

Санкт-Петербургский политехнический университет Петра Великого
Институт компьютерных наук и технологий

Неделя науки 2016

Сжатие изображений при помощи рядов Фурье

Выполнили:

Соболева Александра (гр. 33506/1)

Рязановский Денис (гр. 33506/1)

Научный руководитель:

доцент каф. «Высшая математика»

Филимоненкова Надежда Викторовна

Задачи и цели проекта

- Изучить теорию рядов Фурье и её приложение к сжатию изображений.
- Изучить работу алгоритма сжатия изображений, используемого в технологии JPEG.
- Смоделировать этап сжатия изображения, в котором используются дискретные аналоги рядов Фурье.
- Проанализировать полученные результаты сжатия конкретных изображений.

Жан Батист Жозеф Фурье

(21.03.1768-16.05.1830)

Известный французский математик и физик XIX века. Одно из его достижений – написанные монографии «Аналитическая теория тепла», в которой был дан вывод уравнения теплопроводности в твердом теле, и разработка методов его интегрирования при различных граничных условиях. Метод Фурье состоял в представлении функций в виде тригонометрических рядов.



Ряд Фурье (по косинусам)

Ряд Фурье – разложение элементов бесконечномерного пространства по ортонормированному базису:

$$f = \sum_{k=0}^{\infty} a_k e_k, a_k = (f, e_k),$$

$\{e_k\}_{k=0}^{\infty}$ – ортонормированный базис в гильбертовом пространстве H , $f \in H$, (f, e_k) – скалярное произведение.

На практике часто используются тригонометрические ряды Фурье. Рассмотрим разложение функции f в пространстве $H=L^2(0; l)$ по ортонормированному базису, состоящему из косинусов:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left(\frac{kx\pi}{l}\right), a_k = \frac{2}{l} \int_0^l f(x) \cos\left(\frac{kx\pi}{l}\right) dx.$$

Если ограничить ряд n слагаемыми, то частичная сумма ряда является аппроксимацией элемента:

$$f(x) \approx S_n(x) = \frac{a_0}{2} + \sum_{k=1}^n a_k \cos\left(\frac{kx\pi}{l}\right).$$

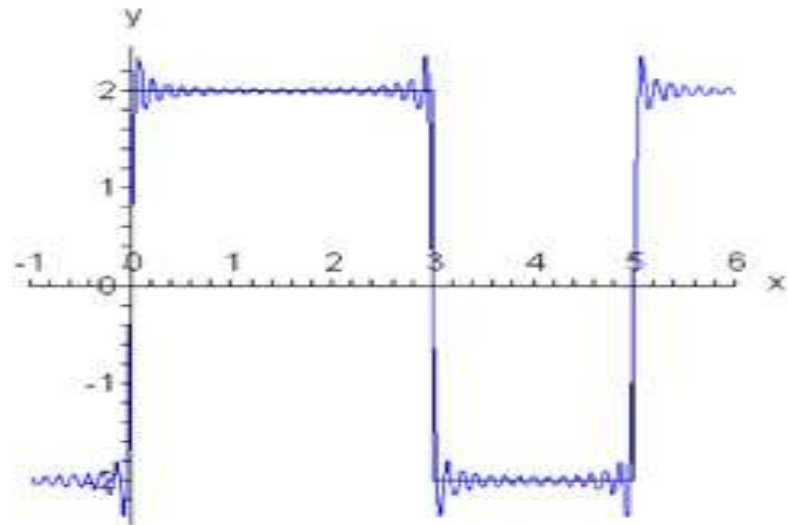
Эффект Гиббса для разрывных функций

Сходимость ряда Фурье: расстояние в метрике $L^2(0;l)$ между функцией f и частичной суммой S_n стремится к нулю:

$$\|f - S_n\| = \sqrt{\int_0^l (f(x) - S_n(x))^2 dx} \rightarrow 0, n \rightarrow \infty$$

Такая сходимость называется среднеквадратичной. Это слабый вид сходимости, при котором нельзя гарантировать приближенное равенство в каждой точке промежутка $(0;l)$. Вблизи точки разрыва наблюдается эффект Гиббса: характерные колебания графика S_n в окрестности точки разрыва.

Область распространения этих колебаний сужается с ростом n , но их амплитуда не уменьшается. Иными словами, эффект Гиббса не подавляется при помощи увеличения порядка частичной суммы ряда Фурье, его только можно сделать визуально почти незаметным.



Дискретный аналог ряда Фурье по косинусам

Пусть функция f задана в натуральных узлах промежутка $(0;l)$. В таком случае функция называется сеточной и является по сути l -мерным вектором.

Тогда определенные интегралы, которые используются при вычислении коэффициентов Фурье, следует заменить на интегральные суммы Римана:

$$f_i = \frac{a_0}{2} + \sum_{k=1}^l a_k \cos\left(\frac{k\left(i + \frac{1}{2}\right)\pi}{l}\right), \quad a_k = \frac{2}{l} \sum_{i=0}^l f_i \cos\left(\frac{k\left(i + \frac{1}{2}\right)\pi}{l}\right),$$

f_i – значение функции в i -м узле сетки, $i = 1..l$.

При этом непрерывные косинусы тоже заменяются на сеточные: для удобства берутся значения косинусов в узлах $i+1/2$. Получается, что l -мерный вектор с координатами f_i раскладывается по системе таких же l -мерных векторов, полученных из косинусов. То есть разложение переходит в конечномерное (l -мерное) пространство. Поэтому уже нет необходимости брать для разложения бесконечную систему косинусов, достаточно использовать первые l штук. Ряд Фурье превращается в конечную сумму. Для аппроксимации используется только часть слагаемых этой суммы, т.е. частичная сумма порядка n :

$$f_i \approx \frac{a_0}{2} + \sum_{k=1}^n a_k \cos\left(\frac{k\left(i + \frac{1}{2}\right)\pi}{l}\right), \quad n < l.$$

Аппроксимация (сжатие) изображения

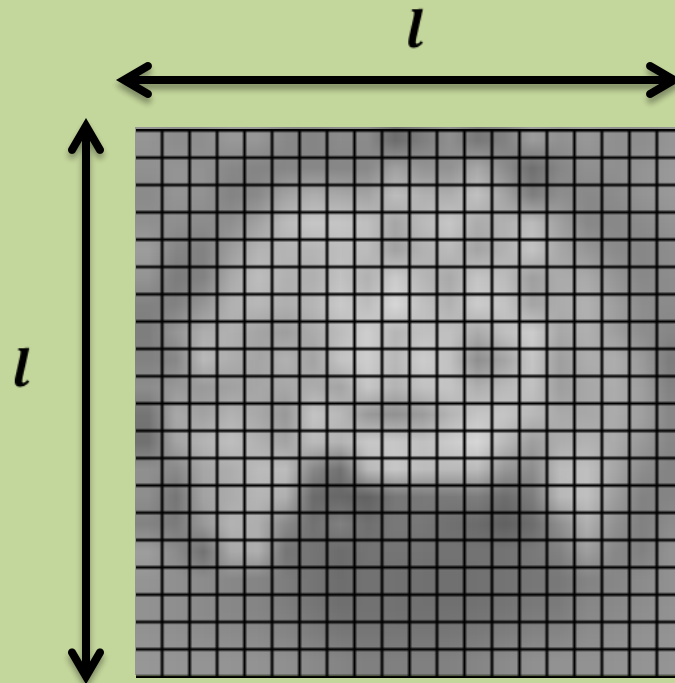
Дискретные аналоги рядов Фурье используются в процессе сжатия изображений, а точнее, на первом этапе сжатия. Монохромное изображение размером l на l пикселей можно представить в виде сеточной функции f_{ij} , где i, j – координаты пикселя, $i, j = 1..l$. Значение функции – яркость пикселя, которая меняется от 0 до 255. При сжатии монохромных изображений функция яркости подвергается аппроксимации частичной суммой дискретного разложения по косинусам.

При этом используется несколько подходов:

1. Аппроксимация всего изображения сразу как сеточной функции двух переменных. Для этого требуется дискретный аналог кратного ряда Фурье. Этот способ не очень практичный, так как затрачивает большое количество машинного времени.

2. Разбиение изображения на строки пикселей, аппроксимация каждой строки в отдельности. Вычисления упрощаются, вместо функции двух переменных обрабатываются l функций одной переменной.

3. Разбиение изображения на прямоугольные блоки и аппроксимация каждого блока в отдельности.



Численная реализация проекта

Практический результат проекта – программа в среде разработки Visual Studio на языке C++, которая считывает изображение в формате bmp, аппроксимирует его частичной суммой разложения по косинусам, и выводит полученное изображение в файл с расширением bmp.

На следующих слайдах продемонстрирована работа программы. Входные данные – два изображения размером 400 на 400 пикселей. На этом слайде представлены оригиналы изображений. Далее приведены результаты аппроксимации монохромного изображения по строкам и по блокам, а также цветного изображения по строкам.



Аппроксимация монохромного изображения по строкам ($l=400$ пикселей)

Порядок	$n = l/2=200$	$n = l/4=100$	$n = l/8=50$
Результат	 A circular logo with a whale silhouette and the text 'КУТ' and 'КОМПЬЮТЕРНЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ' around the perimeter. The reconstruction with n=200 shows a very clear, sharp image with minimal noise.	 A circular logo with a whale silhouette and the text 'КУТ' and 'КОМПЬЮТЕРНЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ' around the perimeter. The reconstruction with n=100 shows a slightly more pixelated image compared to n=200.	 A circular logo with a whale silhouette and the text 'КУТ' and 'КОМПЬЮТЕРНЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ' around the perimeter. The reconstruction with n=50 shows significant horizontal banding artifacts across the entire image.
Разность с оригиналом	 A dark circular image showing the difference between the original and the reconstruction with n=200. The difference is very low, appearing as faint, sparse noise.	 A dark circular image showing the difference between the original and the reconstruction with n=100. The difference is slightly more visible than for n=200, showing some faint horizontal patterns.	 A dark circular image showing the difference between the original and the reconstruction with n=50. The difference is very prominent, showing clear horizontal banding artifacts across the entire image.

Анализ результатов аппроксимации монохромного изображения по строкам

На предыдущем слайде приведены результаты трех аппроксимаций изображения по строкам. Порядок аппроксимации обозначен символом n и означает порядок частичной суммы разложения по косинусам ($n < l = 400$). Видно, что при $n = 200$ аппроксимация визуально уже неотличима от оригинала.

Обратим внимание на участки, где исходное изображение содержит контрастные переходы от черного к белому: при аппроксимации косинусами появляется периодический «шум» вблизи контраста, который постепенно утихает при удалении от линии контраста. Поскольку линия контраста на изображении аналогична разрыву функции яркости пикселей, то так проявляется эффект Гиббса. При увеличении порядка аппроксимации эффект Гиббса становится визуально незаметен, ореол его распространения сужается, однако все равно присутствует, что видно по разности между аппроксимацией и оригиналом.

Аппроксимация монохромного изображения по блокам ($l=8$ пикселей)

Порядок	$n = l/2=4$	$n = l/4=2$	$n = l/8=1$
Результат			
Разность с оригиналом			

Анализ результатов аппроксимация монохромного изображения по блокам

Аппроксимация по блокам – один из способов борьбы с эффектом Гиббса. Изображение разбивают на блоки и аппроксимируют каждый блок по отдельности. Размеры блоков выбирают такие, чтобы в пределах одного блока яркость пикселей менялась по возможности незначительно, но при этом аппроксимация имела смысл. Цель: зажать контрастные фрагменты изображения в небольших по размеру блоках и предотвратить распространение эффекта Гиббса (шума, помех) из этих блоков в другие блоки. Те блоки, где нет контрастных фрагментов, оказываются вообще освобождены от эффекта Гиббса. Кроме того, если внутри какого-то блока яркость изображения имеет постоянное или почти постоянное значение (например, белый фон данного изображения), то для их аппроксимации можно использовать меньший порядок n .

На предыдущем слайде реализовано разбиение изображения по блокам размерами 8 на 8 пикселей и аппроксимация строк по косинусам внутри каждого блока. Следовательно, в данном случае $l=8$. По результатам видно, что зоны распространения эффекта Гиббса сократились до размеров блока.

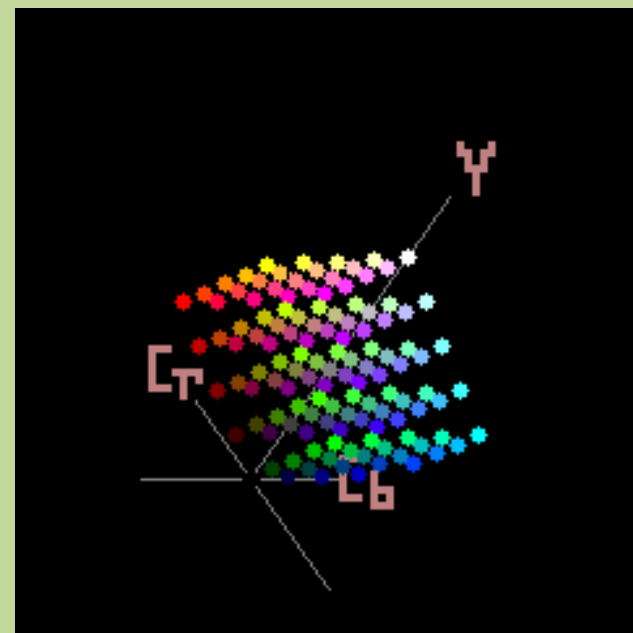
Аппроксимация цветного изображения

Разработанная программа рассчитана и на цветные изображения.

Изначально изображение считывается как цветное, для аппроксимации по яркости оно конвертируется в монохромное. Цветное же конвертируется в формат YCbCr по формуле, которую приведена внизу данного слайда, в которой Y-компонента яркости, Cb и Cr являются синей и красной цветоразностными компонентами.

Далее изображение аппроксимируется по двум каналам Cb и Cr. Яркость в данном случае аппроксимации не подвергается, так как глаз более чувствителен к изменению яркости, а не к цвету, следовательно, именно цвет можно хранить с большими потерями. Таким образом, мы приходим к задаче с монохромным изображением, но аппроксимируем не одну функцию, а две. Затем полученное изображение обратно конвертируется в RGB и записывается в файл с расширением bmp.

В результате аппроксимации цветного изображения также наблюдается эффект Гиббса.



$$Y' = K_R \cdot R' + (1 - K_R - K_B) \cdot G' + K_B \cdot B'$$

$$P_B = \frac{1}{2} \cdot \frac{B' - Y'}{1 - K_B}$$

$$P_R = \frac{1}{2} \cdot \frac{R' - Y'}{1 - K_R}$$

$Y' =$	16 +	$(65.481 \cdot R' +$	$128.553 \cdot G' +$	$24.966 \cdot B')$
$C_B =$	128 +	$(-37.797 \cdot R' -$	$74.203 \cdot G' +$	$112.0 \cdot B')$
$C_R =$	128 +	$(112.0 \cdot R' -$	$93.786 \cdot G' -$	$18.214 \cdot B')$

Аппроксимация цветного изображения по строкам ($l=400$ пикселей)

Порядок	$n = l/8=50$	$n = l/16=25$	$n = l/32 \approx 12$
Результат			
Разность с оригиналом			

Этапы сжатия изображения

Аппроксимация изображения частичной суммой дискретного разложения по косинусам является только первым этапом его сжатия. Сжатие информации получается за счет того, что исходные l^2 значений сеточной функции заменяются на $n \cdot l$ значений коэффициентов Фурье (при аппроксимации по строкам), которые используются в частичной сумме ($n < l$). При этом происходит, разумеется, потеря качества.

Далее аппроксимированное изображение проходит этап квантования, затем группировки и в конце – архивирования. В частности, так устроена технология JPEG.

Литература

1. Филимоненкова Н. В. Конспект лекций по функциональному анализу: учеб. пособие, – СПб., 2015. – 107 с.
2. Володин Ф. Изобретаем JPEG// Хабрахабр. – Дата обновления: 17.12.2013. – URL: <https://habrahabr.ru/post/206264/>.
3. Ткаченко Д.С. Применение ряда Фурье для аппроксимации изображений// – Дата обновления: 12.12.2012. – URL:<http://tkachenko-mephi.narod.ru/pdfs/ia.pdf>.
4. YCbCr// Wikipedia. – Дата обновления: 19.02.2014. – URL: <https://ru.wikipedia.org/wiki/YCbCr>.