

Федеральное агентство связи
Государственное образовательное учреждение высшего профессионального
образования
Московский технический университет связи и информатики

Кафедра телевидения имени С.И. Катаева

Лабораторная работа № 64

ИССЛЕДОВАНИЕ СПЕКТРАЛЬНЫХ ПРЕОБРАЗОВАНИЙ ИЗОБРАЖЕНИЙ

Москва 2011

План УМД на 2010/2011 уч. г.

Лабораторная работа № 64

ИССЛЕДОВАНИЕ СПЕКТРАЛЬНЫХ ПРЕОБРАЗОВАНИЙ ИЗОБРАЖЕНИЙ

Составители: к.т.н., Власюк И.В.,
Поташников А.М.

Издание утверждено советом факультета Р и Т.
Протокол № 9 от 20 мая 2011г.

Рецензент: Безруков В.Н., д.т.н., профессор

Цель работы

Исследовать влияние параметров двумерных спектральных преобразований и квантования их коэффициентов на характеристики изображений, изучить и сравнить объективные и субъективные методики оценки качества изображений.

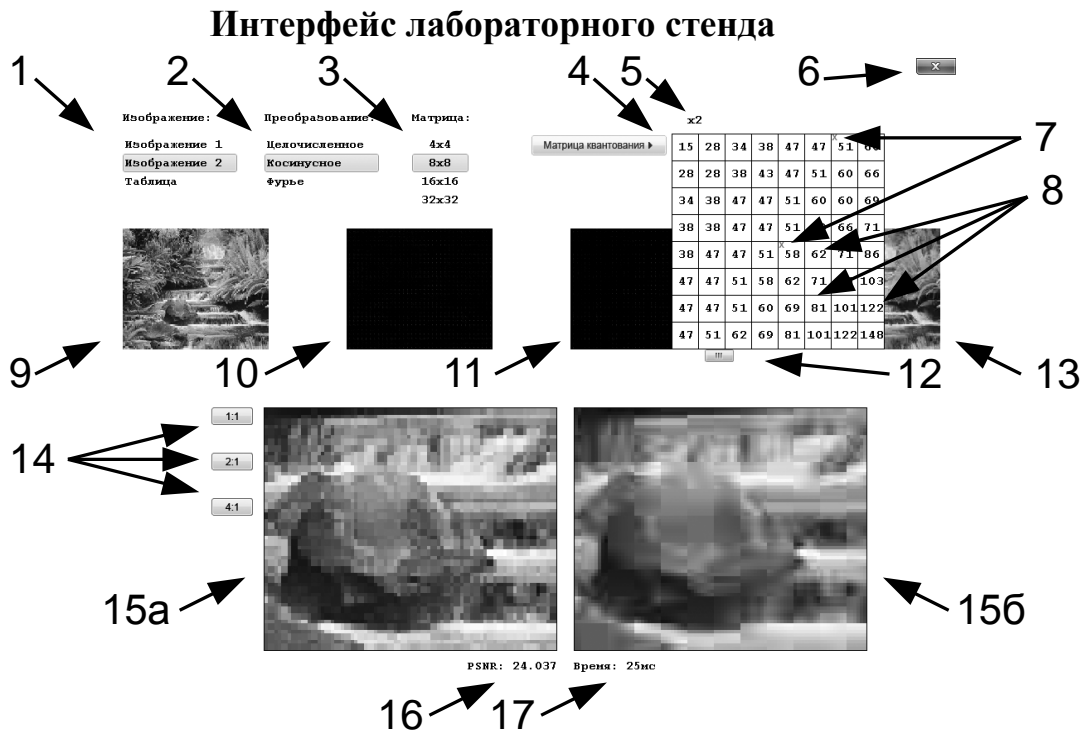


Рисунок 1. Интерфейс лабораторного стенда.

1) Список изображений, 2) список преобразований, 3) список размеров блоков, доступных для выбранного преобразования, 4) кнопка управления матрицей коэффициентов квантования, 5) Коэффициент умножения стандартной матрицы квантования, 6) Выход из программы, 7) указатель на то, что соответствующая спектральная составляющая блока не используются, 8) значения коэффициентов квантования, 9) эскиз исходного изображения, 10) Представление коэффициентов преобразования исходного изображения, яркость пропорциональна значениям отсчетов в спектральной области (чем ярче, тем больше значение), 11) Представление преобразованных коэффициентов, 12) управление коэффициентом масштабирования матрицы квантования, 13) эскиз восстановленного изображения, 14) Выбор масштаба просматриваемых изображений, 15) окна для просмотра увеличенных изображений. Для просмотра требуется перенести эскиз на соответствующее окно. Если изображение масштабировано, его можно двигать с помощью указателя мыши при нажатой левой кнопке, 16) Значение метрики PSNR (пиковое отношение сигнал-шум), 17) время спектрального преобразования.

Теоретическая часть

Компрессией телевизионного сигнала называется устранение из него избыточной информации для уменьшения скорости цифрового потока и, как следствие, более эффективного использования каналов передачи.

Избыточность информации в потоке цифровых данных объясняется, прежде всего, спецификой ТВ изображения. Известно, что, несмотря на равную вероятность любых из возможных значений яркости (цветности) для одного элемента изображения, содержание соседствующих с ним элементов мало отличается или не отличается вовсе. При поэлементной передаче яркости и цветности методом ИКМ в канал таким образом посылаются одна и та же или мало отличающаяся по содержанию информация. Подобную избыточность информации в сигнале называют пространственной или внутрикадровой.

Значительная избыточность содержится и в передаче смежных во времени кадров изображения, в которых, несмотря на движение отдельных фрагментов, существенную роль играют общие для них неподвижный фон или задний план. Такого рода избыточность называют временной или межкадровой. И пространственная, и временная избыточность обусловлены статистическими свойствами телевизионного изображения.

Следует также учитывать, что цифровой ИКМ сигнал, являющийся продуктом преобразования аналогового сигнала, содержит и определенную психофизическую избыточность информации. В этой информации содержатся данные о таких особенностях воспроизводимого изображения, которые визуально не фиксируются и, следовательно, такая информация без ущерба качеству изображения может быть изъята из передачи.

Устранение избыточности информации в цифровом сигнале, в конечном счете, должно выразиться в уменьшении числа отсчетов сигнала и (или) уменьшении разрядности их двоичных кодовых символов. Реализация таких процедур напрямую с ИКМ сигналом неэффективна, так как изъятие из сигнала отдельных отсчетов или замена их значений на более грубо квантованные соответствует изъятию в воспроизводимом изображении элементов или их искаженному представлению. Попытка изъятия отдельных элементов из группы равноправных отсчетов изображения как с информационной, так и с энергетической точки зрения приводит к необратимой потере качества изображения.

Решение задачи по устранению избыточности информации в цифровом сигнале возможно только путем его предварительной обработки, которая должна перераспределить вклад отдельных отсчетов преобразованного сигнала в общее содержание изображения. Перераспределив таким образом функциональную значимость между отдельными отсчетами, можно будет выделить «главные» отсчеты, несущие основной объем информации, обеспечив им наилучшие условия передачи, а на остальных отсчетах «сэкономить», не передавая их или передавая с минимальным числом градаций.

К настоящему времени предложено сравнительно много методов обработки сигнала с таким подходом. В работе будут рассмотрены самые распространенные приемы, и в частности рекомендованные стандартами компрессии MPEG.

Методы преобразования изображения в общем случае основаны на том, что его цифровой эквивалент (сигнал ИКМ) приводится к виду, удобному для сокращения избыточной информации. В этом отношении наиболее эффективным является

преобразование видеоинформации из временной области в спектральную. Это преобразование, как правило, предваряется разбивкой изображения на частичные подобласти, фрагменты (в терминологии MPEG – блоки), которые затем по отдельности подвергаются необходимой обработке. Результат преобразования представляет собой совокупность спектральных коэффициентов, которые характеризуют амплитуды и фазы пространственных частот изображения.

В основу преобразования изображений могут быть положены различные приемы. Наиболее часто используются методы линейных ортогональных преобразований. Линейность преобразований означает, что операции сложения, вычитания и умножения на скаляр действительны и после преобразований, а ортогональность – что преобразуемый фрагмент представляется ограниченным набором ортогональных функций.

Из различных ортогональных преобразований, позволяющих эффективно выявлять избыточную информацию, стандартом MPEG рекомендовано использовать дискретно-косинусное преобразование (ДКП), являющееся частным случаем двумерного преобразования Фурье. В последних стандартах предлагается целочисленное преобразование. Оно отличается от ДКП использованием функций преобразования, состоящих из целых значений и, как следствие, увеличением скорости обработки блоков.

Количество операций для нахождения этих коэффициентов весьма значительно, поэтому преобразования осуществляются над небольшими по размеру фрагментами, обычно 8x8 элементов. Дискретно-косинусное преобразование Фурье в определенной степени минимизирует объем этих вычислений использованием в качестве набора преобразующих (базисных) функций только косинусных составляющих. Аналитически двумерное дискретно-косинусное преобразование описывается следующим образом:

$$F_{(m,n)} = \frac{2}{N} \cdot C_{(m)} C_{(n)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f_{(x,y)} \cos \frac{\pi m(2x+1)}{2N-1} \cos \frac{\pi n(2y+1)}{2N-1}; \quad (1)$$

где $C_{(m)} = 1$ при $m > 0$; $C_{(m)} = \frac{1}{\sqrt{2}}$ при $m = 0$; $C_{(n)} = 1$ при $n > 0$; $C_{(n)} = \frac{1}{\sqrt{2}}$ при $n = 0$; $f_{(x,y)}$

– отсчеты изображения с пространственными координатами x, y (от 0 до $N - 1$); N – размер блока изображения ($N \times N$ элементов); $F_{(m,n)}$ — коэффициенты, характеризующие изображение в спектральной плоскости со спектральными координатами m, n (от 0 до $N - 1$).

ДКП является обратимым: по распределению $F_{(m,n)}$ обратным преобразованием однозначно восстанавливается $f_{(x,y)}$.

Поскольку число отсчетов преобразуемого сигнала равно числу отсчетов преобразованного сигнала, устранения избыточности информации в результате такого преобразования не происходит. Однако, следует обратить внимание на значительное изменение содержания блока коэффициентов ДКП по отношению к блоку преобразуемого изображения. По физическому смыслу блок коэффициентов ДКП представляет собой совокупность значений амплитуд пространственных косинусоидальных гармоник с частотами m и n . При этом значение $f_{(0,0)}$ пропорционально среднему уровню (постоянной составляющей) в блоке и может достигать значения 2040, как

следует из (1), при 256 уровнях квантования для блока 8x8 элементов. Количество разрядов, необходимое для точного представления коэффициентов преобразования, также ограничивает допустимые размеры блока.

59	9	59	60	60	65	64	64
63	62	62	62	62	61	61	62
137	123	111	101	96	89	88	86
237	236	235	233	231	237	238	239
225	229	232	232	231	237	238	239
193	195	197	198	199	204	204	205
182	82	181	181	181	180	180	180
183	182	181	180	179	178	178	177

а)



в)

1255,6	8,1	8 0	1,4	-1,4	1,1	1,4	-1,1
-378,5	10,9	4,8	2,3	1,1	0,5	0,6	0,0
-330,1	0,2	-4,5	1,2	1,8	-0,6	-1,0	0,6
23,7	-35,8	-8,3	-3,2	-3,4	-0,8	1,0	-2,3
163,6	-18,6	-2,5	-2,3	-3,4	0,6	0,9	,9
49,0	6,9	3,8	2,1	0,2	0,3	0,0	-0,3
-44,1	17,0	5,5	3,3	0,7	0,6	0,3	0,6
-55,5	18,4	2,9	1,4	1,2	0,2	-0,7	1,0

б)

Рисунок 2. Пример ДКП для некоторого произвольного изображения.

а) - блок значений яркостей пикселей изображения; б) - блок коэффициентов ДКП. в) Изображение фрагмента

Компоненты $F_{(0,i)}$ и $F_{(i,0)}$ характеризуют изменение яркости в блоке вдоль строки и поперек строк соответственно. Разночастотные изменения яркостей пикселей с диагональными структурами характеризуются диагональными спектральными компонентами $F_{(1,1)}$, $F_{(2,2)}$, $F_{(3,3)}$, ...

В современных стандартах сжатия видеoinформации (H.264/AVC) применяются преобразования, которые используют целочисленные коэффициенты. Такие преобразования осуществляются быстрее. Поскольку из спектральных преобразований для систем сжатия видеoinформации ДКП обеспечивает наилучшие характеристики таких параметров, как отношение Сигнал/Шум, то целочисленные преобразования должны быть наиболее близки к этому типу. Косинусное преобразование можно записать в следующей форме:

$$\begin{aligned}
 Y = AXA^T &= \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix} X \begin{bmatrix} a & b & a & c \\ a & c & -a & -b \\ a & -c & -a & b \\ a & -b & a & -c \end{bmatrix} = \\
 &= (CXC^T) \otimes E = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & d & -d & -1 \\ 1 & -1 & -1 & 1 \\ d & -1 & 1 & -d \end{bmatrix} X \begin{bmatrix} 1 & 1 & 1 & d \\ 1 & d & -1 & -1 \\ 1 & -d & -1 & 1 \\ 1 & -1 & 1 & -d \end{bmatrix} \otimes \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix}. \quad (2)
 \end{aligned}$$

где X – пространственное представление, Y – спектральное представление; A – матрица косинусного преобразования, C – матрица преобразования без учета масштабирования, E – матрица масштабирования $a = \frac{1}{2}$; $b = \sqrt{\frac{1}{2}} \cos\left(\frac{\pi}{8}\right)$; $c = \sqrt{\frac{1}{2}} \cos\left(\frac{3 \cdot \pi}{8}\right)$; $d = \frac{c}{b}$.

Для упрощения реализации преобразования d считают равным 0.5, для ортогональности необходимо b приравнять к $\sqrt{\frac{2}{5}}$. Увеличение второй и четвертой строки матрицы C в два раза уменьшает количество умножение в два раза. Для компенсации вторую и четвертую строку, а также второй и четвертый столбец матрицы E необходимо уменьшить в два раза. Результирующее преобразование имеет следующий вид:

$$Y = (C_f X C_f^T) \otimes E = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \\ a^2 & \frac{ab}{2} & a^2 & \frac{ab}{2} \\ \frac{ab}{2} & \frac{b^2}{4} & \frac{ab}{2} & \frac{b^2}{4} \end{bmatrix}. \quad (3)$$

При этом обратное преобразование изменяется следующим образом:

$$X = C_i^T (Y \otimes E_i) C_i = \begin{bmatrix} 1 & 1 & 1 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -1 \\ 1 & -\frac{1}{2} & -1 & 1 \\ 1 & -1 & 1 & -\frac{1}{2} \end{bmatrix} Y \otimes \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \frac{1}{2} & -\frac{1}{2} & -1 \\ 1 & -1 & -1 & 1 \\ \frac{1}{2} & -1 & 1 & -\frac{1}{2} \end{bmatrix}. \quad (4)$$

Учитывая, что умножения и деления на степени 2 осуществляются в двоичной системе операциями сдвига, при таком преобразовании значительно уменьшается количество операций умножения, и операций с дробными числами.

В случае, когда необходима большая точность при малых вычислительных затратах, пользуются преобразованием Адамара. В отличие от ДКП и Целочисленного преобразования, оно производится по функциям, содержащим только значения 1 и -1 согласно (5). Как видно, функции этого преобразования можно получить из косинусных функций за счет бесконечного усиления и ограничения по -1 и 1. За счет этого, преобразование Адамара не содержит операций умножения, а остальные операции производятся над целыми числами. Таким образом, оно большинством процессоров выполняется быстрее других видов преобразований. Кроме того, только при использовании этого преобразования можно потенциально избежать ошибок округления. Недостатком преобразования является то, что при квантовании коэффициентов резко понижается качество сжимаемого изображения.

$$Y = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} X \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad (5)$$

где X – блок пространственных отсчетов, а Y – блок преобразованных отсчетов.

Обратное преобразование выполняется по той же формуле при замене X на Y , а Y на X .

Обычно для большинства блоков изображения лишь малая часть коэффициентов имеет значительную величину. Это объясняется небольшими размерами блока, внутри которого яркость меняется мало, и, поэтому, относительно большие величины имеют только постоянная составляющая и несколько низкочастотных компонентов, расположенных в левом верхнем углу матрицы коэффициентов спектрального представления.

Мелкие детали изображения имеют спектральное представление, богатое высокими пространственными частотами и коэффициенты, характеризующие их амплитуды, располагаются правее и ниже. Поскольку, как правило, мелкие детали изображения выражены энергетически слабо относительно среднего уровня, то и соответствующие им коэффициенты спектрального представления по сравнению с $F_{(0,0)}$ малы или вообще равны нулю. Таким образом, если передавать вместо значений яркости изображения коэффициенты спектрального представления, то сокращение скорости передачи данных может быть достигнуто уже хотя бы за счет исключения нулевых коэффициентов. Однако эта задача решается уже вне процедуры преобразования.

Выгодное для компрессии различие в амплитудах между отдельными коэффициентами спектрального представления может быть еще более усилено за счет устранения психофизической избыточности в изображении. Как следствие, увеличится число нулевых коэффициентов и коэффициентов с малыми значениями. Эта задача решается в процессе квантования коэффициентов, полученных после преобразования.

Установлено, что глаз более чувствителен к ошибкам передачи яркости и цветности на больших площадях, в то время как при передаче контуров и мелких деталей остаются незамеченными более серьезные ошибки. Отсюда вытекает возможность определенного огрубления значений коэффициентов представления, отвечающих за передачу мелких деталей и контуров, без возникновения заметных для глаза искажений в изображении.

С этой целью производят процедуру квантования коэффициентов спектрального представления блока на разное число уровней. Коэффициенты, расположенные в левом верхнем углу блока, квантуются на максимально большое число уровней (особенно это касается коэффициента $F_{(0,0)}$, отвечающего за среднюю яркость блока); остальные коэффициенты передаются с меньшей точностью, а значит, квантуются на меньшее число уровней. Для тех же из них, что располагаются в правом нижнем углу, шкала квантования может содержать всего несколько уровней. Практическая реализация процесса квантования достигается поэлементным делением матрицы коэффициентов преобразования на матрицу квантования. В приемном устройстве, прежде чем осуществить обратное дискретно-косинусное преобразование для восстановления исходного изображения, матрица коэффициентов спектрального представления умножается на матрицу квантования. Эта операция называется деквантованием. Очевидно, что после деквантования возвратиться к исходному, неквантованному блоку коэффициентов преобразования уже нельзя. Ошибки, возникающие от округления квантуемых величин, и связанные с ними искажения в изо-

бражении необратимы. Отсюда вытекает необходимость отыскания таких матриц квантования, которые не приводили бы к визуально заметным искажениям. Учитывая, что значения большинства коэффициентов спектрального представления в блоке весьма малы, деление их на числа, характеризуемые почти двумя порядками, приводит или к обнулению многих коэффициентов, или к сильному их уменьшению. Это в свою очередь позволит при передаче проквантованных значений коэффициентов спектрального представления по каналу связи значительно уменьшить скорость цифрового потока.

Следующим шагом после квантования коэффициентов спектрального представления является преобразование матрицы этих коэффициентов в одномерную последовательность. Именно здесь окончательно реализуется процесс устранения избыточности, подготовка к которому проводилась на рассмотренных выше этапах преобразования и взвешенного квантования. Данное преобразование предусматривает объединение коэффициентов матрицы в определенные группы и применение затем так называемого энтропийного кодирования.

Алгоритм группирования (упорядочивания) коэффициентов преобразования существенно влияет на эффективность компрессии. Он заключается в том, что в процессе сканирования преобразуемой во временную последовательность чисел матрицы нулевые коэффициенты объединяются в максимально длинные серии. Тогда их описание может сводиться к лаконичной записи длины серии, а так же ее местоположения в матрице. Одним из вариантов такого алгоритма группирования является зигзагообразное сканирование, при котором преобразование начинается с левого верхнего угла матрицы и заканчивается в ее правом нижнем углу. Поскольку именно в правом нижнем углу сосредоточено большинство нулевых коэффициентов, такой порядок сканирования обеспечивает формирование наиболее длинных серий нулей, а, следовательно, и самую компактную форму их передачи.

Полученная в результате сканирования последовательность чисел подвергается упомянутому выше энтропийному кодированию.

Домашнее задание

Выполните двумерное преобразование Адамара для блока 4x4, отсчеты которого рассчитываются по следующей формуле:

$$X_{(i,j)} = 100 + 10a - 3b + i(a - 5) - 2j(b - 3) + j^2 - 2i(i - 2); \quad (6)$$

где i и j – индексы строк и столбцов (первая строка и первый столбец имеют индексы, равные нулю); a и b – предпоследняя и последняя цифра студенческого билета.

Порядок выполнения работы

1. Запустите программу из директории "LabFourier". Если такой файл не найден, то обратитесь за помощью к преподавателю.
2. Выберите размер блоков 4x4.
3. Откройте матрицу квантования нажатием кнопки "Матрица квантования".

4. Установите коэффициенты квантования в минимальное положение, используя движок под матрицей для получения преобразования без искажения.
5. Поместите исходное и конечное изображения в поля для просмотра.
6. Запишите в таблицу PSNR и оцените конечное изображение по пятибалльной шкале.
7. Исключите коэффициенты наибольших из используемых частот (максимальная сумма индексов матрицы).
8. Повторяйте действия 6-7 до тех пор, пока не останется только значение $F_{(0,0)}$, например, для преобразования размером 4x4 на первом шаге следует удалить коэффициент $F_{(3,3)}$, на втором - $F_{(3,2)}$ и $F_{(2,3)}$, на третьем - $F_{(3,1)}$, $F_{(2,2)}$ и $F_{(1,3)}$ и т.д.
9. Повторите действия 3-8 для всех преобразований для двух изображений и таблицы.
10. Для одного любого изображения запишите в таблицу PSNR и оцените изображение по пятибалльной шкале для всех преобразований для всех размеров блоков для 5 различных положений множителя квантования.
11. Для косинусного преобразования и для преобразования Фурье запишите отношение сигнал/шум и оцените конечное изображение для размеров преобразования 4x4, 8x8, 16x16, 32x32 при наличии и отсутствии составляющей $F_{(0,0)}$ (постоянная составляющая).
12. Уберите все вертикальные коэффициенты, кроме постоянных составляющих, для всех преобразований 4x4 ($F_{(x,y)}$, $y \neq 0$) и зарисуйте три квадрата таблицы – 4-й, 5-й, 6-й, считая серые квадраты сверху, во втором столбце.
13. Используя любое изображение, для всех вариантов преобразований и размеров блоков проведите по 10 испытаний, запишите время преобразования в таблицу.

Обработка результатов

1. Зарисуйте графики зависимости PSNR и субъективной оценки от количества отсчетов преобразованных блоков.
2. Зарисуйте графики зависимости PSNR и субъективной оценки от множителя квантования.
3. Рассчитайте среднее время выполнения преобразований и нанесите на один график зависимости времени преобразования от размера блока для всех преобразований (для преобразований Адамара и целочисленного – одна точка).

Содержание отчета

1. Название и цель работы.
2. Домашнее задание
3. Таблица результатов экспериментов.
4. График зависимости пикового отношения сигнал/шум от количества используемых коэффициентов.
5. Зависимость субъективного восприятия от количества используемых коэффициентов.
6. Зависимость времени преобразования от типа и размера блоков.

7. Требуемые рисунки с пояснениями.
8. Выводы.

Контрольные вопросы

1. Для чего применяются спектральные преобразования в обработке изображений?
2. Сравните ДКП, целочисленное преобразование и преобразование Адамара применительно к видеокомпрессии.
3. По каким соображениям преобразование применяется не ко всему изображению, а к его блокам, назовите основные критерии выбор размера блоков.
4. Покажите, как получить выражение для целочисленного преобразования из ДКП.
5. Перечислите основные виды искажений изображений, возникающие при обработке изображений со спектральным преобразованием, укажите причины их возникновения и способы коррекции.
6. Как проводится субъективная оценка качества изображений?
7. Напишите выражение для вычисления метрики PSNR, покажите, по каким причинам результаты оценки качества по этой метрике часто непропорциональны результатам субъективной оценки.

Литература

1. Д.Сэлмон. Сжатие данных, изображений и звука. – М.: Техносфера, 2004. – 368 с.
2. Ричардсон Я. H.264 и MPEG-4 – стандарты нового поколения. – М.: Техносфера, 2005. – 368 с.
3. Recommendation ITU-R BT.500-11, Methodology for subjective assessment of the quality of television pictures. ITU-T. 2002.