

На правах рукописи



Голованов Роман Вячеславович

УВЕЛИЧЕНИЕ КОМПРЕССИИ ИЗОБРАЖЕНИЙ
В АЛГОРИТМАХ С
ФУРЬЕ-ПРЕОБРАЗОВАНИЯМИ

Специальность 05.13.11 — Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Автореферат
диссертации на соискание учёной степени
кандидата физико-математических наук

Москва — 2014

Работа выполнена на кафедре «Высшая математика №1» Национального исследовательского университета «МИЭТ».

Научный руководитель:

Прокофьев Александр Александрович доктор педагогических наук,
заведующий кафедрой
«Высшая математика №1»
МИЭТ

Официальные оппоненты:

Богданов Юрий Иванович доктор физико-математических
наук, заведующий лабораторией
физики квантовых компьютеров
ФТИАН

Лаптева Валентина Владимировна кандидат физико-математических
наук, научный сотрудник,
ЗАО «НТЦ ЭЛИНС»

Ведущая организация: ОАО НПЦ «ЭЛВИС»

Защита состоится «23» сентября 2014 г. в 16 часов 00 минут на заседании диссертационного совета Д212.134.02 при Национальном исследовательском университете «МИЭТ» по адресу: 124498, г. Москва, Зеленоград, проезд 4806, д. 5.

С диссертацией можно ознакомиться в библиотеке МИЭТ.

Автореферат разослан «__»_____ 2014 г.

Ученый секретарь
диссертационного совета Д212.134.02
доктор технических наук

Гуреев А.В.

Общая характеристика работы

Актуальность темы. Возрастающий поток передачи цифровой информации определяет всё более высокие требования к алгоритмам обработки цифровых данных. Переход на цифровые устройства во всех видах создания, использования, передачи и хранения данных, происходящий во всём мире, ещё больше повышает важность оптимального представления цифровых данных. Наиболее актуальными остаются алгоритмы сжатия цифровых изображений. Широкое распространение среди таких алгоритмов получил кодек JPEG, предложенный около 20 лет назад. Общемировая тенденция к совершенствованию этого кодека сохраняется и по сей день. Наиболее перспективным оказывается алгоритмический подход, так как усовершенствования на программном уровне практически исчерпаны. Поэтому разработка эффективных алгоритмов для сжатия цифровых изображений является актуальной.

Разработка алгоритмов сжатия связана с проблемой оценки качества искажённого изображения. Общепринятый критерий сигнал-шум для оценки качества плохо согласуется с субъективным мнением экспертов. Поэтому актуальной является задача разработки лучшего математического критерия. В настоящее время используется более 20 альтернативных критериев. Критерии проверяют на специальных базах тестовых изображений, сопоставляя значения, получаемые критериями, с усреднёнными экспертными оценками. В настоящий момент нет методики сопоставления результатов тестирования критериев на разных базах. Разработка такой методики — одна из актуальных задач.

Предметная область. Исследование ведётся в нескольких смежных областях, соответствующих паспорту специальности 05.13.11:

1. Моделирование, разработка методик и алгоритмов для цифровой обработки изображений.
2. Оценка качества работы алгоритмов сжатия изображений.

Предметами исследования являются алгоритмы дискретного преобразования Фурье для сжатия изображений с потерями и критерии оценки качества получаемого изображения посредством таких алгоритмов.

Цель и задачи. В диссертации разрабатываются улучшенные методы компрессии цифровых изображений на основе контекстного кодирования в области дискретного косинусного преобразования.

Для достижения цели были поставлены и решены следующие научно-технические задачи:

1. Разработка алгоритма адаптивного кодирования в кодеках на основе дискретного косинусного преобразования (ДКП).
2. Обобщение алгоритмов сжатия изображений на основе ДКП на блоки произвольного размера.

3. Построение математического критерия оценки качества искажённого изображения, соответствующего человеческому восприятию.
4. Повышение достоверности методов тестирования алгоритмов на базах изображений с различными видами искажений.

Методы исследования. В ходе работы над диссертацией использовались теоретические и экспериментальные методы. Теоретическими являются теория вероятностей и математической статистики, теории цифровой обработки сигналов и кодирования данных, линейной алгебры, математического анализа, численных методов.

В ходе исследования производилось численное моделирование на компьютере с использованием написанных и доработанных приложений на языках С и С++, а также пакета MATLAB.

Эксперимент состоял в сравнении результатов теоретических расчётов с экспертными оценками, взятыми из современных баз тестовых изображений.

Достоверность результатов, научных положений, выводов и рекомендаций подтверждается использованием для их получения общепринятых научных подходов и методов разработки, а также результатами численных экспериментов. Расчёты, проведённые с использованием наборов стандартных тестовых изображений, подтверждают, что использованные алгоритмы оценки качества искажённого изображения статистически достоверны и соответствуют человеческому восприятию.

Преимущество созданных методов и алгоритмов сжатия цифрового изображения оценивалось с использованием известных критериев качества.

Личный вклад автора. В основу диссертации легли результаты исследований, выполненных лично автором. Личный вклад автора состоял также в непосредственном участии в получении исходных данных, в апробации результатов исследований, в обработке и интерпретации полученных данных, в подготовке основных публикаций по выполненной работе.

Научная новизна. В работе предлагаются следующие новые алгоритмы, методы и программы.

1. Предложен алгоритм BigJPEG, обобщающий кодек JPEG на блоки большего размера, и его программная реализация.
2. Предложена экономичная схема кодирования JPEG-IT для базового кодека JPEG и её программная реализация.
3. Создан критерий оценки качества изображения, построенный на основе нормы Соболева.
4. Предложен метод объединения баз тестовых изображений, построенных на основе разнотипных экспертных оценок.
5. Разработана методика сравнения критериев оценки качества изображения и определение лучшего среди них.

6. Проведено сравнение алгоритмов сжатия изображений на основе лучшего по достоверности современного критерия оценки качества изображения.

Практическая значимость:

1. Разработанные кодеки BigJPEG и JPEG-IT, повышающие сжатие изображений от 9 до 20%, целесообразно использовать для хранения и передачи изображений в условиях технических ограничений аппаратуры.
2. Предложенная методика сравнения критериев оценки качества изображений на объединённой базе, может быть использована для статистически достоверного ранжировать критерии качества и выбора лучшего из них.
3. Предложенный критерий оценки качества изображения SGC, может быть использован для ускорения процесса тестирования алгоритмов обработки изображений.

Основные положения, выносимые на защиту:

1. Разработанные кодеки BigJPEG и JPEG-IT повышают сжатие изображений от 9 до 20% при том же качестве по сравнению с базовым кодеком JPEG.
2. Предложенная методика сравнения критериев оценки качества изображений на объединённой базе позволяет достоверно ранжировать критерии качества и оценивать их доверительную вероятность.
3. Предложенный критерий оценки качества изображения, сопоставим по достоверности с лучшими современными критериями, но существенно превосходит их по простоте реализации.

Апробация. Основные результаты работы докладывались на 15-й Международной конференции «Цифровая обработка сигналов и её применение» (Москва, 2013 г.), на 19-ой, 20-ой и 21-ой Всероссийских межвузовских научно-технических конференциях студентов и аспирантов «Микроэлектроника и информатика» (Москва, МИЭТ, 2012, 2013 и 2014 гг.), на научном семинаре ЗАО «ЭЛВИС-НеоТек» (Москва, Зеленоград, 2013 г.), на научном семинаре ФТИАН (Москва, 2014 г.).

Публикации. По теме диссертации опубликовано 9 работ, в том числе 5 статей в журналах, рекомендованных ВАК: «Доклады Российской академии наук» — 2, «Цифровая обработка сигналов» — 1 и «Математическое моделирование» — 2. Получено 2 свидетельства о государственной регистрации программы для ЭВМ.

Структура и объём диссертации. Диссертация состоит из введения и четырёх глав, изложена на 134 страницах, содержит 101 рисунок и 16 таблиц. Список литературы насчитывает 94 источника.

Содержание работы

Во введении излагается область исследования и даётся обзор научной литературы по данному направлению. Обосновывается актуальность темы диссертации, формулируются цели и задачи работы, научная новизна, практическая значимость полученных результатов, а также положения, выносимые на защиту. Приведены структура и краткое содержание работы.

Глава 1 посвящена проблеме построения математического критерия оценки качества искажённого изображения. Практика показывает, что при передаче изображения часто возникают различные искажения. Часть из них связана с техническими сбоями каналов связи (в числе которых может быть передача по воздуху). Другая часть вынужденно вводится при сжатии изображения, если требуется существенно уменьшить количество передаваемой информации. В обоих случаях в искаженном изображении вместо истинной яркости n -го пикселя x_n будет стоять изменённое число y_n (также целое и в тех же пределах).

Пользователь получает искаженное изображение, и должен по нему судить об исходном изображении. Разработка, как аппаратуры, так и способов сжатия с потерями направлены на получение достаточно хорошего качества переданного изображения. При тестировании новых разработок можно предложить пользователю большой набор исходных изображений и переданных искажённых изображений, чтобы пользователь визуально оценил качество. Для получения точных оценок необходимо привлекать большую группу испытуемых и чётко регламентировать процедуру сравнения. Это очень трудоёмкая и дорогостоящая работа, аналогичная натурным экспериментам в физике и технике. Каждый раз, когда будет предложена новая аппаратура или новый метод обработки изображений, такую процедуру придётся повторять заново.

Для решения этой проблемы на практике создают базу эталонных и искаженных изображений, и один раз выполняют экспертную оценку качества искажённых изображений (MOS — mean opinion score). Далее по массивам яркостей $X = \{x_n\}$ и $Y = \{y_n\}$ пытаются построить некоторый индекс качества – математическую функцию от массивов яркостей $q(X, Y)$. Если эта функция монотонна, то данным критерием качества можно пользоваться при конструкторско-исследовательской работе вместо проведения новых экспертных оценок.

Однако построить такой критерий оказывается не просто. Этот вопрос активно обсуждается в научной литературе. Обычно исследователи исходят из некоторых естественных идей и пишут достаточно простые критерии. На практике они не очень хорошо совпадают с экспертными оценками. Их начинают усложнять, причем довольно эклектичным образом и при этом вводят ряд свободных (подгоночных) параметров. Итоговые формулы

зачастую оказываются весьма громоздкими, а их адекватность человеческому восприятию — не очевидной.

Критерий SGC. В данной работе предложен очень простой критерий сглаженных градиентов (SGC — Smoothed Gradients Criterion), основанный на анализе психофизики зрения. Он содержит всего один свободный параметр, а по точности оценки сопоставим с двумя лучшими критериями используемыми специалистами. Предлагаемый критерий заключается в следующем. Сначала массивы X и Y сглаживаются фильтром Гаусса:

$$\rho \sim \exp(-r^2/r_0^2), \quad (1)$$

где: r — расстояние между центрами пикселя и маски, r_0 — радиус размытия (эта величина является свободным параметром). Множители пропорциональности в (1) подбираются так, чтобы сумма весов по всей маске равнялась единице.

Если вычислять интеграл от квадрата ошибки (разница яркостей каждого пикселя изображений X и Y), то получаем погрешность в норме \mathbb{L}_2 , которую традиционно используют для оценки качества изображений. Если к этому интегралу прибавить ошибку производной в норме \mathbb{L}_2 , то это будет норма Соболева W_2^1 . Для двумерного дискретного сигнала X норма Соболева имеет вид:

$$\|X\|_{W_2^1} = \sqrt{\|X\|^2 + \|X'_{\rightarrow}\|^2 + \|X'_{\downarrow}\|^2}. \quad (2)$$

где X'_{\rightarrow} и X'_{\downarrow} — сеточная производная X по горизонтали или вертикали, а знак нормы $\|X\|$ означает Евклидову норму.

На основе (2) предлагается вычислять относительную погрешность между сигналами X и Y , задающими исходное и искажённое изображения. Ранее эту норму не использовали для оценки качества изображений. Относительная погрешность будет иметь вид

$$\delta^2 = \frac{\|X - Y\|^2}{\|X + Y\|^2} + \alpha \frac{\|X'_{\rightarrow} - Y'_{\rightarrow}\|^2}{\|X'_{\rightarrow} + Y'_{\rightarrow}\|^2} + \alpha \frac{\|X'_{\downarrow} - Y'_{\downarrow}\|^2}{\|X'_{\downarrow} + Y'_{\downarrow}\|^2} \quad (3)$$

где $X - Y$ означает поэлементную разность матриц X и Y . Фактически (3) содержит один свободный параметр α — вес градиентных слагаемых. В ходе наших экспериментальных расчётов на тестовых изображениях было установлено, что результат получается наилучшим, если коэффициент α очень велик. Поэтому достаточно оставить в (3) только градиентный член.

Для массивов сглаженных по (1) погрешность вычисляется согласно (3). За критерий качества принимается величина

$$\Delta_{SGC} = -0.5 \lg(\delta). \quad (4)$$

Здесь использование логарифма позволяет нагляднее оценивать погрешность, меняющуюся на несколько порядков. Видно, что критерий (1)–(4) прост, нагляден и содержит только один свободный параметр r_0 .

Сравнение критериев. Оценка работы критерия давалась традиционным способом в сравнительном анализе с 17-ю критериями качества, используемыми на практике. Для всех искажённых изображений, полученных из оригинала, вычислялось количественное значение каждого критерия, которое сопоставлялось с экспертной оценкой.

Любой теоретический критерий можно считать некоторой функцией экспертной оценки. Экспертная оценка является экспериментальной величиной и содержит случайные ошибки. Поэтому рассматриваемая функция является случайной. В качестве простейшей характеристики этой функции обычно вычисляют коэффициенты парной корреляции Пирсона, Спирмена и Кендалла. Наиболее жёстким является сравнение по Пирсону. Поэтому окончательное ранжирование критериев выполнялось по нему. Тестирование по коэффициентам Спирмена и Кендалла даёт сходные результаты, но менее обоснованные.

В ходе исследования производилось тестирование критериев оценки качества на двух наиболее содержательных и используемых базах тестовых изображений LIVE и TID2008. Была выработана более достоверная методика сравнения. Наборы искажённых изображений группировались по оригиналам, для каждого оригинала вычислялось среднее значение коэффициента парной корреляции. Затем по массиву таких значений определялся стандарт, который характеризует стабильность работы критерия качества. Так для каждой базы и каждого критерия были получены свои характеристики. Окончательное ранжирование делалось по полу сумме средних значений на каждой базе и определялся разброс мест занимаемых критериями на разных базах.

Из Табл. 1 видно, что созданный критерий SGC вошёл в число трёх лучших, причём его отличия от второго места статистически незначимо. Следует отметить, что второй критерий относится к недостаточно стабильным критериям. Критерий SGC стабилен: в базе LIVE он занимает третье место, а в базе TID2008 четвертое, лишь незначимо уступающее третьему.

Критерии с плохой стабильностью рискованно использовать в работе. Такие критерии занимают не близкие места в разных базах тестовых изображений. Они отмечены в таблице звёздочками. Одна звёздочка — критерий не стабильный, а две — очень не стабильный. Отметим, что ряд популярных критериев, таких как IW-SSIM, MSSIM, SSIM и IFC, оказались среди очень нестабильных. Это означает, что они менее надёжны и их использование не целесообразно.

Глава 2 посвящена проблеме сопоставления критериев оценки качества на нескольких экспертных базах. Практически отсутствуют научные публикации, посвящённые решению этой проблемы. Критерии оценки ка-

Таблица 1 — Сравнение критериев по PLCC

№	Критерии качества	Полу сумма баз	База LIVE	База TID2008
1	VIF	0,868 4	0,951 ± 0,033	0,786 ± 0,062
2*	PSNRHA	0,856 7	0,888 ± 0,039	0,824 ± 0,074
3	SGC	0,848 1	0,900 ± 0,039	0,796 ± 0,063
4*	PSNRHMA	0,831 6	0,878 ± 0,039	0,784 ± 0,078
5	SAC	0,817 4	0,900 ± 0,041	0,733 ± 0,077
6	UQI	0,801 4	0,897 ± 0,046	0,705 ± 0,143
7**	IWSSIM	0,796 15	0,756 ± 0,054	0,837 ± 0,077
8*	VIFP	0,787 8	0,928 ± 0,037	0,646 ± 0,085
9	NQM	0,754 0	0,882 ± 0,057	0,626 ± 0,083
10**	SSIM	0,753 10	0,746 ± 0,078	0,760 ± 0,107
11**	MSSIM	0,748 15	0,688 ± 0,073	0,808 ± 0,090
12	PSNRHVS	0,732 2	0,885 ± 0,040	0,580 ± 0,083
13*	VSNR	0,731 6	0,892 ± 0,039	0,570 ± 0,287
14	PSNRHVSM	0,717 1	0,877 ± 0,040	0,556 ± 0,080
15	PSNR	0,683 1	0,837 ± 0,049	0,528 ± 0,101
16	SNR	0,683 1	0,837 ± 0,049	0,528 ± 0,101
17*	WSNR	0,681 9	0,891 ± 0,049	0,471 ± 0,083
18**	IFC	0,548 12	0,895 ± 0,050	0,200 ± 0,106
19	MSE	0,066 0	0,417 ± 0,063	0,285 ± 0,196

чества следует тестировать не на одной базе, а на наибольшем доступном числе баз. Это даст более объективную оценку критерию. Обычно разработчики критериев приводят значения коэффициентов корреляции на каждой из баз и по этим результатам делают выводы (см. Главу 1). Зачастую один и тот же критерий в сравнении с другими оказывается на разных местах в разных базах после ранжирования. Заметное отличие в занимаемых местах (дисперсия номера места) указывает на нестабильность критерия. Поэтому предлагается произвести объединение нескольких баз тестовых изображений и сравнение критериев производить на объединённой базе.

Объединение баз. Требуется проведение тщательного анализа самих баз: насколько хорошо структурированы данные, каковы стандарты приведенных чисел (статистический разброс экспертных оценок), насколько чётко формулировались инструкции экспертам и проверялось правильно ли поняли эксперты эти инструкции.

В работе проводится тщательный анализ двух наиболее представительных баз LIVE и TID2008, используемых на практике. Для каждой базы были подробно описаны представленные типы искажений, разобраны методы получения экспертных оценок и выявлены некоторые недостатки. Другие базы пока ещё не получили широкого распространения, но уже используются узким кругом специалистов. Почти все они уступают базам LIVE и TID2008 по количеству рассматриваемых типов искажений и по количеству тестовых изображений.

Следующий этап работы по объединению баз — это сопоставление экспертной оценки (MOS — mean opinion score) и оценки на основе применения математического критерия (IQA — image quality assesment). Эта зависимость проиллюстрирована на Рис. 1. Чем уже облако точек, тем адекватнее и стабильнее работает критерий. В работе выявлено, что критерий VIF является наилучшим из известных.

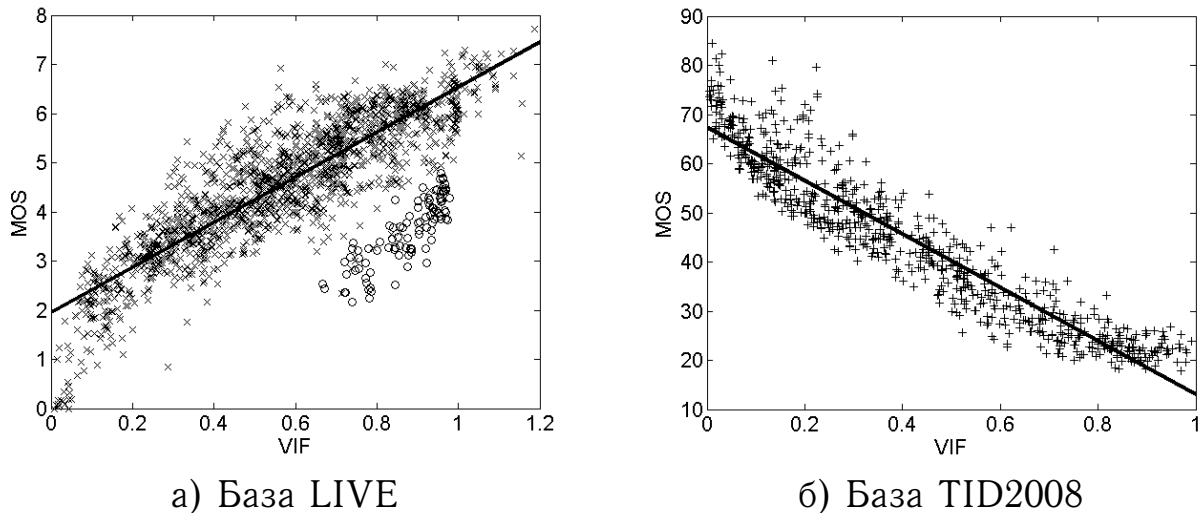


Рисунок 1 — Зависимость усреднённой экспертной оценки MOS от критерия VIF. Прямые — линейные регрессии. Маркеры: \circ — искажения типа №15 в базе TID2008, \times — все остальные типы искажения в базе TID2008, $+$ — все искажения базе LIVE

По неоднородности облака точек можно судить об адекватности критерия на отдельных типах искажений. Для этого достаточно оценить, как накладываются точки данного искажения на точки всех других искажений. По критерию VIF было выявлено, что обособленное скопление соответствует только изображениям с локальными блочными искажениями разной интенсивности (тип искажения №15 из спецификации базы TID2008, см. чёрные точки на Рис. 1б). Точки других искажений лежат в основном облаке.

Далее оценивалось среднеквадратичное отклонение точек от регрессионной кривой. При этом исключались точки, не вошедшие в основное облако. Линейная регрессия является оптимальной для описания облаков точек. Соответственно для баз LIVE и TID2008 получили уравнения линейных регрессий:

$$\begin{aligned} f_{LIVE}(x) &= -54.4225 \cdot x + 67.4189, \\ f_{TID2008}(x) &= 4.5800 \cdot x + 1.9550. \end{aligned} \quad (5)$$

Объединение баз сводится к преобразованию шкалы экспертных оценок нескольких баз к одной. Это преобразование должно быть таким, чтобы оптимальные регрессионные кривые точно совпадали. Для линейных

регрессий из (5) искомое преобразование будет линейным:

$$X' = -0.084157 \cdot X + 7.6287, \quad (6)$$

где: X — экспертные оценки LIVE в “родной” шкале, а X' — в шкале TID2008. Результат объединения показан на Рис. 2.

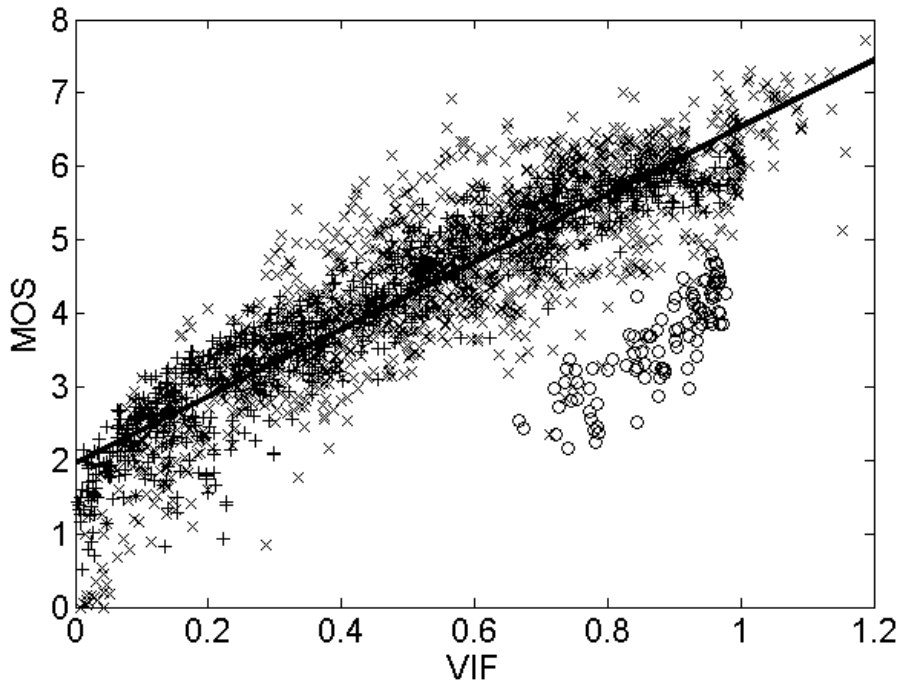


Рисунок 2 — Критерий VIF на объединенной базе. Маркеры см. на Рис. 1

В ходе исследования были проведены аналогичные (очень трудоёмкие) работы по объединению баз на основе других критериев. Для них картины облаков точек оказываются существенно более сложными (например, кометообразными и содержащими несколько боковых выбросов), а плотность распределения точек в облаках является заметно неравномерной. Для ряда критериев облака на базах TID2008 и LIVE сильно отличались друг от друга. Во всех этих случаях не представлялось разумных регрессий, совмещающих облака обеих баз. Объединение баз с помощью таких критериев не разумно. Именно поэтому предлагается объединение баз только по критерию VIF.

Важной характеристикой базы является стандарт экспертных оценок. Зная стандарты исходных баз, не трудно вычислить стандарт объединённой базы по формуле

$$S = \sqrt{\frac{S_1^2 \cdot N_1 + S_2^2 \cdot N_2}{N_1 + N_2}}, \quad (7)$$

где: S_1 и S_2 — стандарт MOS в обеих базах, а N_1 и N_2 — количество тестовых изображений в них.

Сравнение критериев оценки качества изображений проводилось на объединённой базе. В ходе работы выяснилось, что современные оценки работы критериев по коэффициентам парной корреляции не являются показательными. Это легко подтверждается на тестовом абстрактном примере приведённом в работе.

Поэтому был выработан другой подход. Наиболее надёжной оценкой для критерия является значение его стандартного отклонения от регрессионной кривой. Регрессия строится по множеству точек $MOS(IQA)$. Критерий с наименьшим стандартом признаётся лучшим. Для оценки его адекватности нужно сравнивать стандарт регрессии и стандарт экспертных оценок базы.

Таблица 2 — Тестирование критериев на объединённой базе

Критерии	Усечённая								Полная			
	S_C/S_B		PLCC		SRCC		KRCC		S_C/S_B		PLCC	
VIF	3,52	1	0,905	1	0,918	2	0,75	1	4,69	1	0,838	1
PSNRHMA	4,05	2	0,871	2	0,92	1	0,746	2	5,28	4	0,789	4
PSNRHA	4,06	3	0,871	3	0,913	4	0,735	4	4,94	2	0,818	2
PSNRHVSM	4,08	4	0,87	4	0,918	3	0,742	3	6,63	12	0,637	12
PSNRHVS	4,09	5	0,869	5	0,912	5	0,733	5	6,48	11	0,657	11
SGC	4,09	6	0,869	6	0,896	7	0,709	7	5,14	3	0,801	3
VSNR	4,19	7	0,862	7	0,895	8	0,706	9	7,05	16	0,571	16
WSNR	4,19	8	0,862	8	0,892	9	0,706	8	7,11	17	0,563	17
VIFP	4,29	9	0,854	9	0,873	12	0,682	11	5,89	6	0,729	6
NQM	4,53	10	0,836	10	0,844	14	0,646	14	6,64	13	0,636	13
SAC	5,06	11	0,791	11	0,848	13	0,646	13	5,95	9	0,722	9
UQI	5,24	12	0,773	12	0,764	19	0,565	18	5,91	7	0,726	7
IWSSIM	5,33	13	0,763	13	0,907	6	0,723	6	5,69	5	0,75	5
SSIM	5,36	14	0,761	14	0,883	10	0,686	10	5,92	8	0,725	8
PSNR	5,36	15	0,76	15	0,802	17	0,594	17	6,84	14	0,607	14
SNR	5,64	16	0,731	16	0,769	18	0,563	19	7,02	15	0,578	15
IFC	5,64	17	0,73	17	0,839	15	0,645	15	8,3	19	0,261	19
MSSIM	5,84	18	0,708	18	0,878	11	0,68	12	6,15	10	0,698	10
MSE	7,73	19	-0,353	19	-0,802	16	-0,594	16	8,18	18	-0,309	18

В соответствии с разработанным методом оценки критериев, было проведено сравнение наиболее распространённых критериев качества (см. Табл.2). Лучшим оказался критерий VIF. Критерий SGC (см. Главу 1) занимает третье место, причём его отличие от второго места статистически не значимо. Соотношения стандартов критерия и базы показывают, что даже лучшие критерии очень далеки от достоверной передачи человеческого восприятия искажённого изображения. Тем не менее современные критерии дают лучший результат, чем традиционный критерий PSNR. Поэтому при проектировании новых алгоритмов цифровой обработки изображений

следует пользоваться только лучшими представителями современных критериев.

Глава 3 посвящена разработке алгоритмов сжатия изображений. В настоящее время самым используемым алгоритмом сжатия является кодек JPEG, основанный на дискретном косинусном преобразовании (ДКП). Сейчас алгоритм сжатия JPEG является фактическим стандартом, поддерживаемым многочисленными аппаратными и программными системами цифровой обработки, хранения и передачи изображений. Существуют и другие кодеки сжатия, например кодек JPEG2000, основанном на вейвлет-преобразовании. Он не является общепринятым из-за сложности реализации алгоритма, как программной, так и аппаратной и из-за противоречивых результатов сопоставления кодеков JPEG и JPEG2000 по сжатию и получаемому качеству изображения.

Известно, что визуальные отличия артефактов блочности, свойственные кодеку JPEG, и размытия, свойственные кодеку JPEG2000, воспринимаются по-разному. Поэтому человек даёт разную оценку изображениям с одинаковой степенью сжатия после компрессии кодеками JPEG и JPEG2000. В современных работах показано, что средняя экспертная оценка качества для изображений сжатых кодеком JPEG выше, чем для изображений сжатых кодеком JPEG2000. Это означает, что алгоритмы сжатия на основе ДКП по-прежнему остаются более эффективными по совокупности характеристик. Поэтому в данной работе проводились исследования в области совершенствования кодека JPEG, как представляющие наибольшую перспективность.

Стандартный алгоритм JPEG обрабатывает изображения блоками $b_{m,n}$ размером 8×8 пикселей, где $m \geq 0$ — номер строк, а $n \geq 0$ — номер столбца, в которых расположен блок. Каждый блок образует матрицу $X : \{x_{i,j}\}_{i,j=0}^7$, где $x_{i,j}$ — яркость пикселя с индексами i и j в текущем блоке. Эта матрица обрабатывается в несколько этапов:

1. Прямое ДКП, которое даёт матрицу коэффициентов разложения $Y : \{y_{i,j}\}_{i,j=0}^7$.
2. Квантование, которое обеспечивает сжатие за счёт округления коэффициентов ДКП $y_{i,j}$ путём поэлементного деления элементов матрицы Y на соответствующие элементы матрицы квантования $Q : \{q_{i,j}\}_{i,j=0}^7$, что даёт новую матрицу \tilde{Y} .
3. «Зигзаг» упорядочивание, которое формирует из элементов матрицы \tilde{Y} одномерный массив V , модули элементов которого преимущественно, но не строго, убывают.
4. Энтропийное кодирование по Хаффману, которое преобразует вектор чисел V в последовательность битовых кодов.

Матрица квантования. В кодеке JPEG матрица квантования рассчитана на блоки фиксированного размера 8×8 пикселей; она приведена слева

на Рис. 3. Матрица квантования определяет качество и сжатие получаемого изображения. Чтобы влиять на качество и сжатие, элементы матрицы домножают на коэффициент, задаваемый пользователем.

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

11	12	13	15	20	28	38	50
12	13	15	20	28	38	50	64
13	15	20	28	38	50	64	78
15	20	28	38	50	64	78	92
20	28	38	50	64	78	92	104
28	38	50	64	78	92	104	116
38	50	64	78	92	104	116	126
50	64	78	92	104	116	126	135

Стандартная
Новая

Рисунок 3 — Матрицы квантования

Отметим достаточно очевидные, хотя не очень существенные, недостатки этой матрицы:

- она не симметрична,
- значение $q_{0,0}$ в $\sim \sqrt{2}$ раз отличается от соседних элементов,
- значения $q_{i,j}$ в правом нижнем углу матрицы немонотонно зависят от индексов i и j ,
- не приведено способа обобщения матрицы на блоки произвольного размера.

В научной литературе также нет примеров матрицы квантования для блоков большего размера. В данной работе построена матрица квантования, лишённая этих недостатков. Процесс построения включал подбор различных регрессий $q_{i,j}$. Лучшей считалась регрессия, устойчиво дающая большее сжатие изображения при заданном качестве.

Наилучшие результаты мы получили для следующей регрессии. Первая строка и первый столбец вычисляются по формуле

$$q_{0,j} = q_{j,0} = \frac{a + b(j+1)^3}{1 + c(j+1)^3}, 0 \leq j \leq B-1. \quad (8)$$

Коэффициенты a, b, c вычисляются из условий:

$$q_{0,0} = 11, q_{0,B-1} = q_{B-1,0} = 50, q_{B-1,B-1} = q_{0,2B-1} = 135. \quad (9)$$

Внутренним элементам матрицы приписываются значения граничных, расположенных на той же побочной диагонали, как при «зигзаг» упорядочивании:

$$q_{i,j} = q_{0,i+j} = q_{i+j,0}, 1 \leq i \leq B-1, 1 \leq j \leq B-1. \quad (10)$$

На Рис. 3 справа приведена новая матрица квантования. Видно, что новая матрица лишена недостатков, отмеченных нами для стандартной матрицы.

Новая матрица использована в разработанных нами кодеках BigJPEG и JPEG-IT.

BigJPEG. Стандарт JPEG был принят около 20 лет назад и рассчитан на возможности аппаратуры того времени. Как известно увеличение мощности вычислительной техники приводило к пересмотру и совершенствованию численных алгоритмов. Развитие ПЗС матриц (увеличение количества пикселей) также ставит вопрос о пересмотре алгоритмов цифровой обработки изображения. Один из таких вопросов — оптимальный размер блоков для ДКП в кодеке JPEG. Если объём ПЗС матрицы мал, то мелкие детали изображения описываются всего несколькими пикселями, а если объём ПЗС матрицы велик, то мелкие детали описываются большим числом пикселей. Таким образом разумным представляется выбирать размер блока разбиения в зависимости от объёма ПЗС матрицы. В данной работе на примере предложенного алгоритма BigJPEG показано, что в этом случае можно получить более хорошее сжатие изображения.

ДКП в кодеке JPEG реализуют на основе быстрого преобразования Фурье (БПФ), что обеспечивает логарифмическую сложность алгоритма. Напомним, что БПФ успешно реализуется для размеров блоков $B = 2^k, k \in \mathbb{N}$. При других размерах блока БПФ либо вообще не возможно, либо становится гораздо более сложным.

Для квадратного изображения размером $S \times S$ пикселей разбитого на блоки $B \times B$ пикселей трудоёмкость составляет

$$\text{const} \cdot (S/B)^2 B \log(B). \quad (11)$$

Видно, что эта функция убывает, как $\log(B)/B$, то есть почти как B^{-1} . Поэтому переход от стандартных блоков 8×8 к большим блокам даёт экономию, приведённую в Табл. 3. Видно, что это большая экономия, быстро возрастающая с увеличением размера блоков.

Таблица 3 — Уменьшение сложности ДКП для блоков разного размера

Размер блока, пикселей	8×8	16×16	32×32	64×64	128×128
Экономия, раз	1.0	1.5	2.4	4.0	6.8

Таким образом, оценка трудоёмкости алгоритма убедительно указывает на преимущество больших блоков.

Программа BigJPEG, обобщающая алгоритм JPEG на блоки большего размера, была написана нами на основе общедоступной реализации JPEG от группы разработчиков Independent JPEG. Фактически программа BigJPEG поддерживает работу с блоками от 8×8 до 128×128 .

Тестирование кодеков BigJPEG и JPEG производилось на 24 разнородных чёрно-белых изображениях из базы KODAK, где каждое изображение

представлено в четырёх размерах: 0.2, 0.4, 1.6 и 6.3 мегапикселя. Степень сжатия оценивалась по отношению размера исходного файла в байтах к размеру полученного файла. Качество полученного изображения оценивалось по критерию VIF; ранее было показано, что этот критерий является лучшим из современных. Полученные результаты сжатия и качества подвергались статистической обработке с вычислением среднего выигрыша в сжатии изображений кодеком BigJPEG в сравнении с базовым JPEG. Дополнительно вычислялось стандартное отклонение выигрыша по сжатию и доверительный интервал среднего выигрыша.

Результаты некоторых экспериментов показаны на Рис. 4. На нём изображены 2 однотипных графика для блоков 16×16 слева и 32×32 справа. По горизонтали отложено качество по критерию VIF. Вертикальными пунктирами отделены границы условных субъективных оценок качества («отличное» [0.75; 1.00], «хорошее» [0.50; 0.75), «удовлетворительное» [0.25; 0.50) и «плохое» [0.00; 0.25)). Качество ниже 0.2 не показано, так как не представляет интереса для практики. По вертикали отложен процент выигрыша в сжатии при использовании блоков указанного размера в сравнении с блоками 8×8 .

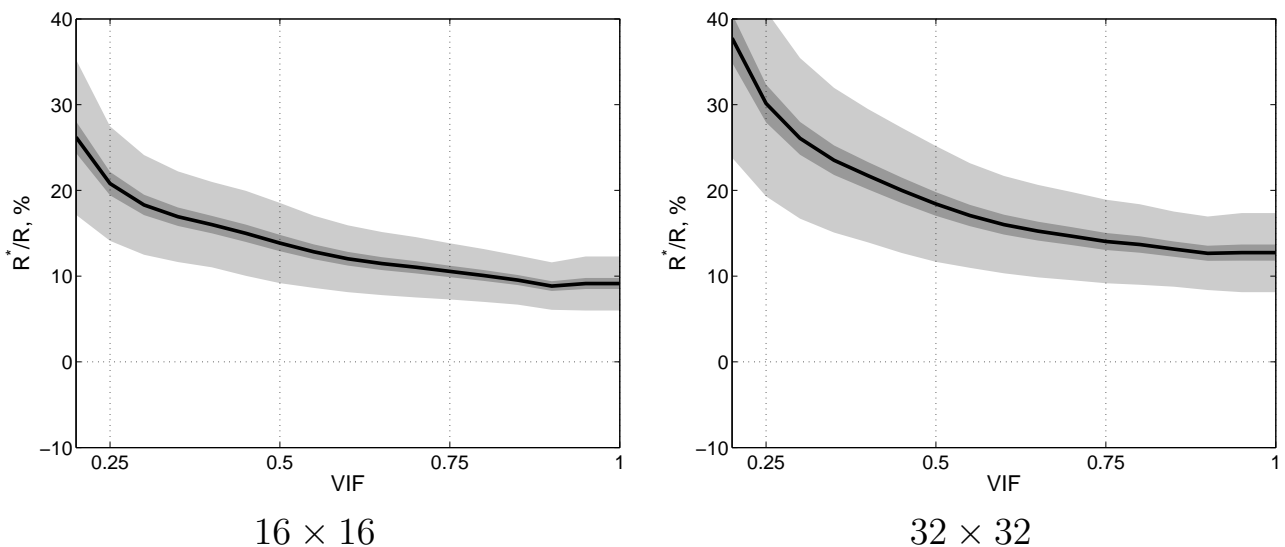


Рисунок 4 — Средний выигрыш в сжатии изображений объёмом 6.3 Мп (сплошная линия), стандартное отклонение (светло-серая область) и доверительный интервал среднего выигрыша (тёмно-серая область).

Важнейшие результаты расчётов представлены на Рис. 5. На нём показан средний выигрыш в сжатии для больших блоков в зависимости от качества сжатого изображения. Семейство сплошных линий на графике соответствует блокам 16×16 , пунктирных — блокам 32×32 (в обоих случаях выигрыш даётся по отношению к блокам 8×8). Каждая линия семейства соответствует определённому объёму исходного изображения; этот объём указан в мегапикселях около каждой кривой. Кривые показывают выигрыш в сжатии, усреднённый по 24-м изображениям.

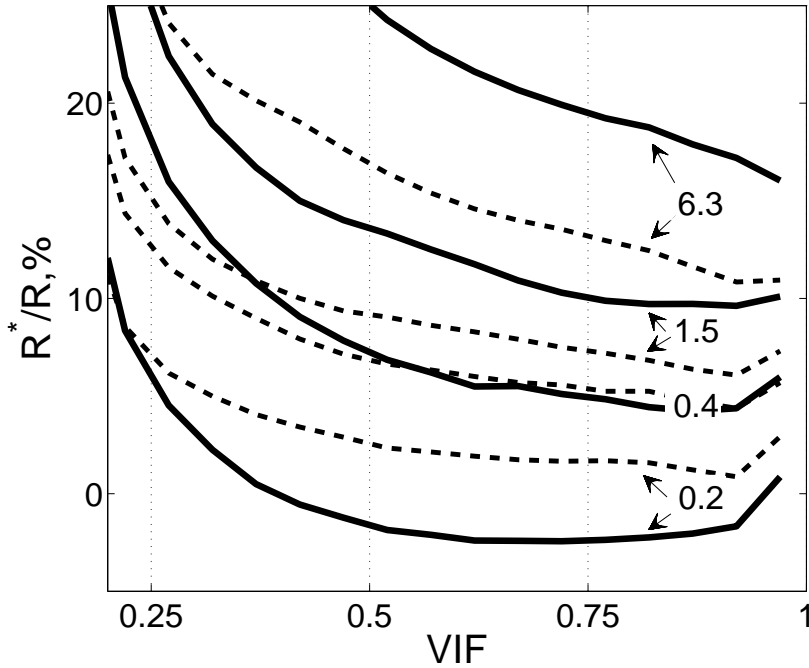


Рисунок 5 — Средний выигрыш в сжатии по 24 изображениям для блоков 16×16 (пунктирная линия) и 32×32 (сплошная линия). Числовые значения около кривых – объём изображения в мегапикселях

Примеры сжатых изображений не приводятся, так как типографская печать в масштабе автореферата даёт гораздо большие искажения, чем рассматриваемые эффекты.

Анализ полученных результатов подтверждает предположение об эффективности перехода на блоки большего размера. Поэтому можно дать рекомендации по выбору размера блока в зависимости от объёма исходного изображения. Они приведены в Табл. 4. Видно, что с учётом перспективного роста объёмов ПЗС матриц, всё большую практическую ценность будут приобретать варианты программы BigJPEG с блоками размерностью до 128×128 .

Таблица 4 — Рекомендуемые размеры блоков в BigJPEG

Объём изображения, Мп	< 0.2	0.2 — 1	1 — 5	5 — 25	> 25
Размер блока	8×8	16×16	32×32	64×64	128×128

JPEG-IT. Нам удалось улучшить характеристики базового кодека JPEG для блоков 8×8 пикселей, использующего встроенные таблицы Хаффмана. Была предложена особая схема кодирования первых элементов $\tilde{y}_{i,j}$ каждого блока $b_{m,n}$, образующие начальный треугольник матрицы \tilde{Y} . Отсюда и предложено название модификации JPEG-IT – initial triangle. Такое изменение позволяет экономить до нескольких бит на блок.

Реализация программы JPEG-IT делалась на основе общедоступной библиотеки, аналогично BigJPEG. Тестирование также производилось на 24 изображения из базы KODAK. Полученные результаты сжатия и качества подвергались статистической обработке аналогично экспериментам с кодеком BigJPEG.

Точные количественные значения среднего выигрыша и стандартного отклонения приведены в Табл. 5. Каждая строка в таблице соответствует правой границе каждого субъективного диапазона качества. Видно, что в области «отличного» качества модификация JPEG-IT даёт наибольший выигрыш при работе с небольшими изображениями (до 0.5 мегапикселей). На границе «хорошего» и «отличного» качества на отдельных изображениях степень сжатия может даже уменьшаться на 0.5%, но в среднем по большому числу изображений проигрыша нет. В области «удовлетворительного» и «плохого» качества наилучший выигрыш в сжатии (до 20 %) достигается на изображениях размером более 1.0 мегапикселя.

Таблица 5 — Средний выигрыш в сжатии и стандартное отклонение для модификации JPEG-IT относительно baseline JPEG (%)

Качество VIF	512×384 (0.2 Мп)	768×512 (0.4 Мп)	1536×1024 (1.6 Мп)	3072×2048 (6.3 Мп)
0.99	3.2 ± 1.2	2.8 ± 1.1	1.9 ± 1.1	1.3 ± 1.0
0.75	0.6 ± 0.5	0.2 ± 0.3	0.0 ± 0.6	0.1 ± 0.8
0.50	1.2 ± 1.6	2.5 ± 2.7	2.5 ± 2.3	3.1 ± 2.5
0.25	8.1 ± 4.9	11.4 ± 6.7	13.5 ± 6.2	17.9 ± 6.3

Анализ результатов позволяет сделать следующий вывод. Предлагаемая модификация JPEG-IT даёт положительный эффект для большинства изображений (от 84%) в области «удовлетворительного» и «плохого» качества ($VIF < 0.50$). В области «хорошего» и частично «отличного» качества ($0.50 < VIF < 0.85$) средний выигрыш незначителен. Возможны колебания отношения степеней сжатия базового кодека JPEG и JPEG-IT $\pm 2\%$ на отдельных изображениях. В оставшемся диапазоне качества имеется незначительный выигрыш в сжатии, сопоставимый с величиной стандартного отклонения; для 84% изображений имеется положительный эффект в сжатии (от 0 до 4%).

Глава 4 вспомогательная. В ней рассматриваются различные способы аппроксимации непрерывных функций рядами Фурье. Как известно, разложение гладких непериодических функций в ряд Фурье может приводить к эффекту Гиббса, существенно снижающему точность вблизи границ отрезка. В 1907 г. академик А.Н. Крылов предложил вычитать из заданной функции $u(x)$ другую функцию, подобранную так, чтобы периодическое продолжение имело достаточно много непрерывных первых производных.

Этот вопрос активно обсуждается в научной литературе. Однако все эти способы требуют использования производных $u(x)$ до p -го порядка на границах и приводят к сравнительно сложным алгоритмам. Для функции двух переменных $u(x, y)$ сложность алгоритмов становится неприемлемой. Поэтому в прикладных расчётах эти способы не получили распространения.

Широкое развитие получила другая идея. В этом случае отрезок задания функции надо считать не периодом, а половиной периода. Чётное продолжение (без всяких вычитаний!) на вторую половину периода даёт непрерывную функцию, разлагающуюся в ряд только по косинусам. Этот метод хорошо известен в теории обработки изображений и на нём основан широко распространённый кодек JPEG. Он удобен при работе с функциями двух и более переменных.

Одномерный случай. Простейшее обобщение этого метода заключается в вычитании линейной функции, нечётном продолжении разности на вторую половину периода и разложении по синусам. Вычитаемая линейная функция должна проходить через граничные точки исходной функции. Поэтому для восстановления исходной функции требуется дополнительно хранить 2 значения с большой разрядностью. Для гладких функций этот способ существенно повышает точность по сравнению с косинус-разложением, поскольку, чем больше непрерывных производных $u^{(p)}(x, y)$ имеет функция, тем быстрее убывают коэффициенты Фурье и быстрее сходится ряд Фурье. Поэтому, в зависимости от p , точность аппроксимации частичной суммой из N членов ряда Фурье будет также заметно отличаться.

На Рис. 6 показаны результаты аппроксимации тестовой функции, из которой предварительно вычитались многочлены степени p и делалось чётное или нечётное продолжение для обеспечения непрерывности p первых производных новой функции. Этот способ хорошо обобщается на функции двух переменных.

Однако, дальнейшее повышение гладкости усложняет алгоритм. В данной работе предложен наиболее эффективный способ повышения гладкости для функции одной переменной. Показано, что для функции двух переменных этот способ даёт алгоритм приемлемой сложности при повышении гладкости на единицу по сравнению с синус-разложением.

Идея метода заключается в том, что при вычитании из исходной p -раз непрерывно дифференцируемой функции $u(x)$ полинома порядка p нет необходимости хранить все значения $u(x)$ и её производных на концах отрезка задания. Вычитаемый полином строится так, чтобы обеспечить непрерывность полученной функции и её p первых производных. Так для обеспечения непрерывности $p - 1$ -ой производной достаточно приравнять нулю значения p -ой, а сами значения $p - 1$ -ой производной нигде не используются. Далее надо приравнять нулю концы $p - 2$ -ой производной, что

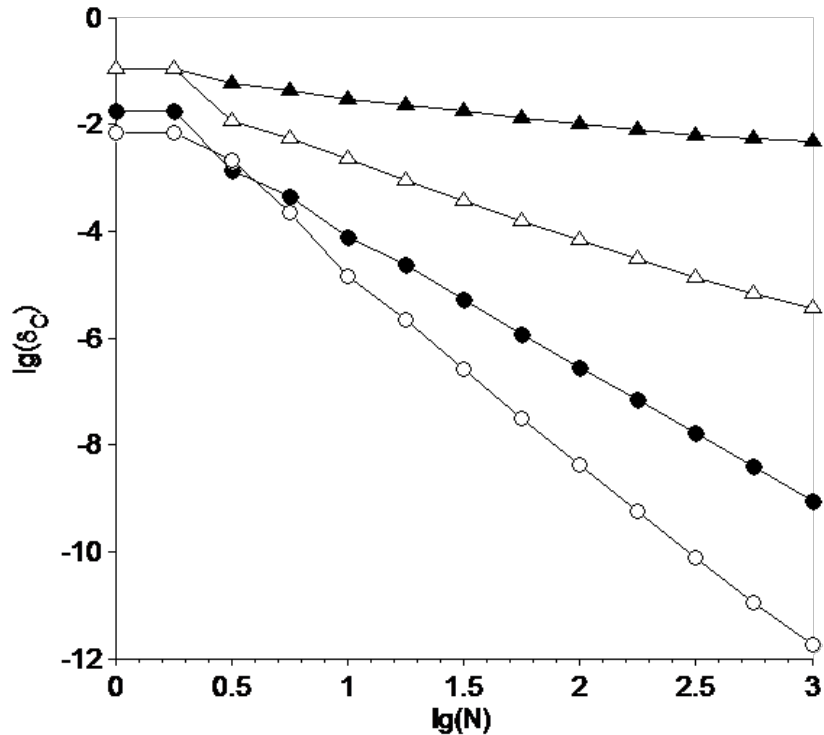


Рисунок 6 — Погрешность аппроксимации одномерной тестовой функции в норме L_2 от числа членов ряда Фурье N : ▲ — разрывное продолжение, △ — чётное продолжение, ● — нечётное продолжение, ○ — продолжение повышенной гладкости

автоматически обеспечивает непрерывность p — 3-ей производной. Так следует продолжать рассуждения до $p = 0$.

Двумерный случай. Отдельно был рассмотрен вопрос аппроксимации двумерной функции и исследован способ оптимального усечения коэффициентов. В одномерном случае ряд Фурье заменялся конечной суммой; оставлялись N членов ряда, а остальные отбрасывались. Для вычисления с требуемой точностью нужно было решить только один вопрос: как выбрать единственный параметр N . В двумерном случае приходится вычислять двойную сумму. Возникает вопрос, как целесообразно выбирать верхние пределы в этих суммах.

В простейшем случае в каждой сумме оставляют одинаковое число членов ряда N , то есть вместо бесконечной матрицы коэффициентов оставляют квадратную порядка N . Общее число учтённых членов при этом будет равно N^2 . Но это далеко не оптимальный способ. На практике используют усечение по треугольнику, которое позволяет отбросить половину коэффициентов без серьёзного ущерба для качества. Однако при больших N трудоёмкость всё равно остаётся чрезмерной.

Оптимальная кривая была найдена в данной работе. Это делалось методом неопределённых множителей Лагранжа и искомая кривая оказалась

гиперболой. При этом число членов двойной суммы будет порядка $N \ln(N)$, что при больших N много меньше, чем N^2 . По ней следует делать усечение частичных сумм при аппроксимации двумерных функций. Каждый способ усечения рассмотрен в экспериментах с тестовой двумерной функцией. Полученные результаты показаны на Рис. 7; они подтверждают, преимущество гиперболического усечения.

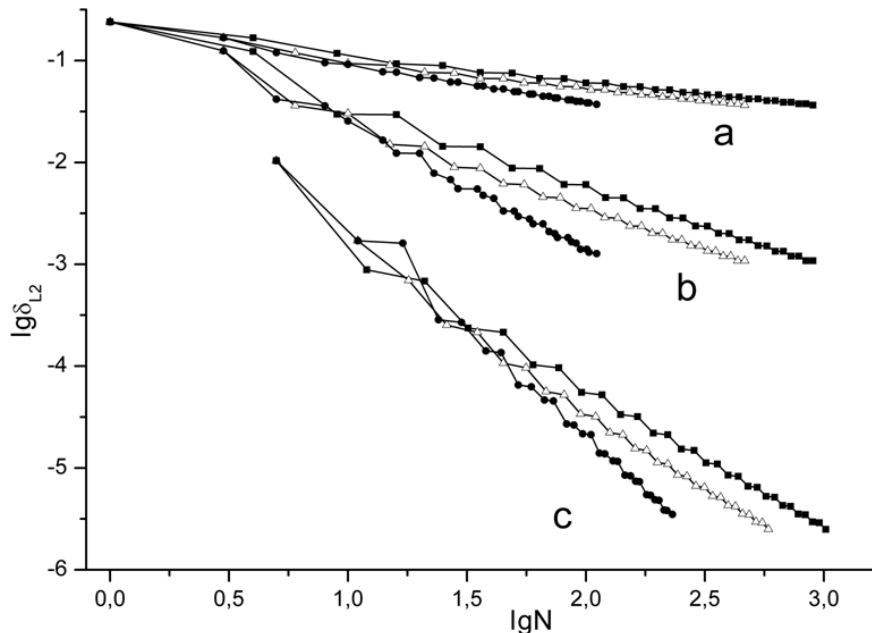


Рисунок 7 — Погрешность двумерной аппроксимации в норме L_2 : «а» — разрывное продолжение; «б» — чётное продолжение; «с» — нечётное продолжение; ■ — усечение по квадрату; △ — усечение по треугольнику; ● — усечение по гиперболе

Описанные методы эффективны для функции высокой гладкости. Однако при обработке изображений функции зачастую могут быть разрывными внутри блока. Кроме того, граничные производные в этом случае не известны. В этом случае данная теория не применима и следует ограничиться первым синус-разложением.

Основные результаты и выводы

1. Разработаны и программно реализованы две модификации кодека JPEG, повышающие степень сжатия от 9 до 20% на современной аппаратуре с ПЗС матрицей объёмом более 6 Мп.
2. На основе обработки большого числа изображений статистически достоверно показано, что чем больше объём изображения, тем большие блоки следует использовать в кодеке BigJPEG, а именно:
 - если объём изображения меньше 0.2 Мп, следует использовать блоки 8×8 ,

- если объём изображения от 0.2 до 1 Мп, следует использовать блоки 16×16 ,
 - если объём изображения от 1 до 5 Мп, следует использовать блоки 32×32 ,
 - если объём изображения от 5 до 25 Мп, следует использовать блоки 64×64 ,
 - если объём изображения больше 25 Мп, следует использовать блоки 128×128 .
3. Разработан общий вид матрицы квантования для блоков произвольного размера, которая включена в программные реализации предложенных модификаций JPEG.
 4. Предложен метод объединения нескольких баз тестовых изображений, построенных на основе неодинаковых систем экспертных оценок. Проведено объединение наиболее распространённых баз TID2008 (Украина, Харьков) и LIVE (США, Техас).
 5. Предложен метод оценки достоверности критериев качества, основанный на отношении стандарта критерия к стандарту экспертных оценок и дающий статистически достоверную оценку в отличие от используемых в настоящее время коэффициентов парной корреляции.
 6. Разработан критерий оценки качества искажённого изображения SGC, построенный с учётом психофизических особенностей человеческого восприятия на основе нормы Соболева и фильтра Гаусса. Критерий SGC имеет достоверность сопоставимую с лучшими современными критериями, но существенно превосходит их по простоте реализации и быстродействию.

Публикации автора по теме диссертации:

1. Голованов Р.В., Калиткин Н.Н. Статистическая недостоверность пространственных критериев оценки качества искажённого изображения // Цифровая обработка сигналов. — 2013. — № 3. — С. 74–79.
2. Голованов Р.В., Калиткин Н.Н. Критерий сглаженных градиентов для оценки качества искаженного изображения // Доклады Российской академии наук. — 2013. — 8. — Т. 451, № 4. — С. 385–388.
3. Голованов Р.В., Калиткин Н.Н., Луцкий К.И. Нечётное продолжение для фурье-аппроксимации непериодических функций // Матем. моделирование. — 2013. — Т. 25, № 5. — С. 67–84.
4. Голованов Р.В., Калиткин Н.Н. Продолжения повышенной гладкости для фурье-аппроксимации непериодической функции // Доклады Российской академии наук. — 2012. — 9. — Т. 446, № 1. — С. 25–29.

5. Голованов Р.В., Калиткин Н.Н. Улучшение сходимости для аппроксимации непериодических функций рядами Фурье // Матем. моделирование. — 2014. — Т. 26, № 2. — С. 108–118.
6. Калиткин Н.Н., Голованов Р.В. Проблема сравнения критериев оценки качества искажённого изображения // Труды конференции DSPA-2013. — Москва: 2013. — 3. — С. 31–34.
7. Голованов Р.В. Ускорение сходимости фурье-аппроксимации для непериодических функций // Труды конференции МЭИнфо-2012. — Зеленоград: 2012. — 4. — С. 110.
8. Голованов Р.В. Современные методы оценки качества искажённого изображения. Базы тестовых изображений // Труды конференции МЭИнфо-2013. — Зеленоград: 2013. — 4. — С. 135.
9. Голованов Р.В. Повышение сжатия в алгоритме JPEG // Труды конференции МЭИнфо-2014. — Зеленоград: 2014. — 4. — С. 107.
10. Голованов Р.В., Калиткин Н.Н. Программный комплекс SGC для оценки качества искажённого изображения на основе сглаженных градиентов // Свидетельство о государственной регистрации программы для ЭВМ № 2014616696. — 02.07.14.
11. Голованов Р.В. Программный комплекс JPEG-IT для компрессии-декомпрессии изображений с особым кодированием начального треугольника // Свидетельство о государственной регистрации программы для ЭВМ № 2014616697. — 02.07.14.