровой видеосистеме, где используется двумерная фильтрация сигналов цветности, протяженность цветовых переходов горизонтальных границ будет зависеть от ПЧХ фильтрующих звеньев, а также от погрешностей алгоритма кодирования. Следует отметить, что при цифровом кодировании на основе дискретных ортогональных преобразований важное значение приобретает точность передачи наклонных границ, т.к. именно такие границы наиболее подвержены искажениям.

Таким образом, оценка резкости вертикальных границ по длительности ПХ каналов яркости и цветности, и горизонтальных границ по величине межстрочного интервала, не дает объективную оценку искажений цветовых переходов в цифровой видеосистеме.

К недостаткам стандартного метода определения протяженности цветового перехода, также, можно отнести следующее. Начало и завершение переходного процесса (см. рис. 2.3) обычно определяется по уровням 0.1 и 0.9 ПХ, что может быть не согласовано со свойствами зрительного восприятия перепадов яркости и цветности. Известно, что наше зрение более чувствительно к перепадам яркости и менее чувствительно к перепадам цветности. Следовательно, пороги зрительного восприятия цветовых перепадов, а также заметности выбросов будут разными.

К недостаткам использования стандартных статических таблиц в анализаторах качества видеоизображений является особенность функционирования кодера MPEG. Фактически передача статических во времени составляющих сигнала ТВ используется опорный I кадр, который не содержит информацию о векторах движения. Таким образом, в сформированных Р и В кадрах отсутствует информация о межкадровой разности, то есть их размер будет равен нулю.

Что дает возможность передавать таблицу в хорошем качестве даже при низких битовых скоростях.

Для анализа качества цифрового видео, таблица или испытательный сигнал должны быть динамическими, а также содержать элементы для проверки цифровых искажений I кадров, таких как блокинг-эффект, обусловленный разбиением изображения на блоки перед ДКП, мозаичный эффект, потеря разрешающей способности или размытие изображения, окантовки на границах, размытие цветов, искажения типа ступеньки, шумы.

В настоящее время существуют много различных вариантов испытательных таблиц, содержащих динамические изображения. Например, такие компании как "Би-би-си" и Snell & Wilcox использовали их при тестировании своего оборудования, а также в вещании. В качестве примера рассмотрим таблицу, разработанную в НИИР. Динамическая таблица (рис. 2.4) представляет собой последовательность из 12 кадров стандартной четкости с соотношением сторон 4:3, содержащие прямоугольные и круглые участки. Частота смены кадров равна 25 Гц. Используемый кодек MPEG-2.

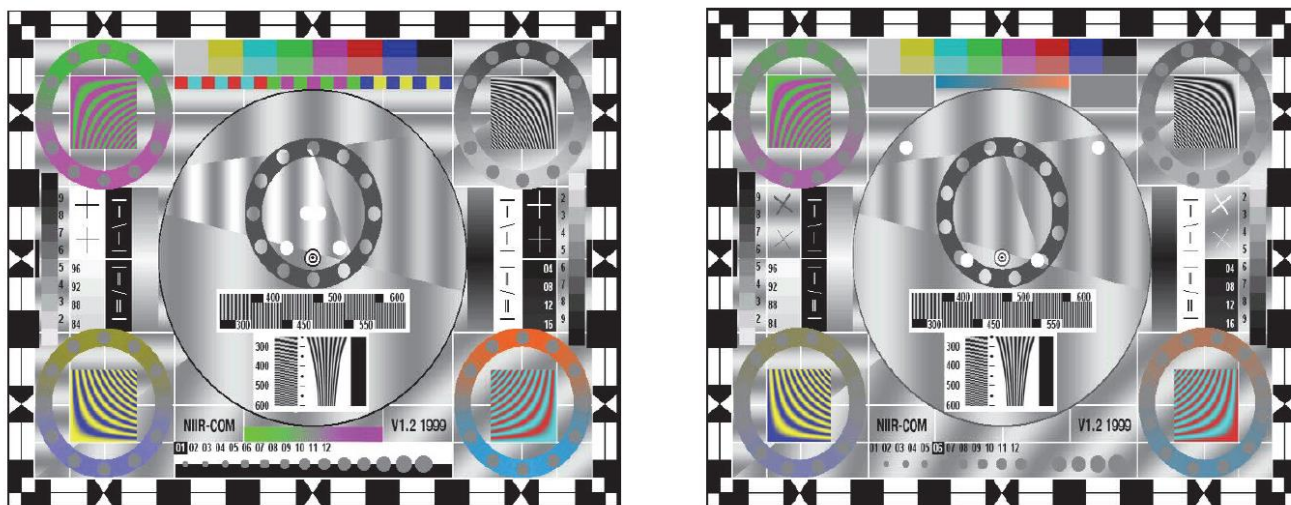


Рис. 2.4. Первый и шестой кадры динамической таблицы НИИР

Элементы данной таблицы предназначены для анализа цифровых и аналоговых искажений. Таблица разбита на прямоугольные и круглые участки, содержит вращающиеся круги, кресты и перемещающиеся белые круги малого диаметра в большой центральной окружности. Векторы движения волн плавно меняются в зависимости от координаты в диапазоне 5-10 пикселей за кадр, что позволяет оценить максимальное значение векторов движения, обрабатывае-

мых кодером, определить, проводится ли кодером MPEG-2 компенсация движения с точностью до пикселя. Также таблица содержит элементы стандартной таблицы, это реперные метки, цветные полосы SECAM, цветные штрихи 75% яркости и 50% насыщенности, косые линии для 105 проверки правильности чересстрочной развертки, штриховые миры для оценки разрешающей способности и др.

2.3. Метрики оценки качества фотореалистичных изображений

В цифровой видеосистеме (см. рис. 1.1.) верность воспроизводимого изображения обычно оценивается по отношению к изображению на входе системы. Существуют два вида мер различий [38]: глобальные и локальные. Глобальные меры позволяют оценивать качество воспроизведения изображений в целом, т. е. с учетом всех искажающих факторов и независимо от характеристик отдельных фрагментов анализируемого изображения. К таким мерам относятся: среднеквадратичное отклонение, нормированный средний квадрат ошибки, отношение сигнал/шум. Локальные меры характеризуют отдельные показатели качества изображения, например, такие как четкость, резкость, точность цветопередачи, зашумленность и др. При выборе меры качества следует учитывать, что количественные результаты оценок должны хорошо согласовываться с результатами субъективных экспертных оценок для широкого класса изображений.

Концепция измерений и оценки качества изображений в цифровом ТВ-вещании, разработанная в России, содержит следующие основные положения.

1. Методы субъективных и объективных испытаний должны учитывать специфические особенности цифровой обработки и передачи сигналов изображения.
2. Для выявления и оценки искажений испытательные изображения должны содержать как неподвижные, так и подвижные объекты с различной формой и параметрами.
3. Точность субъективных и объективных оценок должна обеспечивать возможность выявления даже малых искажений изображения в допороговой области цифровой системы.
4. Количественные критерии оценок и весовые функции должны учитывать особенности зрительного восприятия искажений.

В данном разделе рассматриваются известные методы анализа качества цветных изображений.

Глобальные меры различий. На входе цифровой видеосистемы (см. рис. 1.1) каждый элемент изображения может быть представлен непосредственно в системе (*RGB*) координат основных цветов приемника. В данном случае абсолютная мера различия между координатами цвета каждого пиксела изображения на входе и выходе системы описывается следующим выражением:

$$\Delta_{ij} = \sqrt{(R_{ij} - \tilde{R}_{ij})^2 + (G_{ij} - \tilde{G}_{ij})^2 + (B_{ij} - \tilde{B}_{ij})^2}, \quad (2.1)$$

где i, j – номер элемента кадра изображения соответственно в горизонтальном и вертикальном направлениях; *RGB* – представляют собой цифровые значения сигналов для каждого пиксела оригинального изображения; $\tilde{R}\tilde{G}\tilde{B}$ – цифровые значения сигналов искаженного изображения.

Следует отметить, что при компьютерной обработке статических изображений наиболее часто используется его описание в BMP-формате, где цвет каждого элемента (пиксела) формируется комбинацией трех *RGB* цифровых сигналов основных цветов. Например, 24-х битное представление пиксела означает, что на “красную”, “зеленую” и “синюю” составляющую отводится по 8 разрядов, т.е. по 256 градаций.

Среднеквадратичное отклонение значений пикселей (RMS) определяется как

$$RMS = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sqrt{(R_{ij} - \tilde{R}_{ij})^2 + (G_{ij} - \tilde{G}_{ij})^2 + (B_{ij} - \tilde{B}_{ij})^2}, \quad (2.2)$$

где N_x, N_y – число элементов изображения по горизонтали и вертикали.

Нормированный по отношению к исходному изображению средний квадрат ошибки (MSE), определяется выражением [38]:

$$MSE = \frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [(R_{ij} - \tilde{R}_{ij})^2 + (G_{ij} - \tilde{G}_{ij})^2 + (B_{ij} - \tilde{B}_{ij})^2]}{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [(R_{ij})^2 + (G_{ij})^2 + (B_{ij})^2]}. \quad (2.3)$$

Нормированный по амплитуде средний квадрат ошибки находится как

$$\Delta_A = \frac{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [(R_{ij} - \tilde{R}_{ij})^2 + (G_{ij} - \tilde{G}_{ij})^2 + (B_{ij} - \tilde{B}_{ij})^2]}{N_x N_y A^2}, \quad (2.4)$$

где A – максимальная амплитуда.

Обратное значение (2.4) в логарифмическом масштабе, которое часто используется на практике, называется мерой отношения сигнал-шум (PSNR):

$$PSNR = 10 \cdot \lg \frac{N_x N_y A^2}{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [(R_{i,j} - \tilde{R}_{i,j})^2 + (G_{i,j} - \tilde{G}_{i,j})^2 + (B_{i,j} - \tilde{B}_{i,j})^2]} \text{ [дБ]}. \quad (2.5)$$

Критерии верности (2.2) - (2.5) на практике часто не согласуются с субъективными оценками, т. к. зрительный анализатор неодинаково воспринимает цветовые различия для разных цветов. Например, экспериментально показано, что человеческое зрение более чувствительно к единичным изменениям зеленого и желтого цвета, чем к изменениям синего и красного [38]. Также, данные оценки являются глобальными, и не учитывает искажений локальных участков изображения, к которым зрительный анализатор может быть более чувствителен. Например, при ухудшении четкости малоразмерных объектов значение ошибки может оказаться достаточно малой величиной при низкой детальности изображения. Однако для глаза такие искажения могут оказаться существенными.

Несмотря на указанные недостатки, критерии верности воспроизведения цветных изображений по значениям ошибки (2.2) – (2.4) или по отношению (2.5) сигнал/шум широко используются для сравнительного анализа различных систем кодирования, а также, в прикладных телевизионных системах, когда свойства зрительного восприятия можно не учитывать.

Начиная с 2000 года появляются более современные метрики, такие как UQI, SSIM, MS-SSIM, GSSIM и др. Ниже приводится краткий обзор наиболее популярных метрик.

Универсальный индекс качества (UQI). Этот алгоритм был предложен Zhou Wang и др. в 2002 году [39]. Универсальность данного критерия заключается в том, что он не только отражает некоторую схожесть обработанного изображения по отношению к оригиналу, но и должным образом учитывает различные виды искажений. В отличие от многих других алгоритмов данный метод не привязан к специфике изображения и искажениям, присутствующим в нем. Он основывается на статистическом анализе блоков декодированного изображения (y) и дальнейшем сравнении полученных результатов со значениями эталонного изображения (x).

Индекс качества для сигнала S блока изображения с размером $N \times N$ можно вычислить следующим образом:

$$UQI_l = \frac{4\sigma_{xy}m_xm_y}{(\sigma_x^2 + \sigma_y^2)[m_x^2 + m_y^2]}, \quad (2.6)$$

где $m_x = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} S_{i,j}$ и $m_y = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \tilde{S}_{i,j}$ – средние значения цифровых сигна-

лов яркости или цветности в блоке оригинального (S) и искаженного ($\tilde{S}$) изображений;

$\sigma_x^2 = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (S_{i,j} - m_x)^2$ и $\sigma_y^2 = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (\tilde{S}_{i,j} - m_y)^2$ – дисперсии яркости в бло-

ках;

$\sigma_{xy} = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (S_{i,j} - m_x)(\tilde{S}_{i,j} - m_y)$ – коэффициент взаимной корреляции;

l – номер блока.

Для цветных изображений параметры m_x , σ_x , m_y , σ_y и $\sigma_{x,y}$ вычисляются раздельно по сигналам основных (R , G , B) цветов оригинального изображения и по сигналам ($\tilde{R}$, $\tilde{G}$, $\tilde{B}$) искаженного изображения. Далее параметры усредняются.

Из выражения (2.6) следует, что значение метрики UQI изменяется в интервале $[-1,1]$, при этом значение $UQI = 1$ соответствует максимально возможному качеству. Оно получается только тогда, когда $\tilde{S}_{i,j} = S_{i,j}$ для всех пикселей блока.

Индекс качества учитывает в себе три искажающих фактора: степень коррелированности пикселей двух блоков, изменение значений математического ожидания и среднеквадратичных отклонений (σ) относительно оригинала. Такое представление позволяет наиболее точно оценивать качество восстановленного сигнала и величину искажений, присутствующих в нем. В результате исходное выражение (2.6) может быть представлено в виде произведения трех множителей:

$$UQI_l = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2m_x m_y}{m_x^2 + m_y^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}. \quad (2.7)$$

Первый множитель это коэффициент взаимной корреляции сигналов $S_{i,j}$ и $\tilde{S}_{i,j}$, который определяет степень их линейной подобности и принимает значения в интервале $[-1...1]$. Второй множитель (2.7) определяет соотношение между математическими ожиданиями входного и восстановленного сигнала и принимает значения в интервале $[0...1]$, при этом максимум достигается в случае их равенства. Этот коэффициент отвечает за изменения яркости блока.

Значения среднеквадратичных отклонений σ_x и σ_y определяют динамический диапазон изменения сигналов $S_{i,j}$ и $\tilde{S}_{i,j}$ соответственно. Вследствие этого они могут рассматриваться как характеристики контраста. Таким образом, третий множитель в выражении (2.7) соответствует искажению контрастности восстановленного блока изображения относительно оригинала и принимает значения в интервале $[0...1]$.

Итоговый критерий, характеризующий качество изображения в целом, будет определяться как среднее арифметическое значений UQI_l для всех блоков L изображения:

$$UQI = \frac{1}{L} \sum_{k=1}^L UQI_l. \quad (2.8)$$

Предложенный индекс качества обладает хорошей согласованностью со значениями субъективных экспертных оценок. К тому же, он имеет абсолютное значение, равное единице, то есть при показаниях UQI , близких к единице, можно говорить о высоком качестве восстановленного изображения.

Метрика SSIM. Метрика SSIM сводится к определению степени сходства соответствующих участков (окон) сравниваемых изображений по трем составляющим:

$$\text{яркость } L_{xy} = \frac{2m_x m_y + c_1}{m_x^2 + m_y^2 + c_1}; \text{ контраст } K_{xy} = \frac{2\sigma_x \sigma_y + c_1}{\sigma_x^2 + \sigma_y^2 + c_1}; \text{ структура}$$

$$C_{xy} = \frac{\sigma_{xy} + c_1}{\sigma_x \sigma_y + c_1},$$

где $c_1 = 0,01$ выравнивающий коэффициент, предотвращающий деление на число, близкое к нулю, при высоком качестве оцениваемого изображения.

Разница между двумя блоками x и y с размером $N \times N$ определяется путем перемножения этих трех составляющих [40]

$$SSIM_l = \frac{(2m_x m_y + c_1)(2\sigma_{xy} + c_1)}{(m_x^2 + m_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_1)}. \quad (2.9)$$

Как правило, метрика (2.9) рассчитана на окно с размером 8×8 пикселей. Для анализа всего изображения окно смещается на заданное число пикселей последовательно по горизонтали и по вертикали. Интегральная оценка определяется как

$$SSIM = \frac{1}{L} \sum_{l=1}^L SSIM_l, \quad (2.10)$$

где l – номер окна (блока); L – количество окон.

Значения (2.9) и (2.10), аналогично (2.7) и (2.8) равны единице только при полном совпадении параметров.

Существуют различные модификации метрики SSIM, ниже приводится их краткое описание.

Многомасштабная метрика MS-SSIM. Это улучшенная версия SSIM. Для M ступеней индекса MS-SSIM, процедура включает в себя M итераций. Во время каждой итерации, эталонное и искажённое изображения, проходят через фильтр нижних частот, проводится децимация отсчётов изображения с коэффициентом 2, вычисляются частные показатели по контрастности и структурности и перемножаются. Этот процесс повторяется $M-1$ раз, на M -м этапе, вводится индекс сравнения по светлотной составляющей. MS-SSIM имеет точность прогнозирования более высокого качества чем одноступенчатый SSIM, но вычислительная сложность высока.

Градиентный индекс структурного подобия (G-SSIM). Глаз человека очень чувствителен к информации о краях и контурах изображения. G-SSIM представляет собой усовершенствованный вариант SSIM, где компоненты контраста K_{xy} и структурной составляющей C_{xy} из оригинального изображения (x) и искажённого изображения (y) SSIM заменяются их градиентными значениями [41]. С этой целью для создания градиентного изображения используется оператор Собеля и далее вычисляются параметры метрики SSIM над преобразованными изображениями. GSSIM показывает лучшую производительность по сравнению с SSIM, особенно для размытых изображений.

Метрики F-SSIM и P-SSIM. Эти алгоритмы используются для улучшения SSIM путём включения визуальной важности различных областей изображения. Гипотеза состоит в том, что некоторые регионы в изображении визуально более важные, чем другие. Следовательно, область интересов, оценка качества на основе может повысить производительность SSIM. Три улучшенные версии SSIM в рамках этой категории являются фиксированная SSIM (F-SSIM), процентная SSIM (P-SSIM) и PF-SSIM, который представляет собой комбинацию из этих двух. С помощью этого метода было показано, что соотношение SSIM, относительно субъективных результатов испытаний, было улучшено.

Контент распределение SSIM (4-SSIM). Это улучшенная версия SSIM или MS-SSIM, где изображение сегментируется в четырёх категориях регионов, таких как изменённые границы, сохранившиеся границы, текстуры и гладкие области. Веса применяются к значениям SSIM по этим областям. Взвешенные

значения SSIM объединяют, чтобы получить единый индекс для качества изображения. В зависимости от того как произошло распределение различного содержания, для SSIM используются MS-SSIM, G-SSIM или MS-G-SSIM. Разделённые показатели качества, а именно 4-SSIM, 4-MC-SSIM, 4-G-SSIM или 4-MC-G-SSIM показали улучшенную совместимость с субъективными тестами по сравнению с метриками G-SSIM или MS-SSIM.

Индекс качества на основе локальной дисперсии (QILV). Этот метод основан на предположении, что большое количество структурной информации изображения закодировано в его локальном распределении дисперсии [42]. В этой процедуре, локальная дисперсия изображения вычисляется с использованием взвешенной окрестности. Среднее значение и стандартное отклонение локальной дисперсии рассчитываются для обоих изображений. Наконец, также рассчитываются ковариация локальной дисперсии между эталонным и искажённым изображениями. Индекс качества QILV рассчитывается аналогичным способом, как SSIM с той лишь разницей, что среднее и стандартное отклонение локальной дисперсии используются вместо среднего значения и стандартного отклонения значений элементов изображения. Этот алгоритм работает лучше, чем SSIM, особенно в случае размытого изображения.

Также в литературных источниках [43] можно найти описание других подобных метрик (FSIM, EGSM, VSNR, PIQA и др.). Анализ метрик UQI, SSIM и им подобным показывает, что все эти метрики учитывают различия между окнами двух изображений. Однако отклонения параметров μ_x , σ_x от μ_y , σ_y и параметра $\sigma_{x,y}$, при наличии искажений, определяются только отклонениями сигналов основных цветов (R , G , B) и, следовательно, не учитывают свойств чувствительности зрения к изменениям этих параметров.

Таким образом, метрики UQI и SSIM, как и им подобные метрики, не дают однозначного ответа на следующий вопрос. При каком их значении будут незаметны или малозаметны искажения структур реалистичных изображений?

Методы оценки качества без эталонного изображения. Все описанные выше метрики требуют наличие эталонного изображения. Тем не менее многие исследователи рассматривают и предлагают алгоритмы оценки качества изображения без наличия эталона.

Метод оценки без эталона или, ещё, этот метод носит название метода оценки «в слепую», (No-Reference IQA (NR-IQA or Blind IQA)) относится к оценке качества изображения без эталонного изображения. Эти алгоритмы предсказывают качество изображения без каких-либо знаний об опорном изоб-

ражении и не очень хорошо коррелируют с человеческим восприятием качества [44]. Наша зрительная система может отличить высококачественные изображения от изображений с низким качеством, не видя исходного изображения. В нашем мозговом анализирующем аппарате есть модели высокого качества эталонных изображений, и мы имеем возможность использовать эти модели для оценки качества изображения [45].

Существуют три основных подхода к NR-IQA, основанные на том, каким способом получена объективная оценка качества [46], [47]. К ним относятся:

1) Подход к конкретным искажениям: этот метод использует конкретный объективный показатель искажения для предсказания субъективной оценки качества. Примерами таких типов искажений являются: размытие мелких структур [48], блокинг-эффект, выбросы на переходах и т.д.

2) Выделение признаков и последующее обучение: такой подход выделяет особенности из образов и обучает алгоритм как отличать искажённые изображения от неискажённых.

3) Подход, основанный на статистике естественных сцен (NSS): этот подход основан на гипотезе о том, что естественные изображения (например, изображения, снятые с помощью оптической камеры) обладают определёнными статистическими свойствами между значениями элементов изображения, и что наличие искажений изменяет эти статистические характеристики. Характеризуя эту неестественность и используя статистические данные сцены, можно выделить искажение и выполнить предсказание оценки без эталонным методом NR-IQA [49].

Стоит отметить, что в настоящее время нет стандартного алгоритма NR-IQA, так как не было доказано, что получаемые оценки надёжны. Примеры некоторых популярных алгоритмов для NR-IQA являются BIQL, BLINDS, DIVINE и Brisque [50].

2.4. Оценка качества видеоданных

Как и метрики оценки качества изображений, метрики оценки качества видеоданных можно разделить на две категории: объективные и субъективные. Объективные метрики качества делятся на эталонные метрики (full reference, FR), сокращённые эталонные (reduced reference, RR) и неэталонные (no-reference, NR). Оценка, основанная на эталонной метрике, вычисляется с учётом оригинального видеофрагмента. Сокращённая эталонная метрика использует для сравнения не исходные видеоданные, а некий набор его априорных ха-

рактических, как, например, информацию о его текстуре. Неэталонные метрики не требуют исходного видеоизображения. Они используют только информацию из битового потока и оценивают её на уровне пикселей. Именно такие метрики наиболее востребованы в практических приложениях.

Неэталонные метрики можно разделить на три категории: основанные на пикселях, на битовом потоке и гибридные метрики.

Неэталонные метрики, основанные на пикселях, используют измерение артефактов (размытие, блочность, звон, шум, временные нарушения) и признаков (статистика естественных сцен и другие пространственные признаки).

Неэталонные метрики оценки качества видео, основанные на битовом потоке, используют модели с соответствующими параметрами. Для параметрической модели планирования такими параметрами являются битрейт и частота потери пакетов, которые используются и в двух других моделях. Для модели уровня битового потока к ним добавляется вектор движения, а для параметрической модели пакетного уровня – кадровая частота и другие параметры.

Гибридные метрики, основанные на пикселях и битовом потоке, используют артефакты пиксельного и поточного уровней (пространственные признаки и параметры битового потока), а также статистику коэффициентов преобразования (статистику, основанную на коэффициентах ДКП или вейвлет-преобразования).

При использовании методов, основанных на признаках, оценка качества зависит от наличия на видеоизображении артефактов, отрицательно влияющих на визуальное восприятие.

Неэталонные оценки, основанные на битовом потоке, проще в вычислении. Алгоритм вычисления оценок такого типа зависит от конкретного метода кодирования видео и формата битового потока. Такой подход удобен для контроля потоковой передачи видеоданных.

Неэталонные метрики уровня битового потока делятся на три категории: параметрическая модель планирования, параметрическая модель пакетного уровня и модель уровня битового потока.

Модель планирования не имеет доступа к самому битовому потоку, но производит оценку, основываясь на битрейте, типе кодека и частоте потери пакетов. Такая модель менее сложна, чем модель пакетного уровня. Модель пакетного уровня имеет доступ к битовому потоку и может извлечь такие параметры, как частоту и типы кадров, частоту потери пакетов, текущий битрейт. Модель пакетного уровня также известна, как модель качества сервиса (QoS).

Модели уровня битового потока имеют доступ к большей части данных, используемых для оценки качества видео, кроме информации пиксельного уровня. На вход модели подаются параметры из заголовка и полезной нагрузки пакета.

Гибридный подход является комбинацией двух предыдущих подходов. Он преследует цель достичь как простоты вычисления, свойственной основанной на битовом потоке оценке, так и робастности оценки пиксельного уровня.

Оценки различных артефактов (см. п.1.5), обнаруживаемых на видеоданных, могут быть скомбинированы для предсказания общего воспринимаемого качества видеофрагмента.

Важность таких метрик подчеркивает тот факт, что некоторые преобразования над изображениями, например, сжатие по стандарту JPEG, приводят к появлению сразу нескольких видов артефактов. Другое приложение таких метрик – случай, когда обработка изображения с целью удалить один артефакт (например, шумоподавление) приводит к возникновению другого артефакта.

Существуют разные подходы к тому, как следует комбинировать эффекты от различных типов повреждений. Согласно рекомендации ITU-T P.1202.1, четыре типа повреждений предлагается отсортировать согласно их степени влияния. Значения первых двух наиболее важных типов объединяются при помощи линейной комбинации с большим коэффициентом у первого. Поскольку измерения различных артефактов проводится в различных диапазонах значений, важно привести их к одному масштабу. Помимо использования линейной комбинации, в некоторых статьях предлагается использовать для объединения метрику Минковского.

Выбор конкретной стратегии объединения может зависеть от многих факторов, включая относительную критичность появления артефакта, пространственно-временные характеристики содержимого и наличие маскирующих эффектов. Линейная комбинация лучше подходит, если её составляющие могут быть соотнесены при помощи линейной регрессии. С другой стороны, метрика Минковского исходит из аддитивных свойств зрительной системы человека.

Некоторые исследователи пытаются интерпретировать различие в изображениях на основе существующих представлений о зрительной системе человека. Специалисты корпорации Sarnoff — исследовательского института, разработали методику под названием JND (Just Noticeable Difference — «едва заметное различие»), которая, по их мнению, дает «оценки качества изображения, адекватные зрительному восприятию». В методе JND анализируются простран-

ственные, временные и цветовые изменения в тестовом изображении при прохождении его через исследуемую систему и определяется, влияют ли эти изменения на восприятие изображения зрителем.

Ниже приводится описание методики JND по материалам [51].

На входе анализатора каждое поле исходной тестовой видеопоследовательности представлено в виде цифровых RGB сигналов (рис. 2.5.).

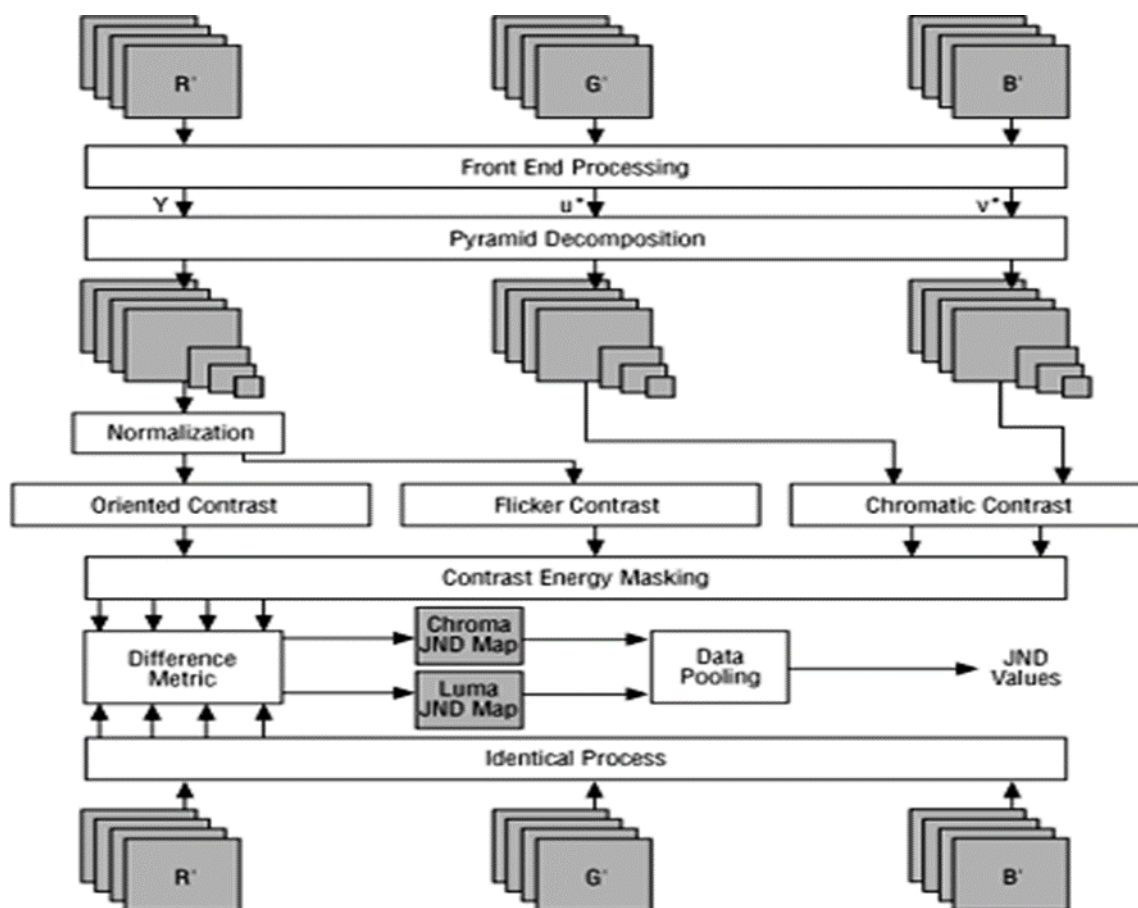


Рис. 2.5. Структура анализатора PQA по методике JND

На первой стадии обработки (Front End Processing) выполняется преобразование RGB сигналов в сигналы яркости (Y) и цветности (u^*, v^*), где сигналы цветности представляют собой значения равноконтрастного цветового пространства $L^*u^*v^*$. На следующем этапе (Pyramid Decomposition) сигналы яркости и цветности каждого поля фильтруются методом «пирамиды Гаусса», что приводит к их разбиению на субдиапазоны с различным пространственным разрешением. Такая фильтрация позволяет выделить ту информацию, которая сильнее всего влияет на зрительное восприятие.

На этапе нормализации (Normalization) вычисляется среднее значение яркости для каждого поля видеопоследовательности, что позволяет учесть такой фактор, как понижение чувствительности зрения, когда яркая сцена сменяется темной.

Затем проводятся три отдельных измерения контраста с учетом пространственной ориентации (Oriented Contrast), мерцания (Flicker Contrast) и цветового смещения (Chromatic Contrast). Контраст рассчитывается по локальным разностям значений пикселей с масштабированием для каждого уровня "пирамиды" и, если численное значение контраста для локальной разности равно 1, считается, что данный результат приблизительно соответствует порогу зрительного восприятия. Тем разностям, которые в соответствии с моделью, будут замечены наблюдателем с вероятностью 75%, присваивается численное значение 1 JND. Различимые отсчеты изображения передаются на следующий этап вычислений.

«Маскирование энергии контраста» (Contrast Energy Masking) позволяет учесть понижение чувствительности зрительной системы человека к искажениям в сложных областях изображения. Например, искаженные элементы изображения, находящиеся в области кроны дерева, учитываются с меньшим весом, чем такие же искажения на гладкой поверхности.

После фильтрации последовательности изображений, в результате которой устанавливаются области наиболее подверженные заметным искажениям, определяется разность (Difference Metric) между данными исходной эталонной последовательности и обработанной, аналогичным образом (Identical Process), тестовой последовательности. При этом формируются две карты JND: одна — для заметных различий яркости (Luma JND Map), вторая — для заметных цветовых различий (Chroma JND Map). Специалисты Sarnoff Corporation считают, что эти карты JND, в отличие от простых разностей изображений, отражают те изменения, которые будут заметны для человека. Чем заметнее изменение, тем ярче соответствующая область на карте.

Разработанный Sarnoff Corporation метод JND произвел большое впечатление на фирму Tektronix, считающуюся лидером в области тестирования видеосигналов. Эта компания настолько высоко оценила новый метод, что включила модель JND в одно из своих изделий – анализатор изображений PQA (Picture Quality Analyzer).

Наиболее современный анализатор PQA-600C и программный пакет PQASW работает с двумя блоками видеоданных: тестовым опорным видеосиг-

налом и версией этого опорного сигнала после компрессии, технической обработки или иного действия, ухудшающего качество изображения. Сначала PQA-600C производит пространственное и временное нормирование этих двух сигналов. Далее с помощью измерений, базирующихся на модели зрительной системы человека, проводится анализ качества изображения по методике JND. Она выдает значения параметров качества сигнала, информацию о пиковом отношении сигнал/шум (PSNR), отображает карты JND для изображений и вычисляет параметр PQR (Picture Quality Rating – рейтинг качества изображения) для исследуемой системы на базе модели JND.

Нулевая оценка PQR означает, что система не обнаружила заметных различий между исходными и обработанными изображениями. Оценка между 0 и 4 свидетельствует о незначительном различии. Если PQR равна 1, то 50% зрителей заметят очень слабое различие между исходными и обработанными изображениями. Оценки выше 10 говорят о существенном (но, возможно, еще приемлемом) различии.

Результаты анализа содержат общие показатели качества, показатели покадровых измерений и карту нарушений качества по каждому кадру, а также измерения общепринятых параметров качества изображения, например, PSNR.

Используемые тестовые видеосигналы могут иметь различное разрешение и скорость кадров. Так, можно измерять изменение качества изображения между HD и SD сигналами и т. д. Прибор может работать в двух режимах - режиме измерений и режиме просмотра. Нужный режим выбирается через конфигурационное меню. В процессе измерений резюмирующие данные и карта результатов отображаются на дисплее и могут быть сохранены на системном жестком диске. Режим просмотра используется для контроля ранее сохраненных резюмирующих данных и карт, созданных либо в режиме измерений или в процессе исполнения XML скриптов. В этом режиме пользователь может выбрать форму многократного отображения и сравнивать результаты, расположив их рядом на дисплее. Таким способом легко сравнивать изменение качества изображения после компрессии.

Главный вопрос состоит в том, действительно ли оценки PQR соответствуют тому, как оценят качество изображения зрители? Конечно, Tektronix и Sarnoff Corporation считают, что да. Конечно, краткий обзор метода JND не дает полного представления о достоинствах и недостатках. К сожалению, более подробная информация о методе JND практически отсутствует в литературных

источниках. Тем не менее, анализ обзорного материала позволяет сделать следующие замечания.

Высокочастотные субдиапазоны сигналов яркости и цветности, полученные после их фильтрации методом «пирамиды Гаусса», содержат информацию о резких границах и текстурах с мелкими деталями. Однако такая информация позволяет лишь локализовать различимые отсчеты для данного уровня «пирамиды» с контрастом более 1 JND и не содержит данных о размерах деталей и протяженности границ. Поэтому, результаты анализа карт JND по интегральным оценкам для яркости и цветности не дают объективных оценок ухудшения визуальной четкости и резкости декодированной видеопоследовательности. Также, методика JND на этапе нормализации не учитывает особенности зрительного восприятия контраста малоразмерных динамических объектов.

На основе метрик качества изображений ряд ведущих фирм предлагает аппаратуру для объективного анализа сжатых цифровых видеопотоков. Например, фирма Rohde & Schwarz аппарат DVQ, фирма Pixelmetrix прибор DVStation, фирма Snell & Wilcox программное обеспечение Mosalina. Все это оборудование позволяет анализировать и объективно оценивать сжатые цифровые потоки (MPEG-2, MPEG-4, H.264/AVC, H.261, H.263, H.263+ и 3GPP) с точностью, в разной степени, совпадающей с субъективными оценками.

Принципы оценки качества изображений после сжатия у этих приборов различны. Например, анализатор фирмы Tektronix PQA -600C, описанный выше, представляет последовательность изображений как тест последовательность с реперными метками и оценивает ее как тест-объект, определяя условную величину PQR (Picture Quality Rating), хорошо совпадающую с субъективными оценками.

В анализаторе фирмы Rohde & Schwarz применяется способ оценки различия между соседними блоками пикселей и макроблоками. При этом предполагается, что в несжатом сигнале различия распределены равномерно. В зависимости от величины сжатия происходит расщепление разностей, и оно возрастает с ростом артефактов.

Анализатор фирмы Pixelmetrix оценивает серию изображений, выбранную из сжатого потока, и определяет ошибки четкости, блочную структуру, «замороженные изображения» и также отношения с/ш в сигналах яркости и цветности.

2.5. Метрика оценки качества передачи мелких структур и границ крупных деталей изображения

В этом разделе приводится описание метрики оценки качества реалистичных изображений, предложенных автором (С.В. Сай) в качестве альтернативных существующим метрикам.

2.5.1. Система N-CIELAB

Четкость относится к основным показателям качества изображения и зависит от разрешающей способности видеосистемы, которая традиционно оценивается количеством различимых глазом штриховых линий тестового изображения (таблицы). При этом используются различные типы таблиц как статические, так и динамические. Увеличение количества пикселей формата видеокadra до стандартов HD приводит к повышению четкости изображения, но при этом для обеспечения необходимой скорости передачи требуется применение более эффективных методов компрессии. В процессе компрессии формат кадра не меняется, но при этом (при высоких коэффициентах сжатия) обычно «размываются» или «теряются» мелкие детали, что в сильной степени ухудшает визуальное качество изображения. Ухудшение четкости в свою очередь приводит к снижению резкости границ крупных деталей изображения.

Для оценки ухудшения четкости изображения необходимо оценить снижение контраста мелких деталей компрессированного изображения по отношению к исходному. В отличие от традиционных способов определения контраста изображения, мы предлагаем использовать равноконтрастные системы координат, в которых результаты измерений мало зависят от цвета передаваемой детали и фона, на котором она расположена.

Метрика оценки цветового контраста мелких деталей эталонного изображения, например, в равноконтрастной системе координат Wyszecki с учетом весовых коэффициентов по яркости и цветности определяется как:

$$\Delta K_{wuv} = 3\sqrt{((\Delta W^*)^2 + (\Delta U^*)^2 + (\Delta V^*)^2)},$$

где $\Delta W^* = (W_o^* - W_f^*) / W_{th}^*$ нормированное значение контраста по индексу яркости и $\Delta U^* = (U_o^* - U_f^*) / U_{th}^*$, $\Delta V^* = (V_o^* - V_f^*) / V_{th}^*$ по индексам цветности; W_o^* , U_o^* и V_o^* цветовые координаты мелкого объекта тестового изображения; W_f^* , U_f^* и V_f^* цветовые координаты фона; W_{th}^* , U_{th}^* и V_{th}^* весовые коэффици-

енты, которые определяются количеством минимальных цветовых порогов (MPCD).

Цветовые координаты по индексам яркости и цветности вычисляются по формулам:

$$W^* = 25 Y^{1/3} - 17; U^* = 13 W^* (u - u_o); V^* = 13 W^* (v - v_o),$$

где Y значение светлоты в интервале от 1 до 100; u и v координаты цветности в диаграмме Mac-Adam; u_o и v_o координаты основного белого цвета $u_o = 0.201$ и $v_o = 0.307$.

Контрастная чувствительность зрения падает с уменьшением размера деталей, следовательно, значения весовых коэффициентов будут возрастать. В общем случае, их значения будут зависеть от яркости фона, шума, маскирующего эффекта, условия наблюдения и др.

Для мелких деталей с размерами не более одного пикселя пороговые значения, полученные экспериментально. В частности для мелких деталей, расположенных на сером фоне ($70 < W_f^* < 90$) пороговые значения приблизительно равны $\Delta W_{th}^* \approx 6$ MPCD, $\Delta U_{th}^* \approx 72$ и $\Delta V_{th}^* \approx 80$ MPCD.

Метрика Wyszecki не оценивает цветовые различия между двумя изображениями, а определяет нормированный цветовой контраст мелких деталей в пределах одного изображения. Если условие $\Delta K_{wuv} > 1$ выполнено, то мелкие детали различаются глазом, и результат слабо зависит от цвета деталей.

Результаты исследования привели к выводу, что применение цветового пространства $W^*U^*V^*$ имеет следующие ограничения:

1. Значения весовых коэффициентов W_{th}^* , U_{th}^* и V_{th}^* в значительной степени зависят от яркости фона, что затрудняет выполнять распознавание и сегментацию мелких деталей в фотореалистичных изображениях с адекватными результатами к визуальному восприятию.

2. Система $W^*U^*V^*$ в настоящее время практически применяется редко, т.к. появились более новые системы CIELAB, CIEDE2000, S-CHIELAB и другие для оценки цветовых различий.

В работе [52] предложена метрика MFSD (Metric of Fine Structures Distortion) искажений мелких структур компрессированных изображений.

К особенностям метрики относится то, что она основана на алгоритме идентификации мелких структур изображения с использованием нормирован-

ной системы N-CIELAB, что позволяет оценивать искажения с учетом свойств контрастной чувствительности зрения.

Колометрическая система CIELAB в настоящее время является международным стандартом и широко используется для оценки цветовых различий между оригинальными и искаженными изображениями. Цветовые различия вычисляются как

$$\Delta E = \sqrt{(L^* - \tilde{L}^*)^2 + (a^* - \tilde{a}^*)^2 + (b^* - \tilde{b}^*)^2}, \quad (2.11)$$

где $(L^* a^* b^*)$ – цветовые координаты объекта оригинального изображения и $(\tilde{L}^* \tilde{a}^* \tilde{b}^*)$ – цветовые координаты объекта искаженного изображения.

В качестве объектов тестовых изображений обычно используются цветовые полосы или поля с постоянными значениями яркости и цветности для всех пикселей внутри объекта. Значение $\Delta E \approx 2,3$ приблизительно соответствует минимальной ощутимой разнице цвета для человеческого глаза.

Цветовые координаты могут быть получены с помощью преобразования основных цветов (RGB) в цветовое пространство (XYZ), а затем с помощью формул:

$$\begin{aligned} L^* &= 116 f(Y/Y_n) - 16; \quad a^* = 500[f(X/X_n) - f(Y/Y_n)]; \\ b^* &= 200[f(Y/Y_n) - f(Z/Z_n)], \end{aligned}$$

где

$$f(t) = \begin{cases} t^{1/3}, & \text{if } t > 0,008856 \\ 7,787t + 16/116, & \text{if } t \leq 0,008856 \end{cases}$$

Значения X_n , Y_n и Z_n являются координатами эталонного белого цвета. Равноконтрастные цветовые пространства CIELAB, CIELUV, CIEDE2000 и другие традиционно используются для оценки искажения цветопередачи крупных объектов, которые имеют внутри равномерное распределение цвета. Преимущество таких цветовых пространств является то, что результаты оценки слабо зависят от цвета объекта.

Очевидно, что оценка цветовых различий, используемая в (2.12) не дает верных результатов с уменьшением размера объекта, поскольку не учитывает снижение контрастной чувствительности зрения по яркости и цветности.

В [53] предложена нормированная система N-CIELAB и метрика для оценки цветового контраста мелких деталей с учетом весовых коэффициентов по яркости и цветности:

$$K_{LAB} = \sqrt{\left(\frac{L_o^* - L_f^*}{L_{th}^*(\delta)}\right)^2 + \left(\frac{a_o^* - a_f^*}{a_{th}^*(\delta)}\right)^2 + \left(\frac{b_o^* - b_f^*}{b_{th}^*(\delta)}\right)^2}, \quad (2.12)$$

где L_o^* , a_o^* и b_o^* – цветовые координаты мелкой детали; L_f^* , a_f^* и b_f^* – цветовые координаты окружающего фона; $L_{th}^*(\delta)$, $a_{th}^*(\delta)$ и $b_{th}^*(\delta)$ – весовые коэффициенты по яркости и цветности, определенные количеством минимальных цветовых порогов MPCD; δ – размер мелких деталей, определяемый числом пикселей изображения.

Формула (2.12), в отличие от формулы (2.11), не оценивает цветовые различия между двумя изображениями, а определяет нормированный цветовой контраст мелкой детали в пределах одного изображения. Для крупной детали весовые коэффициенты будут равны единице ($L_{th}^*(\delta) = a_{th}^*(\delta) = b_{th}^*(\delta) = 1$) и, следовательно, пороговое значение цветового контраста будет соответствовать величине $K_{LAB}^{th} = \Delta E \approx 2,3$. С уменьшением размера детали пороговый контраст будет увеличиваться за счет снижения контрастной чувствительности зрения. Весовые коэффициенты в формуле (2.13) подобраны таким образом, чтобы получить нормированное значение порогового контраста равным единице: $K_{LAB}^{th} = 1$. При этом должно выполняться условие: если контраст мелкой детали превышает пороговое значение

$$K_{LAB} > K_{LAB}^{th} = 1, \quad (2.13)$$

то деталь различается глазом относительно окружающего фона.

Для экспериментальной оценки весовых коэффициентов были разработаны тестовые таблицы (рис. 2.6) с мелкими деталями, расположенными на неокрашенном фоне.

Контраст мелких деталей задавался отдельно по трем координатам L^* , a^* и b^* . Для визуализации изображения в формате *.bmp выполнялось попиксельное преобразование цветовых координат L^* , a^* , b^* в RGB . При проведении экспериментов использовалась следующая методика. Для каждого эксперимента использовались тестовые таблицы с фиксированными размерами мелких деталей – 1 пиксель, 2×2 пикселя и 4×4 пикселя. При этом пространственное положение деталей и их количество задавалось случайным образом. Светлота фона выбрана как среднее значение: $L_f^* = 50$.

Для экспериментальной оценки весовых коэффициентов были разработаны тестовые таблицы с мелкими деталями, которые расположены на неокрашенном фоне. Контраст мелких деталей задавался отдельно для каждой координаты L^* , a^* и b^* . Для визуализации изображения использовалась формат BMP, для которого выполнялось попиксельное преобразование L^* , a^* , b^* в RGB сигналы.

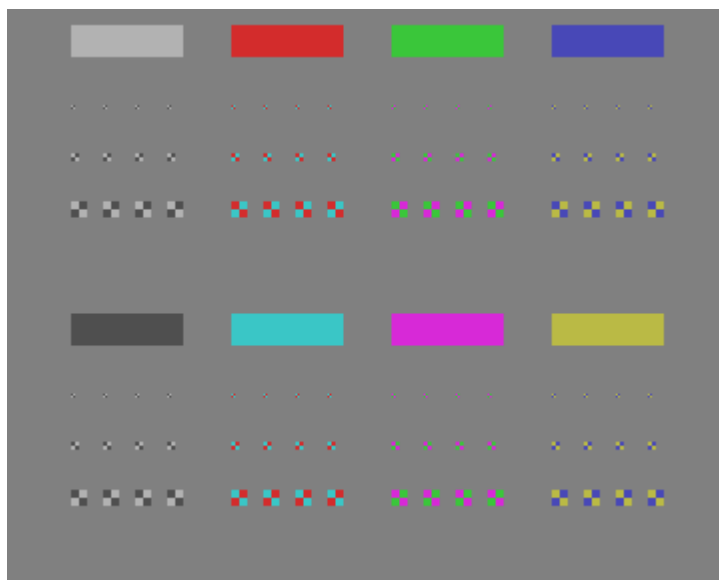


Рис. 2.6. Пример тестовой таблицы

В экспериментах использована следующая методика. Тестовые таблицы с фиксированными размерами мелких деталей (1 пиксель, 2×2 пикселя, 3×3 пикселя и 4×4 пикселя) были использованы для каждого эксперимента. Пространственное положение и количество мелких деталей были определены случайным образом. Яркость фона установлена как среднее значение $L_f^* = 50$.

Наблюдатель через программный интерфейс изменял контраст деталей по координате L^* (a^* или b^*) для выбранной таблицы. Этот контраст изменялся от нуля до некоторого значения при котором мелкие детали начинали различаться наблюдателем. В конце эксперимента наблюдатель фиксировал свои субъективные оценки порогового контраста L_{th}^* (a_{th}^* or b_{th}^*).

В табл. 2.4 представлены средние значения весовых коэффициентов, полученные от 20 наблюдателей. Снижение характеристики контрастной чувствительности зрения в зависимости от размера деталей показано на рис. 2.7.

Таблица 2.4. Зависимость весовых коэффициентов от размера деталей

δ	L_{th}^*	a_{th}^*	b_{th}^*
> 4	1	3	4
4	1	10	15
3	2	16	20
2	3	20	30
1	6	40	55

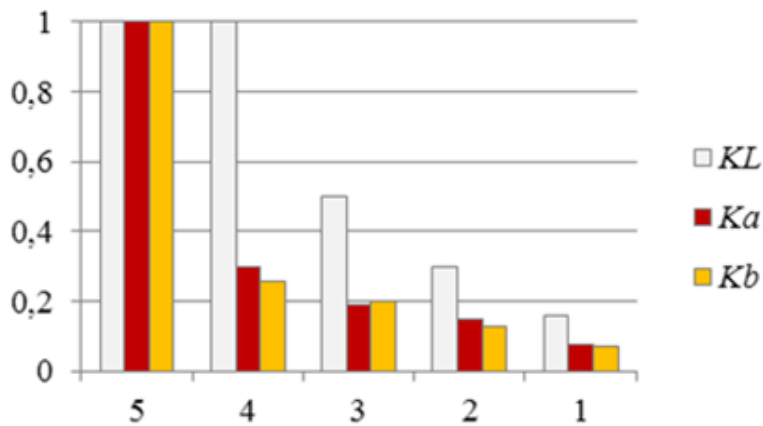


Рис. 2.7. Зависимость контрастной чувствительности от размера деталей:

$$KL = 1 / L^*_{th}(\delta), Ka = 3 / a^*_{th}(\delta) \text{ and } Kb = 4 / b^*_{th}(\delta)$$

В результате экспериментов получено, что при изменении яркости фона весовые коэффициенты меняются несущественно для мелких деталей с размером $\delta > 1$. Однако для мельчайших деталей ($\delta = 1$) значения весовых коэффициентов по координатам $\pm a^*$ и $\pm b^*$ находятся на границе их диапазона. Возникает вопрос – следует ли учитывать при этом цветовые координаты в формуле (2.12)?

Здесь следует отметить, что в фотореалистичных изображениях изменение координат цветности для мелких деталей без изменения яркости происходит крайне редко. Поэтому мы считаем, что правильно будет, при оценке цветового контраста мельчайших деталей, учитывать координаты a^* и b^* в формуле (2.12) с полученными весовыми коэффициентами, так как они вносят свой вклад в окончательное значение контраста.

Для подтверждения корректности формулы (2.12) была проведена серия экспериментов. Мы оценили пороговые значения контраста мелких деталей для различных цветов. В табл. 2.5 приведены экспериментальные результаты для мелких деталей ($\delta = 1$) для первичных и дополнительных цветов. Зависимости порогового контраста от цвета и яркости фона показаны на рис. 2.8.

В программном обеспечении интерфейса тестовой таблицы использованы следующие функции: установка яркости фона (Y_f) в диапазоне 0...255; установка цвета мелких деталей в системе RGB ; изменение контраста мелких деталей в системе RGB относительно фона; преобразование сигналов из RGB в $L^* a^* b^*$ и вычисление контраста с помощью (2.12).

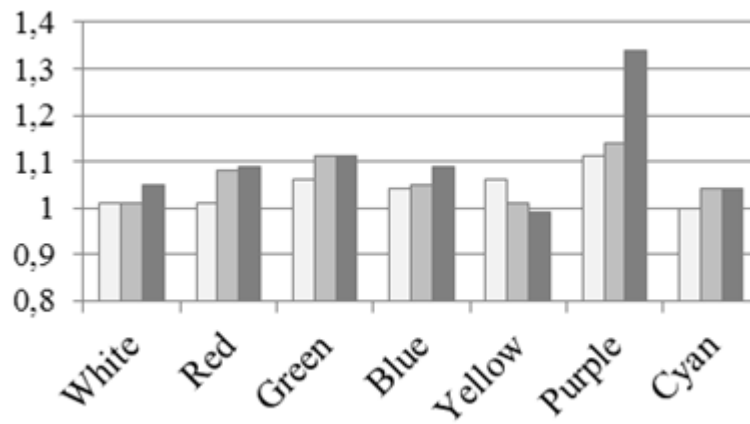
Таблица 2.5. Значения координат цвета и пороговых значений контраста

Color	R	G	B	L^*	a^*	b^*	K_{LAB}	K_{WUV}	Y_f
white	16	16	16	6.0	0	0	1.01	1.00	150
	15	15	15	6.1	0	0	1.01	1.22	100
	14	14	14	6.3	0	0	1.05	1.73	50
red	27	0	0	-5.1	21.0	8.5	1.01	1.30	150
	29	0	0	-5.2	24.7	10.9	1.08	1.85	100
	40	0	0	-3.3	34.8	22.7	1.09	4.59	50
green	0	24	0	4.6	-25.7	19.6	1.06	0.88	150
	0	23	0	5.0	-25.8	20.3	1.11	1.08	100
	0	20	0	5.2	-23.8	19.8	1.11	1.43	50
blue	0	0	18	-5.8	7.9	-19.0	1.04	1.01	150
	0	0	17	-5.8	8.4	-19.3	1.05	1.26	100
	0	0	16	-6.0	9.9	-20.3	1.09	1.94	50
yellow	18	18	0	5.9	-6.1	18.7	1.06	0.95	150
	16	16	0	5.7	-5.6	17.7	1.01	1.09	100
	14	14	0	5.6	-5.1	17.1	0.99	1.46	50
purple	30	0	30	-3.6	33.3	-22.3	1.11	1.06	150
	31	0	31	-3.3	36.1	-23.6	1.14	1.37	100
	48	0	48	1.2	48.3	-29.9	1.34	2.95	50
cyan	0	23	23	5.5	-15.2	-5.0	1.00	1.02	150
	0	22	22	5.8	-15.1	-4.8	1.04	1.23	100
	0	19	19	5.9	-13.4	-4.2	1.04	1.57	50

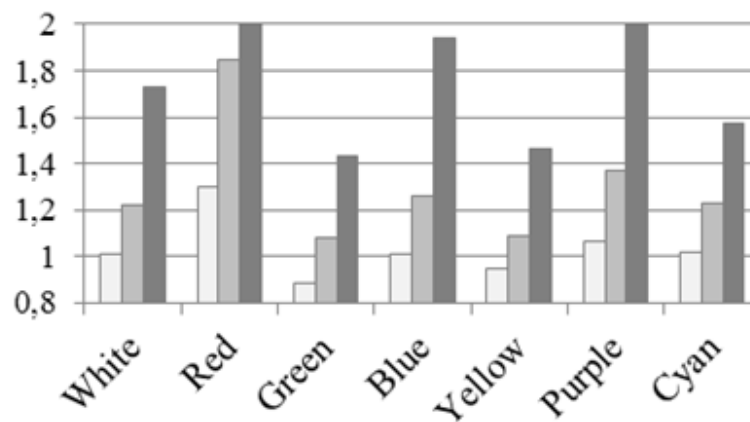
Для заданных параметров цвета и фона наблюдатель изменял контраст мелких деталей, пока они не становились различимыми. В этот момент наблюдатель фиксировал значение контраста K_{LAB} . Значения контраста K_{WUV} в системе $W^*U^*V^*$, приведены для сравнения (рис. 2.8, б).

Из анализа результатов (рис. 2.8, а) можно сделать вывод о том, что пороговый контраст K_{LAB} незначительно отклоняется от единицы ($K_{LAB} = 1$) с изменением цвета детали и яркости фона. Максимальное отклонение порога (+0.34) получается для фиолетовых мелких деталей на темном фоне ($Y_f = 50$).

Пороговое значение контраста K_{WUV} в системе с $W^*U^*V^*$ существенно зависит от яркости фона и имеет большой разброс в зависимости от цвета мелких деталей с уменьшением яркости фона. Максимальное отклонение порога (+3.59) получается для красных мелких деталей на темном фоне ($Y_f = 50$).



а) K_{LAB}



б) K_{WUV}

Рис. 2.8. Зависимость порогового контраста от цвета и яркости фона

Таким образом, пороговые значения системы координат CIELAB более точно соответствуют визуальной модели, и применение системы CIELAB является более предпочтительным в задачах сегментации мелких деталей.

Как показали результаты экспериментов, по сравнению с другими равноконтрастными системами [54], значение порогового контраста K_{LAB}^{th} незначительно отклоняется от единицы при изменении цвета мелких деталей и яркости фона.

Таким образом, использование системы N-CIELAB в задачах идентификации мелких структур изображения является более предпочтительным.

2.5.2. Алгоритм идентификации мелких структур

Алгоритм идентификации мелких структур основан на способе анализа структуры микроблоков изображения с размером 3×3 пикселя. Мелкие структуры в микроблоке можно разделить на следующие типы: точечные объекты, фрагменты тонких линий, фрагменты контуров, фрагменты текстур. На рис. 2.9 показаны примеры мелких структур для бинарного изображения, где цифрами указаны номера пикселей в микроблоке.

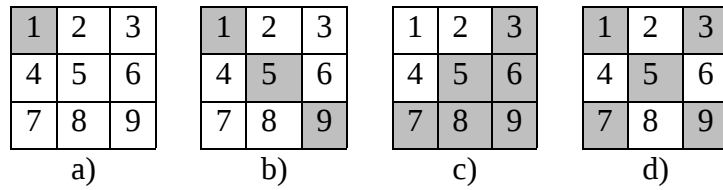


Рис. 2.9. Примеры мелких структур

Из рис. 2.9 видно, что максимальное количество цветовых переходов (n) между соседними пикселями внутри микроблока в горизонтальном и в вертикальном направлениях будет равно $n = 12$ (рис. 2.7 d – фрагмент текстуры). Минимальное количество цветовых переходов равно $n = 2$ (рис. 2.7 а – точечный объект).

Идентификация микроблока выполняется по следующему критерию: если значения цветового контраста между соседними пикселями для количества переходов $n \geq 2$ превышают зрительный порог (2.11), то в микроблоке присутствует мелкая структура различимая глазом. Если данное условие не выполняется, то мелкая структура не различается глазом.

Рассмотрим алгоритм идентификации мелких структур изображения по шагам.

1. Выполняем попиксельно преобразование изображения из RGB в $L^*a^*b^*$.
2. Выделяем первый микроблок с помощью маски 3×3 пикселя, присваиваем номер $m = 0$ и вычисляем двенадцать значений контраста между соседними пикселями:

$$K_{m,j,i} = \sqrt{\left(\frac{\Delta L_{m,j,i}^*}{L_{th}^*}\right)^2 + \left(\frac{\Delta a_{m,j,i}^*}{a_{th}^*}\right)^2 + \left(\frac{\Delta b_{m,j,i}^*}{b_{th}^*}\right)^2}, \quad (2.14)$$

где $L_{th}^* = 6$, $a_{th}^* = 40$ и $b_{th}^* = 55$ – значения весовых коэффициентов из таблицы 1 для мелких деталей с размером в один пиксель; значения $\Delta L_{m,j,i}^* = L_{m,j}^* - L_{m,i}^*$, $\Delta a_{m,j,i}^* = a_{m,j}^* - a_{m,i}^*$ и $\Delta b_{m,j,i}^* = b_{m,j}^* - b_{m,i}^*$ определяют контраст между соседними пикселями по светлоте и по цветности; $(j, i) = (1, 2); (2, 3); (4, 5); (5, 6); (7, 8); (8, 9); (1, 4); (4, 7); (2, 5); (5, 8); (3, 6); (6, 9)$ – номера координат пикселей внутри микроблока (рис. 1).

3. Устанавливаем счетчик цветовых переходов в ноль ($n = 0$) и для каждого значения контраста (4) проверяем условие

$$K_{m,j,i} > 1. \quad (2.15)$$

Если условие (2.15) истинно, принимаем решение о том, что изменение контраста между соседними пикселями заметно глазом и увеличиваем счетчик на единицу $n = n + 1$. Если условие (2.15) ложно, принимаем решение о том, что изменение контраста между соседними пикселями незаметно для глаза и сохраняем значение счетчика $n = n$.

4. После анализа всех 12-ти переходов оцениваем значение n . Если $n \geq 2$, принимаем решение о том, что в микроблоке присутствует мелкая структура, контраст которой различается глазом, и помечаем его с помощью маркера $mark(m) = 1$. Если $n < 2$, принимаем решение о том, что в микроблоке отсутствуют мелкие детали и помечаем его маркером $mark(m) = 0$.

5. Смещаем координаты маски на три пикселя, переходим к следующему микроблоку с номером $m = 1$ и повторяем пункты 2-4.

После анализа всего изображения получим выделенные участки изображения в виде микроблоков с размерами 3×3 пикселя, помеченных маркером $mark(m) = 1$. На рис. 2.10 показан фрагмент (102×114) тестового изображения с выделенными микроблоками.

Почему выбрана система N-CIELAB?

Контраст (2.14) можно вычислять в любой другой системе цветовых координат, например: *RGB*, *YUV*, *HSI*, *WUV*, *Luv*, и др.

В результате многих исследований мы пришли к выводу, что применение системы N-CIELAB имеет следующие преимущества: оценка контраста мелких структур по уравнениям (2.14-2.15) соответствует контрастной чувствительности зрения; оценка не зависит от цвета; оценка мало зависит от яркости окружающего фона.

Таким образом, предложенный алгоритм позволяет идентифицировать фрагменты фотореалистичных изображений с мелкими структурами, контраст

которых превышает пороговую контрастную чувствительность человеческого зрения. Кроме того, алгоритм позволяет оценить уровень детальности изображения (FDL) [55], как

$$FDL = \frac{9 \cdot N_{mark}}{W \cdot H}, \quad (2.16)$$

где W и H – ширина и высота изображения в пикселях; N_{mark} – количество идентифицированных микроблоков. Для нашего примера (см. рис. 2.10) уровень детальности всего изображения равен $FDL = 0,208$.

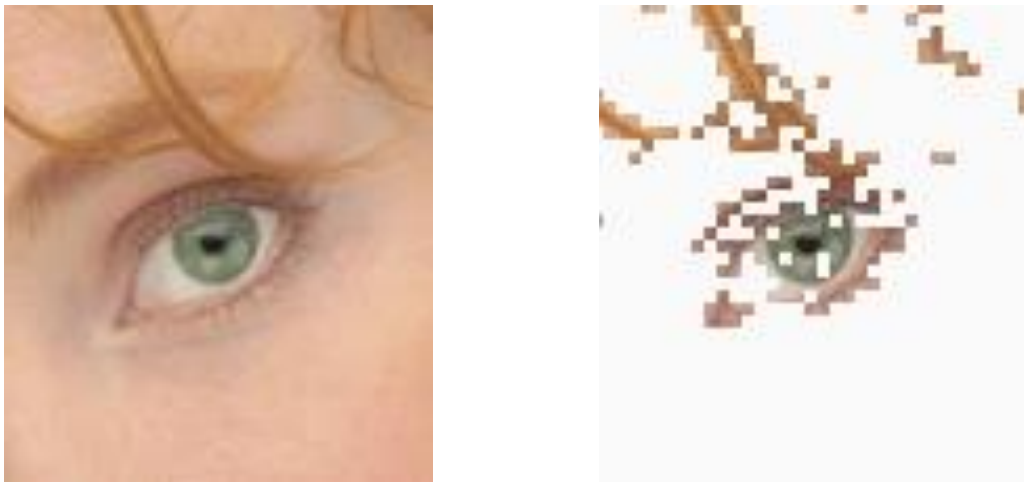


Рис. 2.10. Фрагмент тестового изображения с выделенными мелкими структурами

Значение FDL можно использовать в задачах классификации и сортировки фотореалистичных изображений по уровню детальности. Экспериментальные исследования алгоритма показали, что он дает адекватные результаты при анализе высококачественных цифровых изображений.

Если изображение искажено шумами, то низкое отношение сигнала к шуму (PSNR) приводит к идентификации компонентов шума, как мелких структур и уровень детальности будет расти.

Артефакты алгоритмов сжатия JPEG2000 или JPEG с высоким уровнем сжатия (низкое качество) приводят к снижению значения FDL благодаря размытию мелких деталей.

Экспериментально мы пришли к выводу, что значение FDL незначительно изменяется для $PSNR > 40$ дБ и/или при сжатии JPEG с высокими настройками параметра качества (80-100%).

2.5.3. Метрика MFSD

В настоящей работе предлагается альтернативная метрика *MFSD* (Metric of Fine Structures Distortion) для оценки искажений мелких структур компрессированных изображений. На первом этапе выполняется алгоритм идентификации мелких структур оригинального изображения. На втором этапе оцениваются искажения в выделенных микроблоках m по следующей формуле:

$$dE_m = \max_{j,i} |K_{m,j,i} - \tilde{K}_{m,j,i}|, \quad (2.17)$$

где значения $K_{m,j,i}$ вычисляются по формуле (2.14) для микроблока оригинального изображения и $\tilde{K}_{m,j,i}$ для микроблока искаженного изображения. Выражение (2.17) оценивает максимальное отклонение контраста цветовых переходов оригинального и искаженного изображений в микроблоке с номером m .

После анализа всех микроблоков получим среднюю оценку искажений, которая соответствует метрике *MFSD*:

$$MFSD = \frac{1}{N_{mark}} \sum_{m=1}^{N_{mark}} dE_m. \quad (2.18)$$

Разработанные алгоритмы реализованы в программном анализаторе качества (PQA_FS) в среде C++. Анализатор выполняет следующие функции:

1. Открытие тестового изображения в окне (in_Image);
2. Выполнение алгоритма идентификации мелких структур оригинального изображения и вычисление его уровня детальности (*FDL*);
3. Открытие компрессированного изображения в окне (out_Image) и вычисление искажений по метрике *MFSD* и, также, по традиционным метрикам PSNR и SSIM.

Для оценки искажений компрессированных изображений в кодеках JPEG и JPEG2000 было выбрано приложение Adobe Photoshop CS6. В табл. 2.6 приведены экспериментальные зависимости *MFSD* от параметра качества Q и соответствующего коэффициента (*Ratio*) компрессии кодека JPEG для трех тестовых изображений (рис. 2.11) с различными уровнями детальности (*FDL*).

Тестовые изображения показаны в уменьшенном масштабе от исходного размера 512×512 пикселей. Для сравнения в таблице приведены результаты оценки искажений по метрикам *PSNR* и *SSIM*. В табл. 2.7 приведены результаты оценки искажений компрессированных изображений кодека JPEG2000.



a)



b)



c)

Рис. 2.11. Примеры тестовых изображений

На следующем этапе необходимо было оценить значение $MFSD_{\min}$, при котором искажения мелких структур компрессированных изображений будут незаметны или малозаметны глазом.

Из формул (2.13) и (2.15) следует, что изменение контраста между соседними пикселями в микроблоке оригинального изображения будет заметно глазом, если $K_{m,j,i} > 1$. Формула (2.17) вычисляет максимальную разницу контраста между соседними пикселями оригинального и искаженного изображений.

Следовательно, можно предположить, что если значение $dE_m < 1$, то искажения будут незаметны или малозаметны.

Однако в результате исследований получено, что данное предположение справедливо для мелких структур с высоким контрастом $\Delta K_m > 2$, где

$$\Delta K_m = \max_{j,i} (K_{m,j,i})$$

определяет контраст микроблока оригинального изображения.

Если контраст микроблока невысокий $1 < \Delta K_m < 2$, то искажения с величиной $dE_m \approx 1$, при условии уменьшения контраста и «размытия» микроблока искаженного изображения, могут привести к тому, что мелкие структуры не будут различаться.

Поэтому, для сохранения высокого качества передачи мелких структур фотореалистичных изображений предлагается следующий критерий:

$$MFSD \leq MFSD_{\min} = 0,5. \quad (2.19)$$

Для подтверждения объективности полученного критерия и экспериментальной оценки параметра $MFSD_{\min}$ использовалась субъективная шкала качества кодера JPEG в приложении Adobe Photoshop CS6. Данная шкала качества получена в результате многочисленных экспериментов выполненных разработ-

чиками приложений Adobe Photoshop и содержит следующие градации: Low (1-4), Medium (5-7), High (8-9), Maximum (10-12).

Таблица 2.6. Оценка искажений JPEG

Q	3	4	5	6	7	8	9	10	
Ratio	17,7	16,4	15,4	13,0	12,8	10,5	8,7	6,9	Image(a) FDL = 0,208
MFSD	0,953	0,884	0,811	0,761	0,692	0,627	0,500	0,358	
PSNR	34,00	34,79	35,68	38,80	37,05	38,45	40,14	42,26	
SSIM	0,929	0,938	0,946	0,963	0,957	0,966	0,975	0,983	
Ratio	13,2	13,7	12,2	10,1	10,5	8,5	6,4	5,7	Image(b) FDL = 0,358
MFSD	0,978	0,882	0,758	0,734	0,587	0,552	0,353	0,021	
PSNR	33,03	33,96	35,16	38,00	36,98	38,24	41,84	48,13	
SSIM	0,946	0,953	0,961	0,971	0,970	0,975	0,987	0,999	
Ratio	9,2	8,3	7,4	6,5	6,5	5,0	4,1	3,3	Image(c) FDL = 0,602
MFSD	1,530	1,388	1,235	0,987	1,123	0,837	0,608	0,405	
PSNR	26,61	27,25	27,93	28,87	29,48	31,30	33,24	35,83	
SSIM	0,887	0,899	0,910	0,923	0,927	0,945	0,958	0,972	

Таблица 2.7. Оценка искажений JPEG 2000

Q	30	40	50	60	70	80	90	100	
Ratio	21,4	16,4	13,5	10,3	7,61	5,57	4,32	3,20	Image(a) FDL = 0,208
MRSD	1,019	0,826	0,634	0,461	0,312	0,209	0,135	0,056	
PSNR	33,00	34,98	37,25	39,77	42,61	45,61	48,13	48,21	
SSIM	0,914	0,937	0,957	0,972	0,983	0,990	0,995	0,996	
Ratio	16,4	14,5	11,6	8,6	6,8	5,1	3,9	3,0	Image(b) FDL = 0,358
MFSD	1,026	0,764	0,574	0,417	0,293	0,199	0,133	0,057	
PSNR	32,18	34,63	37,12	39,70	42,48	45,50	48,13	48,12	
SSIM	0,937	0,954	0,967	0,977	0,985	0,991	0,995	0,997	
Ratio	8,3	6,1	4,6	3,5	2,8	2,3	1,9	1,6	Image(c) FDL = 0,602
MFSD	1,398	1,036	0,730	0,490	0,318	0,207	0,134	0,057	
PSNR	28,05	30,78	33,82	37,09	40,56	44,14	47,75	48,08	
SSIM	0,911	0,942	0,965	0,980	0,989	0,994	0,997	0,997	

В процессе экспериментов, в которых принимали участие группы студентов, и анализа более 100 компрессированных JPEG фотореалистичных изображений с различной детальностью мы получили численное среднее значение $MFSD_{\min} \approx 0,5$, при котором искажения становятся незаметными (малозаметными) для глаза. В табл. 2.6 и 2.7 значения $MFSD \leq 0,5$ выделены белым цветом. При этом качество изображений соответствует градациям High и Maximum субъективной шкалы приложения Adobe Photoshop. Сравнение полученного результата с результатами метрик $PSNR$ и $SSIM$ показывает, что эти метрики имеют разброс в оценках от 0,972 до 0,987 ($SSIM$) и от 35,8 до 41,8 дБ ($PSNR$). Следовательно, по этим метрикам, нельзя однозначно определить, при

каких их значениях искажения мелких структур для любых изображений будут незаметны.

У специалистов может возникнуть вопрос: является ли достаточным выполнение критерия (2.19) для оценки высокого качества не только мелких структур, но и высокого качества всего изображения? В частности, шумы в сигналах изображения могут маскироваться на фоне мелких контрастных деталей, но будут заметны глазом на равномерных участках.

Для ответа на данный вопрос выполнен ряд экспериментов, в которых для оценки искажений всего изображения был использован дополнительный параметр:

$$dE_F = \frac{1}{N_F} \sum_{k=1}^{N_F} dE_k, \quad (2.20)$$

где значение

$$dE_k = \frac{1}{9} \sum_{i=1}^9 \sqrt{(L_{k,i}^* - \tilde{L}_{k,i}^*)^2 + (a_{k,i}^* - \tilde{a}_{k,i}^*)^2 + (b_{k,i}^* - \tilde{b}_{k,i}^*)^2}$$

оценивает среднее отклонение цветовых координат оригинального и искаженного изображения в системе CIELAB внутри микроблока (k), в котором отсутствуют различимые глазом мелкие структуры; i – номер пикселей в микроблоке (3×3). Выражение (2.20) дает среднюю оценку по всем «фоновым» микроблокам N_F .

Искажения участков изображения с постоянными значениями цветовых координат ($L^* a^* b^*$) в пределах сегмента с размерами не менее 4×4 пикселей будут незаметными или малозаметными при выполнении условия

$$dE_F < 2,3. \quad (2.21)$$

В результате экспериментов получен вывод о том, что, если выполняется критерий (2.19) высокого качества воспроизведения мелких структур изображения, то всегда будет выполняться условие (2.21) незаметности (малозаметности) искажений остальных частей изображения. Для примера, в табл. 2.8 приведены результаты оценок параметра dE_F при высоком качестве воспроизведения мелких структур тестовых изображений (см. рис. 2.11).

Для дополнительного подтверждения объективности и достоверности метрики $MFSD$ был выполнен анализ качества компрессированных изображений из базы данных [56] лаборатории LIVE. Данная база содержит набор изображений с различным качеством по стандартам JPEG (204 шт.) и JPEG2000 (198 шт.) по-

лученных в результате компрессии 29 оригинальных фотореалистичных изображений. В базе данных для каждого изображения приведены результаты субъективных оценок качества от 20 наблюдателей по 100 бальной шкале, которая конвертирована из 5 бальной шкалы с градациями качества: “плохое”, “низкое”, “среднее”, “хорошее” и “отличное”.

Таблица 2.8. Оценка искажений с учетом параметра dE_F

	Image(a)	Image(b)	Image(c)	
Q	9	9	10	JPEG
$MFSD$	0,500	0,353	0,405	
dE_F	1,264	0,936	2,205	
Q	60	60	60	JPEG2000
$MFSD$	0,461	0,417	0,490	
dE_F	1,226	1,205	1,837	

Для анализа выбирались те изображения, для которых средние субъективные оценки были не менее 80 баллов, т.е. их качество соответствовало оценки «отличное». Результаты экспериментов показали, что для всех выбранных изображений критерий (2.19) выполняется, что подтвердило объективность метрики $MFSD$.

Отметим, что все результаты и выводы данного раздела справедливы при обработке и анализе высококачественных оригинальных изображений, для которых значение PSNR (7) не менее 40 дБ.

Таким образом, в отличие от известных метрик (PSNR, SSIM и др.), новая метрика $MFSD$ оценивает заметность искажений мелких структур компрессированных изображений и позволяет настраивать параметры любого кодера на высокое визуальное качество.

Например, для тестового изображения Image(a) выполнение критерия (2.19) обеспечивается при выборе параметра качества $Q > 8$ (High) кодера JPEG и при выборе $Q > 55$ кодера JPEG2000. При этом кодер JPEG 2000 обеспечивает более высокий коэффициент компрессии, что следует из приведенных данных в таблицах 2.5 и 2.6.

На основе предложенной метрики $MFSD$ оценки искажений получен новый объективный критерий (2.19) высокого качества воспроизведения мелких структур компрессированных изображений.

К особенностям метрики относится то, что она основана на алгоритме идентификации мелких структур изображения с использованием нормирован-

ной системы N-CIELAB и позволяет оценивать искажения с учетом свойств контрастной чувствительности зрения [57].

Объективность метрики подтверждена результатами анализа компрессированных изображений в стандартах JPEG и JPEG2000 с помощью субъективной шкалы качества приложения Adobe Photoshop CS6, а также результатами анализа изображений из базы данных лаборатории LIVE.

К перспективным приложениям предложенной метрики относятся:

1. Оптимизация и настройка параметров кодеров JPEG и JPEG2000 на высокое качество воспроизведения мелких структур изображения.
2. Реализация метрики *MFSD* в системах анализа качества передачи мелких структур видео изображений в системах сжатия MPEG на основе анализа выборочных кадров из видеофайла.
3. Разработка новых адаптивных алгоритмов компрессии фото и видео-изображений с высоким визуальным качеством.
4. Классификация и сортировка изображений по их уровню детальности FDL (2.16) в существующих и новых базах данных тестовых фотореалистичных изображений.
5. Сегментация изображения по заданным интервалам визуального контраста в задачах анализа и распознавания объектов изображения.

Считаем, что применение новой метрики в системах компрессии цифровых изображений позволит обеспечить компромиссное решение между коэффициентом компрессии и качеством.

2.5.4. Алгоритм оценки резкости

На первом этапе для каждого пикселя изображения с пространственными координатами (i, j) выполняется преобразование *RGB* сигналов основных цветов в равноконтрастное пространство $L^*a^*b^*$.

Алгоритм поиска границ и оценки их длительности рассмотрим по шагам.

1. Поиск начинаем по горизонтальной поисковой линии слева направо (цикл по координате j при фиксированной i). Количество поисковых линий (N_y) выбираем как $N_y = H/r$, где H – количество строк изображения и r – коэффициент прореживания. Устанавливаем счетчик переходов в ноль: $k = 0$.
2. При движении точки поиска по поисковой линии (i) вычисляем контраст между соседними точками

$$\Delta K_{i,j} = \sqrt{\left(L_{i,j+1}^* - L_{i,j}^*\right)^2 + \left(a_{i,j+1}^* - a_{i,j}^*\right)^2 + \left(b_{i,j+1}^* - b_{i,j}^*\right)^2} \quad (2.22)$$

и проверяем условие

$$\Delta K > 3. \quad (2.23)$$

3. Если условие (2.23) выполняется, принимается решение о том, что изменения контраста на границе объекта заметны для глаза. Координата j является началом перехода, поэтому фиксируем координаты $x_0 = j$, $y_0 = i$ и инкрементируем счетчик переходов: $k = k+1$. Далее выполняем алгоритм кругового поиска градиента.

4.1. Алгоритм поиска градиента выполняется в окрестности пикселя с координатами (x_0, y_0) .

4.2. На первой итерации, для каждой (n) точки окрестности вычисляем контраст $\Delta K_{i,j}(n)$ относительно центральной точки и находим максимальное значение $\Delta K_{\max} = \max[\Delta K_{i,j}(n)]$. Поиск максимума выполняем в сторону возрастания или только в сторону убывания яркости, т.е. в выражении (2.22) учитываем знак изменения контраста по индексу яркости.

4.3. Вектор (dx, dy) , соответствующий максимальному элементу $\Delta K_{\max}$ массива $\Delta K_{i,j}(n)$, принимается за направление градиента.

4.4. Фиксируем знак (s) изменения контраста и перемещаем поисковую точку в направлении вектора: $(x_0 = x_0 + dx; y_0 = y_0 + dy)$.

4.5. Пункты поиска градиента с учетом знака (s) продолжаются, пока $\Delta K_{\max}$ не станет меньше порогового значения (2.23). Это будет означать, что поисковая точка находится уже на изображении объекта. В итоге получим координаты точки (x_1, y_1) , которая находится на завершении границы перехода (рис. 2.12).

5. Вычисляем расстояние от нулевой до первой точки:

$$\delta_k = \sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2} \quad (2.24)$$

и записываем значение δ_k в массив D .

6. Продолжаем движение точки поиска по поисковой линии (i) и выполняем пункты 2-5 до окончания строки изображения.

7. Переходим к следующей поисковой линии $(i = i + r)$ и повторяем алгоритм поиска вертикальных и наклонных границ.

8. Повторяем алгоритм поиска по вертикальным поисковым линиям сверху вниз (цикл по координате i при фиксированной координате j) и добавляем в массив D значения δ_k для обнаруженных горизонтальных границ.

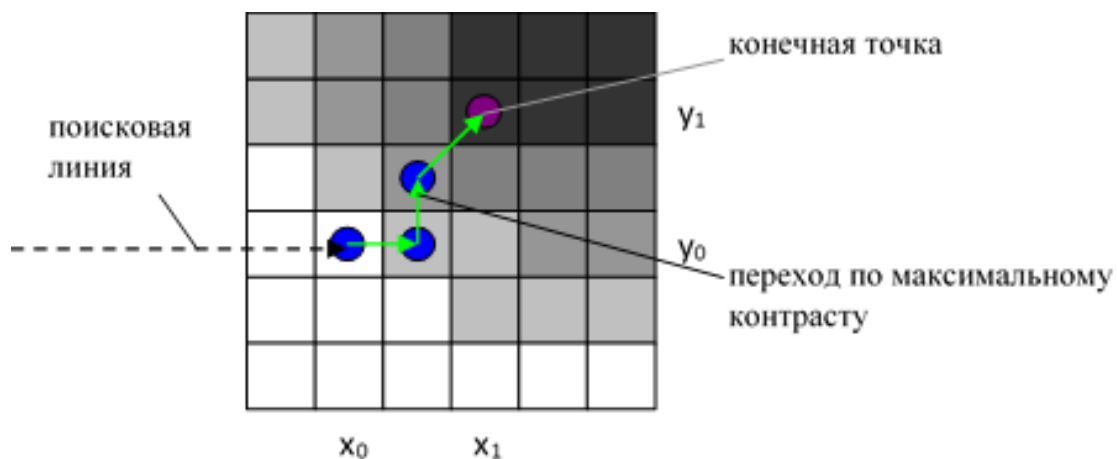


Рис. 2.12. Определение конечной точки перехода

9. Из массива D находим среднее значение $\bar{\delta}$, которые соответствует резкости границ деталей анализируемого изображения.

Предложенный алгоритм позволяет оценивать длительность $\bar{\delta}$ цветовых переходов границ крупных деталей изображения и, следовательно, его резкость. Как показывают результаты экспериментальных исследований, для высококачественных изображений, длительность цветовых переходов: $\bar{\delta} \approx 2..3$. Данный критерий можно использовать для оценки резкости изображения с различным качеством.

В частности, если после выполнения алгоритма значение: $\bar{\delta} \approx 4..6$ – это будет означать, что резкость изображения ухудшилась в два раза.

Отличительной особенностью предложенного алгоритма является то, что в анализируемом изображении не требуется наличие тестовых сигналов в виде максимальных перепадов яркости или цветности. Также, контраст цветового перехода (2.22) оценивается числом нормированных цветовых порогов зрения, что позволяет повысить достоверность обнаружения границ.

Таким образом, предложенный алгоритм можно использовать для оценки качества кодеков без наличия эталона по показателю резкости цифровых изображений. Для сравнения качества разных кодеков можно выполнить анализ по простой методике. Исходное (любое) изображение, закодированное с помощью разных архиваторов, надо сравнить по показателю резкости границ крупных деталей изображения и выбрать наилучший архиватор с минимальной средней длительностью цветовых переходов.

3. Повышение качества компрессированных изображений

В настоящей главе приводится описание способов повышения качества цифровых изображений в стандартах JPEG и JPEG 2000 на основе адаптивного квантования коэффициентов трансформации. Рассматриваются особенности корректирующих фильтров для улучшения качества изображений на выходе видеокодеков. Предлагаются результаты сравнения качества и эффективности сжатия видеофайлов с помощью анализатора VQMT.

3.1. Повышение качества изображений в стандарте JPEG

3.1.1. Идентификации мелких деталей

В настоящее время стандарт JPEG являются наиболее популярным методом цифровой компрессии фотореалистичных изображений. Основные потери качества при выполнении алгоритма JPEG возникают на этапе квантования коэффициентов дискретного косинусного преобразования (DCT):

$$\bar{F}_{i,j,k} = \frac{Q \cdot F_{i,j,k}}{M_{i,j}}, \quad (3.1)$$

где $F_{i,j,k}$ – коэффициенты DCT полученные для k -го блока яркости или цветности с размерами 8×8 пикселей; Q – параметр качества; $M_{i,j}$ – элементы матрицы квантования, целочисленные значения которой зависят от номера элемента (i, j) в блоке и могут изменяться в интервале от 1 до 255. Стандарт JPEG рекомендует разные матрицы квантования для блоков яркости и цветности, где значения элементов матриц увеличиваются с ростом i и j , что обусловлено свойствами пространственно-частотной характеристики зрения.

К особенностям стандарта относится то, что пользователь может использовать свои матрицы квантования и адаптивно изменять их значения в зависимости от структуры блока. Данное обстоятельство является предметом исследований многих специалистов в области цифровой обработки изображений, которые пытаются повысить эффективность компрессии с сохранением высокого качества изображения. Известные методы адаптивного JPEG основаны на способах настройки параметров квантователя на требуемое качество или уровень компрессии для идентифицированных блоков с разными уровнями детальности. Матрицу квантования можно интерпретировать в виде двумерного пространственно-частотного фильтра нижних частот, в котором ослабляются более

высокочастотные коэффициенты DCT блока. Следовательно, уменьшение параметра качества Q в формуле (3.1) будет приводить к более сильному подавлению и обнулению высокочастотных составляющих блока, что приводит к увеличению искажений мелких деталей и резких границ изображения. Увеличение Q приводит к снижению искажений, но при этом будет уменьшаться степень компрессии. Таким образом, для адаптивного квантования на первом этапе обработки необходимо выполнить анализ изображения с целью идентификации блоков с разными характеристиками. К таким характеристикам относятся яркость и контраст блока [58]. В некоторых алгоритмах [59] используются методы выделения контуров деталей изображений. Также используются методы сегментации [60] областей интереса (ROI) в виде объектов и фона, на котором они расположены.

После идентификации и классификации блоков выполняется дискретно косинусное преобразование и адаптивное квантование коэффициентов DCT:

$$\bar{F}_{i,j,k} = \frac{Q_k \cdot F_{i,j,k}}{M_{i,j}}, \quad (3.2)$$

где Q_k – параметр качества для блоков класса k . Диапазон значений Q зависит от числа уровней классификации блоков. Самый простой случай это использование всего двух уровней, например, объект и фон. Для объектов выбирается более высокое значение Q , чем для фона. Отметим, что в некоторых алгоритмах параметр Q не изменяется и, при этом, меняются значения матриц квантования.

В настоящем разделе предлагается адаптивный способ квантования коэффициентов DCT [61] на основе алгоритма поиска и распознавания мелких деталей в нормированном цветовом пространстве N-CEALAB. Описывается алгоритм классификации блоков, и приводятся экспериментальные зависимости качества изображения и степени компрессии от выбранных параметров квантователя.

Рассмотрим алгоритм идентификации самых мелких деталей ($\delta = 1$) изображения по шагам.

1. Выполняем попиксельно преобразование изображения из RGB в $L^* a^* b^*$.
2. Выделяем первый микроблок с помощью маски 2×2 пикселя, присваиваем номер $m = 0$ и вычисляем его контраст:

$$K_m = \sqrt{\left(\Delta \bar{L}_m^* / L_{th}^*\right)^2 + \left(\Delta \bar{a}_m^* / a_{th}^*\right)^2 + \left(\Delta \bar{b}_m^* / b_{th}^*\right)^2}, \quad (3.3)$$

где $\Delta \bar{L}_m^*$, $\Delta \bar{a}_m^*$ и $\Delta \bar{b}_m^*$ – средние отклонения цветовых координат в микроблоке m :

$$\Delta \bar{L}_m^* = \frac{1}{N} \sum_{n=0}^{N-1} |L_n^* - L_{mean}^*|, \Delta \bar{a}_m^* = \frac{1}{N} \sum_{n=0}^{N-1} |a_n^* - a_{mean}^*|, \Delta \bar{b}_m^* = \frac{1}{N} \sum_{n=0}^{N-1} |b_n^* - b_{mean}^*|, \quad (3.4)$$

L_{th}^* , a_{th}^* и b_{th}^* – значения весовых коэффициентов из табл. 2.4 для мелких деталей с размером в один пиксель; L_n^* , a_n^* и b_n^* – значения цветовых координат; L_{mean}^* , a_{mean}^* и b_{mean}^* – средние значения цветовых координат в микроблоке; N – число пикселей в микроблоке ($N = 4$).

3. Проверяем условие

$$K_m > K^{th}, \quad (3.5)$$

где K^{th} – пороговое значение контраста, при котором соседние пиксели в микроблоке начинают различаться глазом.

4. Если условие (3.5) выполняется, принимаем решение о том, что в микроблоке присутствуют мелкие детали и помечаем его с помощью маркера $b_1(m) = 1$, где m – номер микроблока.

5. Если условие (3.5) не выполняется, принимаем решение о том, что в микроблоке нет мелких деталей различимых глазом, и помечаем его как $b_1(m) = 0$.

6. Смещаем координаты маски на два пикселя, переходим к следующему микроблоку с номером $m = 1$ и повторяем шаги 2-5.

После анализа всего изображения получим выделенные участки в виде микроблоков с размерами 2×2 пикселя, помеченных маркером $b_1(m) = 1$.

На следующем этапе переходим к идентификации мелких деталей с размерами в два пикселя ($\delta = 2$).

1. Выделяем первый блок с помощью маски 4×4 пикселя.

2. Проверяем наличие микроблоков (2×2) с маркером $b_1 = 1$. Если в блоке (4×4) такие микроблоки отсутствуют, выполняем его анализ: вычисляем контраст блока (3.3), где L_{th}^* , a_{th}^* и b_{th}^* – значения весовых коэффициентов из табл. 2.4 для мелких деталей с размером $\delta = 2$ и количество элементов в блоке $N = 16$.

3. Проверяем условие (3.5).

4. Если условие (3.5) выполняется, принимаем решение о том, что в блоке присутствуют мелкие детали с размером $\delta = 2$, маркируем блок с помощью маркера $b_2(m) = 2$ и переходим к следующему блоку.

5. Если условие (3.5) не выполняется или в блоке присутствуют микроблоки с маркером ($b_1 = 1$) помечаем маркер $b_2(m) = 0$ и переходим к анализу следующего блока.

Аналогично выполняем сегментацию деталей с размерами 4×4 пикселя. Выделяем блоки с размером 8×8 пикселей, и проверяем наличие микроблоков с маркерами $b_1 = 1$ и $b_2 = 2$. Если такие микроблоки отсутствуют, вычисляем контраст блока по формуле (5), где L^*_{th} , a^*_{th} и b^*_{th} – значения весовых коэффициентов из таблицы 1 для $\delta = 4$. Проверяем условие (3.5). Если условие (3.5) выполняется, маркируем блок с помощью маркера $b_4(m) = 4$ и переходим к следующему блоку.

После идентификации микроблоков с размерами 2×2 , 4×4 и 8×8 пикселей помечаем оставшиеся блоки (8×8) маркером $b_F(m) = F$ и классифицируем их как фоновые блоки с незначительными изменениями цветовых координат внутри каждого блока.

Таким образом, после выполнения алгоритма идентификации, с помощью маркеров b_1 , b_2 и b_4 мы выделяем три множества участков (микроблоков) изображения с мелкими деталями, а также фоновые блоки с помощью маркера b_F .

Выбор значения порога K^{th} в формуле (3.5) зависит от следующих факторов. Из метрики MFSD (п. 2.4) следует, что если значение цветового контраста равно единичному пороговому значению $K_{LAB} = K^{th}_{LAB} = 1$, то мелкие детали тестовой таблицы начинают различаться глазом в виде изображения одиночных объектов расположенных на однородном неокрашенном фоне.

Для реальных изображений визуальный контраст микроблока будет снижаться в зависимости от неоднородности фона и от маскирующего эффекта соседних микроблоков. В результате экспериментов, оценивающих влияние выбора порога K^{th} на зрительную идентификацию микроблоков реальных изображений, было предложено для анализа фотореалистичных изображений использовать значение порога $K^{th} = 1,5$.

3.1.2. Классификация блоков и оптимизация параметров квантования

Рассмотрим способ классификации блоков изображения (8×8) для адаптивного квантования коэффициентов DCT в компрессоре JPEG. На первом этапе выполняем идентификацию микроблоков с мелкими деталями по описанному выше алгоритму и получаем микроблоки изображения с маркерами b_1 , b_2 , b_4 и b_F . На втором этапе выполняем классификацию блоков (8×8) по следующему алгоритму:

- в каждом блоке (8×8) выполняем поиск микроблоков с маркерами b_1 , b_2 и b_4 ;

- если в блоке имеется хотя бы один микроблок с маркером b_1 , то отмечаем этот блок с помощью маркера B_1 ;
- если в блоке отсутствуют микроблоки с маркером b_1 и имеется хотя бы один микроблок с маркером b_2 , то отмечаем его с помощью маркера B_2 ;
- если в блоке отсутствуют микроблоки с маркером b_1 и b_2 , и имеется микроблок с маркером b_4 , то отмечаем его с помощью маркера B_4 ;
- остальные блоки, в которых отсутствуют микроблоки с маркерами b_1 , b_2 и b_4 отмечаем с помощью маркера B_F .

Отметим, что маркировка блоков требует два дополнительных бита информации, которые могут быть расположены в заголовке блока или, после сжатия модифицированным методом Хаффмана, в конце сжатого блока.

На рис. 3.1 показан фрагмент (512×512) тестового изображения и выделенные фрагменты с блоками B_1 , B_2 и B_4 .

На следующем этапе необходимо выбрать параметр качества (Q) квантователя в зависимости от класса блока. Для решения данной задачи в среде C++ была разработана программная модель кодека JPEG с возможностью установки значения Q для каждого класса блоков изображения по яркости (Y) и по цветности (U , V). В программе реализованы все функции стандарта JPEG: преобразование изображения из RGB формата в формат YUV с использованием цветовой модели 4:4:4; разбиение матриц Y , U и V на блоки (8×8); выполнение блочного DCT; квантование блоков с параметром Q ; зигзаг-сканирование и статистическое сжатие.

Квантование блоков в матрицах яркости и цветности выполнялось с помощью следующих формул:

$$\bar{Y}_{i,j,m} = \frac{Q_{Bk}^Y \cdot Y_{i,j,m}}{M_{i,j}^Y}; \bar{U}_{i,j,m} = \frac{Q_{Bk}^{UV} \cdot U_{i,j,m}}{M_{i,j}^{UV}}; \bar{V}_{i,j,m} = \frac{Q_{Bk}^{UV} \cdot V_{i,j,m}}{M_{i,j}^{UV}},$$

где m – номер блока; i, j – пространственные координаты коэффициентов DCT внутри блока (8×8); M^Y и M^{UV} – матрицы квантования по яркости и по цветности рекомендованные стандартом JPEG; Q_{Bk}^Y и Q_{Bk}^{UV} – параметры квантования блоков, которые зависят от их класса (B_1 , B_2 , B_4 и B_F).

Известно, что артефакты компрессии JPEG оказывают большее влияние на искажения мелких деталей и такие искажения наиболее заметны глазом. Поэтому мы предлагаем убывающую прогрессию параметра Q в зависимости от класса блока, где максимальное значение Q_{max} будет соответствовать блокам с маркером B_1 .

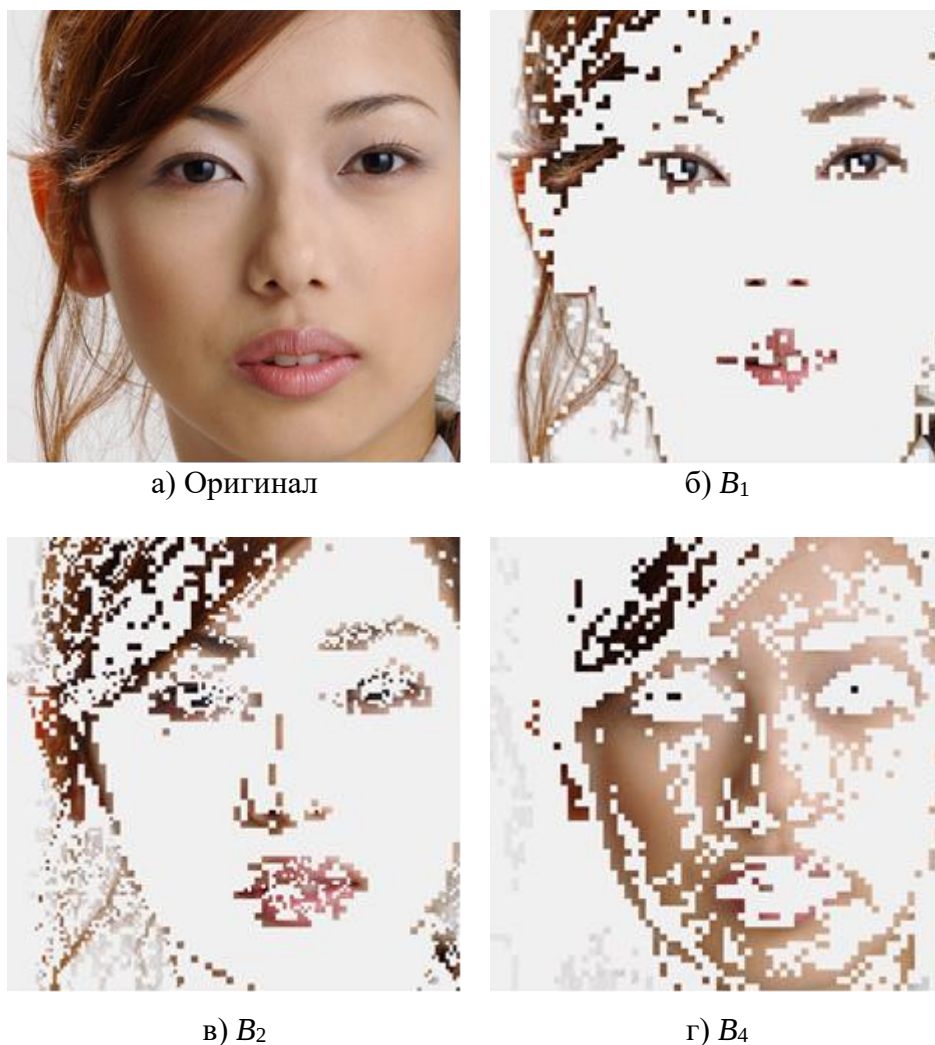


Рис. 3.1. Результат выделения блоков тестового изображения

Для определения значений параметров адаптивного квантования мы использовали следующую методику.

На первом этапе необходимо было определить максимальные значения Q_{B1}^Y и Q_{B1}^{UV} , при котором искажения в блоках B_1 будут незаметны или малозаметны глазом. В процессе эксперимента наблюдатель выбирал тестовое изображение и изменял параметр Q_{B1}^Y от минимального значения до значения, при котором искажения становились незаметными. При этом остальные параметры квантования Q_{B2}^Y , Q_{B4}^Y , Q_{BF}^Y и Q_{B1}^{UV} , Q_{B2}^{UV} , Q_{B4}^{UV} , Q_{BF}^{UV} устанавливались в максимальное значение.

Для визуализации искажений в программе реализован модуль, выполняющий следующую функцию. Для каждого сегментированного блока вычислялись отклонения цветовых координат оригинального (RGB) и искаженного ($\tilde{R}\tilde{G}\tilde{B}$) изображения:

$$\begin{aligned}
\Delta R_1(i, j) &= R(i, j) - \tilde{R}(i, j); \\
\Delta G_1(i, j) &= G(i, j) - \tilde{G}(i, j); \\
\Delta B_1(i, j) &= B(i, j) - \tilde{B}(i, j).
\end{aligned}
\tag{3.6}$$

Далее, отклонения цветовых координат отображались в блоках B_1 на неокрашенном сером фоне с яркостью Y_F .

На рис. 3.2 показаны фрагменты тестового изображения и выделенные блоки B_1 с визуализацией искажений для разных значений параметра Q_{B1}^Y .

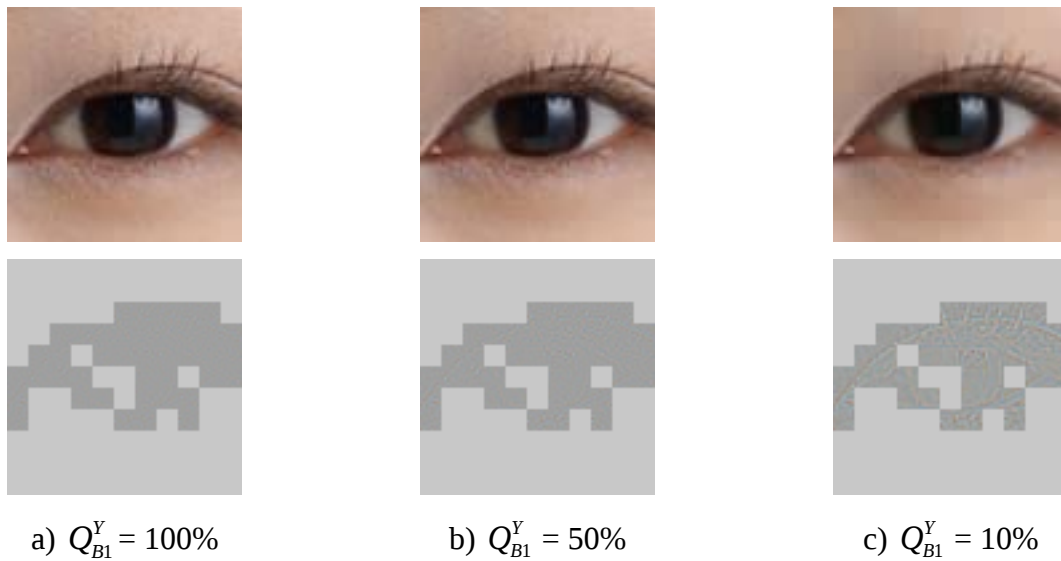


Рис. 3.2. Визуализация искажений в блоках B_1

После определения максимального значения $Q_{B1}^Y = 100\%$ наблюдатель изменял параметр Q_{B1}^{UV} в блоках цветности до тех пор, пока искажения становились незаметными. При этом параметр Q_{B1}^Y устанавливался в максимальное значение. В процессе экспериментов и анализа ряда фотореалистичных изображений с различной детальностью мы получили численное среднее значение Q_{B1}^Y , при котором искажения в блоках B_1 становятся незаметными (малозаметными) для глаза и присвоили этому значению максимальный параметр $Q_{B1}^Y = Q_{max} = 100\%$.

Для подтверждения достоверности значения Q_{max} были проведены аналогичные эксперименты в известных приложениях Adobe Photoshop CS и ACD See Pro Photo Manager. В процессе эксперимента с помощью программного модуля, реализующего функцию (3.6), наблюдатель оценивал визуальные искажения сжатого изображения для разных параметров качества JPEG. В результате экспериментов получено, что визуальные оценки искажений по критерию неза-

метно (малозаметно) соответствуют: для разработанной программы при значении параметра $Q_{B1}^Y = Q_{max}$; для приложения Adobe Photoshop CS при максимальном значении параметра качества $Quality = 12$; для приложения ACD See Pro Photo Manager при максимальном значении параметра качества $Quality = 100\%$.

Также было определено численное значение параметра квантования для блоков цветности Q_{B1}^{UV} . Экспериментально установлено, что это значение составляет 70% от значения Q_{max} .

Для количественной оценки искажений сегментированных блоков B_1 , B_2 , B_4 и B_F в программе вычислялось среднее отклонение цветовых координат в системе CIELAB по светлоте (L^*) и по цветности (a^* , b^*) с помощью метрики:

$$\varepsilon_{Bk}^L = \frac{1}{M_{Bk}} \sum_{m=0}^{M_{Bk}-1} |\Delta \bar{L}_m^*|; \varepsilon_{Bk}^{ab} = \frac{1}{M_{Bk}} \sum_{m=0}^{M_{Bk}-1} \sqrt{(\Delta \bar{a}_m^*)^2 + (\Delta \bar{b}_m^*)^2}, \quad (3.7)$$

где M_{Bk} – количество выделенных блоков с маркером B_k ; k – параметр определяющий тип микроблока (B_1 , B_2 , B_4 и B_F); $\Delta \bar{L}_m^*$, $\Delta \bar{a}_m^*$ и $\Delta \bar{b}_m^*$ – средние отклонения цветовых координат оригинального и декодированного изображения в блоке с номером m :

$$\Delta \bar{L}_m^* = \frac{1}{N} \sum_{n=0}^{N-1} |L_{n,m}^* - \tilde{L}_{n,m}^*|, \Delta \bar{a}_m^* = \frac{1}{N} \sum_{n=0}^{N-1} |a_{n,m}^* - \tilde{a}_{n,m}^*|, \Delta \bar{b}_m^* = \frac{1}{N} \sum_{n=0}^{N-1} |b_{n,m}^* - \tilde{b}_{n,m}^*|, \quad (3.8)$$

где $L_{n,m}^*$, $a_{n,m}^*$ и $b_{n,m}^*$ – значения цветовых координат оригинального изображения; $\tilde{L}_{n,m}^*$, $\tilde{a}_{n,m}^*$ и $\tilde{b}_{n,m}^*$ – значения цветовых координат декодированного изображения; N – число пикселей в блоке ($N = 64$).

Отметим, что в отличие от (3.3) в выражение (3.8) не используются весовые коэффициенты L_{th}^* , a_{th}^* и b_{th}^* . Данное решение обусловлено тем, что

1. Выражение (3.3) оценивает визуальный контраст мелких деталей, при котором они различаются глазом в микроблоке. Метрика (3.8) оценивает отклонения цветовых координат оригинального и декодированного изображения, т.е. искажения блока, а не его контраст.

2. В блоках (8×8) могут присутствовать мелкие детали из разных микроблоков. Например, в блоках с маркером B_1 кроме микроблоков b_1 могут присутствовать микроблоки b_2 , что приводит к неоднозначности выбора весовых коэффициентов.

Таким образом, в отличие от известных метрик оценки искажений (PSNR, SSIM и др.) мы имеем четыре параметра искажений по светлоте (ε_{B1}^L , ε_{B2}^L , ε_{B4}^L ,

ε_{BF}^L) и четыре параметра по цветности (ε_{B1}^{ab} , ε_{B2}^{ab} , ε_{B4}^{ab} , ε_{BF}^{ab}) для четырех классов выделенных блоков изображения. При этом значения этих параметров соответствуют зрительным оценкам цветовых различий между оригинальным и искаженным изображением.

В табл. 3.1 приведены параметры искажений для фрагментов (512×512) трех тестовых изображений (рис. 3.3), где для всех блоков выбрано одинаковое значение параметра $Q_{max} = 100\%$. Здесь же приведены значения коэффициента компрессии и PSNR (dB).



a) Image 1



b) Image 2



c) Image 3

Рис. 3.3. Примеры тестовых изображений

Таблица 3.1. Оценка искажений JPEG (Q = 100%)

	ε	B_1	B_2	B_4	B_F	PSNR	compress
Image 1	ε_{Bk}^L	0,39	0,38	0,33	0,29	43,8	4,25
	ε_{Bk}^{ab}	1,32	1,16	1,02	0,90		
Image 2	ε_{Bk}^L	0,43	0,37	0,33	0,28	42,7	2,18
	ε_{Bk}^{ab}	1,54	1,32	1,05	0,92		
Image 3	ε_{Bk}^L	0,41	0,39	0,31	0,27	44,9	4,47
	ε_{Bk}^{ab}	0,89	0,78	0,56	0,43		

Анализ искажений тестовых изображений приводит к следующим результатам:

1. При значении параметра $Q = 100\%$ искажения в сегментированных областях не превышают допустимых значений и практически незаметны глазом.

2. Искажения по светлоте ε_{Bk}^L незначительно отличаются для разных типов изображений.

3. Искажения по цветности ε_{Bk}^{ab} зависят от цветовой насыщенности (S) изображений и увеличиваются для тестовых изображений с более высоким значением S , например для Image 2.

4. Искажения максимальны в сегментированных блоках B_1 , т.е. где присутствуют самые мелкие детали, и уменьшаются в блоках B_2 , B_4 и B_F , что дает возможность применить адаптивное квантование.

После обработки ряда тестовых фотореалистичных изображений мы получили, что в среднем искажения в блоках B_2 практически незаметны глазом, если параметр квантования $Q_{B_2}^Y$ будет не менее 70%. Для определения параметра квантования $Q_{B_2}^{UV}$ блоков по цветности наблюдатель изменял это значение, при установленном значении $Q_{B_2}^Y = 70\%$. Аналогично оценивались параметры квантования в блоках B_4 и B_F .

В результате проведения экспериментальных исследований мы получили, что высокое визуальное качество изображения обеспечивается при следующих параметрах квантования блоков сегментированных областей изображения:

$$Q_{B_1}^Y = 100\%; Q_{B_2}^Y = 70\%; Q_{B_4}^Y = 50\%; Q_{B_F}^Y = 30\%; \quad (3.9)$$

для блоков по яркости (Y) и

$$Q_{B_1}^{UV} = 70\%; Q_{B_2}^{UV} = 50\%; Q_{B_4}^{UV} = 30\%; Q_{B_F}^{UV} = 10\%; \quad (3.10)$$

для блоков по цветности (U и V).

В табл. 3.2. показаны результаты анализа трех тестовых изображений (рис. 2) после сжатия JPEG с параметрами квантования (3.9) и (3.10).

Анализ результатов в таблицах 2-3 показывает, что для наших примеров применение адаптивного квантования позволяет повысить коэффициент компрессии тестовых изображений на: 38% (Image 1); 16% (Image 2); 29% (Image 3). При этом сохраняется высокое визуальное качество изображений.

Из данных табл. 3.2 следует, что для полученных параметров адаптивного квантования искажения по светлоте ε_{Bk}^L в среднем не превышают 0,5 и по цветности ε_{Bk}^{ab} не превышают 1,5:

$$\bar{\varepsilon}_{Bk}^L \leq 0,5; \bar{\varepsilon}_{Bk}^{ab} \leq 1,5. \quad (3.11)$$

В процессе анализа более 100 фотореалистичных изображений с различной детальностью и цветовой насыщенностью мы получили, что условия (3.11) выполняются практически для всех изображений. Таким образом, условия (3.11) являются объективными критериями высокого качества изображения, т.к. соответствуют зрительным оценкам и, следовательно, могут быть использованы для настройки параметров любых других алгоритмов компрессии.

Таблица 3.2. Оценка искажений адаптивного JPEG

	ε	B_1	B_2	B_4	B_F	PSNR	compress
Image 1	ε_{Bk}^L	0,39	0,46	0,44	0,39	42,9	5,86
	ε_{Bk}^{ab}	1,57	1,43	1,22	1,09		
Image 2	ε_{Bk}^L	0,43	0,48	0,45	0,31	41,7	2,53
	ε_{Bk}^{ab}	2,03	2,05	1,61	1,05		
Image 3	ε_{Bk}^L	0,41	0,50	0,42	0,36	44,1	5,75
	ε_{Bk}^{ab}	0,99	0,93	0,74	0,66		

В этом разделе описан новый алгоритм компрессии изображения на основе стандарта JPEG, который основан на выполнении следующих функций:

1. Разбиение изображения на четыре множества микроблоков (b_1 , b_2 , b_4 и b_F) в зависимости от контраста и размера мелких деталей с использованием нормированной системы N-CIELAB.
2. Разбиение блоков (8×8) изображения на классы B_1 , B_2 , B_4 и B_F .
3. Реализация алгоритма JPEG с адаптивным квантованием блоков по яркости и по цветности в зависимости от их класса.

На основе предложенной метрики оценки искажений (3.7) получены объективные критерии (3.11) высокого качества изображений. В результате исследований нового алгоритма получено, что он обеспечивает повышение коэффициента компрессии фотореалистичных изображений в среднем на 30% с сохранением их высокого визуального качества. Мы считаем, что применение нового алгоритма в системах компрессии цифровых изображений позволит обеспечить компромиссное решение между коэффициентом компрессии и качеством.

В заключение отметим, что предлагаемый алгоритм может быть использован в мультимедийных системах передачи высококачественных видеофайлов на этапе компрессии опорных (I) кадров видеоряда.

3.2. Повышение качества изображений в стандарте JPEG 2000

3.2.1. Способ оценки искажений

Процесс кодирования и квантования сигналов изображения описан в п. 1.3. При разработке адаптивного алгоритма квантования JPEG 2000 исследователи обычно предлагают свою шкалу коэффициентов квантования [62], которая, по их мнению, должна повысить качество изображения.

В некоторых работах предлагается использовать разные материнские вейвлеты в зависимости от номера субдиапазона [63]. Анализ современного состояния исследований в данной области показывает, что авторы обычно оценивают качество изображения для разных параметров квантователя с помощью интегральных метрик, таких как PSNR, MSE, SSIM и др.

К недостатку таких оценок относится то, что эти метрики не учитывают изменений зрительного восприятия цветовых различий и, следовательно, искажений отдельных участков изображения в зависимости от их цвета, контраста и детальности.

На первом этапе выполняется алгоритм идентификации мелких структур оригинального изображения (см. п. 2.4), который основан на результатах анализа структуры микроблоков с размером 3×3 пикселя.

На втором этапе оцениваются искажения компрессированного изображения по метрике MFSD. Для оценки искажений в выделенных участках с мелкими структурами предлагается следующая методика. В каждом идентифицированном микроблоке m вычисляем 12 отклонений контраста между цветовыми переходами оригинального $K_{m,p}$ и искаженного $\tilde{K}_{m,p}$ изображений

$$dE_{m,p} = |K_{m,p} - \tilde{K}_{m,p}| \quad (3.12)$$

и находим максимальное отклонение

$$dE_m = \max(dE_{m,p}).$$

Далее проверяем условие

$$dE_m > 0,5. \quad (3.13)$$

Если условие (3.13) выполняется, то принимается решение о том, что отклонения цветового контраста в микроблоке m заметны глазом и такой блок идентифицируется как искаженный $\tilde{m}$.

После анализа всех микроблоков получим объективную оценку искажений в виде отношения:

$$\varepsilon = \frac{N_{\tilde{m}}}{N_m} 100\%, \quad (3.14)$$

где значение $N_{\tilde{m}}$ определяет количество искаженных микроблоков.

После вычислений значения ε выполняем анализ качества компрессированного изображения по следующему критерию:

$$\varepsilon < 1\%. \quad (3.15)$$

В результате экспериментальных исследований доказано, что при выполнении этого условия, искажения мелких структур практически незаметны для глаза.

Дополнительно можно использовать среднюю оценку искажений по всем микроблокам с мелкими структурами:

$$dE_N = \frac{1}{N_m} \sum_{m=1}^{N_m} dE_m. \quad (3.16)$$

3.2.2. Алгоритм адаптивного квантования

В качестве основы шкалы квантования выбрано оптимальное соотношение параметров квантования для субдиапазонов в соответствии с табл. 1.1 п. 1.3.

Регулировка качества передачи изображения заключается в автоматической настройке параметра квантования Δ_1^{HH} на допустимые искажения в соответствии с критерием (3.15). По умолчанию выбрано начальное значение $\Delta_1^{HH} = 32$. В процессе настройки этот параметр изменяется дискретно на заданный интервал $\pm\sigma$, пока не будет выполнено условие (3.15).

Таким образом, мы используем классическую систему автоматического регулирования по заданному критерию.

Основная задача исследований – это выбор оптимального значения интервала $\pm\sigma$ с целью минимизации вычислительных затрат при условии обеспечения заданного качества изображений.

Для решения этой задачи реализован программный модуль в среде C++, выполняющий следующие функции:

1. Открытие тестового изображения и линейное преобразование цифровых сигналов оригинального изображения из RGB в YUV формат.

2. Преобразование цифровых сигналов оригинального изображения из RGB в $L^*a^*b^*$ формат.

3. Идентификация микроблоков (3×3) оригинального изображения с мелкими структурами.

4. Выполнение прямого ДВП для составляющей яркости (Y) и для составляющих цветности (U и V) с заданным количеством итераций $J = 5$.

5. Квантование коэффициентов ДВП, где на первом шаге использованы параметры квантования по умолчанию (табл. 1.1).

6. Выполнение обратного квантования и обратного ДВП с преобразованием цифровых сигналов изображения из YUV в $L^*a^*b^*$ формат;

7. Вычисление параметра искажений ε .

8. Проверка качества декодированного изображения и настройка оптимального параметра $\check{\Delta}_1^{HH}$:

а) При выполнении критерия (3.15) увеличиваем Δ_1^{HH} на интервал σ и выполняем п. 5-7 до тех пор, пока ε не станет больше 1%. Фиксируем оптимальное значение $\check{\Delta}_1^{HH} = \Delta_1^{HH} - \sigma$;

б) При невыполнении (3.15) уменьшаем Δ_1^{HH} на интервал σ и выполняем п. 5-7 до тех пор, пока ε не станет меньше 1%. Фиксируем оптимальное значение $\check{\Delta}_1^{HH} = \Delta_1^{HH} - \sigma$;

9. Сжатие квантованных коэффициентов ДВП на основе модифицированного алгоритма кодирования длин серий нулей [64] с переменной длиной кода.

Таким образом, после выполнения алгоритма мы получим оптимальное значение $\check{\Delta}_1^{HH}$ для анализируемого изображения, при котором будет обеспечиваться его высокое качество. Это значение можно передать в заголовке файла или кадра видеоданных с целью обратного квантования и декодирования.

К особенностям реализации разработанного алгоритма относится следующее: в качестве материнского вейвлета выбран CDF 9/7; ширина «мертвой» зоны квантования равна нулю; для каждого квантованного коэффициента ДВП использовано 12 бит, включая один знаковый разряд.

Алгоритм сжатия в отличие от стандартного алгоритма выполнен по упрощенной схеме. Размер тайла соответствует размеру изображения. В процессе сжатия последовательно по субдиапазонам каждый коэффициент ДВП

проверяется на выполнения условия $\hat{S}_j = 0$. Если это условие выполняется, вычисляется длина серии нулей. Если не выполняется, то кодируется ненулевой коэффициент. Серии нулей и ненулевые коэффициенты кодируются кодами переменной длины в соответствии с кодовыми таблицами.

Комбинированный кодер длин серий нулей с переменной длиной кода подобран экспериментально в соответствие со статистикой вейвлет-трансформант реальных изображений. Он работает следующим образом. Вейвлет-трансформанта просматривается построчно. Если в исходном потоке обнаруживается от одного до семи нулевых коэффициентов, то в выходной поток записывается на единицу большее число нулевых битов. Если обнаруживается восемь или более нулей, то в результирующем битовом потоке записывается префикс серии нулей (2 бита – 01), префикс длины серии (2 бита) и код длины серии. Префиксы и коды длин серий перечислены в табл. 3.3.

Таблица 3.3. Префиксы длины серии

Префикс длины серии	00	01	10	11
Длина кода длины серии	3	6	12	24
Минимальная длина серии	8	16	80	4176
Максимальная длина серии	15	79	4175	16781391

Для кодирования отличных от нуля значений используется код, длина которого увеличивается с увеличением абсолютного значения коэффициента и соответственно уменьшением его частоты появления. Перед ненулевым символом следует префикс переменной длины, определяющий длину следующего за ним поля для хранения значения коэффициента. Префиксы и коды ненулевых коэффициентов перечислены в табл. 3.4.

Таблица 3.4. Префиксы ненулевых коэффициентов

Префикс ненулевого коэффициента	10	110	1110	11110	11111
Длина кода коэффициента	1	2	4	8	16
Минимальный коэффициент	-1	-3	-11	-139	-32907
Максимальный коэффициент	1	3	11	139	32907

В результате экспериментов получено, что эффективность сжатия нашего алгоритма практически не уступает стандартному алгоритму сжатия JPEG2000 реализованного в приложении Adobe Photoshop CS6.

3.2.3. Результаты экспериментов

Для получения оптимального значения σ был выполнен анализ качества оригинальных фотореалистичных изображений из базы данных лаборатории LIVE, а также из других баз данных.

На рис. 3.4 показаны примеры тестовых изображений: «Женщина в шляпе» (Woman_hat), «Байкеры» (Bikes) и «Попугаи» (Parrots). В табл. 3.5 приведены результаты их анализа.

Для каждого изображения приведен уровень детальности (FDL) который показывает процент участков изображения с мелкими структурами. Серым цветом выделен оптимальный параметр $\check{\Delta}_1^{HH}$, при котором выполняется критерий (3.15). Для сравнения приведена оценка искажений по метрике SSIM.



Рис. 3.4. Примеры тестовых изображений из базы данных лаборатории LIVE

Таблица 3.5. Зависимости искажений от параметров квантователя

Женщина в шляпе (FDL = 21%)				
Δ_1^{HH}	$\varepsilon\%$	dE_N	$k_{сж}$	SSIM
20	0,0	0,147	5,9	0,989
24	0,0	0,172	6,6	0,987
28	0,1	0,197	7,3	0,984
32	0,3	0,222	8,0	0,982
36	1,3	0,248	8,7	0,979
40	3,3	0,272	9,3	0,977
44	6,1	0,296	10,0	0,975
Байкеры (FDL = 56%)				
Δ_1^{HH}	$\varepsilon\%$	dE_N	$k_{сж}$	SSIM
20	0,0	0,173	4,2	0,995
24	0,0	0,201	4,6	0,994
28	0,2	0,231	5,0	0,993
32	1,2	0,261	5,3	0,992
36	3,8	0,290	5,7	0,990
40	7,9	0,319	6,0	0,989
44	13,8	0,348	6,3	0,988
Попугаи (FDL = 9%)				
Δ_1^{HH}	$\varepsilon\%$	dE_N	$k_{сж}$	SSIM
20	0,0	0,153	8,2	0,987
24	0,0	0,180	9,4	0,985
28	0,0	0,203	10,7	0,983
32	0,3	0,225	12,0	0,981
36	1,2	0,250	13,3	0,979
40	3,2	0,274	14,7	0,977
44	6,8	0,299	16,1	0,976

В результате анализа более ста оригинальных изображений из разных баз данных получены следующие результаты.

1. Для большинства тестируемых изображений с уровнем детальности $FDL < 50\%$ высокое качество обеспечивается при значении параметра $\Delta_1^{HH} \leq 32$. При этом средняя оценка искажений (3.16) не превышает значения 0,25.

2. Для настройки оптимального значения $\tilde{\Delta}_1^{HH}$ оптимальным интервалом регулировки будет значение $\sigma = \pm 4$. В этом случае достаточно выполнить одну итерацию изменения параметра Δ_1^{HH} .

3. Оценка искажений по метрике SSIM не дает однозначного решения. В частности, если для обеспечения высокого качества выбрать условие $SSIM > 0,98$, то для изображений с высокой детальностью и контрастностью (пример «Байкеры») значение $\tilde{\Delta}_1^{HH}$ может превышать начальное значение Δ_1^{HH} почти в два раза.

В заключении раздела отметим отличительные особенности предложенного способа управления качеством передачи мелких структур изображений. Анализ качества изображений и настройка квантователя выполняются не по интегральным оценкам (SSIM, PSNR и др.), а по критерию (3.15) определяющему допустимые визуальные искажения выделенных участков с мелкими структурами в системе N-CIELAB. При этом искажения оцениваются относительным количеством микроблоков (3×3) в которых максимальное отклонение контраста превышает зрительный порог. С помощью разработанной программной модели кодера JPEG2000, экспериментально определен начальный параметр квантования $\Delta_1^{HH} = 32$ и его шаг настройки $\sigma = \pm 4$, при которых обеспечивается высокое качество изображений с минимальным количеством шагов настройки.

К перспективным исследованиям относится внедрение и анализ эффективности разработанного алгоритма в мультимедийных системах сжатия видеоданных на основе стандартов MPEG-2 и MPEG-4.

3.3. Корректирующие фильтры для улучшения качества изображений видеокодеков

Практически во всех видеокодеках применяются корректирующие фильтры для улучшения субъективного качества изображений на выходе.

Под «корректирующими фильтрами» подразумеваются методы, которые не кодируют информацию об изображении или улучшают эффективность кодирования, но они улучшают зрительное восприятие.

Корректирующие фильтры созданы для того, чтобы обойти или устранить специфические артефакты, незаметные для объективных метрик, но очевидные человеческому глазу. И даже если считать фильтры формой искусственного улучшения качества, хорошему видеокодеку всё равно нужны все практические и эффективные способы улучшения изображения, какие он может использовать.

Фильтры делятся на несколько категорий. Во-первых, они могут быть нормативными или ненормативными. Нормативный фильтр – обязательная часть кодека; без него невозможно правильно декодировать видео. Ненормативный фильтр необязателен.

Во-вторых, фильтры отличаются по месту применения. Существуют фильтры предварительной обработки (препроцессинг), применяемые к

входным данным до начала кодирования, фильтры постобработки (постпроцессинг), применяемые к выходным данным после завершения декодирования, а также контурные фильтры (in-loop filters, loop filters), интегрированные в процесс кодирования.

Фильтры предварительной обработки и постобработки обычно ненормативные и не включаются в кодек. А контурные фильтры являются нормативными практически по определению. Это часть кодека; они используются в процессе оптимизации кодирования и применяются к опорным кадрам или для межкадрового кодирования.

Рассмотрим особенности корректирующих фильтров видеокodeка AV1 [65].

AV1 использует три нормативных корректирующих фильтра (рис. 3.5). Первый – деблокирующий фильтр для удаления «блочности» – очевидных артефактов по краям закодированных блоков. Хотя DCT относительно хорошо подходит для сжатия естественных изображений, но всё равно имеет тенденцию скапливать ошибки на краях блоков. Но AV1 – более традиционный кодек с жёсткими границами блоков. В результате здесь требуется традиционный деблокирующий фильтр, чтобы сгладить артефакты на краях блоков (рис. 3.6)

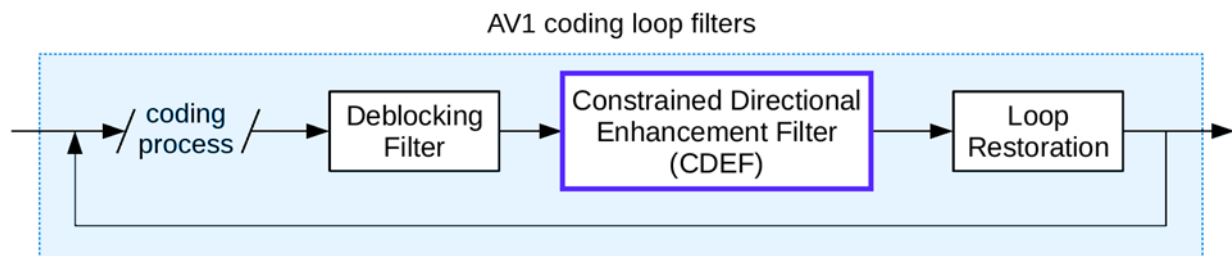
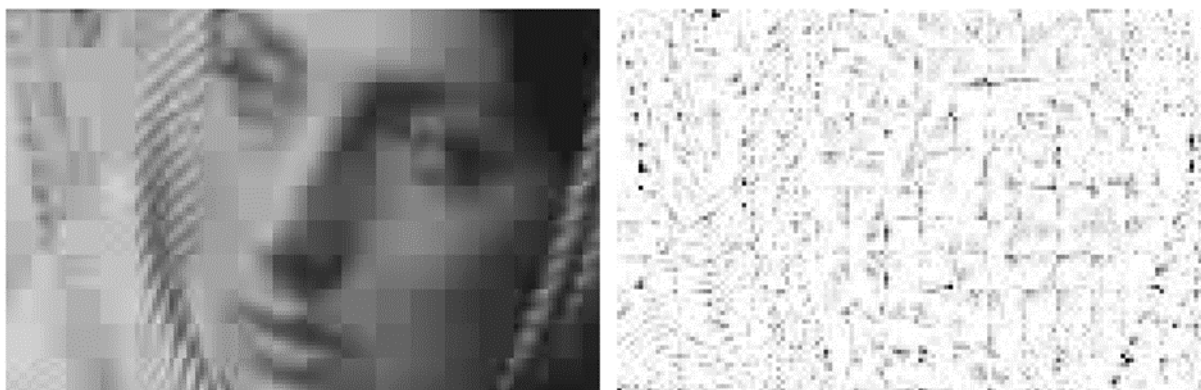


Рис. 3.5. Корректирующие фильтры видеокodeка AV1

Последний из трёх фильтров – это контурный фильтр восстановления (Loop Restoration filter). Он состоит из двух конфигурируемых и заменяемых фильтров: винеровского фильтра и «самонаводящегося» фильтра (Self-Guided filter). Это два свёрточных фильтра, которые пытаются построить ядро для частичного восстановления потерянного качества исходного входного изображения. Обычно они используются для шумоподавления и/или коррекции на краях блоков. В случае AV1 они выполняют общую задачу шумоподавления, удаляя базовый шум DCT через регулируемое размытие.



Closeup of reconstructed image

Normalized error distribution within each block

Рис. 3.6. Пример артефактов на границах блоков в традиционном блочном кодеке DCT

Между ними находится ограниченный направленный корректирующий фильтр (Constrained Directional Enhancement Filter, CDEF). Как и контурный фильтр восстановления, он удаляет артефакты и базовый шум на стыках, но отличие от контурного фильтра восстановления, это направленный фильтр. Отличие от большинства фильтров, он применяется не ко всему подряд, а конкретно находит границы блоков. Поэтому CDEF особенно интересен: это первый практичный и полезный направленный фильтр, который применяется кодировании видео.

Весь смысл преобразования блоков изображения с помощью DCT и DCT-подобных преобразований заключается в том, чтобы представить блок пикселей как можно меньшим количеством коэффициентов трансформации. DCT достаточно хорошо сжимает энергию в большинстве изображений.

Но есть исключения в эффективности сжатия DCT. Например, DCT не очень хорошо преобразует направленные границы или паттерны. Если посмотреть на DCT-выдачу острого диагонального края, то выходные коэффициенты тоже образуют острую диагональ. Она отличается после преобразования, но всё ещё присутствует в изображении, хотя обычно в более сложном виде, чем в начале.

Резкие границы – основная проблема для кодеков на основе DCT, поскольку они не очень хорошо сжимаются, если вообще сжимаются. На рис. 3.7 показана резкая граница (слева) и коэффициенты преобразования DCT (справа). Энергия исходной границы распространяется через DCT в рисунок направленной ряби.

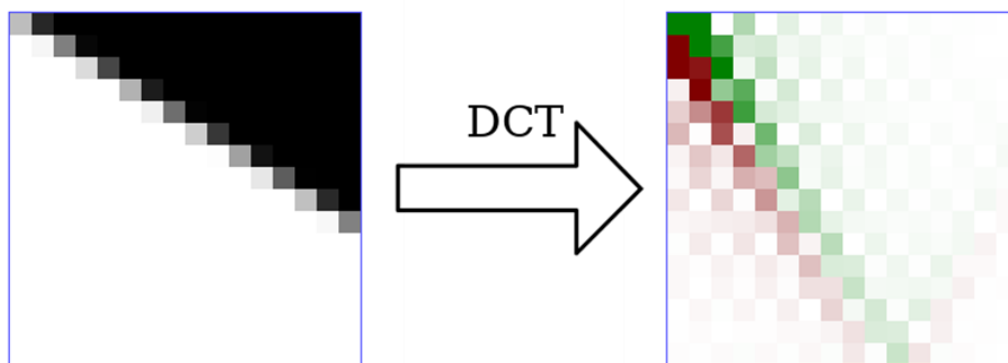


Рис. 3.7. Преобразование резкой границы в кодеке DCT

За два последних десятилетия в исследованиях видеокодеков всё чаще рассматривали преобразования, фильтры и методы прогнозирования, которые по своей сути являются направленными. Исследователи искали способ лучшей обработки подобных границ и шаблонов, чтобы исправить это фундаментальное ограничение DCT.

CDEF предназначен для удаления или уменьшения основного шума и артефактов окантовки (ringing) у резких границ изображения без их размытия или повреждения. Фильтр в настоящее время используется в кодеке AV1 и даёт сглаживающий эффект (рис. 3.8), где на первой врезке показана область без обработки CDEF, на второй – та же область после обработки CDEF.

Количественное значение эффективности любого корректирующего фильтра должно определяться субъективными тестами. Объективные метрики не стоит сбрасывать со счетов, но разработка CDEF создана для получения результата, который выходит за рамки возможностей инструментов объективного тестирования, таких как PSNR и SSIM.

Разработчиками было проведено несколько раундов субъективного тестирования, сначала во время разработки CDEF, а затем провели более полное тестирование объединённого фильтра CDEF. Поскольку это новый фильтр, который вообще отсутствует в предыдущих поколениях кодеков, тестирование в основном состояло из AV1 с включённым CDEF и AV1 без CDEF.

Субъективные результаты показывают статистически значимое ($p < 0.05$) улучшение в трёх из шести клипов. Как правило, это соответствует повышению эффективности кодирования на 5–10%. Это довольно большой выигрыш для одного корректирующего фильтра, добавленного к кодеку [66].

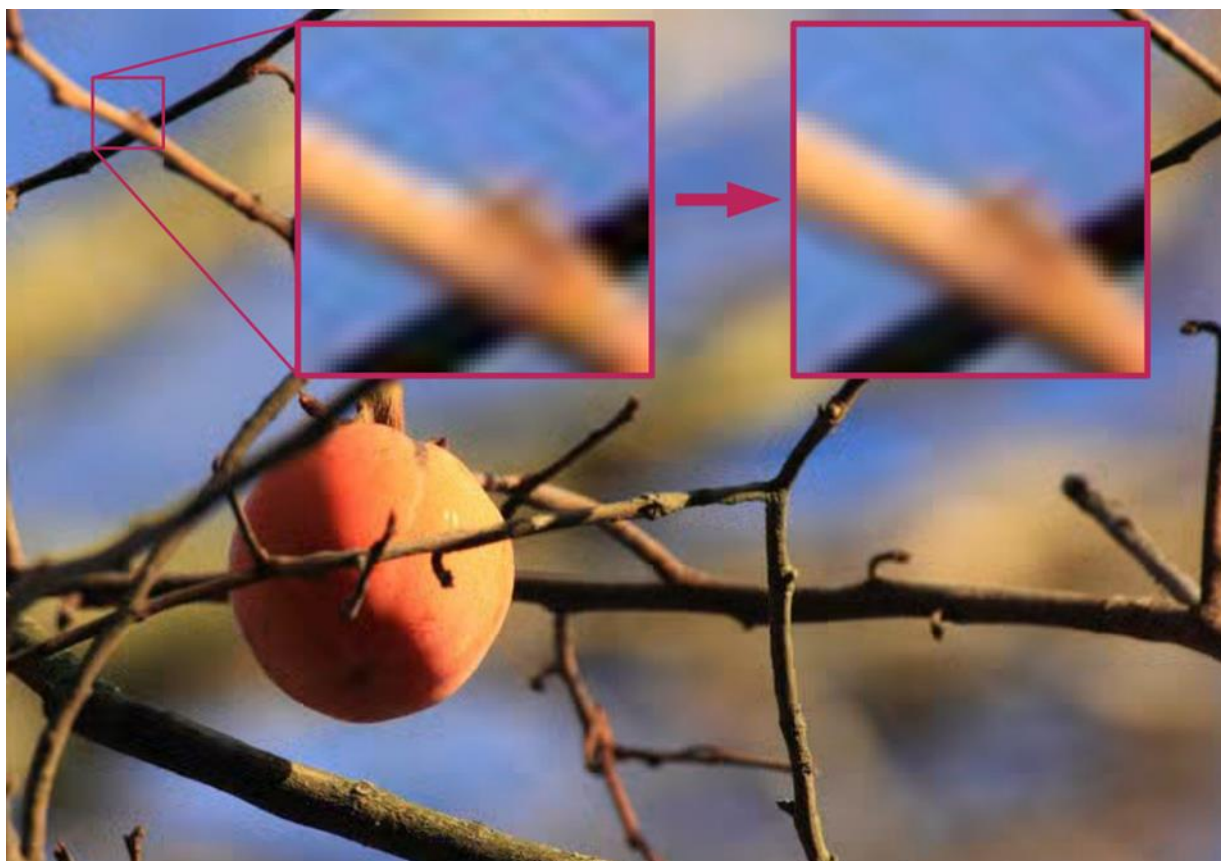


Рис. 3.8. Пример уменьшения шума и артефактов при кодировании изображения.

Тестирование также показало, что CDEF лучше проявляет себя, если во время кодирования отключен дополнительный «инструментарий» кодека. Как и направленные преобразования, CDEF конкурирует с другими, более сложными методами AV1. Поскольку CDEF — простой и быстрый фильтр, в будущем он может найти применение для уменьшения сложности кодера AV1.

3.4. Сравнение качества видеофайлов на основе MSU VQMT

В этом разделе приводятся результаты сравнения качества и эффективности сжатия видеофайлов с помощью анализатора VQMT для различных кодеков.

MSU VQMT - программа для измерения качества видео.

Основные характеристики программы [67]:

1. Поддержка 19 объективных метрик и 5 плагинов. Среди них стандартные метрики (PSNR, VQM, SSIM и др.) и современные метрики с хорошей корреляцией с субъективными оценками (MS-SSIM, PSNR HVS, PSNR HVSM, VIFP).

2. Поддержка более 30 форматов видео, в том числе с 10 -, 14 - и 16-битной глубиной цвета. Поддерживаются форматы AV1, YUV и Y4M (YUV, YV12, UYVY и т.д.), BMP, MOV, VOB, WMV и другие.

3. Визуализация. Программа может генерировать визуализацию искажений видео для каждого видеоформата. Это видео содержит визуальную информацию об искажениях в каждом пикселе кадра с использованием меж-кадровой разности. Визуализация видео позволяет судить о характере качества видео по заметным искажениям.

4. Высокая скорость расчета метрик: поддерживаются многоядерные процессоры, MMX, SSE и оптимизации Open MP, а так же расчёт на графических ускорителях.

5. Результаты сохраняются в *. CSV формате (рис. 3.9).

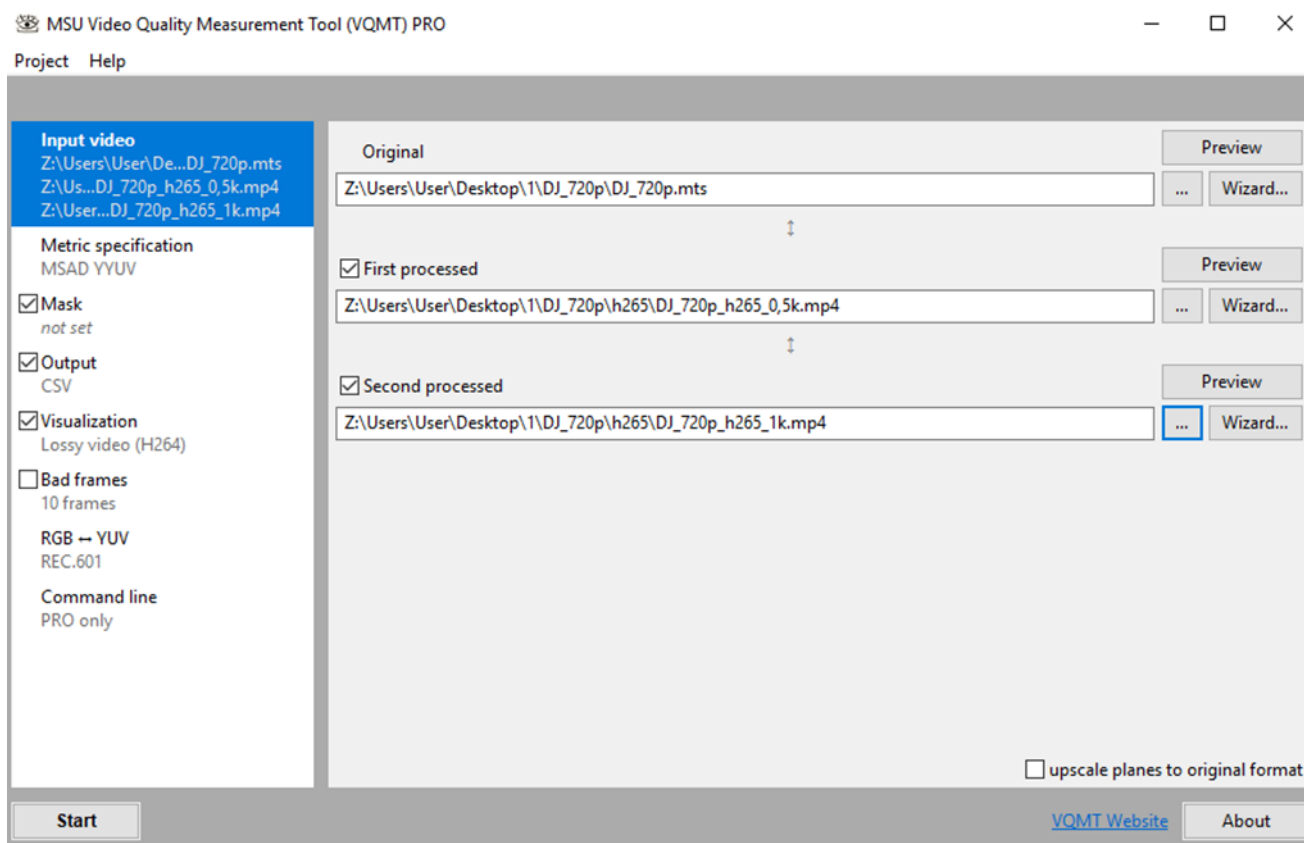


Рис. 3.9. Скриншот программы MSU VQMT

В экспериментальных исследованиях были использованы оригинальные видеофайлы сжатые без потерь качества, над которыми было применено пять основных алгоритмов сжатия видео: AV1, H.264, H.265, MPEG-2 и MPEG-4.

Для сравнения кодеков, далее мы приводим результаты оценок по трем тестовым видеофайлам с разрешением кадра HD (1280×720) и с различной частотой кадров (25 fps – тестовое видео №1, 50fps – тестовое видео №2, 60fps – тестовое видео №3) рис. 3.10 – 3.12. Характеристики тестовых видеофайлов приведены в табл. 3.6 – 3.8.

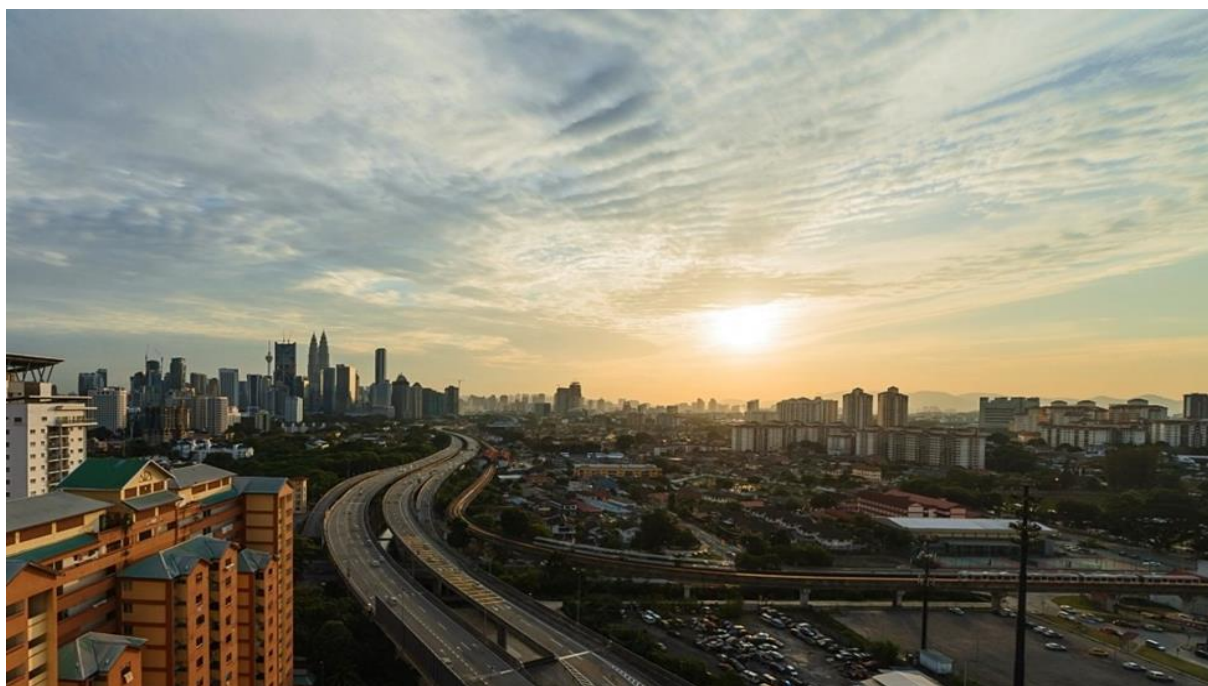


Рис. 3.10. Двенадцатый кадр тестового видео № 1

Таблица 3.6. Характеристики тестового видео № 1

Контейнер	MPEG-4
Профиль формата	Base Media
Идентификатор кодека	isom (isom/iso2/avc1/mp41)
Размер файла	37,4 Мбайт
Продолжительность	15 с. 224 мс.
Общий поток	20,6 Мбит/сек
Формат/Информация	Advanced Video Codec
Профиль формата	High@L3.1
Настройки формата	CABAC / 4 Ref Frames
Параметр CABAC формата	Да
Параметр RefFrames формата	4 кадра
Идентификатор кодека	avc1
Идентификатор кодека	Advanced Video Coding
Продолжительность	15 с. 224 мс.
Битрейт	20,6 Мбит/сек
Ширина	1280 пикселей
Высота	720 пикселей

Продолжение табл. 3.6.

Соотношение сторон	16:9
Режим частоты кадров	Постоянный
Цветовое пространство	YUV
Субдискретизация насыщенности	4:2:0
Битовая глубина	8 бит
Тип развёртки	Прогрессивная
Бит/(Пиксели*Кадры)	0.932
Размер потока	37,4 Мбайт (100%)
Библиотека кодирования	x264 core 138 r2358 9e941d1

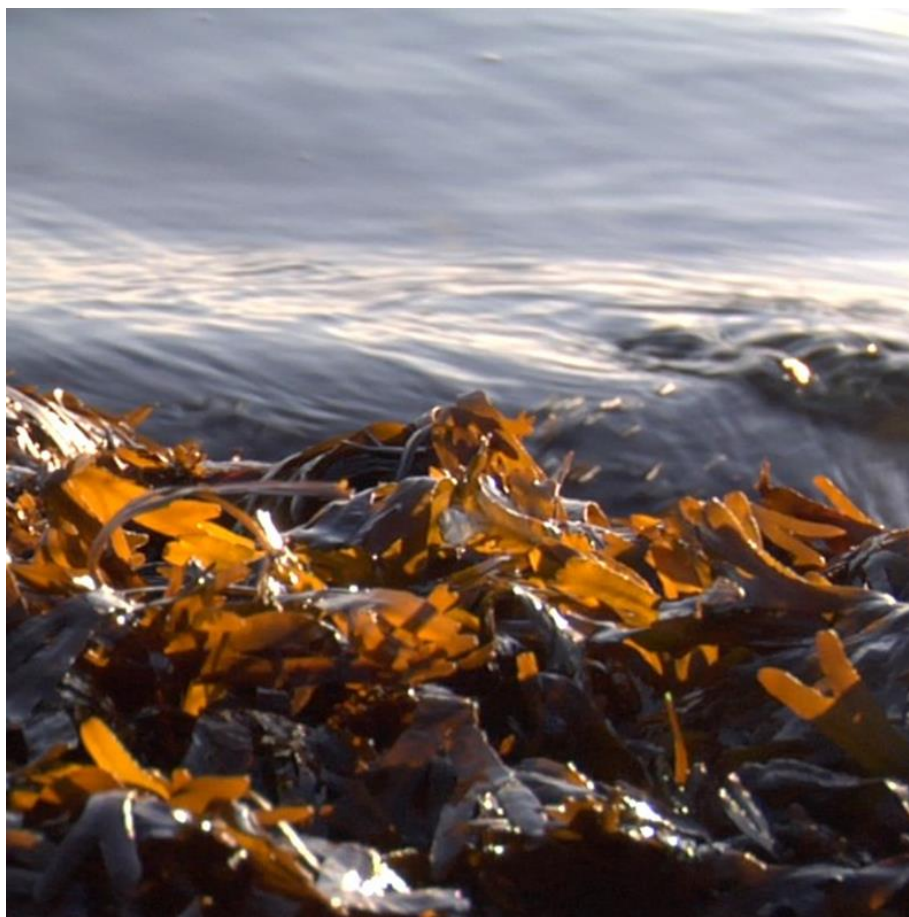


Рис. 3.11. 328-й кадр тестового видео № 2

Таблица 3.7. Характеристики тестового видео № 2

Формат	QuickTime
Формат/Информация	Original Apple specifications
Размер файла	175 Мбайт
Продолжительность	8 с. 0 мс.
Общий поток	183 Мбит/сек
Библиотека кодирования	Apple QuickTime
Формат	ProRes

Продолжение табл. 3.7.

Профиль формата	422 HQ
Идентификатор кодека	arsh
Продолжительность	8 с. 0 мс.
Битрейт	182 Мбит/сек
Ширина	1280 пикселей
Высота	720 пикселей
Соотношение сторон	16:9
Режим частоты кадров	Постоянный
Частота кадров	50,000 кадров/сек
Цветовое пространство	YUV
Субдискретизация насыщенности	4:2:0
Бит/(Пиксели*Кадры)	3.946

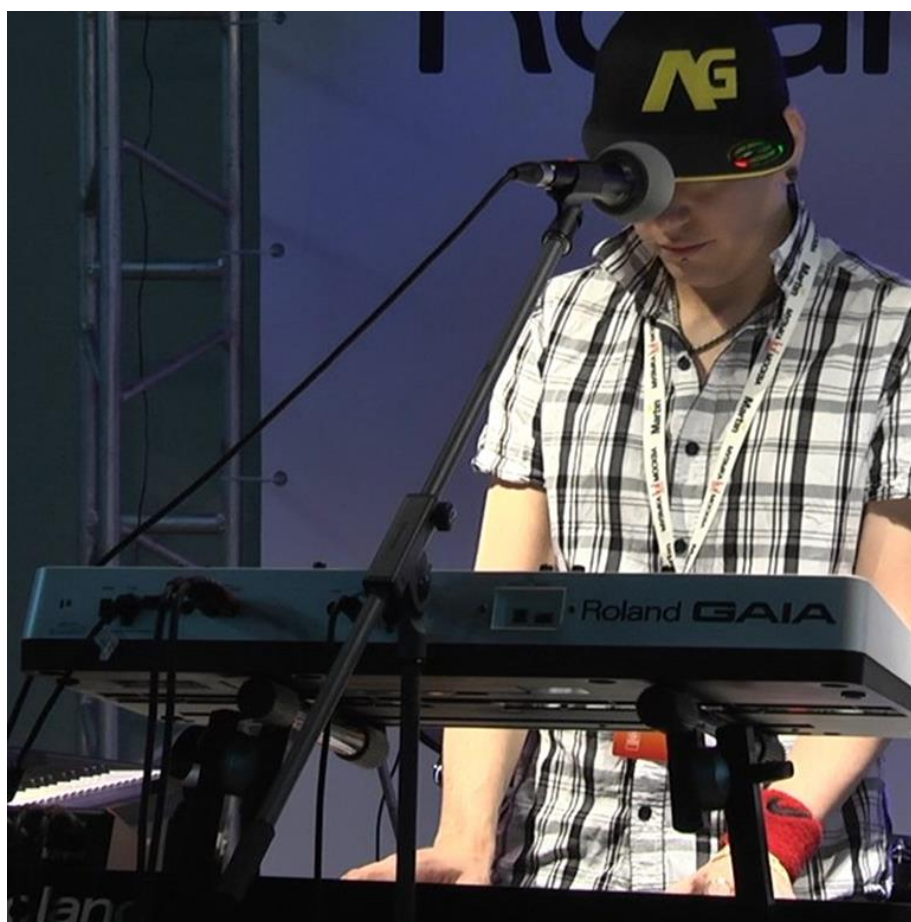


Рис. 3.12. 1604-й кадр тестового видео № 3

Таблица 3.8. Характеристики тестового видео № 3

Формат	BD-AV
Формат/Информация	Blu-ray Video
Размер файла	126 Мбайт
Продолжительность	49 с. 824 мс.

Продолжение табл. 3.8.

Режим общего битрейта	Переменный
Общий поток	21,2 Мбит/сек
Максимальный общий битрейт	35,5 Мбит/сек
Формат	AVC
Формат/Информация	Advanced Video Codec
Профиль формата	High@L4
Продолжительность	49 с. 866 мс.
Вид битрейта	Переменный
Битрейт	20,0 Мбит/сек
Ширина	1280 пикселей
Высота	720 пикселей
Соотношение сторон	16:9
Частота кадров	59,940 (60000/1001) кадров/сек
Цветовое пространство	YUV
Субдискретизация насыщенности	4:2:0
Битовая глубина	8 бит
Тип развёртки	Прогрессивная
Бит/(Пиксели*Кадры)	0.361

Видеофайлы были закодированы 5 разными алгоритмами сжатия с потерями: AV1, H.264, H.265, MPEG-2, MPEG-4, при различных битрейтах (0,5 МБ/с; 1 КБ/с; 1,5 МБ/с; 2 МБ/с; 2,5 МБ/с; 3 МБ/с; 4 МБ/с; 5 МБ/с; 6 МБ/с; 7 МБ/с; 8 МБ/с; 9 МБ/с; 10 МБ/с).

Все закодированные видео сравнивались по следующим метрикам: SSIM, MS-SSIM, PSNR, PSNR HVS, PSNR HVSM, VIFP. Как показали результаты исследований, метрики SSIM, MS-SSIM, PSNR, PSNR HVS и PSNR HVSM дают примерно похожие зависимости.

Зависимости метрики VIFP немного отличаются, т.к. эта метрика основана на использовании специального дискретно-косинусного преобразования (ДКП), которое выявляет в изображении визуальные артефакты блоками 8x8 пикселей.

Результаты сравнения качества трех тестовых видео в зависимости от выбранных кодеков по метрике SSIM показаны на рис. 3.13-3.15. Из полученных объективных оценок следует, что наихудшим кодеком является MPEG-2, далее по качеству следуют MPEG-4, H.264 и H.265. Лучшим кодеком является AV1.

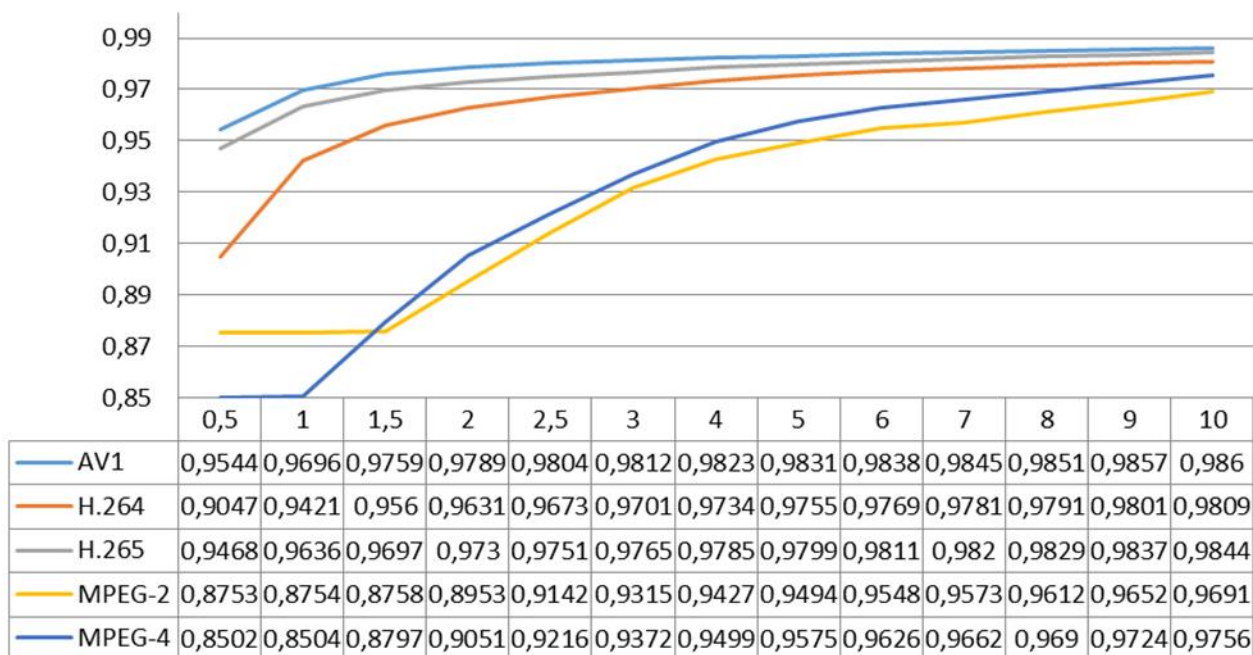


Рис. 3.13. Результаты сравнения качества тестового видео №1 по метрике SSIM

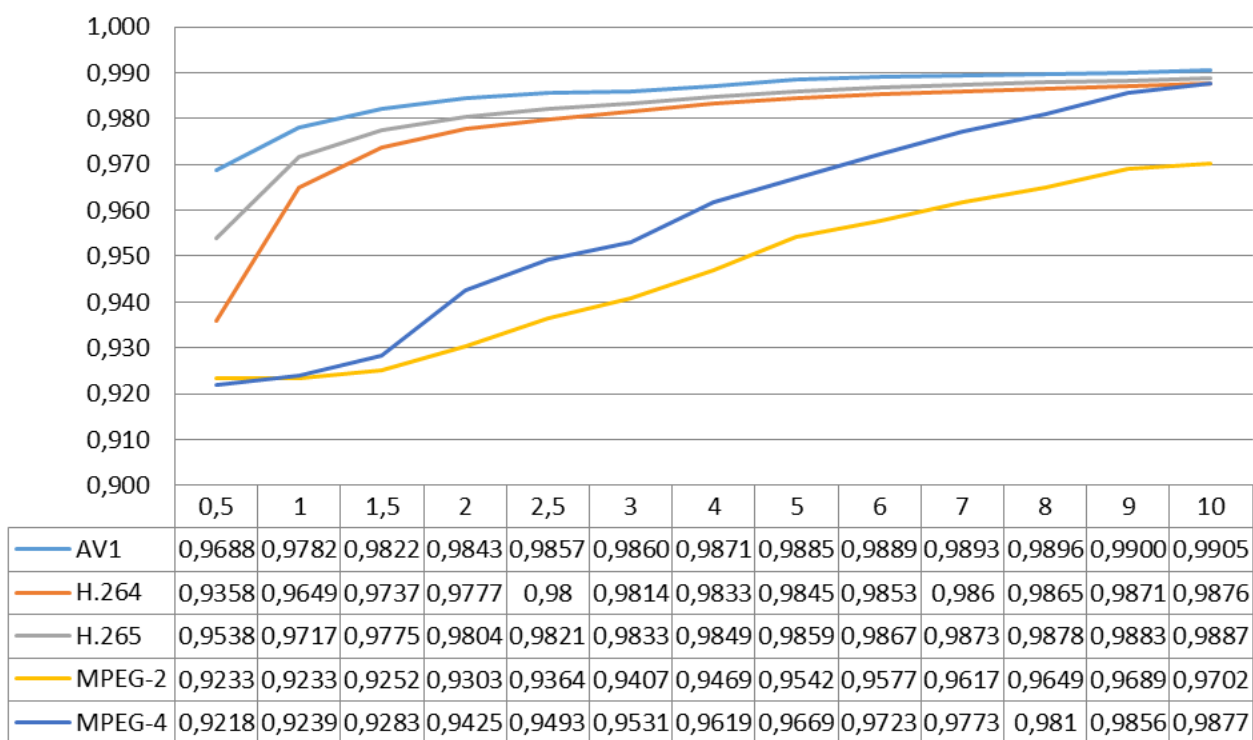


Рис. 3.14. Результаты сравнения качества тестового видео №2 по метрике SSIM

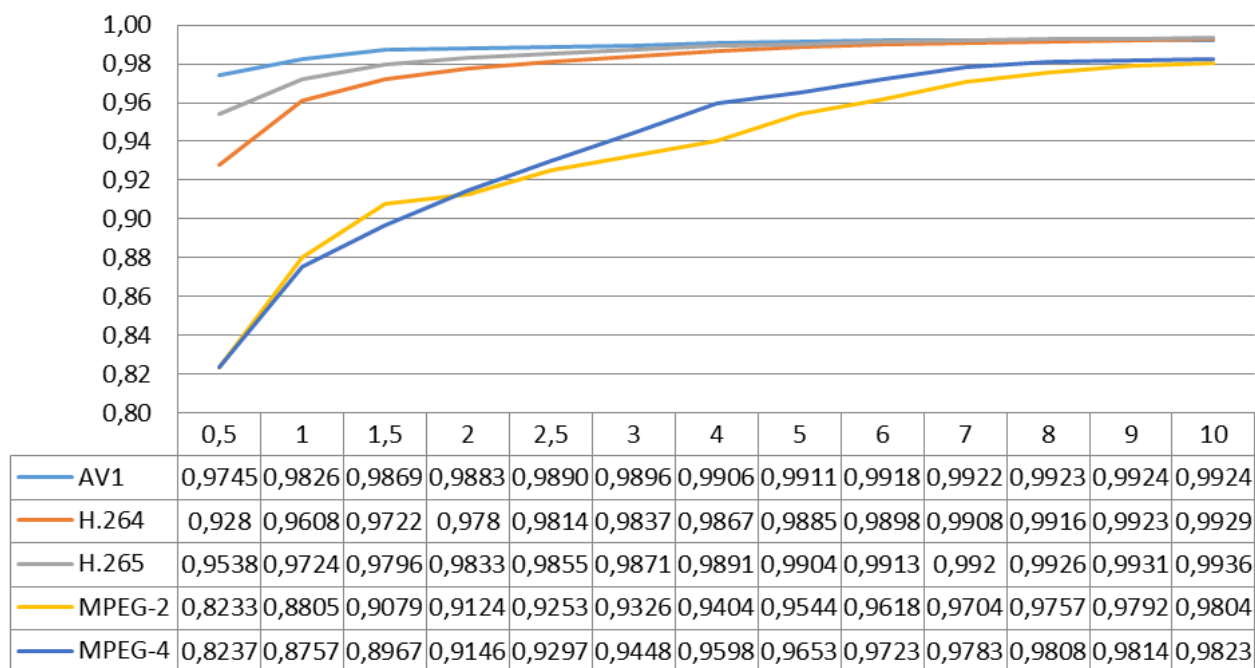


Рис. 3.15. Результаты сравнения качества тестового видео №3 по метрике SSIM

Особенно эти закономерности проявляются при низких значениях битрейта. Для визуального сравнения качества кодеков на рис. 3.16 -3.21 приведены визуальные «карты» искажений, полученные в программе MSU VQMT при битрейте 500 Кб/с для лучшего и наихудшего видеокodeка.



Рис. 3.16. Кадр закодированного видеофайла № 1 кодеком AV1



Рис. 3.17. Кадр закодированного видеофайла № 1 кодеком MPEG-2

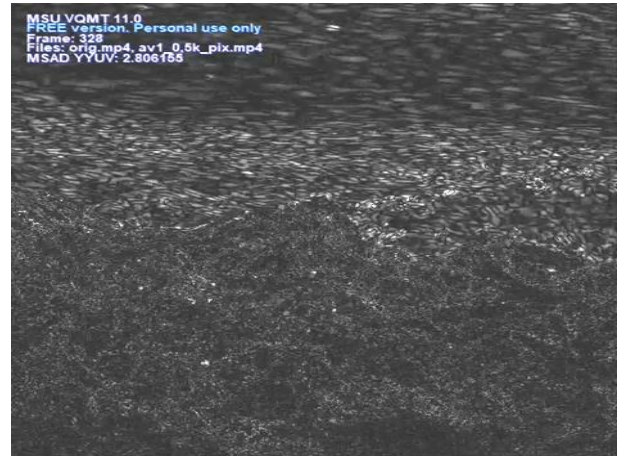


Рис. 3.18. Кадр закодированного видеофайла № 2 кодеком AV1

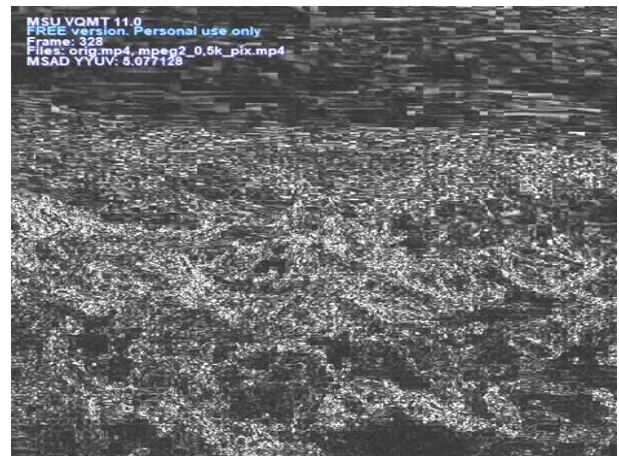


Рис. 3.19. Кадр закодированного видеофайла № 2 кодеком MPEG-2



Рис. 3.20. Кадр закодированного видеофайла № 3 кодеком AV1



Рис. 3.21. Кадр закодированного видеофайла № 3 кодеком MPEG-2

Таким образом, использование инструментов MSU VQMT позволяет выполнять выбор лучшего кодека для решения задач компрессии видеоданных с повышенным качеством передачи цифровых изображений.

Тем не менее отметим, что использованные в программе метрики дают интегральные оценки. Поэтому перспективным направлением будет внедрение локальных оценок качества передачи мелких структур, подобных метрики MFSD (п. 2.5), с целью получения компрессированного видео с качеством близким к зрительному восприятию оригинала.

3.5. Повышение точности фокусировки цифровой камеры

Точность фокусировки камеры является определяющим фактором для получения высококачественных фото и видео изображений. В настоящее время существуют различные способы фокусировки камер, как активные, включающие в себя излучающие дальномеры, так и пассивные. В пассивной фокусировке может использоваться контрастный метод и метод фаз. Самой простой способ это ручная фокусировка.

Автофокус по контрасту используется в большинстве цифровых камер, включая камеры в смартфонах. Контрастная система отличается медлительностью и эффектом «рысканья». Специфика ее в том, что процессор дает команды камере перемещать линзы туда-обратно, находя положение максимального контраста в деталях, аналогично ручной фокусировке на высокую резкость изображения. В отличие от ручной фокусировки и визуальной оценки качества снимка, процессор анализирует гистограмму автоматически и выбирает нужный фокус. Минусов контрастного автофокуса два: 1) медленная скорость работы и 2) неуверенная фокусировка в сложных условиях освещения. Плюс – простота исполнения и как следствие низкая себестоимость производства.

В случае с фазовым автофокусом датчики находятся на самой матрице камеры. Эти датчики определяют, куда падают лучи света, проходящие через противоположные края объектива. Если объект находится в фокусе, отраженные от него лучи света попадают на датчик на определенном расстоянии друг от друга (находятся в фазе), если объект не в фокусе, это расстояние другое (не в фазе), меньше или больше. Таким образом система точно знает, в какое положение надо передвинуть линзу, чтобы достичь необходимого расстояния между лучами. Именно поэтому автоматике важно знать оптические параметры объектива, этот объектив должен иметь быстрый и точный мотор, так как речь идет о микронах и миллисекундах.

Минусов фазового автофокуса тоже два: 1) сложная конструкция и как следствие удорожание техники, 2) необходим совместимый объектив. Плюсы – точность и высокая скорость.

Возникает вопрос – какой режим лучше и как оценить его качество? Самый достоверный способ это визуальная оценка качества полученных фотоснимков или видео для выбранных режимов фокусировки. В этом случае критерием качества будут четкость и резкость полученных изображений. Так

как, не все режимы обеспечивают требуемое качество при различных условиях съемки, то выбор режима фокусировки по визуальным оценкам представляется достаточно трудоемким процессом.

В настоящем разделе мы предлагаем автоматизировать данный процесс с помощью объективного анализатора на основе метрики оценки уровня мелких деталей FDL (2.16). Суть данной метрики подробно описана в п. 2.5.

Рассмотрим применение метрики FDL в системе автоматической фокусировки камеры.

Способ автоматической пассивной фокусировки камеры на малоразмерных объектах изображения реализован с помощью следующего алгоритма.

1. На первом этапе настройки устанавливаем начальное (минимальное) значение фокусного расстояния (F) объектива и вычисляем FDL для выделенного фрагмента.

2. Изменяем значение F с заданным шагом до максимального значения и на каждом шаге i вычисляем FDL_i .

3. Из полученного массива FDL_i находим максимальное значение FDL_{max} и определяем оптимальное фокусное расстояние F_o .

Шаг изменения фокусного расстояния ΔF и его диапазон ($F_{max} - F_{min}$) выбираются программным способом в зависимости от типа камеры. При этом количество анализируемых фрагментов изображения определяются следующим выражением:

$$N_{image} = (F_{max} - F_{min}) / \Delta F.$$

Для реализации способа фокусировки и оценки его эффективности была выбрана цифровая камера Sony ILCE-6000 с объективом E 3.5-5.6/PZ 16-50 OSS, которая по сравнению с аналогами имеет несколько режимов фокусировки и считается наиболее высококачественной камерой в настоящее время.

Способы фокусировки, реализуемые в камере Sony ILCE-6000 с объективом E 3.5-5.6/PZ 16-50 OSS перечислены ниже.

AF-S – покадровый автофокус. При нажатии кнопки спуска затвора наполовину камера пытается сфокусироваться на объекте и в случае удачи удерживает фокусировку. Покадровый автофокус отлично подходит для съемки неподвижных объектов, например пейзажей, архитектуры и т.п.

AF-A – автоматический автофокус, установка по умолчанию. То же самое, что и покадровый, но он переключается в непрерывный при перемещении

объекта. При нажатии кнопки спуска затвора наполовину: если камера поймет, что объект неподвижен, она заблокирует фокус, если объект станет перемещаться, камера будет продолжать выполнять фокусировку. Во время непрерывной съемки камера автоматически переключается в режим непрерывной фокусировки, начиная со второго кадра в серии.

AF-C – непрерывный автофокус. В этом режиме камера продолжает выполнять фокусировку, пока кнопка затвора удерживается наполовину нажатой. Используется для съемки движущихся объектов.

DMF – прямая фокусировка. Этот режим при половинном нажатии на кнопку спуска переводит кольцо на объективе из режима зума в режим подстройки фокуса, позволяя таким образом одновременно использовать и ручную фокусировку, и автофокусировку.

MF – регулировка фокуса вручную. Помощь при ручном фокусе – на экране камера показывает увеличенный фрагмент объекта в центральной области и окрашивает зоны, которые попали в фокус, белым цветом.

Пользователь может также выбрать следующие области фокусировки: широкая АФ, зонная фокусировка, центр, регулируемое пятно АФ.

На следующем необходимо было выбрать устройство дистанционного управления и способ подключения камеры к ПК. Для управления камерой было выбрано программное приложение Remote Camera Control Ver.3.5 которое устанавливает USB-подключение между камерой и компьютером и позволяет управлять камерой с компьютера. С помощью данного приложения пользователь может: задавать режимы фокусировки; изменять фокусное расстояние в ручном (MF) и в полуавтоматическом режиме (DMF); задавать качество, размер и формат изображения; выполнять фото и видео съемку для установленных параметров и записывать файлы в ПК.

Для передачи реального видео с камеры в ПК в режиме Live View выбрана плата видеозахвата Aver Media Dark Crystal HD Capture Pro которая позволяет захватывать HD-видео с HDMI-входа без компрессии в формате 1920×1080 пикселей.

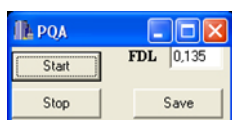
Для анализа качества изображений, на основе метрики *FDL* разработан программный модуль – анализатор качества изображения (PQA) позволяющий выбирать и настраивать режимы фокусировки камеры на высокую четкость изображений по объективному критерию. Описание программного модуля приведено в ниже.

Методика эксперимента состоит из следующих этапов:

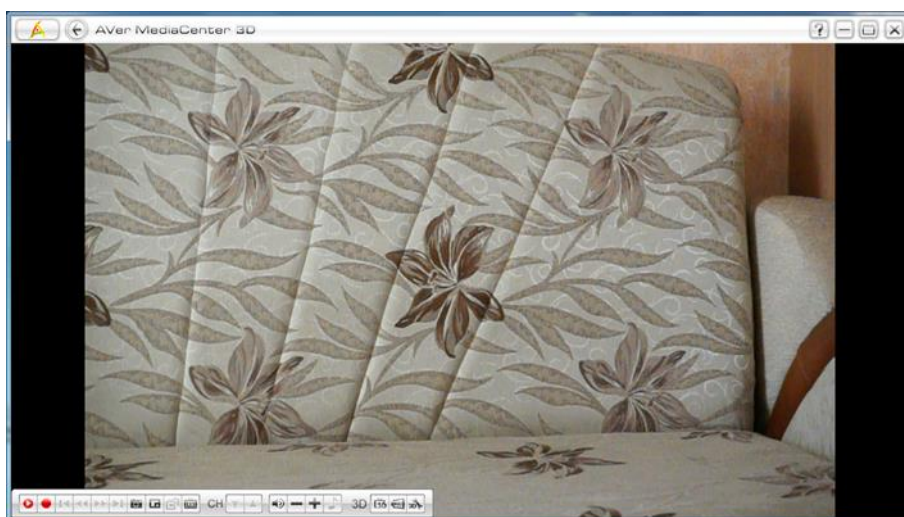
1. На первом этапе подключаем камеру к ПК с помощью интерфейсов HDMI и USB.

2. На втором этапе запускаем приложения: Remote Camera Control, Aver Media и PQA. На экране монитора ПК появляются интерфейсные окна всех приложений (рис. 3.22).

3. На третьем этапе: устанавливаем Aver Media в режим реального видео (Live View), выбираем с помощью Remote Camera Control режим фокусировки и оцениваем значение *FDL* по изображению с камеры.



а)



б)



в)

Рис. 3.22. Интерфейсные окна приложений: а) модуль PQA; б) модуль Aver Media; в) модуль Remote Camera Control

Для решения задачи проверки точности фокусировки камеры была разработан программный модуль в среде C++ реализующий следующие функции.

При нажатии кнопки «Start»:

1. Захват фрагмента изображения в центре экрана с размерами 512×512 пикселей.

2. Вычисление значения *FDL* и вывод в окно.

3. Повторение п. 1-2 в цикле с заданным интервалом времени.

При нажатии кнопки «Stop»:

4. Сохранение значения *FDL* и вывод текущего фрагмента изображения в интерфейсное окно модуля (рис. 3.23).

При нажатии кнопки «Save»:

5. Сохранение текущего фрагмента изображения в виде файла *.bmp.

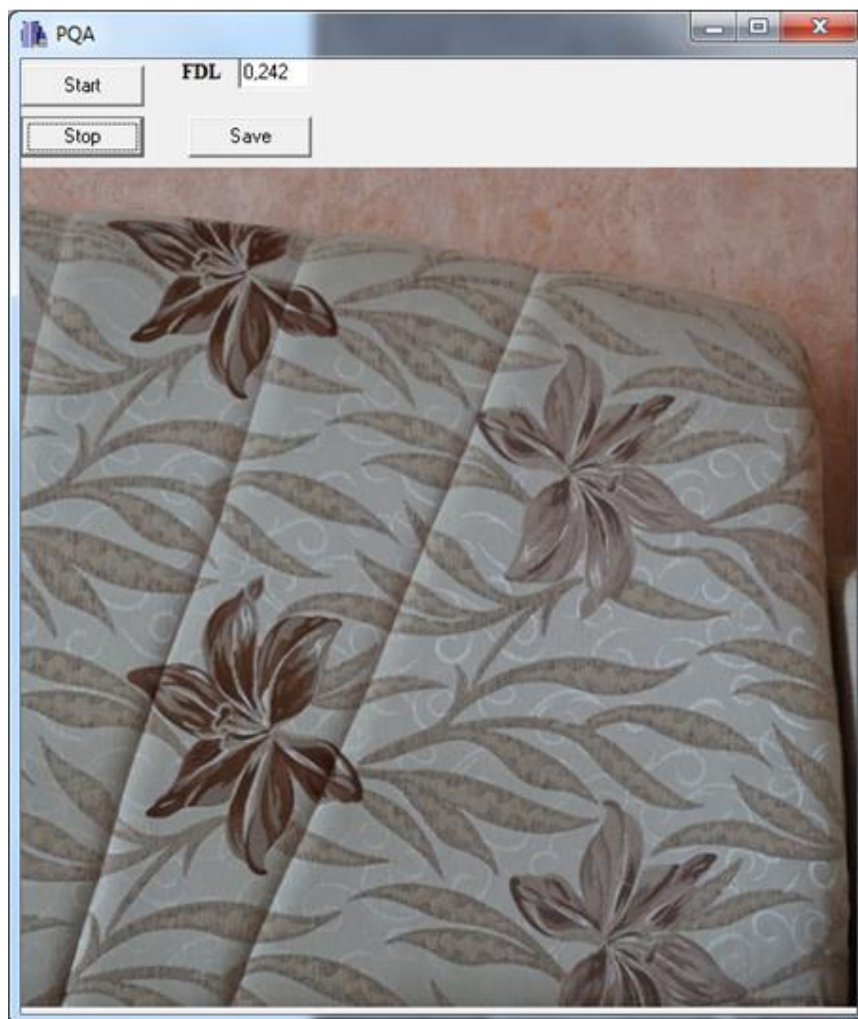


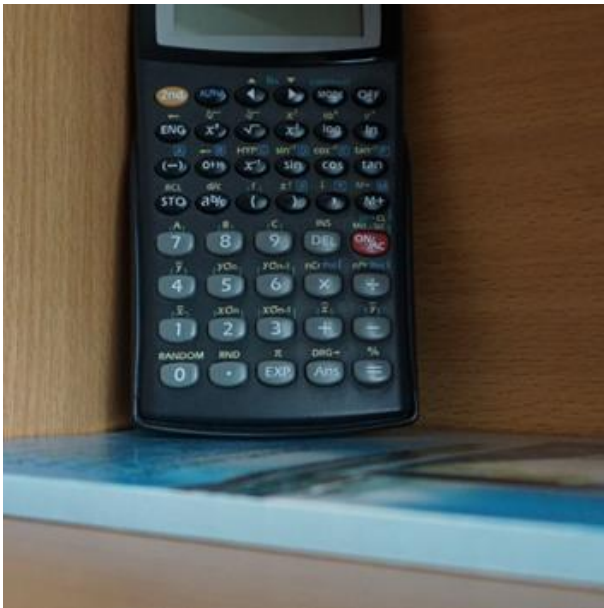
Рис. 3.23. Интерфейсное окно модуля PQA

Для оценки эффективности разработанного способа был проведен ряд экспериментов с использованием камеры Sony ILCE-6000. В задачу экспериментов входило:

1. Оценка точности автоматической фокусировки для режимов AF-S, AF-A, AF-C и DMF при выбранных условиях фотосъемки (расстояние до объекта, неподвижность объекта, освещенность и др.) и при заданных параметрах камеры (размер и качество изображений, значение ISO, область фокусировки и др.).

2. Оценка эффективности настройки камеры на высокую четкость и резкость изображения по метрике *FDL* в полуавтоматическом режиме.

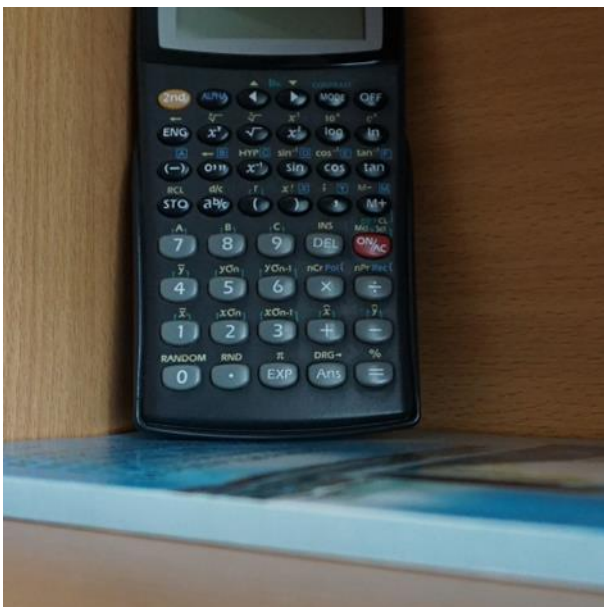
На рис. 3.24 показаны результаты оценки точности автоматической фокусировки для тестового изображения «Калькулятор» и на рис. 3.25 для тестового изображения «Диван» при выбранном пороге $Th = 2$ NPCD.



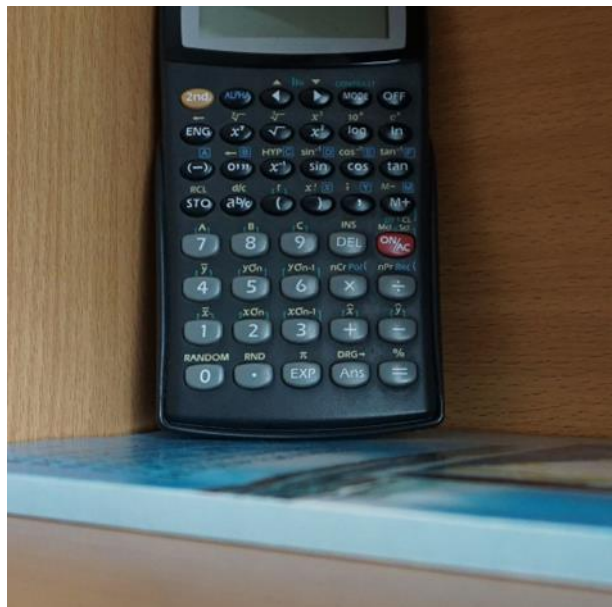
а) AF-A: $FDL = 0,202$



б) AF-C: $FDL = 0,195$



в) AF-S: $FDL = 0,190$



г) DMF: $FDL = 0,211$

Рис. 3.24. Тестовое изображение «Калькулятор»

Из результатов экспериментов следует, что при автоматической фокусировке камеры визуальное качество изображений достаточно высокое и при этом затруднительно выбрать наиболее предпочтительный режим по субъективным оценкам.



а) AF-A: $FDL = 0,354$



б) AF-C: $FDL = 0,360$



в) AF-S: $FDL = 0,353$

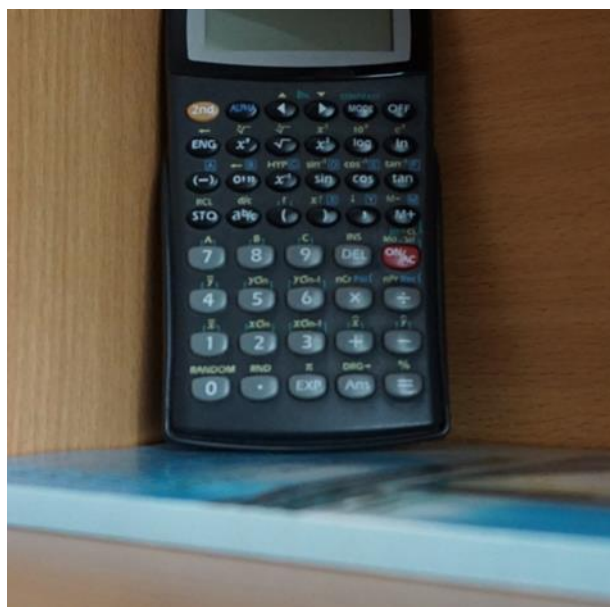


г) DMF: $FDL = 0,351$

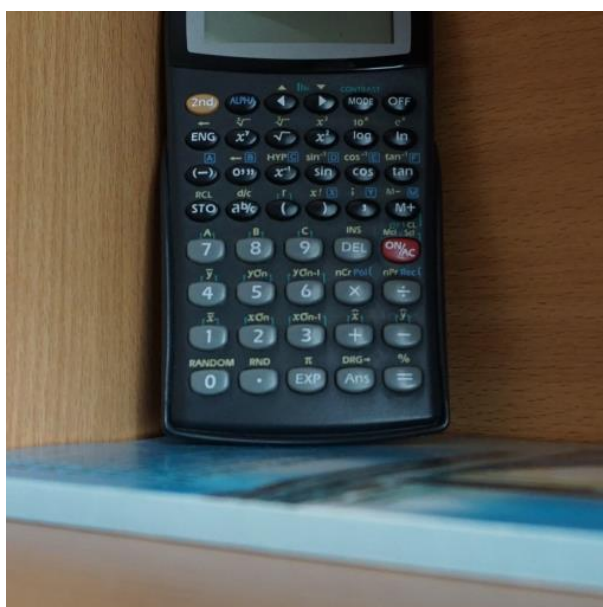
Рис. 3.25. Тестовое изображение «Диван»

Использование метрики FDL позволяет решить данную проблему. В частности, для первого примера (см. рис. 3.24) максимальное значение FDL соответствует режиму DMF и для второго примера (см. рис. 3.25) режиму AF-C.

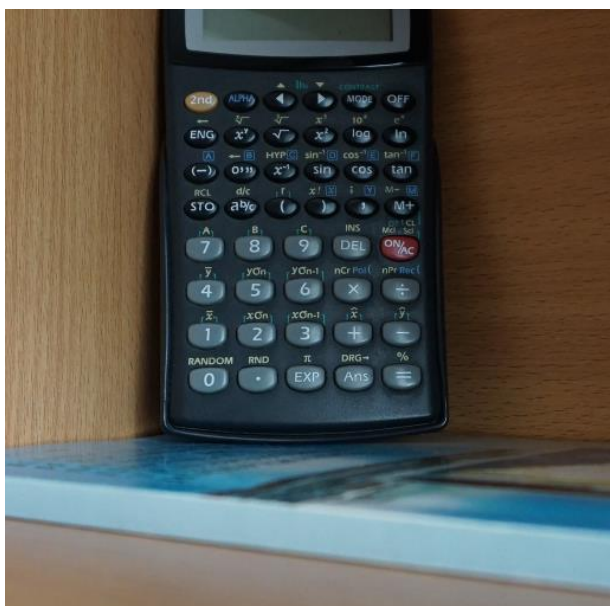
Для исследования способа фокусировки в полуавтоматическом режиме, использовался ручной режим камеры MF с применением функции микро регулятора фокусного расстояния реализуемой в приложении Remote Camera Control. На рис. 3.26 показаны результаты оценки точности фокусировки при разных значениях фокусного расстояния F для тестового изображения «Калькулятор».



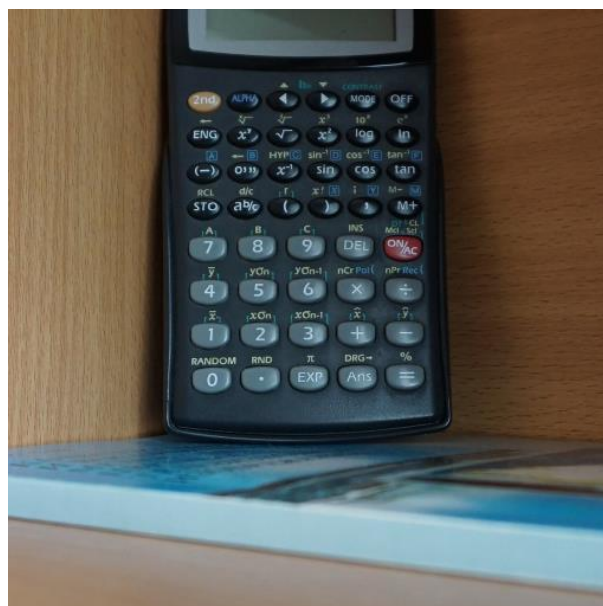
а) $F1$: $FDL = 0,152$



б) $F2$: $FDL = 0,174$



в) $F3$: $FDL = 0,209$



г) $F4$: $FDL = 0,200$

Рис. 3.26. Тестовое изображение «Калькулятор» в режиме MF

Из примера следует, наилучшая четкость и резкость тестового изображения обеспечивается при фокусном расстоянии $F3$. При этом качество изображения не хуже чем для автоматических режимов (см. рис. 3.24).

В результате ряда экспериментов получено, что разработанной способ фокусировки не уступает по качеству автоматическим режимам фокусировки камеры, а в некоторых случаях превосходит.

4. Цифровые фильтры повышения четкости изображений

4.1. Особенности измерения разрешения по полю изображения

Цифровые изображения подвержены воздействию различных искажений, которые могут появиться в зависимости от способа получения изображений или способа передачи информации, а также и методов оцифровки видеоданных. Соответственно, в цифровой обработке изображений решаются задачи обработки изображений с целью обеспечения их хранения, передачи, визуализации в электронном виде и последующего анализа заложенной в них информации.

Для решения задач в области цифровой обработки изображений требуется применять специальные методы и алгоритмы, а также неоднократное тестирование с привлечением широкой базы различных изображений для улучшения их визуального восприятия и повышения информативности для систем технического зрения. Одни алгоритмы направлены на устранение одного из типов искажений в изображении, а другие учитывают положение объектов интереса на изображении.

Задача повышения четкости изображений возникает, как правило, при решении большинства практических задач.

Главная цель повышения четкости (разрешения) изображений заключается в том, чтобы подчеркнуть мелкие детали или улучшить те детали, которые оказались расфокусированы вследствие ошибок или несовершенства самого метода получения информации.

Так же отметим, что при решении задачи улучшения визуального качества изображений, как правило, не возникает задача восстановить изображение, то есть вернуть его к “оригиналу”. При повышении резкости иногда следует проинвестировать перекомпенсацию искажений, то есть избыточно повысить уровень высокочастотных составляющих пространственного спектра изображения [68].

В любом случае обработка включает в себя дискретизацию, квантование, выполнение различных ортогональных преобразований и сверток изображения. Однако более глубокий анализ демонстрирует, что переход от одномерного случая к двумерному, является не только количественным, но и качественным. Проблемы, которые возникают при обработке многомерных сигналов, просто не существуют в обработке одномерных сигналов и, соответственно наоборот, многие трудности цифровой обработки сигналов не существуют или легко снимаются при переходе от одномерных к многомерным сигналам [69].

Под полем изображения понимается, пространство попадающие в поле зрения оптического прибора (ОП), с помощью которого формируется видео-изображение.

Полем зрения ОП называется та часть пространства предметов, которая видна или изображается с помощью ОП. Поле зрения принято измерять в угловой мере (2ω), если наблюдаются удаленные предметы, и в линейной мере ($2y$), если предметы расположены близко. В последнем случае величина поля дублируется и в угловом мере. Так поле зрения телескопических систем, определяется в угловой мере (градусной). Поле зрения в приборе ограничивается специальной диафрагмой, называемой диафрагмой поля зрения. Ограничение поля делается потому, что на краях понижается качество изображения, сильно падает освещенность изображения, а в некоторых случаях это ограничение накладывает и фотоприемник [70].

Для понимания процесса распределения разрешения по полю изображения необходимо определить некоторые термины из оптики, которые помогают понять процессы распределения разрешения изображения по его полю. Начнем с того что введем определение оптической системы.

Оптическая система – это совокупность оптических сред, разделенных оптическими поверхностями, которые ограничиваются диафрагмами. Оптическая система предназначена для формирования изображения путем перераспределения в пространстве электромагнитного поля, исходящего из предмета (преобразования световых пучков). Оптическая ось в геометрическом смысле – прямая, проходящая через центры сферических поверхностей, составляющих центрированную оптическую систему (линзу, фотографический объектив). Она часто является осью симметрии в оптической системе [71].

Фокусное расстояние – физическая характеристика оптической системы, характеризующая её основные свойства и, главным образом, увеличение и угловое поле [72]. Для системы линз, как и для простой линзы конечной толщины, фокусное расстояние связано с радиусами кривизны поверхностей, а также показателями преломления оптических материалов и толщин элементов используемой системы.

Фокусное расстояние – это расстояние от передней главной точки до переднего фокуса (для переднего фокусного расстояния), или расстояние от задней главной точки до заднего фокуса (для заднего фокусного расстояния). При этом под главными точками понимаются точки пересечения передней (задней) главной плоскости с оптической осью. Фокус оптической системы – это точка,

в которой пересекаются («фокусируются») первоначально параллельные лучи после прохождения через собирающую систему (либо где пересекаются их продолжения, если система рассеивающая). Изображение бесконечно удалённой точки находится в фокусе оптической системы [73]. Множество фокусов идеальной оптической системы определяет её фокальную плоскость. Главный фокус системы является точкой пересечения её главной оптической оси и фокальной поверхности. Часто вместо термина «главный фокус» (передний или задний) используются термины задний фокус и передний фокус [74]. Термин «фокус» в его современном понимании ввёл Кеплер в 1604 году [75].

Передний фокус – точка на оптической оси в пространстве предметов, сопряжённая с бесконечно удалённой точкой, расположенной на оптической оси в пространстве изображений, а расстояние от передней главной точки до переднего фокуса – переднее фокусное расстояние [76]. Задний фокус – точка на оптической оси в пространстве изображений, сопряжённая с бесконечно удалённой точкой, расположенной на оптической оси в пространстве предметов. Соответственно, расстояние от задней главной точки до заднего фокуса называют задним фокусным расстоянием.

Четкость изображения описывается воспроизведением мелких деталей и определяется разрешающей способностью системы формирования изображения. Разрешающая способность оптической системы численно выражается количеством пар черно-белых линий на 1 мм изображения, которое формируется объективом системы. Если плоскость формируемого изображения находится в фокусе объектива, то пучок лучей, исходящий от точки на объекте, сходится в точку на изображении. При расфокусировке точка воспроизводится на формируемом изображении в виде некоторого пятна (кружка размытия), и две близко расположенные точки на исходном изображении сливаются воедино на наблюдаемом изображении. Величина кружка размытия зависит от фокусного расстояния объектива, а также от расстояний от объектива до объекта и до плоскости формируемого изображения. Дискретное изображение будет четким (сфокусированным), если диаметр кружка размытия не превышает шага дискретизации наблюдаемого изображения. В противном случае линейные искажения становятся заметными [77]. Примеры изображения находящегося в фокусе камеры и изображения, которое не попало в фокус камеры, представлены на рис. 4.1.

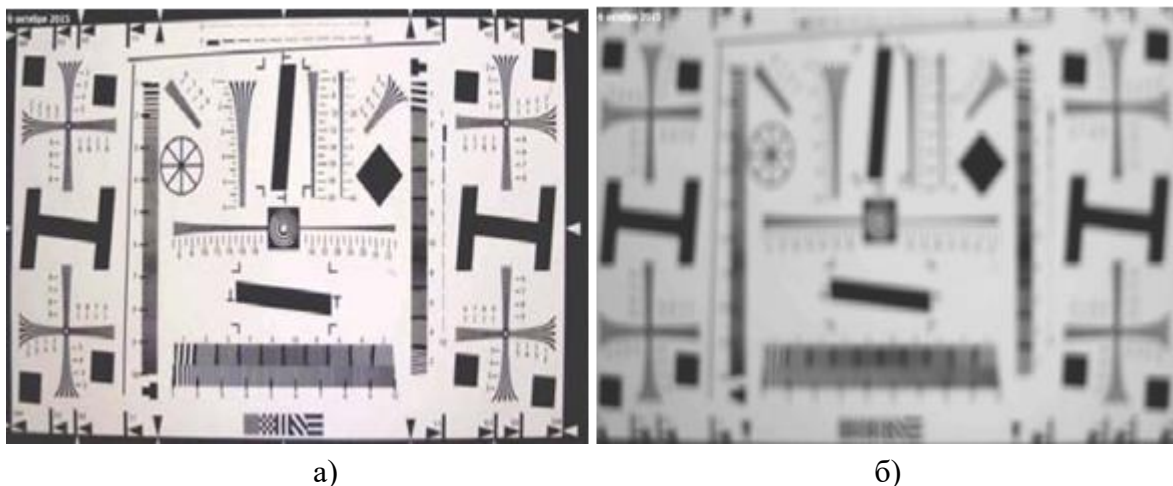


Рис. 4.1. Пример сфокусированного (а) и расфокусированного (б) изображений

Глубина резко изображаемого пространства (глубина резкости) – это расстояние вдоль оптической оси объектива между двумя плоскостями в пространстве предметов, в пределах которого объекты отображаются в сопряжённой фокальной плоскости субъективно резко. Глубина резкости непосредственно зависит от важнейших характеристик оптической системы: главного фокусного расстояния и относительного отверстия, а также от дистанции фокусировки. При этом абсолютно резко отображаются только объекты, расположенные в одной плоскости предметного пространства, соответствующей дистанции фокусировки [78].

Аберрация оптической системы – это ошибка или погрешность на изображении в оптической системе, появляющаяся из-за отклонения луча от того направления, по которому он должен был пройти в идеальной оптической системе. Аберрацию характеризуют различного вида нарушения гомоцентричности в структуре пучков лучей, выходящих из оптической системы. Можно выделить несколько видов аберраций оптической системы:

- монохроматические геометрические аберрации;
- сферическая аберрация;
- аберрация кома;
- астигматизм;
- кривизна поля изображения;
- дисторсия;
- хроматические аберрации;
- дифракционная аберрация.

Из терминологии, представленной выше можно описать процесс распределения разрешения по полю изображения. На рис. 4.2 представлены области

распределения разрешения на изображении, которые описываются ниже. Зеленым цветом выделена область, расположенная в фокусе системы. Синие области с разрешением, ухудшенным за счет aberrаций оптической системы. Красные области, расположенные не в фокусе оптической системы и подверженные aberrациями, ухудшающими разрешение в них.

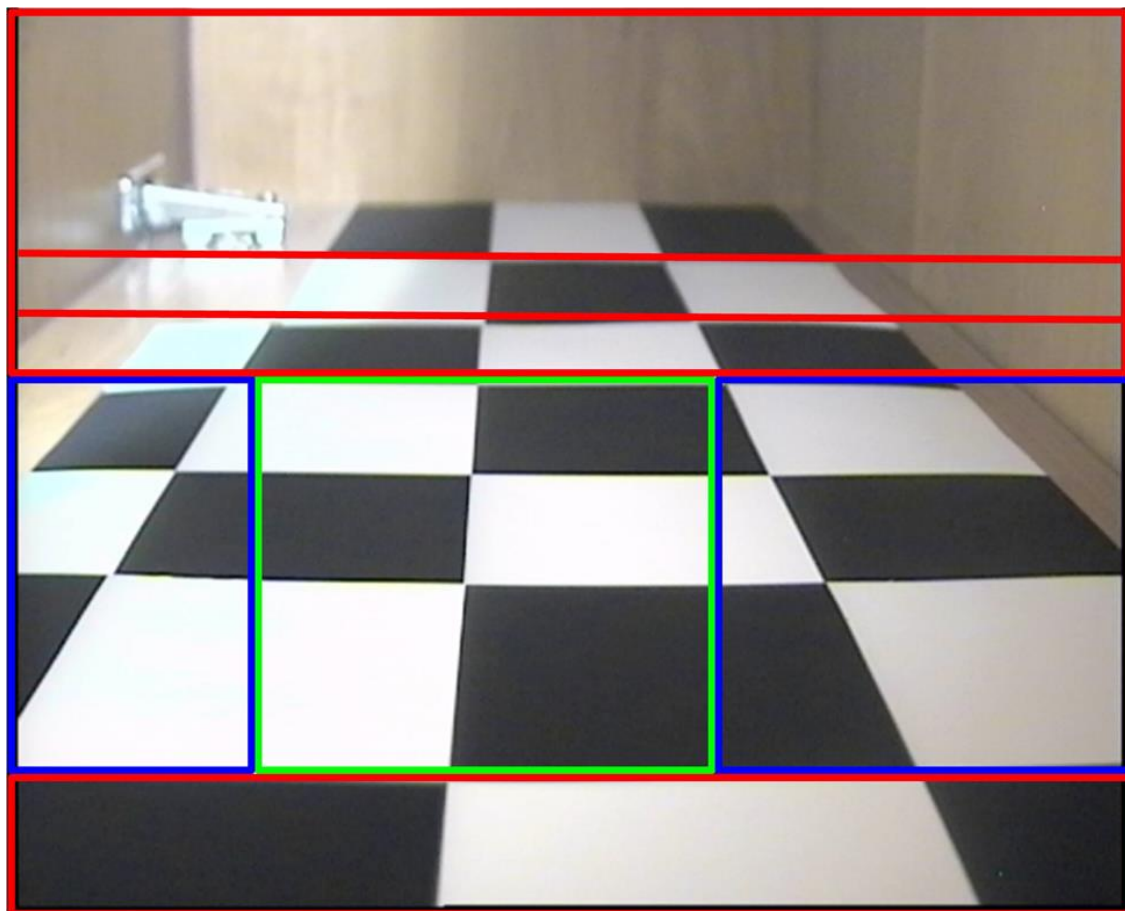


Рис. 4.2. Области распределения разрешения в поле изображения

Часть изображения, которая находится прямо на оптической оси в пространстве изображения и попадает в фокус системы, будет областью с самым высоким уровнем разрешения. Поскольку оптические системы подвержены различным видам aberrаций, вызываемых отклонениям луча светового потока от истинного направления, по которому он должен был следовать в идеальной оптической системе. Из-за эффектов, оказываемых aberrациями по краям изображения возникают искажения, что приводит к снижению уровня разрешения, и соответственно, возникают еще две зоны с небольшим понижением уровня разрешения от области расположенной в фокусе системы. Можно назвать дан-

ные области первой зоной обработки, поскольку ухудшение разрешения на краях изображения не так велико из-за близости к фокусу системы.

Следующей областью изображения можно выделить зону, которая расположена между фокусом системы и объективом. В данной области наблюдается размытие изображения, вызванное тем фактом, что она расположена не в фокусе системы, а также на качество изображения в ней оказывают влияние аберрации, вносимые самой оптической системой.

Условно третьей областью можно определить ту зону изображения, которая находится за оптическим фокусом системы (красная область выше области находящейся в фокусе выделенной синим цветом). В ней на качество изображения влияет тот факт, что область изображения находится на удалении от оптического фокуса системы и существуют все те же эффекты, оказываемые аберрациями. В ней так же наблюдается падение разрешения от края области, расположенной в фокусе к краю изображения.

Так же учитывая тот факт, что фокусировка камеры можно происходить не только в области, представленной на рис 4.2, но и смещаться по оптической оси системы, следовательно, будет изменяться положение остальных областей в аналогичном направлении движения фокуса оптической системы.

С целью подтверждения выдвинутых утверждений по распределению разрешения по полю изображения, была определена следующая методика подтверждения данной гипотезы. На первом этапе рассматривалось изменение разрешения на тестовом изображении, расположенном перпендикулярно оптической оси системы, приведенном на рис. 4.3. Так же на данном изображении обозначены зоны, в которых производилось измерение разрешения.

Проведя измерение разрешения в каждой обозначенной зоне с использованием ПО Imatest [79], была составлена таблица усредненных значений разрешения по столбцам изображения, которые получены суммированием всех значений в столбцах (1+10+19+28+37+46) и делением на общее количество измерений в столбце (6 измерений). Результаты усреднения приведены в табл. 4.1.

Таблица 4.1. Усредненные результаты измерения разрешения

Столбец	1	2	3	4	5	6	7	8	9
ТВЛ	488	550	526	558	574	555	510	514	409

По экспериментальным результатам, полученным после проведения процедуры усреднения получена зависимость разрешения по полю изображения, приведенная на рис. 4.4. В этом случае из графика видно, что максимальный уровень разрешения наблюдается в центральном (5-ом) столбце изображения «шахматное поле» и снижается к его краям.

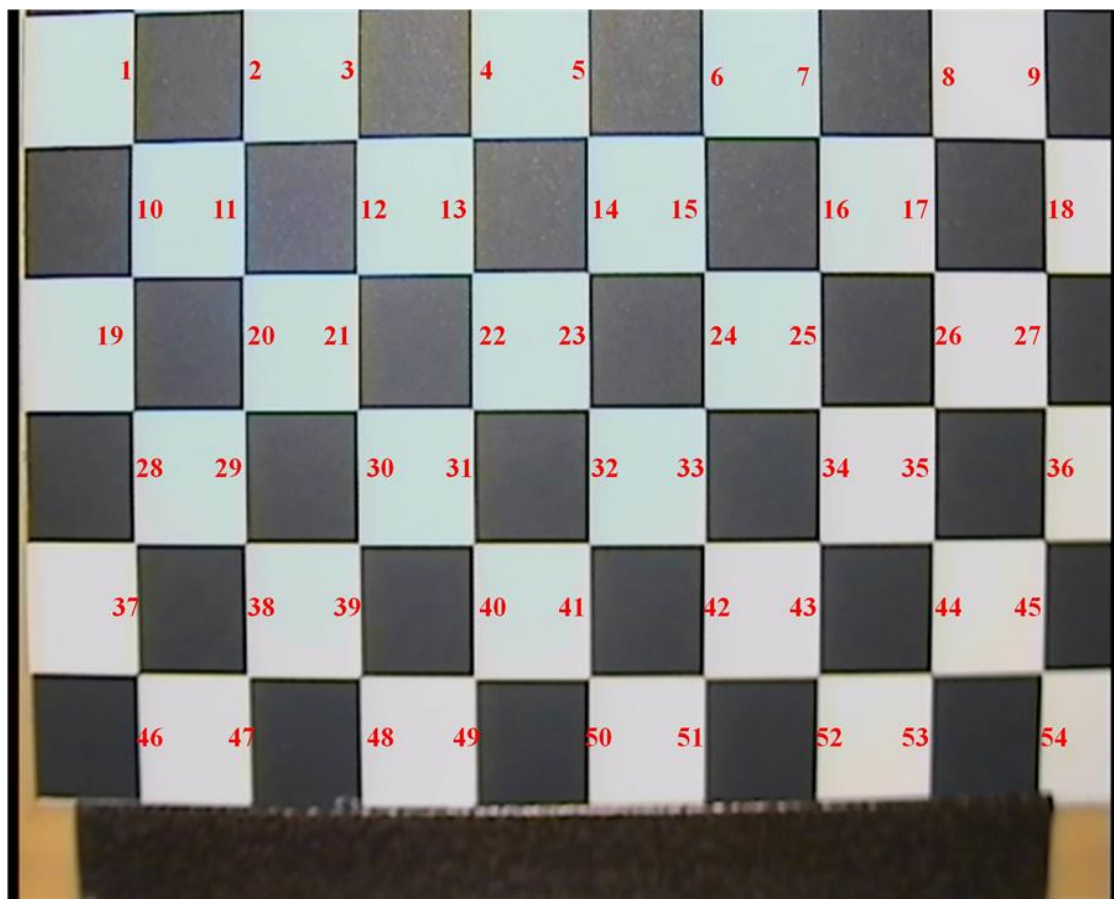


Рис. 4.3. Тестовое изображение с обозначенными зонами измерения разрешения

При измерении разрешения по полю изображения следует учитывать тот факт, что при изменении угла наблюдения за объектом и отсутствии повторной фокусировки объектива произойдет и смещение зон различного разрешения на изображении.

Для подтверждения данного утверждения изменим угол наблюдения за объектом интереса и произведем повторное измерение.

Пример поворота изображения и точек измерения в обоих направлениях представлен на рис. 4.5. Рассмотрим вариант отклонения объекта на 60° в обе стороны от оптической системы наблюдения.

Проведя подобную процедуру усреднения значений разрешения получаем следующие значения, которые представлены в табл. 4.2, а график для данных значений приведен на рис. 4.6.

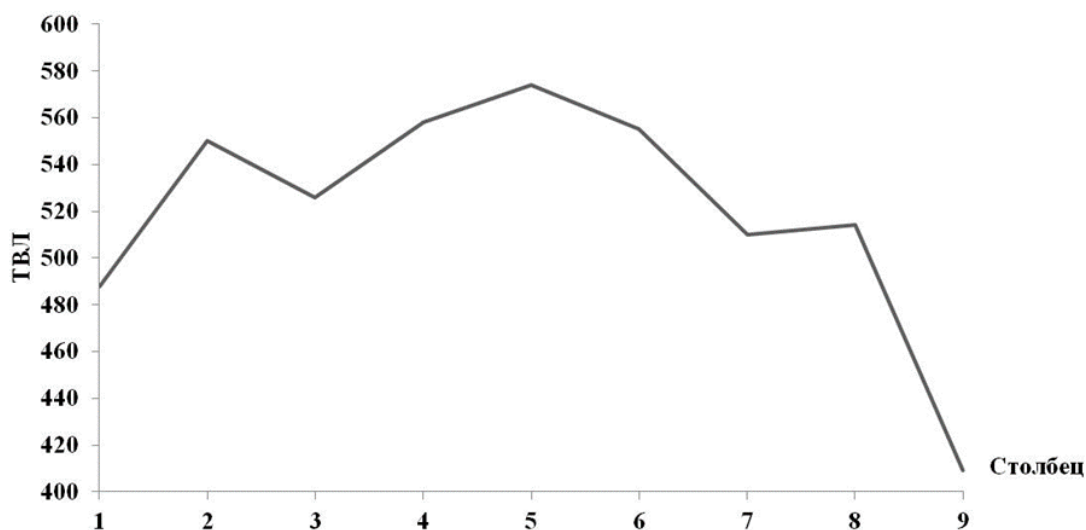


Рис. 4.4. Усредненный график изменения разрешения

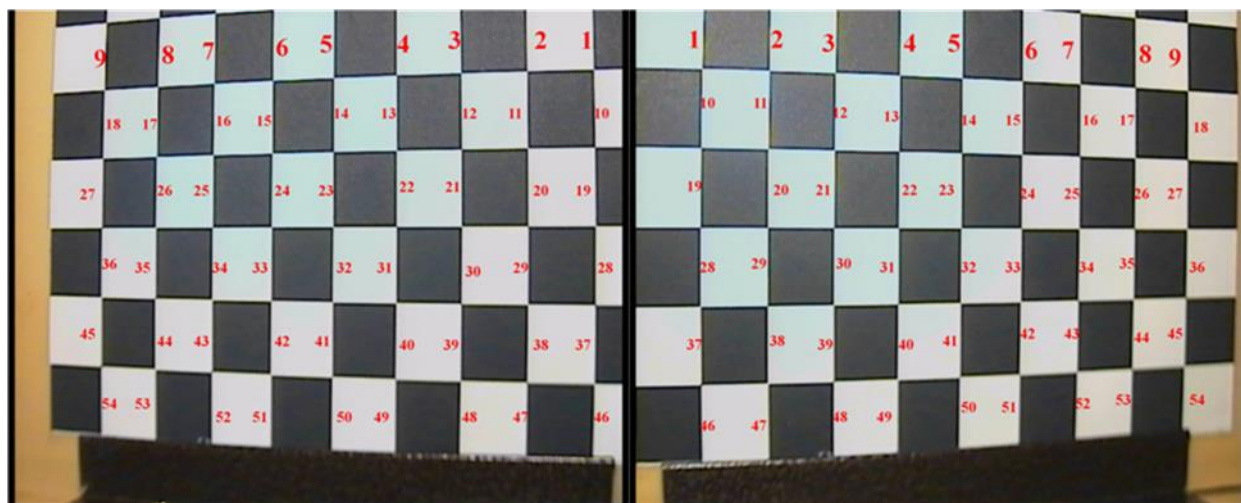


Рис. 4.5. «Левое» (60°) и «Правое» (-60°) положение тестовой таблицы

Таблица 4.2. Усредненные результаты измерения разрешения

Столбец	1	2	3	4	5	6	7	8	9
60°, ТВЛ	623	610	572	527	494	413	336	302	277
-60°, ТВЛ	591	576	546	474	484	408	367	303	272

Из графика усредненных значений разрешения видно, что передний край тестовой таблицы ближе всего находится к оптическому фокусу объектива и от

него идет постепенное ухудшение изображения, а соответственно и уменьшение значений разрешения.

Так же можно сделать вывод о том, что не имеет особого значения в какую сторону “поворачивать” тестовую таблицу относительно центральной оси, поскольку график на рис. 4.6 подтверждает схожесть распределения разрешения по полю изображения.

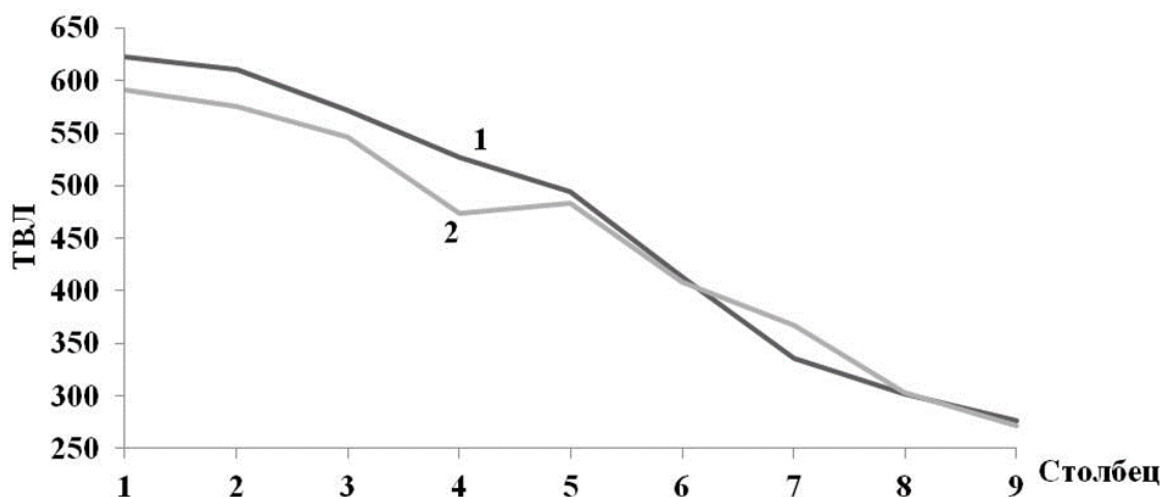


Рис. 4.6. Усредненный график распределения разрешения по полю изображения:
1 – угол 60° ; 2 – угол -60°

Подводя итоги по особенностям измерения разрешения по полю изображения, можно выделить следующие:

- важно учитывать угол съемки при котором получено изображение, поскольку это влияет на расположение фокуса системы формирования изображения;
- уровень разрешения падает от области находящей в фокусе системы формирования изображения и, соответственно, нельзя делать вывод об уровне разрешения на изображении по одному измерению в поле зрения телевизионного датчика либо указывать, что уровень разрешения был получен в конкретной точке на изображении;
- поскольку в каждой точке измерения невозможно получить одинаковые значения разрешения, то в каждой области определения разрешения на которые “разбито” изображение, следует проводить усреднение полученных результатов.

4.2. Методы и алгоритмы повышения качества изображений на основе цифровой фильтрации

4.2.1. Линейные алгоритмы обработки изображений

Существуют различные методы цифровой обработки изображений, такие как:

- линейная обработка изображений;
- нелинейная обработка изображений;
- рекурсивная реализация линейных и нелинейных алгоритмов обработки изображений.

В работе [80], которая посвящена направлениям развития методов, алгоритмов и аппаратных средств повышения качества изображений оптоэлектронных систем, рассматриваются подходы к повышению качества изображений в распознающих системах технического зрения, построенных на новых способах и алгоритмах обработки изображений.

Основными направлениями развития оптоэлектронных систем и распознающих систем технического зрения являются.

1. Углубленное знание объектов исследований и условий работы систем, создание адекватных моделей данных объектов.
2. Модернизация методов расчета и общей теории и проектирования (прежде всего – теории приема и обработки многомерного сигнала, общих схемотехнических решений, унификации методов расчета основных узлов и характеристик прибора в целом) систем с учетом их специфики.
3. Рост и внедрение методов адаптации алгоритмов, адаптации конструктивных параметров и характеристик систем, учитывающих внешние изменяющиеся условия их работы.
4. Использование современной элементной базы, например, приемников изображений на основе КМОП технологий, выполненных по технологии „активный пиксель“, с применением микроболометрических матриц.
5. Расширение методов обработки оптических сигналов, строящихся на основе вейвлет-преобразований, быстром преобразовании Фурье.
6. Создание новых методов, алгоритмов и аппаратных средств обработки изображений на основе системного подхода и комплексного учета основных факторов, оказывающих влияние на качество работы систем.

Одним из самых распространенных методов фильтрации изображений является обработка их линейными алгоритмами. Для понимания процессов в реа-

лизации различных фильтров рассмотрим основные понятия линейной фильтрации.

Одномерный линейный цифровой фильтр (ЛЦФ) – устройство, в котором выходные отсчеты сигнала представлены в виде линейной комбинации предыдущих отсчетов входного и выходного сигналов и текущего отсчета входного сигнала.

Основное разностное уравнение для одномерного ЛЦФ имеет вид:

$$y(nT) = \sum_{i=0}^M a_i \cdot x(nT - iT) + \sum_{j=1}^N b_j \cdot y(nT - jT), \quad n \geq 0, \quad (4.1)$$

где T – период (интервал) дискретизации; n – текущий номер отсчета сигнала, включая нулевой отсчет ($n = 0, 1, \dots$); j – начинается с 1, чтобы уравнение было приведенным.

Первая сумма описывает нерекурсивную часть фильтра, вторая – рекурсивную часть фильтра.

Если a_i и $b_j = \text{const}$, то ЛЦФ с постоянными параметрами, иначе ЛЦФ с переменными параметрами.

Большее из чисел M и N – является порядком фильтра.

Если $b_j = 0$, то получаем нерекурсивные фильтры или фильтры с конечной импульсной характеристикой, так называемые КИХ-фильтры.

Фильтры без обратных связей, т.е. $b_j = 0$ называют также однородными, трансверсальными или фильтрами с многоотводной линией задержки [81]. Если $b_j \neq 0$, то получаем класс рекурсивных фильтров, основная (прямая) форма реализации которого приведена на рис. 4.7.

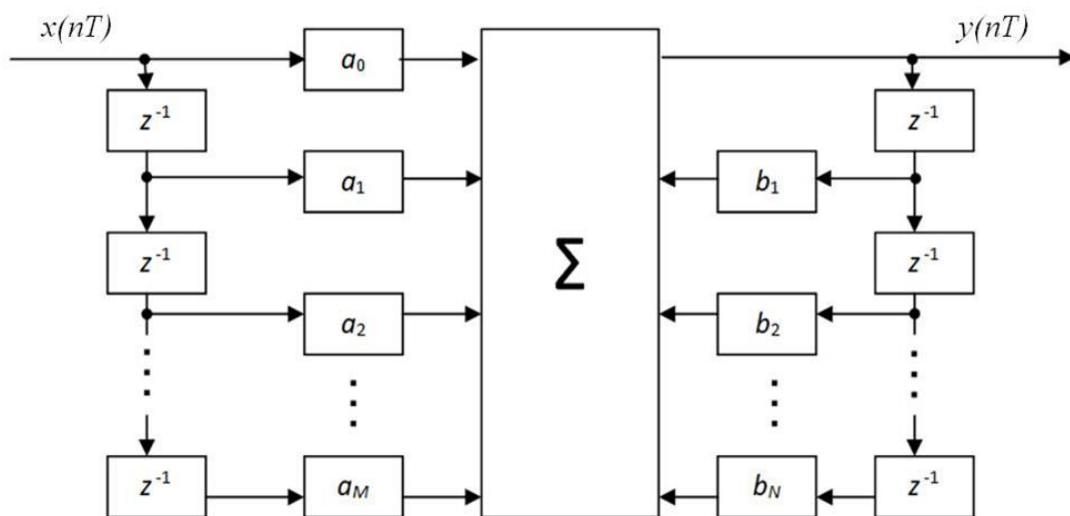


Рис. 4.7. Основная (прямая) форма реализации

Схематически реализация алгоритма двумерной линейной фильтрации приведена на рис. 4.8.

Здесь $(M_1+1)(M_2+1)$ – размер опорной области по входным данным – $x(n_1, n_2)$, а $(N_1+1)(N_2+1)$ – размер опорной области по выходным данным – $y(n_1, n_2)$.

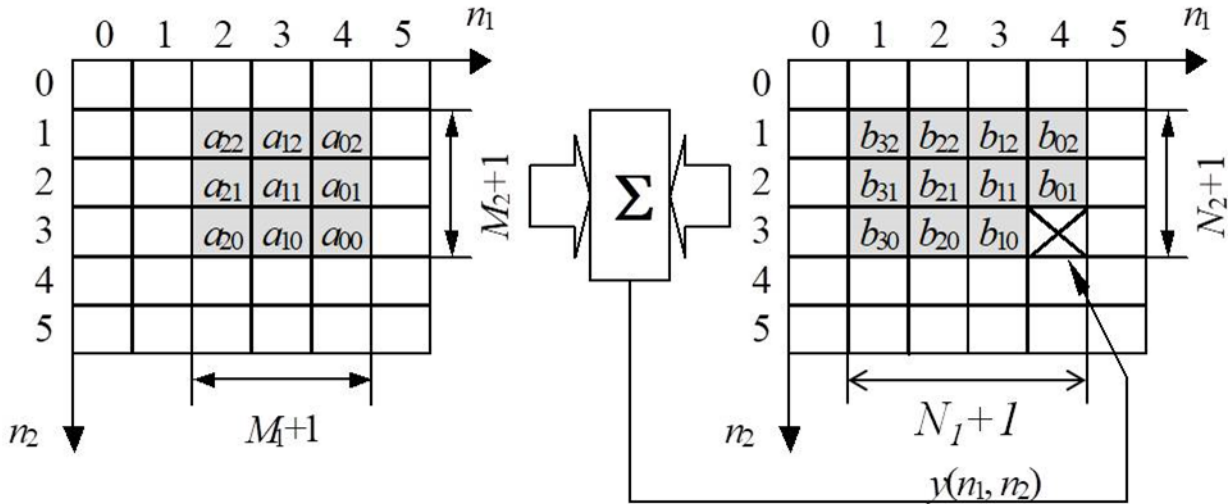


Рис. 4.8. Алгоритм двумерной линейной фильтрации

Разностное уравнение для двумерного ЛЦФ имеет вид:

$$y(n_1, n_2) = \sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} a_{i_1, i_2} x(n_1 - i_1, n_2 - i_2) + \sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} b_{j_1, j_2} y(n_1 - j_1, n_2 - j_2), \quad (4.2)$$

где $(j_1, j_2) \neq 0$.

Пример сглаживающего фильтра:

$$y(n_1, n_2) = x(n_1, n_2) + k y(n_1 - 1, n_2 - 1), \quad k = b_{11}, \quad n_1 \geq 0, \quad n_2 \geq 0, \quad (4.3)$$

Импульсная характеристика двумерного ЛЦФ имеет вид:

$$h(n_1, n_2) = y(n_1, n_2) |_{x(n_1, n_2) = \delta(n_1, n_2)} \text{ – при нулевых начальных условиях.}$$

Системная функция $H(z_1, z_2) = \frac{Y(z_1, z_2)}{X(z_1, z_2)}$ – определяется как отношение z-образов входной и выходной последовательностей; здесь z_1^{-1} – задержка на 1 шаг по строке (координата n_1); z_2^{-1} – задержка на 1 шаг по кадру или столбцу (координата n_2) [82]. При этом исходя из двумерного Z-преобразования

$$F(z_1, z_2) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} f(n_1, n_2) z_1^{-n_1} z_2^{-n_2}$$

можно определить взаимосвязь системной функции двумерных ЛЦФ с его импульсной характеристикой в виде

$$H(z_1, z_2) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} h(n_1, n_2) z_1^{-n_1} z_2^{-n_2} - \text{системная функция в Z-форме.}$$

Если заданы коэффициенты a_{i_1, i_2} и b_{j_1, j_2} в разностном уравнении (РУ), то системная функция (СФ) примет вид

$$H(z_1, z_2) = \frac{\sum_{i_1=0}^{M_1} \sum_{i_2=0}^{M_2} a_{i_1, i_2} z_1^{-i_1} z_2^{-i_2}}{1 - \sum_{j_1=0}^{N_1} \sum_{j_2=0}^{N_2} b_{j_1, j_2} z_1^{-j_1} z_2^{-j_2}}; \text{отсчет } (j_1, j_2) = 0 \text{ исключается.}$$

Рассмотрим типы наиболее часто используемых двумерных ЛЦФ на примере масок размером 3x3, приведенных на рис. 4.9:

- а) скользящее среднее;
- б) лапласиан;
- в) двойное дифференцирование;
- г) оператор выделения вертикальных линий;
- д) оператор выделения малоразмерных деталей из шумов;
- е) градиентный оператор выделения перепада.

a

1	1	1
1	1	1
1	1	1

b

-1	-1	-1
-1	8	-1
-1	-1	-1

c

1	-2	1
-2	4	-2
1	-2	1

d

-1	2	-1
-1	2	-1
-1	2	-1

e

1	2	1
2	4	2
1	2	1

f

1	1	-1
1	-2	-1
1	1	-1

Рис. 4.9. Типы масок

Широкое использование линейной фильтрации для обработки изображений связано с тем, что она в значительной мере позволяет скорректировать целый ряд искажений, появляющихся при формировании, передаче и воспроизведении изображений. К данным искажениям, прежде всего, относятся апертурные искажения, связанные с несовершенством устройств, при помощи которых

формируются изображения, а также ошибками, допущенными при съемке (недостаточностью фокусировки объектива, “смазом” изображения вследствие недостаточно короткой экспозиции при съемке движущегося объекта и т.д.).

Линейная фильтрация используется для ослабления заметности растровой структуры на изображениях, полученных в результате сканирования типографских репродукций. Так же линейная фильтрация используется для уменьшения помех, обусловленных наложением побочных спектров, появляющихся при пространственной дискретизации изображений. Кроме того, использование линейной фильтрации позволяет сократить число ошибок интерполяции, то есть заметность растровой структуры при воспроизведении изображения, например, при их печати. Посредством линейной фильтрации в известных пределах так же удастся сократить заметность флуктуационного шума и других дефектов на воспроизводимых изображениях [82]. Линейные фильтры могут быть предназначены не только для уменьшения уровня шума, а но и для выделения перепадов яркости и контуров. Выделение вертикальных перепадов осуществляется дифференцированием по строкам, горизонтальных – по столбцам.

Выделение перепадов по диагонали можно получить, вычисляя разности уровней диагональных пар элементов. Для выделения перепадов используются следующие наборы весовых множителей, реализующих двумерное дифференцирование (см. рис. 4.10).

„север”	„северо-восток”	„восток”	„юго-восток”
1 1 1	1 1 1	-1 1 1	-1 -1 1
1 -2 1	-1 -2 1	-1 -2 1	-1 -2 1
-1 -1 -1	-1 -1 1	-1 1 1	1 1 1
„юг”	„юго-запад”	„запад”	„северо-запад”
-1 -1 -1	1 -1 -1	1 1 -1	1 1 1
1 -2 1	1 -2 -1	1 -2 -1	1 -2 -1
1 1 1	1 1 1	1 1 -1	1 -1 -1

Рис. 4.10. Градиентные операторы выделения перепадов яркости в различных пространственных направлениях

Оператор Лапласа (лапласиан) – это есть вторая пространственная производная. После применения масок на основе лапласиана фон изображения устраняется и становится «серым». Данное изменение изображения можно разделить на два этапа: обработка изображения сглаживающим фильтром; вычитание из исходного изображения обработанного изображения. Данная операция реализует метод «нерезкого маскирования», при котором величина весового коэффициента подбирается экспериментально, но следует помнить, что выбор слишком большого значения коэффициента приводит к возникновению на изображении «окантовок», а также к росту уровня шума. Апертуру используемого фильтра можно интерпретировать как сумму разностей центрального элемента с каждым из восьми его ближайших соседей. Таким образом, в равной степени учитываются возможные перепады яркости во всех направлениях [83].

На рис. 4.11 показан дифференцирующий оператор, реализованный на основе маски (3x3).

Разреженные маски имеют дифференцирование не такое «равномерное».

Их используют для выделения малоразмерных объектов, находящихся на гладких фонах. В них размер маски согласован с размером объекта, а именно: размер маски выбирается в 2 раза больше, чем размер объекта. Пример незаполненной маски 3x3 элементов показан на рис. 4.12.

-1	-1	-1
-1	8	-1
-1	-1	-1

Рис. 4.11. Лапласиан для «восьми соседей»

0	-1	0
-1	4	-1
0	-1	0

Рис. 4.12. Незаполненная маска 3x3 (лапласиан для «четырёх соседей»)

Разностное уравнение лапласиана для «четырёх соседей» имеет вид:

$$y(n_1, n_2) = -x(n_1 - 1, n_2) - x(n_1, n_2 - 1) + 4x(n_1 - 1, n_2 - 1) - x(n_1 - 2, n_2 - 1) - x(n_1 - 1, n_2 - 2). \quad (4.4)$$

Лапласианы получаются путем суммирования дифференцирующих масок. Сложив первые два блока, реализуем лапласиан для «четырёх соседей», а про-

суммировав все четыре маски, реализуем лапласиан для «восьми соседей» (см. рис. 4.13).

Прибавив к исходному сигналу выходной сигнал после обработки лапласианом, получаем обработку, аналогичную двумерному апертурному корректору, используемому в телевидении для повышения резкости изображения.

Вариант реализации лапласиана для «восьми соседей» (см. рис. 4.14).

Принцип работы фильтра лапласиан и выборка значений коэффициентов для данного фильтра, который является типичным примером фильтрации изображений с подъемом верхних частот, поясняется на рис. 4.15.

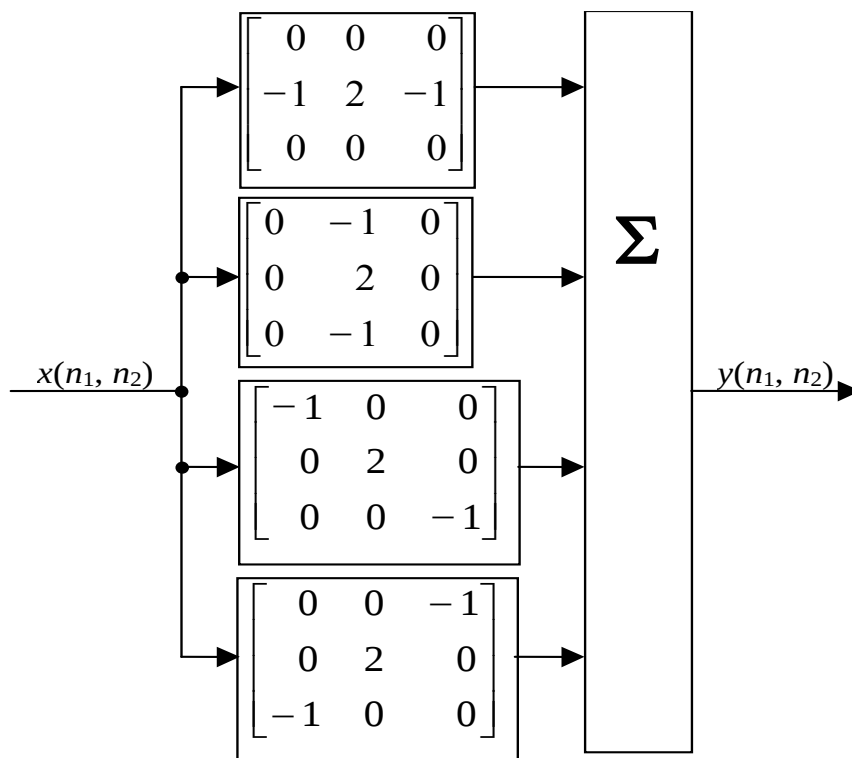


Рис. 4.13. Структурная схема реализации лапласиана для «восьми соседей»

Варьированием коэффициента усиления высоких частот обычно удается добиться общего увеличения контраста изображения и получить желаемый результат. Рис. 4.16 иллюстрирует применение такого метода. Часть рис. 4.16, *а* есть исходное «темное» изображение. На рис. 4.16, *б* показано «изображение-лапласиан», полученное при использовании маски на рис. 4.15, *б* с $A=0$.

Изображение на рис. 4.16, *в* было получено при той же маске, но с $A=1$. Как и ожидалось, изображение стало более резким, но все еще остается столь же «темным», что и исходное.

Наконец на рис. 4.16, z показан результат с $A=1,7$. Этот результат более информативен, поскольку увеличен средний уровень яркости, следовательно, изображение стало более “светлым” и естественным.

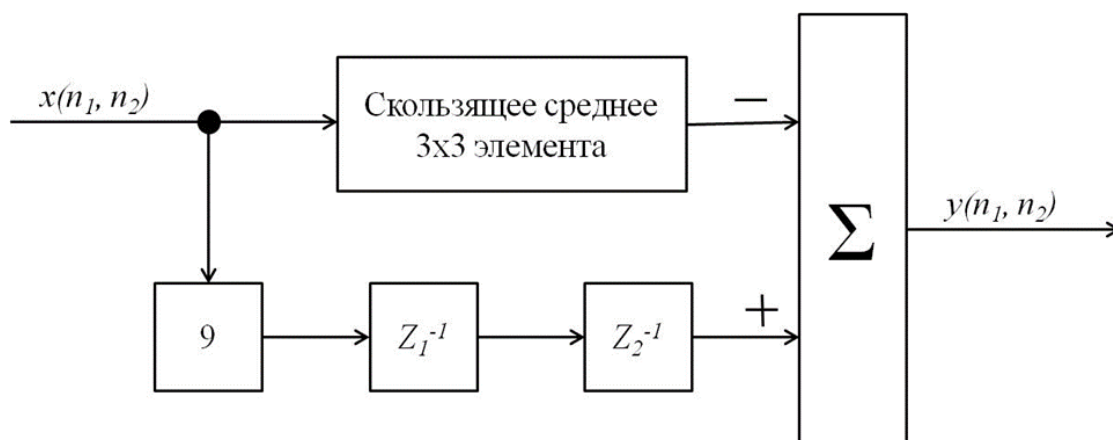


Рис. 4.14. Лапласиан для «восьми соседей»

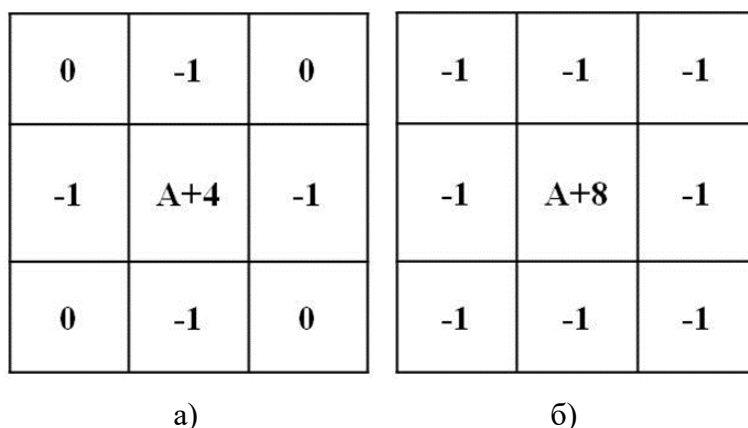


Рис. 4.15. Метод фильтрации изображения с подъемом верхних частот может быть реализован при использовании любой из этих масок с $A>1$

Влияние масок малого размера неізотропно, так как расстояние между центральным элементом маски и вертикальными, и горизонтальными элементами условно равно 1, а между центральным элементом маски и диагональными элементами равно 1,41. Очевидным решением становится увеличение размера маски для обеспечения изотропности. Когда методы цифровой обработки изображений только осваивались, и время, расходуемое на обработку, было критично, рост размера маски вел к росту объема вычислений, поэтому часть обрабатываемых элементов было принято заполнять нулями в разреженных масках. Разреженные маски используются для уменьшения количества вычислительных операций в двумерном фильтре. Размеры маски должны быть, по

крайней мере, в два раза больше размера выделяемого малоразмерного объекта. Однако эффект размытия в разы сильнее, так как в формировании центрального элемента обработанного изображения участвуют более далекие «соседи» [84, 85].

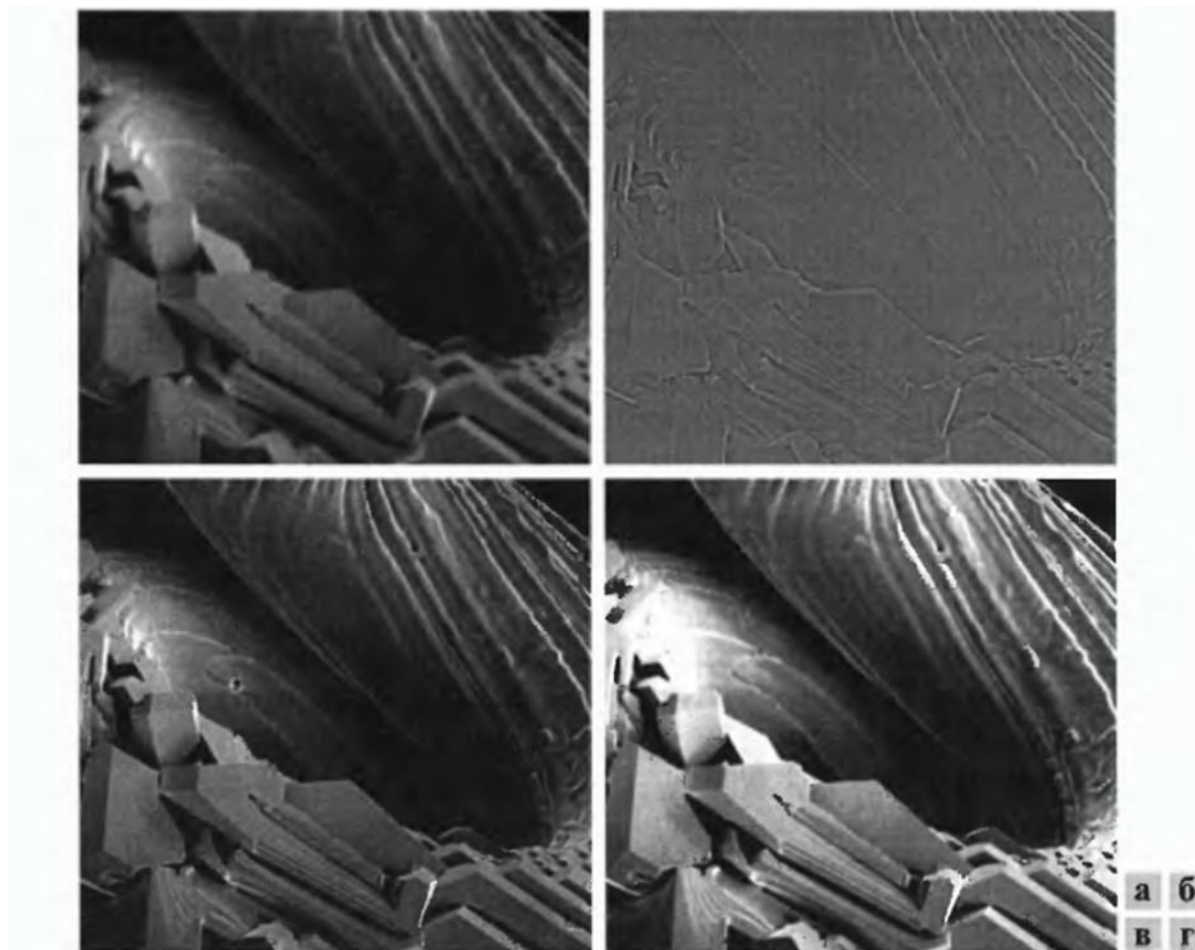


Рис. 4.16. Пример работы фильтра лапласиан: *а* - исходное изображение; *б* – лапласиан изображения, полученный при использовании маски (рис. 4.15, *б*) с $A=0$; *в* – изображение, улучшенное наложением лапласиана с маской (рис. 1.10, *б*) с $A=1$; *г* – то же, что и (*в*) но $A=1,7$

В работе [86] приведены результаты исследования по оптимизации параметров двумерных фильтров повышения качества изображений.

4.2.2. Рекурсивные алгоритмы обработки изображений

Рекурсивные методы фильтрации изображений являются самыми перспективными и широко используемыми в решении задач по уменьшения вычислительных затрат при обработке изображений. Данный критерий, а именно затрачиваемое время на обработку изображения (видеопотока) очень важен для си-

стем, работающих в режиме реального времени. Так же данный тип фильтрации используется в различных модификациях на основе рекурсивных свойств.

Рекурсивные цифровые фильтры (РЦФ) – фильтры в которых количество коэффициентов фильтра может быть уменьшено по сравнению с нерекурсивными цифровыми фильтрами (НЦФ). Рекурсивные фильтры называют фильтрами с бесконечной импульсной характеристикой (БИХ-фильтры), так как "память" по значениям предыдущих отсчетов, которую имеют эти фильтры, в пределе может быть бесконечной. Реакция рекурсивного фильтра на сигнал с учетом "памяти" исключает возможность реализации фильтров с четным импульсным откликом, и частотные характеристики рекурсивных фильтров всегда являются комплексными. Создание рекурсивных частотных фильтров с заданными частотными характеристиками реализуется с использованием z -преобразований [87].

Рекурсивная фильтрация базируется на рекуррентном соотношении между входными и выходными переменными системы. Для одномерных сигналов подобное рекуррентное соотношение имеет следующий вид:

$$r(m) = \sum_{j=1}^J a(j)f(m-j+1) - \sum_{k=2}^K b(k)r(m-k+1), \quad (4.5)$$

где $f(n)$, $n=1,2,\dots,N$ - отсчеты входной последовательности, $r(m)$, $m=1,2,\dots,M$ - отсчеты выходной последовательности, а $b(k)$ и $a(j)$ - весовые множители. Ключевой момент здесь заключается в том, что m -й элемент выходной последовательности зависит не только от последнего и $j-1$ предпоследних элементов входной последовательности, но и от $K-1$ предыдущих элементов выходной последовательности.

Существуют различные методы рекурсивной обработки изображений, например:

- интервальное интегрирование;
- интервальное дифференцирование;
- квазиоптимальная фильтрация малоразмерных объектов из шумов;
- трапецеидальная импульсная характеристика;
- двухкаскадный рекурсивно - сепарабельный цифровой фильтр.

В настоящее время существует достаточно большое количество видов рекурсивных фильтров и их модификаций. Наиболее известный из них - фильтр Баттерворта. Он обладает монотонной, гладкой АЧХ во всем частотном диапазоне. При том же порядке многочленов фильтров (равном количестве полюсов) большую крутизну обеспечивают фильтры Чебышева (прямой и инверсный),

однако при этом в полосе пропускания (для инверсного – в полосе подавления) у фильтров Чебышева появляются равноволновые пульсации (с одинаковой амплитудой пульсаций). Еще более крутые срезы характеристик (при равноволновых пульсациях как в полосах пропускания, так и в полосе подавления) реализуются с использованием эллиптических функций [88].

Рассмотрим несколько вариантов реализации рекурсивных фильтров. Начнем с таких фильтров как БИХ-фильтры и КИХ-фильтры.

Для обработки двумерных сигналов можно использовать фильтры с бесконечной импульсной характеристикой (БИХ-фильтры), так и фильтры с конечной импульсной характеристикой (КИХ-фильтры). БИХ-фильтры обладают значительным преимуществом относительно КИХ-фильтров, так как для того чтобы достичь тех же частотных свойств, им потребуются меньший объем вычислений. Более того, БИХ-фильтры имеют более узкие переходные полосы от области пропускания до области подавления. Но чем ближе к нулю полином в знаменателе частотной характеристики, тем активней меняется частотная характеристика и с этим связана важная проблема устойчивости БИХ-фильтров [89].

Касательно устойчивости данного типа фильтров можно сказать следующее, что для проверки устойчивости БИХ-фильтров по двумерному z -преобразованию и их импульсных характеристик было предложено несколько методов. Одним из таких методов является применение теоремы Шенкса. Хотя применение данной теоремы в практических ситуациях сопряжено с трудностями. Для проверки устойчивости необходимо единичный круг из плоскости z_1 отобразить на плоскость z_2 , решив неявное уравнение $B(z_1, z_2)=0$ относительно переменной z_2 . При этом фильтр будет устойчивым тогда и только тогда, когда отображение плоскости d_1 на плоскость z_2 не пересекается с единичным кругом на плоскости z_2 . Подводя итог можно сказать, что использование данного критерия устойчивости требует решения уравнения относительно плоскости z_2 для большого набора значений, что на практике всегда затруднительно (за исключением простейших случаев) [90].

Об устойчивости фильтра с бесконечной импульсной характеристикой судят по его передаточной функции. Для дискретного фильтра достаточно, чтобы все полюса его передаточной функции по модулю были меньше единицы (т.е. лежали внутри единичного круга на z -плоскости). Все критерии устойчивости, применимые в теории линейных стационарных систем, например, критерий

устойчивости Найквиста или критерий устойчивости Рауса применимы и в случае БИХ-фильтров [91].

КИХ-фильтр – один из видов линейных цифровых фильтров, имеющий характерную особенность как ограничение по времени его импульсной характеристики. Достоинства: не требуют проверки на устойчивость, так как все КИХ-фильтры устойчивы; КИХ-фильтры имеют линейную фазовую характеристику.

Недостаток: имеют больший порядок, чем БИХ-фильтры со схожими характеристиками.

БИХ-фильтр – один из видов линейных фильтров, использующие один или несколько своих выходов в качестве входа, образующих обратную связь.

Достоинство: имеют меньший порядок, чем КИХ-фильтры со схожими характеристиками.

Недостатки: требуется проверка на устойчивость.

В работе [92] предлагается метод построения алгоритма рекурсивного вычисления свертки изображения и двумерного фильтра с неразделимой конечной импульсной характеристикой (КИХ). Данный метод базируется на представлении конечной импульсной характеристики фильтра путем вертикальных и горизонтальных рекуррентных соотношений. Каждое из рекуррентных соотношений сводит к полурекурсивной операции вычисления свертки изображения и двумерного КИХ-фильтра.

Так же исследованиям БИХ и КИХ-фильтров с различной формой их реализации посвящены работы [93] и [94], в том числе прототипы подобных фильтров в среде программирования Matlab [95].

Решение многих задач в цифровой обработке изображений и распознавания образов основано на использовании простой вычислительной процедуры «скользящим окном», которая может быть реализована двумерным цифровым фильтром, получившим широкое распространение в области обработки изображений. При этом большие размеры «окна» потребуют чрезмерно большого числа арифметических операций. Для обработки изображений в режиме «скользящего окна», наиболее удобны алгоритмы параллельно-рекурсивного вычисления свертки, при надлежащей импульсной характеристики КИХ-фильтра [96].

В работе [97] рассматривается быстрое вычисление момента Чебышева и его обратное преобразование на основе z-преобразования. Авторами выводится упрощенное рекуррентное соотношение для вычисления полиномов Чебышева

на основе свойств z -преобразований, что открывает путь для реализации цифрового фильтра второго порядка с ускоренным вычислением полиномов Чебышева.

В работе [98] представлены рекурсивные алгоритмы согласованной фильтрации со сложными бинарными фазоманипулированными сигналами по сравнению с нерекурсивной реализацией, которые дают хороший вычислительный выигрыш. Следует отметить, что несмотря на рекурсивную реализацию, данные фильтры остаются КИХ-фильтрами, и это реализует их потенциальную устойчивость. Число вычислительных операций (за исключением элементов задержки) зависит от структуры сложного сигнала для которого строится согласованный фильтр.

Рекурсивные КИХ-фильтры занимают промежуточное место между нерекурсивными КИХ-фильтрами и рекурсивными БИХ-фильтрами: по способу построения они рекурсивны, то есть имеют обратные связи, а по длительности импульсной характеристики – КИХ-фильтры [99].

В настоящее время цифровая обработка сигналов для подавления случайных помех является необходимым условием для обеспечения высокого качества управления сложными объектами. Для сокращения времени наблюдения за входным сигналом путем сохранения заданного качества подавления помех предлагается использовать совместно рекурсивные и нерекурсивные фильтры. Применение системы из двух фильтров дает возможность уменьшить время наблюдения входного сигнала [100] с сохранением качества подавления помех в выходном сигнале [101].

Большинство алгоритмов обработки изображений и сигналов являются локальными, независимо от конкретной физической природы изображений и сигналов. Из-за сильной корреляции операций, необходимых для определения двух соседних элементов, обработка изображений может быть существенно ускорена с помощью быстрых алгоритмов. В работе [102] приведено рассмотрение быстрых алгоритмов, в том числе линейной и медианной фильтрации с использованием рекурсивных вычислений. Поскольку линейная фильтрация и алгоритмы с линейной функцией часто используются для обработки изображений и сигналов в различных областях. Сделан вывод о том, что общая форма линейной фильтрации может быть определена с использованием кусочно-линейной аппроксимации, которая также обеспечивает независимость числа операций от размера окна, хотя и за счет некоторой потери точности. Допуская некоторую неточность при линейной фильтрации, мы ускоряем обработку, то-

гда как более высокая точность достигается за счет снижения скорости обработки.

Фильтрацию неподвижного изображения несложно представить в виде обработки случайной функции двух переменных (яркости в прямоугольной системе координат) на фоне шума. В таком варианте решения задачи изображение, шум и смесь изображения с шумом представляют собой неподвижные случайные поля. Решение данной задачи затрудняется тем, что по причине большого количества связей каждого элемента изображения с другими элементами, изучаемые поля точно рекуррентными моделями с ограниченным количеством связей не описываются, что препятствует применению хорошо развитой в настоящее время марковской теории фильтрации. Поэтому следует вначале вычислить нижнее граничное значение среднего квадрата ошибки для нереализуемого оптимального нерекурсивного фильтра, а далее реализовывать квазиоптимальные рекурсивные фильтры для рекуррентных моделей. Чтобы осознанно выбрать фильтр с учетом точности фильтрации и сложности реализации, необходимо сравнить характеристики оптимального нерекурсивного и квазиоптимальных рекурсивных фильтров [103].

Операции обработки изображений, такие как размытие, обратная свертка и таблицы зоны суммирования часто эффективно вычисляются как последовательность 1D рекурсивных фильтров. В то же время, много ученых исследовали параллельную рекурсивную фильтрацию, так как предыдущие методы не оптимизируются по всем фильтрам последовательности. В работе [104] описана эффективная алгоритмическая структура, которая сокращает память в пропускной способности по ряду рекурсивных фильтров. Это метод делит вход на блоки, которые обрабатываются параллельно современной архитектурой графического процессора, и перекрывает обработку фильтра причинно-следственных связей, «анти-каузалов», «строк и столбцов». Полная очередность проходов фильтра требует прочтения входного изображения только дважды и записывает данные только один раз. Уменьшенное количество доступов к глобальной памяти приводит к значительному росту производительности цифрового фильтра.

Оценка собственного шума структур рекурсивных цифровых фильтров представлен в работе [105], где предлагается эквивалентная линейная модель с векторным представлением соотношения вход/выход, на основе которой созданы универсальные эквивалентные линейные модели структур рекурсивных цифровых фильтров.

Такой тип изображений как интегральное изображение, является представление промежуточного изображения, нашел обширное использование в многомерных алгоритмах обнаружения локальных функций, таких как ускоренные надежные функции (SURF), позволяющий быстро вычислять прямоугольные элементы с постоянной скоростью, независимо от размера фильтра. Для встроенных систем видения в реальном времени с ограниченными ресурсами вычисления и хранения целостного изображения представляет несколько проблем из-за ограничений по времени и аппаратных ограничений. Хотя расчет интегрального изображения состоит только из простой операции сложения, общее количество операций велико из-за размера данных изображения. Рекурсивные уравнения позволяют существенно уменьшить число операций, но требуют расчета в последовательном порядке. В работе посвященной рекурсивным системам обработки интегральных изображений, предложена модернизация двух аппаратных алгоритмов, основанных на разложении Виолы-Джонса, что позволило уменьшить вычислительные затраты и используемый объем внутренней памяти системы [106].

В серии работ, посвященных алгоритмам шумоподавления в видео-данных представлена модернизация алгоритма, основанного на нелокальных средствах (non-local means, NLM) при помощи временной рекурсии [107]. Этот тип рекурсивной обработки позволяет использовать временную корреляцию с небольшими дополнительными вычислительными затратами [108].

Метод синтеза комплексных рекурсивных цифровых фильтров реализован методом инвариантных переходных характеристик. Комплексные фильтры имеют два квадратурных канала, которые позволяют сравнить изменение частоты настройки фильтра, что делает их весьма производительными при создании адаптивных и когерентных систем обработки информации. В работе [109] приведены конкретные примеры построения комплексных дискретных фильтров.

Применение рекурсивных алгоритмов фильтрации в реальном времени описано в работе [110]. Система способна обнаруживать и отслеживать объекты из видеокadra в реальном времени, используя фильтр Калмана. Фильтр Калмана – это прогнозирующий фильтр, который зависит от пространственной техники и рекурсивных алгоритмов с использованием инструментов обработки изображений в среде Matlab.

В работе [111] обсуждаются некоторые моменты проблемы устойчивости в фильтрах при двумерной рекурсивной фильтрации. В основном предлагаемый

метод, представляет упрощенную версию теоремы устойчивости Шенкса и показывает, что она эквивалентна некоторым результатам Анселла.

Подавление шумов является главной задачей для улучшения качества видеоизображений и цифровых фотографий. Исследование рекурсивной и сепарабельной реализаций фильтров нелокального усреднения рассматриваются в работе [112].

Билатеральная фильтрация осуществляет фильтрацию с пространственным ядром в виде одномерного вектора сначала по строкам, а затем по столбцам. Как и с линейными сепарабельными фильтрами, этот фильтр тоже называют сепарабельным, который реализует похожий способ обработки изображений.

Применение билатерального фильтра рассмотрено при решении задач распознавания тех или иных объектов на снимках, которое как правило, затруднено ввиду наличия разного рода искажений. В данной работе рассмотрен метод предобработки изображений с использованием билатерального фильтра и «шок-фильтра» для последующего распознавания образов методом контурного анализа. Приведена оценка качества распознавания при использовании указанных фильтров. Качество распознавания оценивалось при разных параметрах фильтров. В работе показано, что в случае изображения, обработанного указанным способом, вероятность правильного распознавания объектов существенно повышается [113].

В работе [114] приведен анализ эффективности алгоритмов билатеральной фильтрации со стандартными методами шумоподавления и сделаны следующие выводы:

- билатеральный фильтр сохраняет границы изображения, что упрощает последующую обработку изображения, в отличие от медианного фильтра, который часто используется на практике;
- фильтрация по методу билатеральной решетки является наиболее эффективной для практического использования по соотношению критериев быстродействия и качества;
- прямой метод билатеральной фильтрации целесообразно применять в том случае, если требуется максимально точный результат без временных ограничений;
- билатеральную фильтрацию по методам разделяемого ядра и локальных гистограмм целесообразно применять при небольших ядрах, что позволит произвести быструю обработку, а ошибки в вычислениях не будут заметны для визуального восприятия.

В работе [115] предложен алгоритм, с помощью которого время реализации билатеральной фильтрации существенно уменьшается по сравнению с непосредственным вычислением по стандартной формуле реализации для данного фильтра. Были обработаны изображения различного размера. Продемонстрировано, что время фильтрации изображения составляет в пределах 1 минуты.

В работе [116] рассмотрено влияние билатеральной фильтрации на специфический тип изображений, таких как аэрофотоснимки. В работе был продемонстрирован быстрый билатеральный фильтр на основе декомпозиции на простые пространственные фильтры с использованием распараллеливания по компонентам (PBFIC). Результаты, полученные в ходе эксперимента, продемонстрировали очевидное преимущество метода по быстродействию за счёт дополнительного использования памяти. Несмотря на это, реализована возможность работы билатерального фильтра в реальном времени, что является существенным преимуществом для использования фильтра в качестве предварительной обработки видеок кадров.

В работе [117] приводятся доказательства того, что сепарабельные аппроксимации билатерального фильтра дают возможность добиться не только увеличения быстродействия, но и лучшего отношения сигнал/шум для большинства видеопоследовательностей. На отдельных последовательностях пространственно-адаптивный фильтр демонстрирует существенное преимущество, но в основном, он не позволяет добиться существенного улучшения отношения сигнал/шум по сравнению с обычной сепарабельной фильтрацией.

4.2.3. Быстродействующие алгоритмы обработки изображений

Для понимания сути алгоритмов данного типа введем определение быстродействия. Быстродействие – это способность приборов, устройств, механизмов и т.п. выполнять с необходимой скоростью работу, для которой они предназначены [118]. Так же данное свойство алгоритмов обработки можно сформулировать как возможности реализации канонических алгоритмов в ином представлении за счет каких-либо преобразований, что ведет к уменьшению вычислительных операций, упрощению алгоритма и в итоге приводит к выигрышу по временным затратам, что очень важно в различных областях жизнедеятельности человека, где востребовано уменьшения вычислительных затрат.

В работах [119, 120] посвященных цифровой обработке изображений рассматриваются три основные области обработки изображений: эффективное ко-

дирование, реставрация и улучшение визуального качества изображений, распознавание образов. Многие методы реставрации и улучшения визуального качества изображений основаны на использовании линейных пространственно-инвариантных фильтров. Для решения подобных задач часто требуются и нелинейные методы. Одним из таких методов является медианная фильтрация [121].

В работе [122] рассматриваются вопросы построения системы высокоскоростной обработки изображений с применением цифровых фильтров и концепции System-On-Chip. Применение цифровых фильтров позволяет реализовать высокоскоростную предварительную обработку изображения, регулировку экспозиции, что позитивно влияет на репрезентативность признаков на изображении.

Рассмотрим методы уменьшения количества вычислительных операций при выполнении цифровой фильтрации изображений, основанных на операциях с использованием двумерной свертки. Используемая при фильтрации двумерная свертка имеет особенности – возможность предварительной обработки коэффициентов фильтра и большую разницу между количеством отчетов во входном сигнале и количеством коэффициентов в фильтре. В статье рассмотрены известные методы разложения двумерной свертки в несколько сверток с меньшим числом элементов применительно к задаче цифровой фильтрации. По результатам [123] сделан вывод о том, что основанные на разложении быстрые алгоритмы вычисления двумерной свертки имеет смысл использовать, начиная с размера ядра 4×4 элемента (практически применяются ядра нечетных размеров). Уменьшение числа необходимых операций при применении разложения при ядре 5×5 элементов составляет 16 %, 7×7 элементов – 27 %, 9×9 элементов – 31 %. Ввиду сложности алгоритма разложения на свертки меньших размеров при практической реализации следует ожидать сокращение выигрыша в быстродействии, тем не менее, следует сделать вывод, что рассмотренные в статье алгоритмы имеют практическую ценность для вычисления коротких сверток для ядер, начиная с размера 5×5 элементов [123].

В работе [124] описывается реализованный иерархический алгоритм кодирования многомерных сигналов с помощью неразделимой системы. Разработанный метод является перспективным для использования в таких направлениях, как трехмерное телевидение, медицинская техника, телекоммуникации и т.п. Предложенные алгоритмы реализуются на массовых и недорогих вычислительных платформах – видеокартах персональных компьютеров. Это доказы-

ет, что современный уровень достижений в области цифровой обработки многомерных сигналов – теоретических, алгоритмических, программно-аппаратных – позволяет реализовать принципиально новые цифровые телекоммуникационные и мультимедийные системы, в том числе – системы трехмерного телевидения.

В работе [125] эффективный гауссов фильтр с постоянным временем обработки, который обеспечивает высокую точность, стабильность и минимизирует вычислительные затраты. Экспериментальные результаты по фильтрации изображений продемонстрировали превосходство использования свойство сдвига второго порядка DST-5. Предложенное алгоритмическое решение для анализа многомасштабных изображений позволяет преодолеть проблему увеличения «вычислительной» гауссовой свертки.

В исследовании, описанном в работе [126] авторы показывают, что параллельные и локально-ориентированные реализации IIR фильтрующие конвейеры могут быть получены с помощью программных преобразований, которые механизмируются с помощью доменного компилятора. В работе представлены DSL и компилятор для автоматической оптимизации конвейеров рекурсивных фильтров. Основным вкладом является автоматическое разбиение на фрагменты общих конвейеров, достигаемое с помощью набора простых преобразований кода.

В работе [127] представлен новый рекурсивный алгоритм отслеживания (GReTA) в трех измерениях, способный восстанавливать непрерывные траектории для большого числа объектов и больших временных интервалов, даже с частыми оптическими «окклюзиями». Данный рекурсивный алгоритм основывается на идее глобальной оптимизации решения как в пространстве, так и во времени. Применимость глобальной оптимизации ограничена вычислительной сложностью, которая увеличивается экспоненциально быстро с продолжительной последовательностью. Здесь реализуется значительное уменьшение сложности вычислений, используя рекурсивную стратегию «разделяй и властвуй», позволяющую сначала оптимизировать совпадения «глобально» (по более коротким временным интервалам), а затем выполнить итерацию, чтобы охватить всю временную последовательность. Таким образом, вычислительная сложность резко уменьшается при сохранении глобальной области видимости, что позволяет контролировать большие наборы данных (с точки зрения количества объектов и продолжительности захвата видеоданных). Авторы также предло-

жили несколько модификаций, делающих алгоритм устойчивым к ошибочным или отсутствующим данным.

Таким образом существуют различные классы и типы цифровых фильтров повышения четкости изображений. Одним из них является класс рекурсивно-сепарабельных фильтров повышения четкости изображений, поскольку за счет принципов построения фильтра и используемых в нем свойств “рекурсивности” и “сепарабельности” достигается значительное сокращение вычислительных операций, требуемых на процесс фильтрации изображения.

4.3. Рекурсивно-сепарабельные фильтры повышения четкости изображений

Рассматривая данный тип фильтрации, следует отметить, что он является одним из самых перспективным и широко используемых в задачах уменьшения вычислительных затрат, при обработке изображений. Данный критерий, а именно затрачиваемое время на обработку изображения (видеопотока) очень важен для систем, работающих в режиме реального времени. Так же данный тип фильтрации используется в различных модификациях на основе рекурсивно-сепарабельных свойств.

С ростом размерности формируемых изображений в ВС, связан и рост вычислительных затрат, которые используются на процесс обработки изображений. Поэтому важно сохранить эффективность обработки изображений, но при этом сократить количество вычислительных операций, требуемых для процесса обработки, для чего и предназначены разрабатываемые рекурсивно-сепарабельные алгоритмы их обработки. При этом важно учитывать и значения коэффициентов, используемых для процесса обработки изображения, поскольку использование неоптимальных коэффициентов может привести к негативным эффектам, возникающим на обработанном изображении.

Шум изменяется независимо от пикселя к пикселю и, при условии, что математическое ожидание значения шума равно нулю, шумы соседних пикселей при суммировании будут складываться векторно (алгебраически). Чем больше окно фильтрации, тем меньше будет усредненная интенсивность шума, однако при этом будет происходить и соответствующее размытие значащих деталей изображения. Выходным изображением “белой точки” на “черном фоне” при фильтрации (реакция на двумерный единичный импульс) будет “серый квадрат” [128]. Реализация рекурсивно-сепарабельных алгоритмов продемонстрирована на рис. 4.17, на примере фильтра “скользящего среднего”.

Маска интегрирующего типа имеет вид, приведенный на рис. 4.18, которая может быть реализована сепарабельно. Физический смысл – выделение малоразмерных объектов на фоне «белого» шума.

Вид маски согласован с видом сигнала, то есть маска есть образ сигнала от изображения “малоразмерного объекта”.

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline \end{array} ** \begin{array}{|c|} \hline 1 \\ \hline 1 \\ \hline 1 \\ \hline \end{array}$$

Рис. 4.17. Процесс сепарабельной реализации фильтра “скользящего среднего”

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline \end{array} ** \begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline 1 \\ \hline \end{array}$$

Рис. 4.18. Процесс сепарабельной реализации фильтра “маска выделения малоразмерных деталей из шумов”

Для построения рекурсивно-сепарабельных алгоритмов требуются первообразующие рекурсивные ячейки (рециркуляторы). На рис. 4.19 представлены первообразующие рекурсивные ячейки (рециркуляторы), рис. 4.19, а по n_1 -строке (строчный рециркулятор – СР) и рис. 4.19, б по n_2 -кадру (кадровый рециркулятор – КР), реализующие соответствующие ортогональные направления обработки “скользящее среднее” [129-131].

Разностное уравнение для строчного рециркулятора имеет вид:

$$y(n_1, n_2) = x(n_1, n_2) - x(n_1 - M_1, n_2) + y(n_1 - 1, n_2), \quad (4.6)$$

а для кадрового рециркулятора:

$$y(n_1, n_2) = x(n_1, n_2) - x(n_1, n_2 - M_2) + y(n_1, n_2 - 1). \quad (4.7)$$

Для реализации двумерного цифрового фильтра лапласиан “усеченная пирамида” (ЛУП) получим системную функцию $H_1(z_1, z_2)$, для случая его рекурсивно-сепарабельной реализации в апертуре 7×7 элементов.

$$H_1(z_1, z_2) = -\frac{(1-z_1^{-5})(1-z_1^{-3})(1-z_2^{-5})(1-z_2^{-3})}{(1-z_1^{-1})(1-z_1^{-1})(1-z_2^{-1})(1-z_2^{-1})} + 25z_1^{-2}z_2^{-2}\frac{(1-z_1^{-3})(1-z_2^{-3})}{(1-z_1^{-1})(1-z_2^{-1})}. \quad (4.8)$$

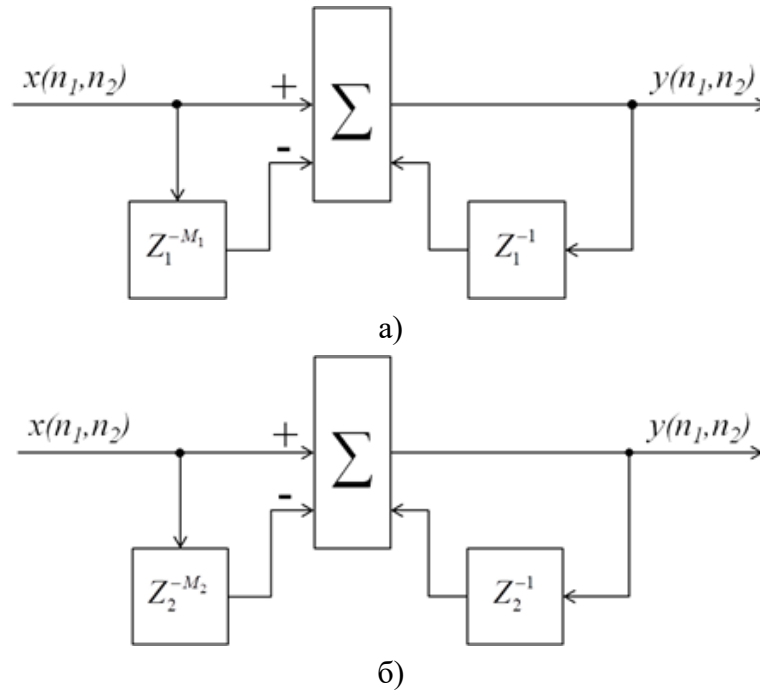


Рис. 4.19. Рециркуляторы: а – по n_1 -строке, б – n_2 -кадру

Структурная схема реализации рекурсивно-сепарабельного фильтра лапласиан “усеченная пирамида” для данного случая приведена на рис. 4.20.

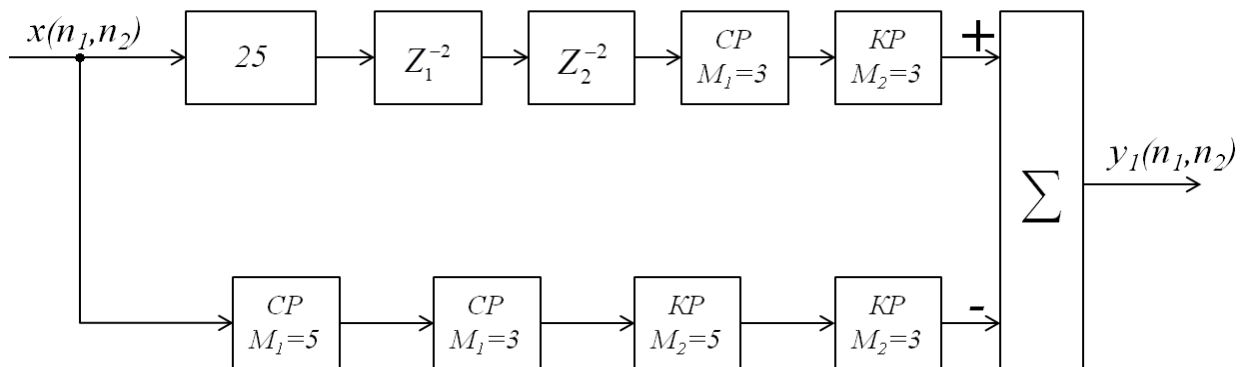


Рис. 4.20. Структурная схема двумерного фильтра ЛУП

Маска для фильтра лапласиан “усеченная пирамида” представлена на рис. 4.21. Ее трехмерный вид приведен на рис. 4.22.

-1	-2	-3	-3	-3	-2	-1
-2	-4	-6	-6	-6	-4	-2
-3	-6	16	16	16	-6	-3
-3	-6	16	16	16	-6	-3
-3	-6	16	16	16	-6	-3
-2	-4	-6	-6	-6	-4	-2
-1	-2	-3	-3	-3	-2	-1

Рис. 4.21. Маска ЛУП 7х7 элементов

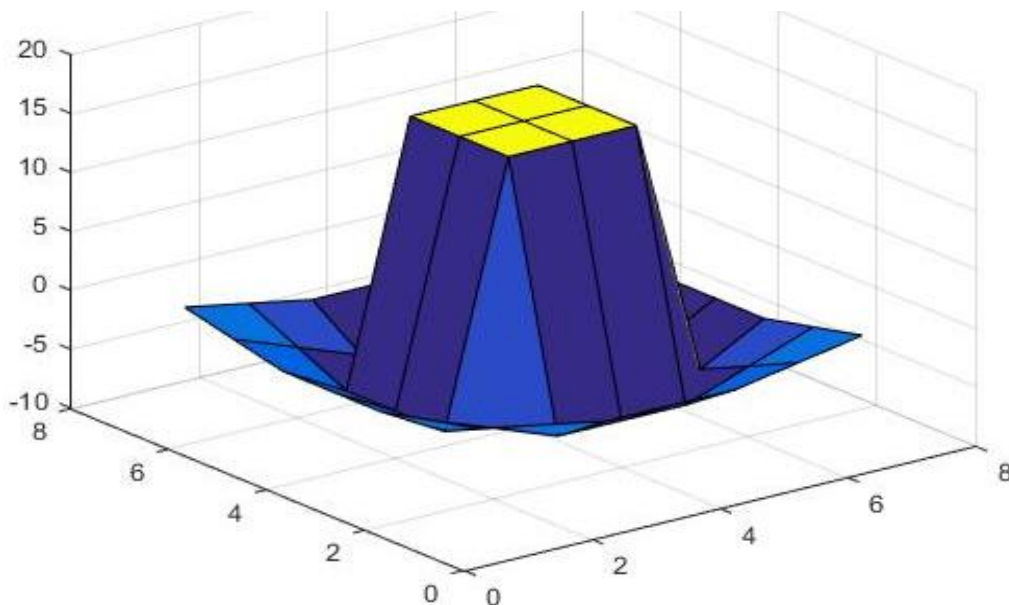


Рис. 4.22. Трёхмерный вид апертury фильтра ЛУП

Для апертury 7х7 элементов в общем случае требуется 49 операций умножения и 48 операций сложения/вычитания при нерекурсивном варианте построения фильтра, а при рекурсивно-сепарабельном варианте построения фильтра ЛУП требуется 13 операций сложения/вычитания и одна операция умножения, что сокращает количество вычислительных операций почти в 6,93 раза. При этом следует отметить, что при увеличении размера апертury обработки выигрыш по вычислительным операциям в рекурсивно-сепарабельных фильтрах по сравнению с нерекурсивным вариантом построения будет возрастать.

При реализации двумерного цифрового фильтра лапласиан “двойная пирамида” (ЛДП) в апертуре 7х7 элементов, её центральная часть “заменяется” на пирамидальную апертuru 3х3 элемента с соответствующими значениями коэф-

фициентов a_{i_1, i_2} , которая реализуется рекурсивно-сепарабельно. При этом системная функция данного фильтра будет иметь вид:

$$H_2(z_1, z_2) = -\frac{(1-z_1^{-5})(1-z_1^{-3})(1-z_2^{-5})(1-z_2^{-3})}{(1-z_1^{-1})(1-z_1^{-1})(1-z_2^{-1})(1-z_2^{-1})} +$$

$$+14z_1^{-2}z_2^{-2}\frac{(1-z_1^{-2})(1-z_1^{-2})(1-z_2^{-2})(1-z_2^{-2})}{(1-z_1^{-1})(1-z_1^{-1})(1-z_2^{-1})(1-z_2^{-1})} + z_1^{-3}z_2^{-3}. \quad (4.9)$$

Приведем структурную схему для рекурсивно-сепарабельного двумерного фильтра лапласиан “двойная пирамида”, представленную на рис. 4.23.

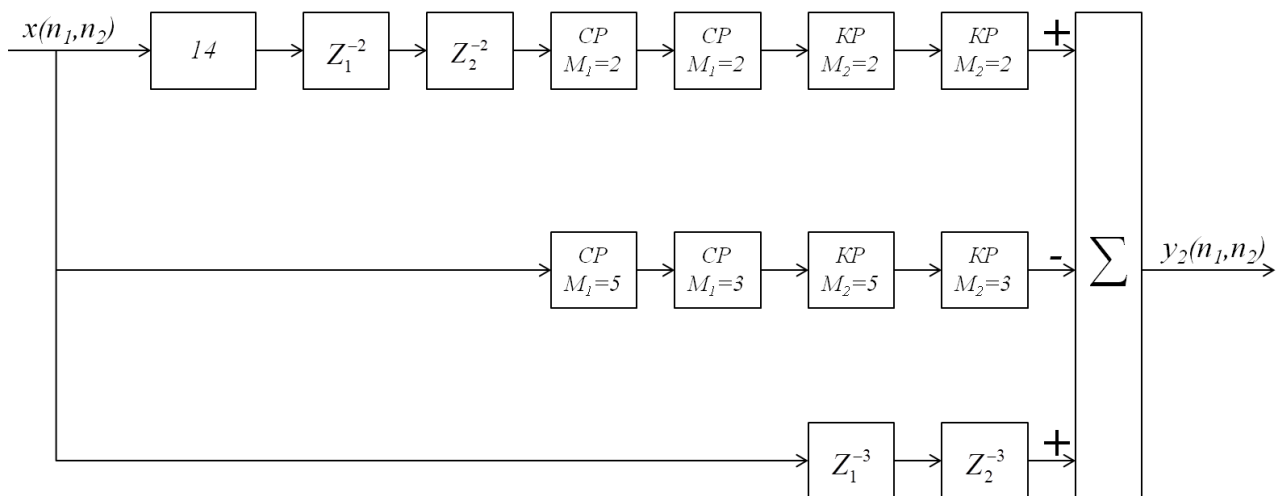


Рис. 4.23. Структурная схема двумерного фильтра ЛДП

На рис. 4.24 приведена маска фильтра лапласиан “двойная пирамида”. Ее трехмерный вид приведен на рис. 4.25.

-1	-2	-3	-3	-3	-2	-1
-2	-4	-6	-6	-6	-4	-2
-3	-6	5	19	5	-6	-3
-3	-6	19	48	19	-6	-3
-3	-6	5	19	5	-6	-3
-2	-4	-6	-6	-6	-4	-2
-1	-2	-3	-3	-3	-2	-1

Рис. 4.24. Маска ЛДП 7x7 элементов

При рекурсивно-сепарабельной реализации фильтра ЛДП требуется всего 18 операций сложения/вычитания и одна операция умножения, тогда как для апертуры 7x7 элементов в общем случае требуется 49 операций умножения и 48 операций сложения/вычитания при не рекурсивном варианте построения фильтра. При данном варианте построения двумерного фильтра количество вычислительных операций сократится в 5,11 раз.

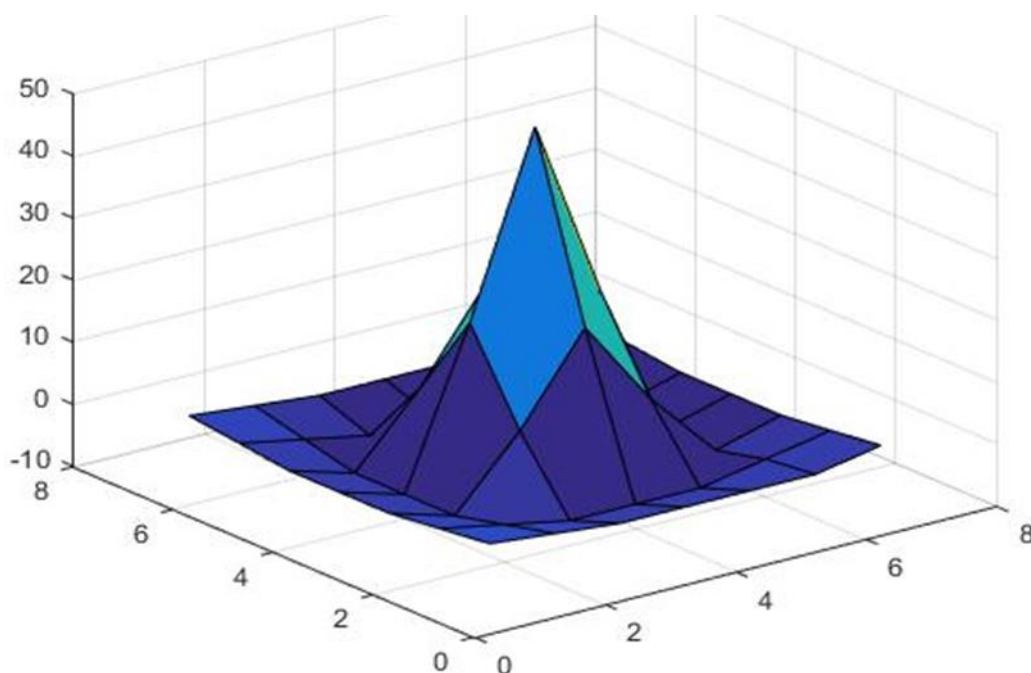


Рис. 4.25. Трехмерный вид апертуры фильтра ЛДП

Реализация рекурсивно-сепарабельных фильтров в ПО Matlab

Для моделирования и оценки эффективности работы рекурсивно-сепарабельных двумерных фильтров повышения четкости изображений рационально разрабатывать программные модули, например, в ПО Matlab [132, 133].

Для работы данного алгоритма требуется задать ряд входных параметров: X – обрабатываемое изображение; A_1 – коэффициент для подъема центральной апертуры размером 3x3 элемента; A_2 – коэффициент для увеличения центрального элемента матрицы на заданное значение.

На первом шаге происходит определение суммы коэффициентов матрицы, для определения коэффициента, на который будет поделено итоговое изображение для нормализации яркости. Далее идет процесс обработки по первой ветви структурной схемы, представленной на рис. 4.20. Исходное изображение умножается на 25, а в случае подъема центральной апертуры на коэффициент A_1 к множителю 25 добавляется необходимое значение. После чего изображе-

ние обрабатывается строчным рециркулятором с размером апертуры $M_1=3$ и кадровым рециркулятором с размером апертуры $M_2=3$.

В конце процедуры фильтрации изображение дополняется “нулями”, поскольку для корректной работы требуется суммирование матриц одной размерности. Далее процедура обработки происходит по второй ветви структурной схемы (рис. 4.20). По ней изображение обрабатывается двумя строчными рециркуляторами с размерами апертур $M_1=5$ и $M_1=3$, после чего происходит процесс обработки двумя кадровыми рециркуляторами с размерами апертур $M_2=5$ и $M_2=3$. После выполнения данной обработки, происходит процедура умножения исходного изображения на коэффициент A_2 , если данный коэффициент задан. После чего происходит процедура суммирования промежуточных изображений и формирование итогового изображения после обработки.

Так же, как и при работе с любыми цифровыми фильтрами повышения четкости изображений следует учитывать, что при своей исходной форме, когда сумма коэффициентов маски фильтра равна нулю, происходит оконтуривание обрабатываемого изображения. Поэтому при фильтрации необходимо добавлять определенный коэффициент, как в данном случае, это коэффициенты A_1 и A_2 , при использовании которых и происходит процесс повышения четкости обрабатываемого изображения.

4.4. Повышение качества изображений при их позонной обработке

Позонная обработка изображений подразумевает под собой способ повышения качества изображения с учетом распределения разрешения по полю изображения. Можно выделить три типа позонной обработки изображений.

1. Предельная четкость изображения. Осуществляется обработка в каждой выделенной зоне до предельного уровня разрешения системы формирования изображения.

2. Выравнивание четкости по полю изображения. Определяется максимальный уровень разрешения в одной из зон на изображение и после чего путем обработки остальных зон на изображении в них достигается максимально близкий уровень разрешения к максимальному разрешению, определенному ранее.

3. Обработка определенной зоны или нескольких зон изображения. Производится обработка только тех областей, в которых находится или должен находиться объект интереса.

Для понимания вышеизложенного рассмотрим первые два случая более подробно. Проведен эксперимент, в котором произведена обработка тестового изображения (см. рис. 4.26) при помощи фильтра ЛУП и ЛДП с подъемом центрального значения апертуры.

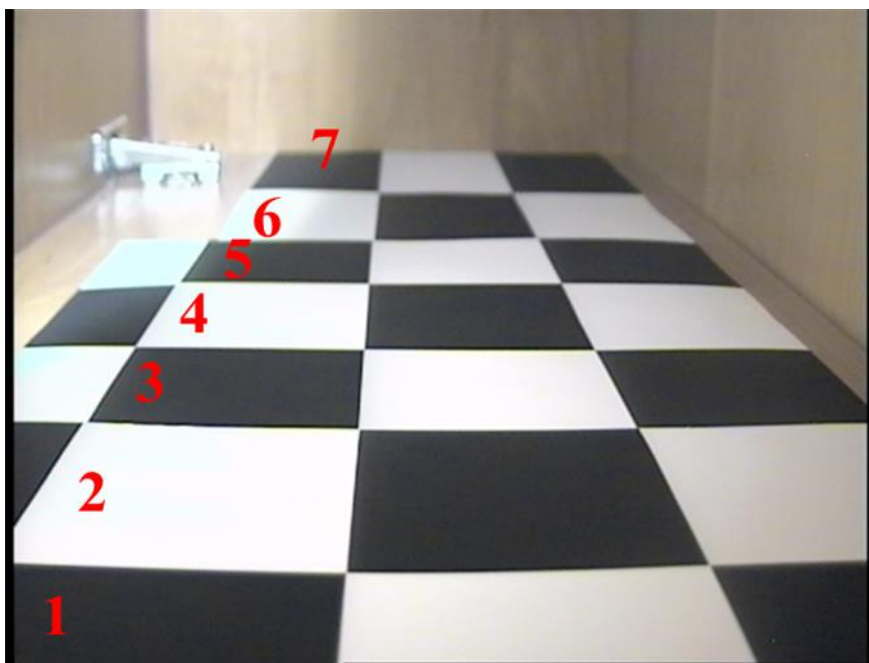


Рис. 4.26. Тестовое изображение

В табл. 4.3 представлены результаты измерения разрешения, по полю тестового изображения, представленного на рис. 4.26. Измерение производилось в зонах, пронумерованных на изображении.

Таблица 4.3. Усредненные результаты измерения разрешения

Зона обработки	1	2	3	4	5	6	7
ТВЛ	160	202	256	345	387	376	200

Произведя обработку зон изображения в отдельности, были определены оптимальные коэффициенты фильтрации для двух вариантов использования позонной внутрикадровой обработки: принципу близости полученного разрешения к предельному разрешению, выдаваемому системой формирования изображения либо выравнивание поля изображения к максимальному значению разрешения. Результат определения оптимальных коэффициентов и значения разрешения при их использовании представлены в табл. 4.4.

Выходное изображение после применения метода позонной внутрикадровой обработки представлено на рис. 4.27.

Таблица 4.4. Усредненные результаты измерения разрешения

Зона обработки	1	2	3	4	5	6	7
Увеличение до предела при помощи фильтра ЛДП							
ТВЛ	347	471	888	1270	1234	1132	524
Коэффициент «А»	12	18	14	18	23	25	15
Выравнивание при помощи фильтра ЛУП							
ТВЛ	347	387	381	345	387	376	387
Коэффициент «А»	12	25	50	-	-	-	23



Рис. 4.27. Тестовое изображение после позонной обработки до предельного разрешения

На рис. 4.28 представлено выравнивание поля разрешения изображения. Для полученного изображения на рис. 4.29 представлен сводные графики контрастно-частотной характеристики (КЧХ) для каждой зоны обработки.



Рис. 4.28. Тестовое изображение после позонной обработки с выравниванием разрешения

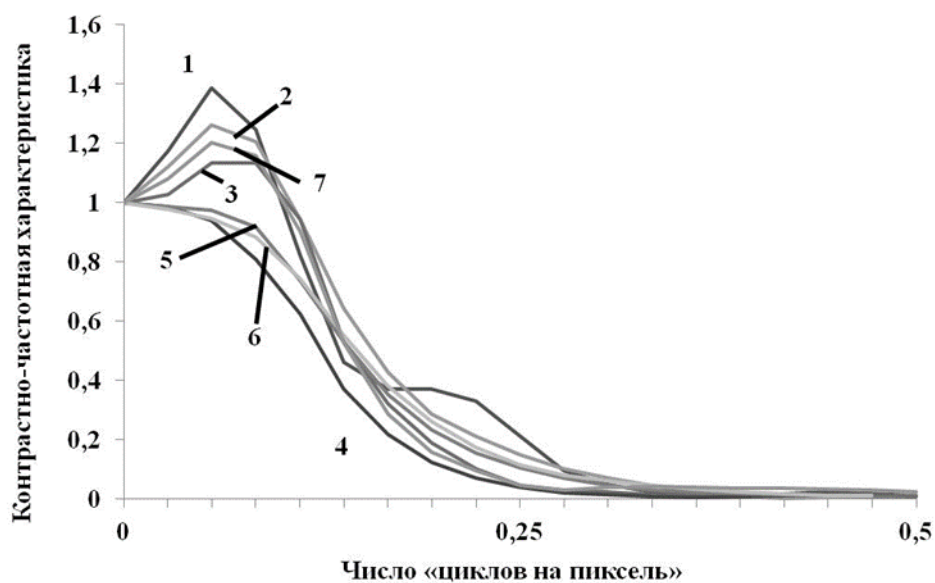


Рис. 4.29. Графики КЧХ изображений после позонной внутрикадровой обработки: 1 – зона обработки “один”; 2 – зона обработки “два”; 3 – зона обработки “три”; 4 – зона обработки “четыре”; 5 – зона обработки “пять”; 6 – зона обработки “шесть”; 7 – зона обработки “семь”.

Делая вывод по применению позонной обработки изображений можно сказать, что ее использование на практике позволяет проводить более качественную обработку изображений, поскольку при обработке всего изображения фильтром с одним коэффициентом часть областей, которая была высокого качества, может быть ухудшена.

Заключение

Анализ современных систем видеокомпрессии показывает, что наибольшее применение имеют алгоритмы сжатия без адаптации к визуальному восприятию разных структур изображения. При этом для настройки параметров и характеристик кодеков используются традиционные метрики оценки качества изображения, что не позволяет объективно оценить ухудшение четкости в зависимости от коэффициента сжатия.

Фундаментальной научной задачей при решении указанной проблемы является поиск интеллектуальных методов и разработка алгоритмов управления качеством передачи больших цифровых видеоданных, позволяющих приблизить качество воспроизводимых изображений к зрительному восприятию оригинальных изображений. При этом необходимо обеспечивать достаточную степень компрессии для передачи больших видеоданных по цифровому каналу связи или для их хранения. Решение поставленной задачи позволит автоматически регулировать соотношение качество – компрессия передаваемых видеоданных в соответствии с выбранными объективными критериями визуальных оценок искажений.

К отдельным задачам исследований относится разработка алгоритмов идентификации и классификации участков изображения по признакам их детальности и по степени заметности искажений после декомпрессии изображения. Решение этой задачи позволит выполнять классификацию объектов изображения и/или отдельных кадров видеоданных с целью оптимизации параметров кодека на требуемое качество. Развитием исследований в области цифровой фильтрации изображений является создание адаптивных двумерных рекурсивно-сепарабельных фильтров и их использование в телевизионных измерительных системах реального времени с повышением четкости изображений.

ЛИТЕРАТУРА

1. Сай С.В. Качество передачи и воспроизведения мелких деталей цветных телевизионных изображений. – Владивосток: Дальнаука, 2003. – 160 с.
2. Грязин Г.Н. Системы прикладного телевидения. – СПб.: Политехника, 2000. – 277 с.
3. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика). – М.: Техносфера, 2012. – 1008 с.
4. Татарников О. Современные цифровые видеокамеры [Электронный ресурс]. – URL: <https://compress.ru/article.aspx?id=19836>
5. Фрейзер Б. Реальный мир Camera Raw и Adobe Photoshop CS2 = World Camera Raw with Adobe Photoshop CS2. – М.: «Вильямс», 2007. – 304 с.
6. Новаковский С.В. Цвет в цветном телевидении. – М.: Радио и связь, 1988. – 288 с.
7. Fairchild M.D. Color Appearance Models. – John Wiley & Sons. 2013. – 450 p.
8. Judd D., Wyszecki G. Color in business, science and industry. – John Wiley & Sons, 1975. – 553 p.
9. Красильников Н.Н. Теория передачи и восприятия изображений и ее приложения. – М.: Радио и связь, 1986. – 248 с.
10. Дворкович В.П., Дворкович А.В. Измерения в видеоинформационных системах (теория и практика). – М.: Техносфера, 2015. – 762 с.
11. Старовойтов В.В., Голуб Ю.И. Цифровые изображения: от получения до обработки. – Минск: ОИПИ НАН Беларуси, 2014. – 202 с.
12. Гонсалес Р., Вудс Р. Цифровая обработка изображений: пер. с англ. – М.: Техносфера, 2005. – 1072 с.
13. Прэтт У. Цифровая обработка изображений: пер. с англ. – М.: Мир, 1982. Кн. 2. – 480 с.
14. Pennebaker W.B., Mitchell J.L. JPEG Still Image Data Compression Standard. – Springer Science & Business, 1992. – 638 p.
15. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М.: ДИАЛОГ-МИФИ, 2002. – 384 с.
16. Wallace G.K. The JPEG still picture compression standard // IEEE transactions on consumer electronics. – Vol. 38, – No. 1, Feb. 1992. – P. 1-67.
17. Стандарт JPEG2000: базовые алгоритмы, примеры реализации и перспективы применения / С.Н. Дроздов, А.А. Жиглатый, П.П. Кравченко и др. – Ростов-на-Дону: Изд-во ЮФУ, 2014. – 255 с.
18. Taubman D., Marcellin M.V. JPEG2000 Image Compression Fundamentals, Standard and Practice. – Kluwer Academic Publishers, 2002. – 779 p.
19. Воробьев В.И., Грибунин В.Г. Теория и практика вейвлет-преобразования. – СПб.: ВУС, 1999. – 208 с.
20. Balster E.J., Fortener B.T., Turri W.F. Post-compression rate-distortion development for embedded block coding with optimal truncation in JPEG2000 imagery // International Journal of Image and Graphics. – 2011. – Vol. 11 (4) – P. 611–627.
21. Цифровое преобразование изображений / Р.Е. Быков, Р. Фрайер, К.В. Иванов К.В., А.А. Манцветов. – М.: Горячая линия-Телеком, 2003. – 228 с.

22. Ричардсон Ян. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. – М.: Техносфера, 2005. – 368 с.
23. Overview of the High Efficiency Video Coding (HEVC) Standard / Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, Thomas Wiegand // IEEE Transactions on Circuits and Systems for Video Technology. – 2012. – Vol. 22 (12) – P. 1649-1668.
24. Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC) / Jens-Rainer Ohm, Gary J. Sullivan, Heiko Schwarz and etc. // IEEE Transactions on Circuits and Systems for Video Technology, – 2012. – Vol. 22 (12) – P. 1669-1684.
25. Испытательные таблицы для измерения качества цифрового и аналогового телевизионного вещания / А.В. Дворкович, В.П. Дворкович, Д.Г. Макаров и др. // «625» – 1999. – № 8. – С. 36-42.
26. Оценка качества цифровых изображений и видеоданных: учебно-методическое пособие / Д.В. Матвеев, А.Г. Седов, В.В. Хрящев и др. – Ярославль: ЯрГУ, 2018. – 76 с.
27. Lin W., Jay Kuo C-C. Perceptual Visual Quality Metrics: A survey // Visual Communication and Image Representation. – 2011. – Vol. 22(4) – P. 297-312.
28. Bovik A., Mittal A. No-Reference Image Quality Assessment in the Spatial Domain // IEEE Transactions on Image Processing. – 2012. – Vol. 21(12). – P. 4695-4708.
29. Wang Z., Li Q. Video quality assessment using a statistical model of human visual speed perception // J. Opt. Soc. Am. – 2007. – Vol. 24 (12). – P. 61-69.
30. Лаборатория изображений и видео инженерии [Электронный ресурс]. – URL: <http://live.ece.utexas.edu/research/Quality/evaluation>
31. О количественной оценке эффективности алгоритмов анализа изображений / П.П. Кольцов, А.С. Осипов, А.С. Куцаев и др. // Компьютерная оптика. – 2015. – Т. 39, № 4. – С. 542-556.
32. Казанцев Г.Д., Курячий М.И., Пустынский И.Н. Измерительное телевидение. – М.: Высш. шк., 1994. – 288 с.
33. Фисенко В.Т., Фисенко Т.Ю. Компьютерная обработка и распознавание изображений: учеб. пособие. – СПб: СПбГУ ИТМО, 2008. – 192 с.
34. РЕКОМЕНДАЦИЯ МСЭ-R BT.500-13 Методика субъективной оценки качества телевизионных изображений. Электронная публикация. – Женева, 2012 г.
35. Латыпова С.М. Испытательные таблицы и сигналы для систем вещательного телевидения // Т-Comm: Телекоммуникации и транспорт. – 2016. – Том 10. – №6. – С. 13-16.
36. Безруков В.Н., Севальнев Л.А. Измерение характеристик телевизионных систем при цифровом многопрограммном вещании // Broadcasting: телевидение и радио. – 2004. – №2 – С.52-55.
37. Кривошеев М.И. Основы телевизионных измерений. - М.: Радио и связь, 1989. – 608 с.
38. Прэтт У. Цифровая обработка изображений: В 2 кн. – М.: Мир, 1982. Кн. 1. – 312 с.
39. Wang Z., Bovik A.C. A universal image quality index // IEEE Signal Processing Letters. – 2002. – Vol. 9(3) – P. 81–84.

40. Image Quality Assessment: From Error Visibility to Structural Similarity / Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli // IEEE Transactions on Image Processing. – 2004. – Vol. 13(4). – P. 600-612.
41. Chen G.H., Yang C.L., Xie S.L. Gradient-based structural similarity for image quality assessment // IEEE International Conference on Image Processing. – 2006. – P. 2929-2932.
42. Image quality assessment based on local variance / S. Aja-Fernandez, R. San-José-Estépar, C. Alberola-Lopez, C.F. Westin // Proceedings of 28th IEEE International Conference, Eng. Med. Biol. Soc.(EMBC). – 2006, August. – P. 4815-4818.
43. Ошаровская Е.В. Оценка показателей качества ТВ изображений // Цифровые технологии. – 2016. – №19. – С. 91-102.
44. Moorthy A.K., Bovik A.C. A two-step framework for constructing blind image quality indices // IEEE Signal Processing Letters. – 2010. – Vol. 17 (5). – P. 513-516.
45. Wang Z., Bovik A.C. Reduced-and noreference image quality assessment // IEEE Signal Processing Magazine. – 2011. – Vol. 28 (6) – P. 29-40.
46. Saad M.A., Bovik A.C. A DCT statistics-based blind image quality index // IEEE Signal Processing Letters. – 2010. – Vol. 17 (6) – P. 583-586.
47. Saad M.A., Bovik A.C., Charrier C. Blind image quality assessment: A natural scene statistics approach in the DCT domain // IEEE Trans. IP. – 2012. – Vol. 21 (8) – P. 3339-3352.
48. Сай С.В., Сорокин Н.Ю. Метод оценки четкости фотореалистичных изображений без использования эталона // Доклады ТУСУР. – 2012. – №2 (26), ч.1 – С.78-82.
49. Moorthy A.K., Bovik A.C. Blind image quality assessment: From natural scene statistics to perceptual quality // IEEE Trans. IP. – 2011. – Vol. 20 (12). – P. 3350–3364.
50. Mittal A., Moorthy A.K., Bovik A.C. Noreference image quality assessment in the spatial domain // IEEE Trans. IP. – 2012. – Vol. 21 (12) – P. 4695-4708.
51. Анализаторы качества изображений Tektronix [Электронный ресурс]. – URL: <http://ru.tek.com/picture-quality-analyzer/>.
52. Сай С.В. Метрика искажений мелких структур компрессированных изображений // Компьютерная оптика. – 2018. – №5 (42). – С. 829-837.
53. Sai S.V., Sorokin N.Yu., Shoberg A.G. Segmentation of Fine Details in the CIELAB // 24th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision. WSCG 2016. Plzen, Czech Republic. May 30 – June 3, 2016. Short Papers Proceedings. – P. 155-162.
54. Wyszecki G. Uniform Color Scales: CIE 1964 U*V*W* Conversion of OSA Committee Selection // JOSA. – 1975. – Vol. 65. – P. 456-460.
55. Sai S.V. Fine-Detail Level of Photorealistic Images: Application in the Multimedia System // International Siberian conference on Control and Communications (SIBCON-2015). Omsk. IEEE Xplore. – 2015.
56. LIVE Image Quality Assessment Database [Электронный ресурс]. – URL: <http://live.ece.utexas.edu/research/quality>
57. Barten P.G.J. Contrast sensitivity of the human eye and its effects on image quality. – Knegsel: HV Press, 1999. – 232 p.

58. Ernawan F., Nugraini S.H. The optimal quantization matrices for JPEG image compression from psychovisual threshold // Journal of Theoretical and Applied Information Technology. – 2014. – Vol. 70 (3) – P. 566-572.
59. Aruna B., Ramesh Ch. A method for signaling block-adaptive quantization in baseline sequential JPEG // International Journal of Modern Engineering Research (IJMER). – 2012. – Vol. 2 (4). – P. 2829-2831.
60. Implementation of ROI Based Baseline Sequential Adaptive Quantization / G. Gowripushpa, G. Santoshi, B. Ravikiran and ect. // International Journal of Emerging Technology and Advanced Engineering. – 2014. – Vol. 4(2) – P. 361-367.
61. Sai S.V. Optimizing the Quantization Parameters of the JPEG Compressor to a High Quality of Fine-Detail Rendition // Pattern Recognition and Image Analysis. – 2018. Vol. 28 (1) – P. 71-78.
62. A novel direction adaptive wavelet based image compression / Liu Guojin, Zeng Xiaoping, Tian Fengchun and ect. // International Journal of Electronics and Communications. – 2010. – Vol. 64 (6). – P.531-539.
63. Дворкович В.П., Дворкович А.В. Расчёт банков фильтров дискретного вейвлет-преобразования и анализ их характеристик // Цифровая обработка сигналов. – №2. – 2006. – С. 2-10.
64. Сай С.В., Савенков И.В. Выбор коэффициентов порогового отбора для вейвлет-трансформант цветного изображения // Информатика и системы управления. – 2001. – № 2. – С. 112-117.
65. Concolato C., Finegan T. AV1 Codec ISO Media File Format Binding // Copyright 2018, The Alliance for Open Media [Электронный ресурс]. – URL: <https://aomediacodec.github.io/av1-isobmff/>
66. Trudeau N., Egge N.E., D. Barr D. Predicting Chroma from Luma in AV1 // cs.MM, 17 Jan 2018. [Электронный ресурс]. – URL: <https://arxiv.org/pdf/1711.03951v2.pdf>
67. MSU Video Quality Measurement Tool [Электронный ресурс]. – URL: http://www.compression.ru/video/quality_measure/video_measurement_tool.html
68. Цифровая обработка изображений в информационных системах: Учебное пособие / И.С. Грузман, В.С. Киричук, В.П. Косых и др. – Спектр. – Новосибирск: Изд-во НГТУ, 2000. – 168 с.
69. Курячий М.И. Цифровая обработка сигналов: Учебное пособие для вузов. – Томск: Томск. гос. ун-т систем упр. и радиоэлектроники, 2009. – 190 с.
70. Митрофанов С.С. Оптические измерительные приборы и устройства: Учебное пособие. – СПб: Университет ИТМО, 2018. – 34 с.
71. Иофис Е.А., Шебалин И.Ю. Фотокинотехника. – «Советская энциклопедия», 1981. – 447 с.
72. Гордийчук О.Ф. Пелль В.Г., Жердецкая Н.Н. Раздел III. Киносъёмочные объективы. Справочник кинооператора. – М.: «Искусство», 1979. – 440 с.
73. Волосов Д.С. Фотографическая оптика. 2-е изд. – М.: «Искусство», 1978. – 543 с.
74. ГОСТ 7427-76. Геометрическая оптика. Термины, определения и буквенные обозначения.

75. Льюцци М. История физики. – М.: «Мир», 1970. – 464 с.
76. Вознесенская А.О., Вознесенский Н.Б., Иванова Т.В. Основы оптики. Конспект лекций. – СПб.: СПбГУ ИТМО, 2009. – 162 с.
77. Гейхман И.Л. Волков В.Г. Основы улучшения видимости в сложных условиях. – М.: ООО «Недра-Бизнесцентр», 1999. – 286 с.
78. Фомин А.В. Общий курс фотографии: Учеб. для техн. 3-е изд., перераб. и до. – М.: Лег-промбыт, 1987. – 256 с.
79. Iamatest [Электронный ресурс]. – URL: www.imatest.com.
80. Титов В.С. Труфанов М.И. Направления развития методов, алгоритмов и аппаратных средств повышения качества изображений оптико-электронных систем // Известия высших учебных заведений. Приборостроение. Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики. Санкт-Петербург. – 2013. – Т.56(6) – С. 7-10.
81. Гольденберг Л.М. Матюшкин Б.Д., Поляк М.Н. Цифровая обработка сигналов: Учебное пособие для вузов. – М.: Радио и связь, 1990. – 256 с.
82. Красильников Н.И., Цифровая обработка 2D- и 3D- изображений: учеб. пособие. – СПб.: БХВ-Петербург, 2011. – 608 с.
83. Каменский А.В. Методы повышения четкости телевизионных изображений высокого качества линейными алгоритмами // Доклады ТУСУР. – 2017. – Т. 20, № 1. – С. 46-49.
84. Конюхов А.Л., Костевич А.Г. Двумерная апертурная коррекция в разреженных масках для повышения качества изображения // Научная сессия ТУСУР–2013: матер. Всерос. техн. конф. студентов, аспирантов и молодых ученых, Томск, 14–16 мая 2013 г.: в 5 ч. – Ч. 1. – Томск: В-Спектр, 2013. – С. 119-122.
85. Каменский А.В., Маланин М.Ю. Измерение разрешающей способности в телевизионных изображениях с использованием программных средств // Научная сессия ТУСУР–2014: материалы Всероссийской научно-технической конференции студентов, аспирантов и молодых ученых, Томск, 14–16 мая 2014 г. – Томск: В-Спектр, 2014: В 5 частях. – Ч. 1. – С. 79-81.
86. Malanin M.Yu., Kamenski A.V., Kuryachiy V.I. Optimization of parameters of two-dimensional filters of increase of clearness of television images on contrastly to frequency characteristics // Control and Communications (SIBCON), IEEE 2015 International Siberian Conference on. IEEE Xplore. – 2015.
87. Рекурсивные цифровые фильтры [Электронный ресурс]. – URL: https://ru.bmstu.wiki/Рекурсивные_цифровые_фильтры, свободный.
88. Давыдов А.В. Цифровая обработка сигналов: Тематические лекции. – Екатеринбург: УГГУ, ИГиГ, ГИН, Фонд электронных документов, 2005 г. – 185 с.
89. Балусов И.Л. Исследование двумерных рекурсивных цифровых фильтров второго порядка: дис. канд. тех. наук: 05.12.04. – Ярославль, 2009. – 171 с.
90. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. – М.: Мир, 1978. – 849 с.
91. Фильтр с бесконечной импульсной характеристикой [Электронный ресурс]. – URL: https://dic.academic.ru/dic.nsf/ruwiki/181332#Фильтр%20с%20бесконечной_импульсной_характеристикой.

92. Мясников В.В. О рекурсивном вычислении свертки изображения и двумерного неразделимого КИХ-фильтра // Компьютерная оптика. – 2005. – С. 117-122.
93. Никитин Д.А. Рекурсивные цифровые фильтры с импульсной характеристикой, описываемой синусоидальной функцией, и свойства их коэффициентов // Известия Тульского государственного университета. Естественные науки. – 2013. – Вып. 2. Ч.1. – С. 161–171.
94. Сергеев В.В. Одномерные параллельно-рекурсивные КИХ-фильтры // Компьютерная оптика. – 1992. – С.186-201.
95. Солонина А. Моделирование цифровой обработки сигналов ЦОС в Matlab. Часть 3. Описание структур КИХ- и БИХ-фильтров в Matlab // Компоненты и технологии – журнал об электронных компонентах. – 2009. – №1 – С. 108-113.
96. Глумов Н.И. Мясников В.В., Сергеев В.В. Применение полиномиальных базисов для обработки изображений в скользящем окне // Компьютерная оптика. – 1995. – Выпуск 14-15. – С. 55-68.
97. Honarvar B., Paramesran R., Lim C. The fast recursive computation of Tchebichef moment and its inverse transform based on Z-transform // Digital Signal Processing. – 2013. – Vol. 23 (5) – P. 1738-1746.
98. Филимонов В.А. Цифровой рекурсивный фильтр, согласованный со сложным сигналом // Информация и космос. – 2016. – №1 – С. 54-61.
99. Турулин И.И. Основы теории рекурсивных КИХ-фильтров. – Таганрог:Южный федеральный университет, 2016. – 264 с.
100. Губин А.Н., Литвинов В.Л. Особенности совместного использования цифровых рекурсивных и нерекурсивных фильтров // Актуальные проблемы инфотелекоммуникаций в науке и образовании. – 2015. – С. 479-482.
101. Губин А.Н., Литвинов В.Л., Филиппов Ф.В. Выбор параметров при разработке рекурсивных цифровых сглаживающих фильтров // Актуальные проблемы инфотелекоммуникаций в науке и образовании. – 2018. – С. 258-262.
102. Palenichka R.M., Lutsyk A. Yu. Fast local image processing algorithms using recursive computation // Cybernetics and Systems Analysis. – 1994. – Vol. 30 (1). – P. 121-130.
103. Синтез квазиоптимального рекурсивного фильтра телевизионных сигналов / В.А. Величкин, В.А. Завьялов, В.П. Бережной, А.М. Мугаттырова // Научное обозрение. – 2015. – С. 177-181.
104. Nehad D., Lima R.S., Hoppe H. GPU-Efficient Recursive Filtering and Summed-Area Tables // Proceeding of the 2011 SIGGRAPH Asia Conference/ Hong Kong, China – December 12–15, 2011.
105. Солонина А.И. Оценки собственного шума структур рекурсивных цифровых фильтров и их вычисление в Matlab // Цифровая обработка сигналов. – 2015. – №2. – С. 39-45.
106. Ehsan S., Clark A.F., Rehman N., McDonald-Maier K.D. Integral Images: Efficient Algorithms for Their Computation and Storage in Resource-Constrained Embedded Vision Systems // Sensors. – 2015. – №15 – P. 16804-16830.
107. Ali R.A., Hardie R.S. Recursive non-local means filter for video denoising // EURASIP Journal on Image and Video Processing (2017) 2017:29.

108. Almahdi R.A., Hardie R.C. Recursive non-local means filter for video denoising with Poisson-Gaussian noise // 2016 IEEE National Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS), Dayton, OH, USA, 25-29 July 2016. – P. 318-322.
109. Зиятдинов С.И., Соколова Ю.В. Синтез комплексных рекурсивных цифровых фильтров // Научная сессия ГУАП. – 2018. – С. 333–335.
110. A Real Time Image Processing Based System to Scaring the Birds from the Agricultural Field / S. Maheswaran, M. Ramya, P. Priyadharshini and ect. // Indian Journal of Science and Technology. – August 2016, Vol. 9(30) – P.5.
111. Anderson B.D., Jury E. I. Stability Test for Two-Dimensional Recursive Filters // IEEE Trans. Audio and Electroacoustics. – AU-21, 4, (August 1973). – P. 366-372
112. Егорова М.А., Литвинов В.Е. Исследование рекурсивной и сепарабельной реализации фильтра нелокального усреднения // 11-я международная научно-техническая конференция и выставка Цифровая обработка сигналов и её применение. – DSPA-2009. – 24-26 марта 2009 г. в ИПУ РАН, Москва. – С. 427-429.
113. Коваль Ю.А., Филиппов М.В. Метод предварительной фильтрации изображений для повышения точности распознавания образов // Инженерный журнал: Наука и инновации. – 2014. – № 12(36).
114. Крапчатова Т.В., Филиппов М.В. Анализ эффективности алгоритмов билатеральной фильтрации // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. – 2012. – № 2.
115. Филиппов М.В., Ермилова О.С. Быстрый метод билатеральной фильтрации изображений // Новые информационные технологии в автоматизированных системах. – 2017. – № 20. – С. 29-36.
116. Беляева О.В., Пащенко О.Б., Филиппов М.В. Быстрая билатеральная фильтрация аэрофотоснимков на основе параллельной декомпозиции на пространственные фильтры // Машиностроение и компьютерные технологии. – 2017. – № 12. – С. 45-56.
117. Рудникович А.С. Сепарабельная билатеральная фильтрация в задачах предобработки видеоданных // Научная сессия ТУСУР–2010: Материалы докладов Всероссийской научно-технической конференции студентов, аспирантов и молодых ученых, Томск, 4–7 мая 2010 г. – Томск: В-Спектр. – 2010. Ч. 1. – С. 169-172.
118. Ефремова Т.Ф. Новый словарь русского языка. Толково-словообразовательный. – М.: Русский язык, 2000.
119. Huang T.S. Two-Dimensional Digital Signal Processing I: Linear Filters // Topics in Applied Physics. – Vol. 42, 1981. – 223 p.
120. Huang T.S. Picture Processing and Digital Filtering. – Springer; 2nd edition (April 1, 1979). – 313 p.
121. Быстрые алгоритмы в цифровой обработке изображений / Т.С. Хуанг, Дж.-О. Эклунд, Г.Дж. Нуссбаумер и др.: Пер. с англ. – М.: Радио и связь, 1984. – 224 с.
122. Захожай О.И., Солошенко А.Н. Высокоскоростная обработка изображений с применением цифровых фильтров на базе FPGA // Сборник научных трудов Донбасского государственного технического университета. – 2013. – Вып. 39. – С. 250-254.

- 123.Альтма Е.А. Захаренко Е.И., Васеева Т.В. Применение метода разложения двумерной свертки при реализации цифровых фильтров // Научный вестник НГТУ. – том 69. – №4. – 2017. – С. 95-104.
- 124.Чобану М., Волеов М. Цифровая обработка многомерных сигналов: предпосылки прорыва // Электроника: Наука, Технология, Бизнес. Микропроцессорная и вычислительная техника. – 2007. – №3 – С. 64-74.
- 125.Sugimoto K. Kamata S. Fast Gaussian filter with second-order shift property of DCT-5 // Proc. IEEE Int. Conf. Image Process. (ICIP), Sept. 2013. – P. 514-518.
- 126.Compiling High Performance Recursive Filters / G. Chaurasia, J. Ragan-Kelley, S. Paris and ect. // HPG '15 Proceedings of the 7th Conference on High-Performance Graphics, Los Angeles, California – August 07-09. 2015. – P. 85-94.
- 127.GReTA – a novel Global and Recursive Tracking Algorithm in three dimensions / A. Attanasi, A. Cavagna, L. D. Castello and ect. // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2015. – Vol. 37 (12) – P. 2451-2463.
- 128.Цифровое телевидение в видеоинформационных системах: моногр. / А.Г. Ильин, Г.Д. Казанцев, А.Г. Костевич и др. – Томск: Томск. гос.ун-т систем упр. и радиоэлектроники, 2010. – 465 с.
- 129.Каменский А.В. Двумерный рекурсивно-сепарабельный фильтр повышения четкости телевизионных изображений // сб. материалов XIV междунар. науч.-тех. конф. /ред. кол.: С.Г. Емельянов [и др.]; Юго-Зап. гос. ун-т. Курск, 2018. – С. 136-138.
- 130.Каменский А.В. Головина Д.Е., Курячий М.И. Высокопроизводительные рекурсивно-сепарабельные алгоритмы повышения качества изображений // материалы XIV Международная научно-практическая конференция «Электронные средства и системы управления». Томск: В-Спектр, 2018. – Ч. 2. – С. 162-165.
- 131.Каменский А.В., Курячий М.И. Рекурсивно-сепарабельные двумерные цифровые фильтры обработки телевизионных // ИЗВЕСТИЯ Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение: г. Курск: 2018. Т.-8, №. 4(29). – С. 25-33.
- 132.Свидетельство о государственной регистрации программы для ЭВМ № 2019660074. MIR: Множественное измерение разрешения. Автор: А.В. Каменский. Заявка № 2019618804. Дата поступления 17 июля 2019 г. Зарегистрировано в Реестре программ для ЭВМ 30.07.2019 г.
- 133.Свидетельство о государственной регистрации программы для ЭВМ № 2019660075. SFI: Построчная фильтрация изображений. Автор: А.В. Каменский. Заявка № 2019618808. Дата поступления 17 июля 2019 г. Зарегистрировано в Реестре программ для ЭВМ 30.07.2019 г.

Оглавление

Введение.....	3
1. Системы и методы передачи изображений	5
1.1. Общие принципы передачи изображений	5
1.2. Особенности сжатия изображений на основе JPEG	12
1.3. Особенности сжатия изображений на основе JPEG 2000	18
1.4. Цифровые видеокодеки.....	23
1.4.1. Стандарт MPEG-2	23
1.4.2. Стандарт MPEG-4	29
1.4.3. Стандарт H.264	32
1.4.4. Стандарт H.265/HEVC	35
1.5. Характеристики визуальных искажений.....	37
1.5.1. Искажения JPEG	37
1.5.2. Искажения JPEG 2000	39
1.5.3. Искажения видеокодеков	41
2. Метрики качества цифровых изображений.....	43
2.1. Общие характеристики качества изображений	43
2.2. Оценка качества изображений по испытательным таблицам	46
2.3. Метрики оценки качества фотореалистичных изображений	54
2.4. Оценка качества видеоданных	61
2.5. Метрика оценки качества передачи мелких структур и границ крупных деталей изображения	68
2.5.1. Система N-CIELAB	68
2.5.2. Алгоритм идентификации мелких структур	76
2.5.3. Метрика MFSD	79
2.5.4. Алгоритм оценки резкости	84
3. Повышение качества компрессированных изображений	87
3.1. Повышение качества изображений в стандарте JPEG	87
3.1.1. Идентификации мелких деталей	87
3.1.2. Классификация блоков и оптимизация параметров квантования	90
3.2. Повышение качества изображений в стандарте JPEG 2000	98
3.2.1. Способ оценки искажений	98
3.2.2. Алгоритм адаптивного квантования	99
3.2.3. Результаты экспериментов	102
3.3. Корректирующие фильтры для улучшения качества изображений видеокодеков.	104
3.4. Сравнение качества видеофайлов на основе MSU VQMT.....	108
3.5. Повышение точности фокусировки цифровой камеры	118
4. Цифровые фильтры повышения четкости изображений	126
4.1. Особенности измерения разрешения по полю изображения	126
4.2. Методы и алгоритмы повышения качества изображений на основе цифровой фильтрации	135

4.2.1. Линейные алгоритмы обработки изображений	135
4.2.2. Рекурсивные алгоритмы обработки изображений	143
4.2.3. Быстродействующие алгоритмы обработки изображений	151
4.3. Рекурсивно-сепарабельные фильтры повышения четкости изображений	154
4.4. Повышение качества изображений при их позонной обработкой обработке.....	160
Заключение	164
Литература	165

Научное издание

САЙ Сергей Владимирович
КАМЕНСКИЙ Андрей Викторович
КУРЯЧИЙ Михаил Иванович

Современные методы анализа и повышения качества цифровых изображений

Дизайнер обложки *И. Л. Тюкавкина*

Отпечатано с авторского оригинала-макета

Подписано в печать 20.07.20. Формат 60х84/16.
Усл. печ. л. 10,23. Тираж 500 экз. Заказ 147.

Издательство Тихоокеанского государственного университета.
680035, Хабаровск, ул. Тихоокеанская, 136.

Отдел оперативной полиграфии
издательства Тихоокеанского государственного университета.
680035, Хабаровск, ул. Тихоокеанская, 136