Частотная фильтрация

Высшая школа цифровой культуры
Университет ИТМО
dc@itmo.ru

Содержание

1 Частотное представление и частотная фильтрация

В прошлой лекции мы обсудили несколько методов обработки изображений в пространственной области, но существует гораздо больше методик, которые лучше всего работают в частотной области. В частотной области местоположение пикселя представлено его частотами x и y, а его значение представлено амплитудой. Частота в изображении говорит о скорости изменения значений пикселей.

Одномерное изображение Чтобы лучше понять, что такое частотное представление, давайте взглянем на одномерное изображение – один ряд пикселей из исходного двухмерного изображения. На графике мы видим, что в начале идут серые пиксели, которые соответствуют значениям интенсивности около 150. Затем следует белая область с высокими значениями интенсивности, близкими к максимальным – 255. Затем идет черная область с очень низкими значениями интенсивности, затем белый, снова черный и снова белый.

Теперь посмотрим на значения интенсивности в ряду пикселей для другого изображения, Лены. Как видите, здесь уровни интенсивности меняются гораздо чаще. Как описать это свойство, показывающее, как часто меняется интенсивность?

Частотное представление – основная идея Основная идея частотного представления принадлежит французскому математику Жану Батисту Жозефу Фурье. Он заявил, что любая функция, которая периодически повторяется, может быть выражена как сумма синусов и косинусов разных частот, каждая из которых умножена на некий коэффициент. Теперь мы называем эту сумму рядом Фурье. Неважно, насколько сложна функция; если она является периодической и удовлетворяет по крайней мере некоторым математическим условиям, она может быть представлен такой суммой.

Частотное представление – пример На экране мы видим еще один пример функции, разложенной на синусы. Ряд Фурье представляет периодический сигнал как сумму гармоник. По сути, он отображает заданную функцию времени (или пространства) в частотный спектр. Частотный спектр – это простой способ продемонстрировать амплитуду для каждой из частот, составляющих заданную функцию. Частотный спектр показывает, какие частоты присутствуют в исходном сигнале, и какие значения у соответствующих им амплитуд. Он сохраняет информацию о частоте, но отбрасывает информацию о фазе. Так, например, частотный спектр синусоидальной волны будет таким же, как у косинусоидальной волны той же частоты, несмотря на то, что полные преобразования Фурье синусоидальной и косинусоидной волн различаются по фазе.

На экране показан частотный спектр заданной функции. На этом примере мы видим, что исходная функция состоит только из двух частот, 1 и 3 с амплитудами 1 и 0.5.

2 Преобразование Фурье

Но что, если рассматриваемая нами функция не периодическая? Даже непериодические функции, площадь под кривой которых конечна, могут быть выражены как интеграл синусов и косинусов, умноженных на некоторые веса. Этот процесс разложения функции на составляющие ее частоты, называется преобразованием Фурье. Разница между рядами Фурье и преобразованием Фурье заключается в том, что ряд Фурье применяется к периодическим сигналам, а преобразование Фурье применяется к непериодическим сигналам. В нашем случае изображения являются функциями конечной длительности, поэтому мы можем использовать для них преобразование Фурье. Преобразование Фурье для функции времени или пространства – это комплексная функция частоты, величина (абсолютное значение) которой отражает как много данный частоты присутствует в исходной функции, и аргументом которой является фазовый сдвиг основной синусоиды на этой частоте. Существует также обратное преобразование Фурье, которое математически синтезирует исходную функцию из ее представления в частотной области без какой-либо потери информации.

2.1 Одномерное преобразование Фурье: прямое и обратное

Преобразование Фурье непрерывной функции $f(x)$ для непрерывной переменной x определяется уравнением, показанным на экране, где u также является непрерывной переменной. Поскольку мы интегрируем по x, результирующая функция F является функцией от u.

$$F(u) = \int_{-\infty}^{\infty} f(x) e^{-\imath 2\pi u x} \mathrm{d}x \tag{1}$$

Используя формулу Эйлера, мы можем выразить $F(u)$ иначе.

$$F(u) = \int_{-\infty}^{\infty} f(x) \left[\cos\left(2\pi u x\right) - \imath \sin\left(2\pi u x\right)\right] \mathrm{d}x \tag{2}$$

Если функция $f(x)$ вещественная, ее преобразование обычно содержит ненулевую комплексную составляющую. Обратите внимание, что преобразование Фурье представляет собой интеграл функции $f(x)$, умноженной на синусоидальные составляющие, частоты которых определяются значениями u. Поскольку после интегрирования остается только частота, мы говорим, что область преобразования Фурье – это частотная область.

И наоборот, зная $F(u)$, мы можем получить $f(x)$, используя обратное преобразование Фурье.

$$f(x) = \int_{-\infty}^{\infty} F(u) e^{\imath 2\pi u x} \mathrm{d}u \tag{3}$$

2.2 Дискретное преобразование Фурье

До сих пор мы рассматривали преобразование Фурье для непрерывных функций. Но цифровые изображения не являются непрерывными – их можно представить как дискретные 2D-функции. Исходный непрерывный аналоговый сигнал преобразуется в последовательность дискретных значений с помощью дискретизации и квантования. Преобразование Фурье для дискретной функции может быть получено из преобразования Фурье для непрерывной функции с использованием понятия дискретизации. Дискретное преобразование Фурье (ДПФ) преобразует конечную последовательность равноотстоящих отсчетов функции в одинаковой длины последовательность равноотстоящих отсчетов дискретного преобразования Фурье, согласно представленной на экране формуле.

$$F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-\imath 2\pi u x / M}, \; u = 0, 1, 2, \ldots, M-1 \tag{4}$$

Пусть нам дана выборка f_x, состоящая из M выборок функций $f(x)$, при помощи дискретного преобразования Фурье мы получаем выборку F_u из M комплексных дискретных значений. Каждая составляющая преобразования Фурье (то есть значение if $F(u)$ для каждого значения u) состоит из суммы всех значений функции $f(x)$. Значения $f(x)$ умножаются на синусы и косинусы различных частот. Каждая из M составляющих $F(u)$ называется частотной составляющей преобразования.

И наоборот, при данном наборе отсчетов $F(u)$, мы можем восстановить набор отсчетов f_x с помощью обратного дискретного преобразования Фурье (ОДПФ):

$$f(x) = \frac{1}{M} \sum_{u=0}^{M-1} F(u) e^{\imath 2\pi u x / M}, \; x = 0, 1, 2, \ldots, M-1 \tag{5}$$

Можно показать, что как прямое, так и обратное дискретное преобразование бесконечно периодичны с периодом M.

Множитель $1/M$, который стоит перед преобразованием Фурье, иногда ставится перед обратным преобразованием, а не перед прямым. Иногда (хотя и нечасто) оба уравнения умножаются на $1/\sqrt{1/M}$. Расположение множителя значения не имеет. Если используются два множителя, единственное требование состоит в том, чтобы их произведение было равным $1/M$.

2.3 Фурье-спектр

В общем, преобразование Фурье содержит комплексные составляющие, и для целей наглядного отображения принято работать с амплитудой преобразования (вещественной характеристикой), которая называется спектром Фурье или частотным спектром, или иногда амплитудой преобразования Фурье:

$$|F(u)| = \left(R^2(u) + I^2(u)\right)^{1/2} \tag{6}$$

В этом уравнении $R(u)$ и $I(u)$ являются, соответственно, вещественной и мнимой частями $F(u)$.

При обработке изображений мы в основном будем иметь дело со свойствами спектра. Другая величина, которая иногда используется, это энергетический спектр, определяемый как квадрат спектра Фурье:

$$P(u) = |F(u)|^2 = (R^2(u) + I^2(u) \tag{7}$$

Еще одна величина, которая иногда используется, это это фазовый спектр или фазовый угол:

$$\phi(u) = \arctan\left(\frac{I(u)}{R(u)}\right) \tag{8}$$

2.4 Преобразование Фурье для простой функции

Прежде чем продолжить, давайте рассмотрим простой одномерный пример преобразования Фурье. Преобразование Фурье для функции, показанной на экране, может быть получено следующим образом:

$$F(u) = \int_{-\infty}^{\infty} f(x)e^{-\imath 2\pi ux}\mathrm{d}x = \int_{-W/2}^{W/2} Ae^{-\imath 2\pi ux}\mathrm{d}x =$$

$$\frac{-A}{\imath 2\pi u}\left(e^{-\imath\pi uW} - e^{\imath\pi uW}\right) = \frac{A}{\imath 2\pi u}\left(e^{\imath\pi uW} - e^{-\imath\pi uW}\right) =$$

$$AW\frac{\sin(\pi uW)}{\pi uW}$$

В этом уравнении на последнем шаге, мы использовали тригонометрическое выражение $sin\theta = (e^{\imath\theta} - e^{-\imath\theta})/2\imath$. В этом случае комплексные составляющие преобразования Фурье прекрасно сочетаются с вещественной синусоидальной функцией.

Давайте посмотрим на график этой функции и график ее амплитуды. Обе функции стремятся к бесконечности в обоих направлениях. Следует отметить следующие ключевые свойства (1) положение нулей как в $F(u)$, так и $|F(u)|$ обратно пропорционально ширине W "коробочной"функции; (2) высота лепестков уменьшается в зависимости от расстояния до начала координат; и (3) функция стремится к бесконечности как для положительных, так и для отрицательных значений u. Позже мы увидим, что эти свойства весьма полезны при интерпретации спектров преобразований Фурье для двумерных изображений.

2.5 Преобразование Фурье для изображений

Как и любая другая дискретная двумерная функция, изображение может быть представлено как взвешенная сумма простых синусоид. Независимо от того, насколько нерегулярным может быть изображение, его можно разложить на набор синусоидальных составляющих, каждая из которых имеет четко определенную частоту. Функции синуса и косинуса для разложения называются базисными функциями разложения.

Компоненты спектра дискретного преобразования Фурье (ДПФ) определяют амплитуды синусоид, которые вместе образуют изображение. На любой заданной частоте изображения большая амплитуда ДПФ подразумевает большую выделенность синусоиды, соответствующей этой частоте в изображении. Если изображение имеет высококонтрастные участки с неоднородной интенсивностью, его преобразование Фурье будет иметь ненулевые компоненты, соответствующие высокочастотным синусоидам. Если изображение сглаженное, с равномерной интенсивностью, то низкочастотные синусоиды будут иметь ненулевые коэффициенты.

2.6 Преобразование Фурье для двумерного случая

Пусть $f(x,y)$ – непрерывная функция двух непрерывных переменных, x и y. Пара двумерных непрерывных преобразований Фурье задается выражениями:

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-\imath 2\pi(ux+vy)}\mathrm{d}x\mathrm{d}y \qquad (9)$$

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{\imath 2\pi(ux+vy)}\mathrm{d}u\mathrm{d}v \tag{10}$$

где u и v – частотные переменные. Когда мы говорим об изображениях, x и y интерпретируются как непрерывные пространственные переменные. Как и в одномерном случае, область значений переменных u и v определяет непрерывную частотную область.

Базисные функции – это двумерные волны.

Подобно одномерному случаю, существует дискретное преобразование Фурье для двумерного случая, которое используется при обработке изображений.

2.7 Визуализация Фурье-спектра

Поговорим немного о том, как можно интерпретировать визуализацию спектра Фурье в двумерном случае.

Спектр Фурье – это массив величин всех коэффициентов Фурье $|F(u,v)|$. Его часто рассматривают как самостоятельное изображение для дальнейшей визуализации и интерпретации. Он показывает наличие в изображении определенных базовых образов. Каждая позиция в визуализированном спектре соответствует определенной базисной функции, определяемой ее частотами u и v. Чем больше значение $F(u,v)$, соответствующее этой базисной функции, тем светлее или ярче пятно в позиции (u,v).

Центр спектра соответствует нулевым частотам u и v. Таким образом, если центр визуализированного спектра светлый, то это означает, что исходное изображение в основном содержит области с однородной интенсивностью без сильных перепадов яркости.

Если же края спектра светлые, это означает, что исходное изображение имеет много областей с перепадами яркости.

На правой части экрана вы видите визуализацию базисных функций, помещенных над соответствующими им позициями в спектре. Таким образом, если правый верхний угол спектра светлый, это означает, что исходное изображение содержит край или перепады яркости, которые проходят примерно под $45°$. Если правый нижний угол спектра светлый, это означает, что исходное изображение содержит границы в направлении $-45°$.

Практические моменты Прежде чем мы рассмотрим примеры изображений и их спектры Фурье, давайте рассмотрим некоторые свойства преобразования Фурье для изображений и типичные приемы, используемые для визуализации Фурье-спектра изображений.

Одним из часто используемых свойств ДПФ является то, что преобразование Фурье для вещественной функции $f(x,y)$ является симметричным:

$$F(u,v) = F^*(-u,-v) \tag{11}$$

Это означает, что и Фурье-спектр изображения симметричен относительно начала координат:

$$|F(u,v)| = |F(-u,-v)| \tag{12}$$

Как и в одномерном случае, преобразование Фурье в двумерном случае и обратное к нему бесконечно периодичны в направлениях u и v. В случае ДПФ, для одномерного преобразования данные в интервале от 0 до M, где M – это количество значений дискретной функции, состоят из двух полупериодов, встречающихся в точке $M/2$. Для целей отображения и фильтрации удобнее иметь в этом интервале полный период преобразования, в котором данные непрерывны и упорядочены должным образом. Для этого нам нужно умножить исходную пространственную функцию $f(x)$ на $(-1)^x$. В двумерном случае, чтобы сдвинуть данные так, чтобы $F(0,0)$ находился в центре результирующего частотного прямоугольника, мы можем умножить исходную функцию $f(x,y)$ на $(-1)^{x+y}$. Это обычная практика.

Стоит отметить еще одно интересное наблюдение, что составляющая ДПФ с нулевой частотой

$$F(0,0) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) \tag{13}$$

по сути, представляет собой среднее значение $f(x,y)$, то есть среднее значение интенсивности исходного изображения.

Обычно изображения состоят в основном из низкочастотных компонентов. Это приводит к появлению светлых пятен в середине спектра и очень ограниченной видимости на высоких частотах. Для улучшения визуализаци полного спектра часто применяют логарифмическое преобразование спектра перед визуализацией.

Визуализация Фурье-спектра: реальное изображение Теперь давайте посмотрим на некоторые примеры изображений и соответствующие им Фурье-спектры. На экране показано реальное изображение и его спектр. По визуализации спектра можно сказать, что изображение не имеет резких краев (светлых пятен почти нет нигде кроме центральной части спектра). А также можно отметить, что это изображение содержит края во всех направлениях – яркое пятно в центре выглядит как круг без деформаций в одном конкретном измерении. Горизонтальная линия в середине спектра говорит нам, что в исходном изображении должны быть вертикальные края, и это связано с тем, что вертикальные края исходного изображения (правый и левый) имеют разную интенсивность. Если предположить, что исходное изображение представляет собой всего лишь один период бесконечной периодической функции, можно увидеть резкое изменение интенсивности между одной копией изображения и другой по вертикальным краям – от очень светлой дыни вдоль левого края до темных яблок справа. Интенсивности верхнего и нижнего краев не так сильно различаются, и в результате вертикальная линия в середине изображения спектра не такая яркая.

Визуализация Фурье-спектра: синтетические примеры Теперь давайте посмотрим на пару синтетических изображений, одно из которых является размытой версией другого. Вы можете видеть, что на первом верхнем изображении есть черные и белые полосы по горизонтали с очень резким переходом от черного к белому. И соответственно в его спектре присутствуют высокочастотные составляющие – вы видите яркую вертикальную линию через все значения v. Это соответствует резким изменениям интенсивности по вертикали на исходном изображении. Больше нигде светлых пятен нет, потому что на исходном изображении нет изменений интенсивности в любом другом направлении. Также, в отличие от предыдущего слайда, мы не видим горизонтальной линии в середине спектра, потому что есть разрыв между левым и правым краями этого изображения.

Теперь посмотрим на пример внизу экрана. Это размытая версия изображения, показанного выше. Резких переходов между черными и белыми полосами нет, переход плавный. В результате, спектр этого изображения больше не имеет высокочастотных составляющих. На изображении спектра мы видим только фрагмент вертикальной линии, который соответствует плавным изменениям интенсивности в вертикальном направлении на исходном изображении. Как и на верхнем изображении, в исходном изображении нет других краев, и, следовательно, на спектральном изображении нет других светлых пятен.

Визуализация Фурье-спектра: еще примеры На экране вы видите еще несколько примеров изображений и их спектров. Напомним, что естественное изображение имеет более плавные переходы интенсивности по сравнению с синтетическим, и, следовательно, его спектр концентрируется больше вокруг центральной части, соответствующей низким частотам.

Перенос и вращение Этот слайд демонстрирует еще несколько свойств ДПФ. В верхнем ряду показаны изображения прямоугольника, в нижнем ряду – соответствующие спектры (после применения логарифмического преобразования). Здесь очевидны две вещи. Спектр нечувствителен к переносу изображения (изображения спектра первых двух изображений идентичны), но он вращается на тот же угол, что и повернутое изображение. Напомним, что спектр описывает направления и частоты изменений интенсивности в исходных изображениях. При переносе прямоугольника ни направления, ни «резкость» его краев не меняются. Следовательно, мы имеем одинаковые спектры. Когда прямоугольник вращается, направления краев на изображении меняются – они поворачиваются. В результате мы видим светлые пятна в спектре, соответствующие другим направлениям изменения интенсивности.

Теорема о сверстке Преобразование Фурье важно для обработки сигналов не только потому, что оно дает нам инструмент для представления любого изображения в виде взвешенной суммы одних и тех же базовых функций и, таким образом, обеспечивает другой способ сравнения изображений, но также потому, что оно позволяет фильтровать изображения в частотной области. И теорема о свертке играет здесь ключевую роль.

Рассмотрим три пространсвенные функции: f, g и h, такие что g является результатом пространственной свертки f и h. Теперь предположим, что F, G и H – соответствующие им преобразования Фурье. Тогда $G = FH$. Другими словами, преобразование Фурье свертки двух функций в пространственной области равно произведению преобразований Фурье этих двух функций в частотной области. И наоборот, если у нас есть произведение двух преобразований, мы можем получить свертку в пространственной области, вычислив обратное преобразование Фурье.

А свертка в частотной области аналогична умножению в пространственной области, причем они связаны прямым и обратным преобразованием Фурье, соответственно. Теорема о свертке – это основа для фильтрации в частотной области.

Свертка изображения с определенным ядром оказывает на это изображение тот же эффект, что и умножение спектра этого изображения на Фурье-преобразование ядра. Следовательно, линейная фильтрация всегда может выполняться либо в пространственной, либо в спектральной областях. Фильтрация в спектральной области проще с точки зрения вычислений, поскольку свертка в пространственной области заменяется умножением точки на точку в частотной области. Таким образом, вместо выполнения дорогостоящей в вычислительном отношении операции пространственной свертки быстрее вычислить преобразования Фурье исходного изображения и фильтра, умножить результаты, а затем применить обратное преобразование Фурье.

Фильтрация в частотной области Методы фильтрации в частотной области основаны на модификации преобразования Фурье для достижения конкретной цели и последующем вычислении обратного ДПФ, чтобы вернуться в пространственную область.

Подводя итог, можно сказать, что фильтрация в частотной области состоит из следующих шагов:

1. Исходное изображение умножить на $(-1)^x + y$, чтобы его Фурье-преобразование стало центрированным.
2. Вычислить ДПФ $F(u, v)$ изображения, полученного после шага 1.
3. Вычислить произведение $G(u, v) = H(u, v)F(u, v)$ используя поэлементное умножение, где $H(u, v)$ является фильтром – симметричной вещественной функцией.
4. Вычислить обратное ДПФ для $G(u, v)$.
5. Выделить вещественную чсать результата шага 4.
6. Умножить результат шага 5 на $(-1)^x + y$.

Примеры фильтрации в частотной области Верхнее левое изображение на этом слайде представляет собой изображение интегральной микросхемы, полученное с помощью сканирующего электронного микроскопа, а рядом с ним справа – его спектр Фурье. Если мы посмотрим на исходное изображение, мы можем заметить резкие границы, которые проходят примерно под $\pm 45°$ и два белых выступа. Спектр Фурье показывает заметные компоненты по направлениям $\pm 45°$, которые соответствуют этим границам. Также на исходном изображении видны два белых пятна окисления, возникшие в результате температурного разрушения. Внимательно посмотрев вдоль вертикальной оси изображения спектра, мы можем увидеть вертикальный компонент преобразования, который находится не на оси, а немного левее. Этот компонент показывает края окислившихся элементов. Обратите внимание, что угол частотной составляющей относительно вертикальной оси соответствует наклону (относительно горизонтальной оси изображения) длинного белого элемента. Также обратите внимание на нули в вертикальной частотной составляющей, соответствующие узкому вертикальному размаху пятен окисления.

Это типичные примеры зависимостей между частотной и пространственной областями, которые мы можем увидеть.

Один из простейших фильтров, которые мы можем построить, – это функция $H(u, v)$, которая равна 0 в центре (центрированного) преобразования и единицам в остальных местах. Этот фильтр отклонит компонент, соответствующий $(u, v) = (0, 0)$, и «пропустит» (т.е. оставит без изменений) все остальные компоненты $F(u, v)$, когда мы вычисляем произведение $H(u, v)$ и $F(u, v)$. Теперь вспомните, что компонент $F(0, 0)$ отвечает за среднюю интенсивность изображения, поэтому установка его на ноль снизит среднюю интенсивность выходного изображения до нуля. На рисунке внизу экрана показан результат этой операции. Как и ожидалось, изображение стало намного темнее. Среднее значение равное нулю подразумевает наличие отрицательной интенсивности. Таким образом, хотя этот пример и иллюстрирует принцип действия, результирующее изображение не вполне представляет оригинал, так как все отрицательные значения интенсивности были обрезаны до 0.

Как упоминалось ранее, низкие частоты в преобразовании соответствуют плавно меняющимся компонентам интенсивности на изображении, например таким, как стены комнаты или безоблачное небо. С другой стороны, высокие частоты показывают резкие перепады интенсивности, такие как границы и шум. Следовательно, мы ожидаем, что функция $H(u, v)$, которая обнуляет высокие частоты при пропускании низких частот (называемая фильтром нижних частот, как уже упоминалось ранее), даст размытое изображение; в то время как фильтр с противоположным свойством (называемый фильтром верхних частот) улучшит четкость деталей, но приведет к снижению контрастности изображения. На экране показаны примеры таких эффектов. Слева показан фильтр нижних частот и соответствующее отфильтрованное изображение. В центре показаны аналогичные результаты для фильтра верхних частот. Обратите внимание, что в результате применения фильтра верхних частот, мы получаем очень темное изображение. Причина в том, что фильтр верхних частот удаляет компонент $F(0, 0)$, что приводит в основном к тому же эффекту, который был продемонстрирован ранее. В третьем столбце показано применение аналогичного фильтра, но с добавлением небольшой константы. Константа не влияет на повышение резкости, но предотвращает удаление компонента $F(0, 0)$ и, таким образом, сохраняет тональность изображения.

Обнаружение разрывов, перепадов яркости Фильтры (как пространственные, так и спектральные) можно использовать для улучшения изображений, а также для выделения некоторых интересующих нас свойств. Для многих методов обработки и анализа изображений ключевыми являются два основных свойства значений яркости изображения – неоднородность и сходство. На них основаны традиционные алгоритмы сегментации изображений.

Резкие, локальные изменения яркости являются обычными индикаторами края, границы объекта. Области изображения с разрывами содержат больше семантической информации, чем однородные области. Часто, основываясь только на границах объекта, можно получить семантический класс этого объекта. Итак, неоднородности важны, и на протяжении многих лет традиционные подходы к обнаружению неоднородностей были фундаментальными для анализа изображений.

Нас интересуют три типа характеристик изображения: изолированные точки, линии и края. Краевые пиксели – это пиксели, в которых яркость изображения резко меняется, а края (или краевые сегменты) представляют собой наборы связанных краевых пикселей. Детекторы краев – это инструменты локальной обработки изображений, предназначенные для обнаружения краевых пикселей. Линию можно рассматривать как тонкий краевой сегмент, в котором яркость фона по обе стороны от линии либо намного выше, либо намного ниже, чем яркость пикселей линии. Фактически, как мы обсудим позже, линии образуют так называемые «края крыши». Наконец, изолированную точку можно рассматривать как пиксель переднего плана (фона), окруженный пикселями фона (переднего плана).

Использование производных Итак, мы хотим обнаружить резкие локальные изменения яркости. Интуитивно понятно, что их можно обнаружить с помощью производных – отличного математического инструмента для обнаружения перепадов и локальных изменений функции. Для этой цели особенно хорошо подходят производные первого и второго порядка.

Рассмотрим фрагмент края склона, яркость которого плавно меняется. Под фрагментом виден график его яркости. А справа первая и вторая производные графика яркости. Двигаясь слева направо по профилю яркости, отметим, что первая производная положительна в начале склона и в точках на наклонной плоскости и равна нулю в областях с постоянной интенсивностью. Вторая производная положительна в начале склона, отрицательна в конце склона, равна нулю в точках на наклонной плоскости и нулю в точках постоянной яркости. Пересечение оси нулевой интенсивности и линии, проходящей между экстремумами второй производной, отмечает точку, называемую пересечением нуля второй производной.

Производные цифровой функции могут быть определены в терминах конечных разностей. Производную первого порядка можно аппроксимировать как разницу в интенсивности между соседними пикселями:

$$f'(x) = f(x+1) - f(x) \tag{14}$$

Это приближение может быть получено с использованием ряда Тейлора $f(x+1)$.

Аналогичным образом мы можем получить приближение производной второго порядка:

$$f''(x) = f(x+1) - 2f(x) + f(x+1) \tag{15}$$

Пример Давайте посмотрим на фрагмент изображения и его упрощенный профиль яркости ниже. Это поможет нам понять, как ведут себя производные первого и второго порядка, если встречается точка, линия и краяй объекта. Давайте проследим этот упрощенный профиль слева направо. Первоначально производная первого порядка отлична от нуля в начале и по всей кривой яркости, а производная второго порядка отлична от нуля только в начале и в конце кривой. Поскольку края цифровых изображений имеют подобный переход, можно заключить, что производные первого порядка определяют «толстые» края, а производные второго порядка – гораздо более тонкие. Затем мы сталкиваемся с изолированной точкой шума. Здесь величина отклика в точке намного сильнее для производной второго порядка, чем для производной первого порядка. Таким образом, можно ожидать, что производные второго порядка усилят мелкую детализация (включая шум) намного больше, чем производные первого порядка. Линия в этом примере довольно тонкая, поэтому это тоже мелкая деталь, и снова, вторая производная имеет большую величину. Наконец, обратите внимание, что на краях склона и ступенчатых переходах вторая производная меняет знак на противоположный (от отрицательного к положительному или от положительного к отрицательному) при переходе на склон и выходе из него. Этот эффект «двойного края» является важной характеристикой, которую можно использовать для определения краев. Когда мы продвигаемся к краю, знак второй производной также используется, чтобы определить, является ли край переходом от светлого к темному (отрицательная вторая производная) или от темного к светлому (положительная вторая производная).

В заключении, мы можем сделать следующие выводы:

1. Производные первого порядка показывают "толстые"края.
2. Производные второго порядка сильнее реагируют на мелкие детали, такие как тонкие линиее, изолированные точки и шум.
3. Производные второго порядка дают двойной отклик на перепаде и ступенчатых изменениях яркости.
4. Знак второй производной можно использовать, чтобы определить, происходит ли переход на краю от светлого к темному или от темного к светлому.

Для вычисления первой и второй производных для каждого пикселя изображения обычно используют пространственную свертку со специально разработанными фильтрами. Давайте посмотрим на несколько примеров.

Обнаружение точек Изолированная точка характеризуется резкими изменениями яркости по всем направлениям. Ее можно обнаружить с помощью фильтра, аналогичного представленному на экране. С помощью этого фильтра мы просто измерим взвешенные различия между пикселем и его 8 соседями. Идея состоит в том, что яркость изолированной точки будет сильно отличаться от ее окружения, и, таким образом, точка будет легко обнаруживаться с помощью этого типа ядра. Обратите внимание, что коэффициенты фильтра в сумме равны нулю, указывая на то, что отклик фильтра будет нулевым в областях с постоянной яркостью.

Различия в яркости, которые будут учитываться, – это различия, превышающие определенный порог. Таким образом, точка будет обнаружена в позиции (x, y), в которой ядро центрировано, если абсолютное значение отклика фильтра в этой позиции превышает заданный порог. Такие точки помечены 1, а все остальные помечены 0 в итоговом изображении, что дает двоичное изображение.

На экране показано исходное изображение, отклик фильтра и изображение двоичного результата после применения порога.

Лапласиан Фильтр, использованный на предыдущем слайде для обнаружения изолированных точек, представляет собой ядро Лапласа – простейший оператор производной, который для функции (изображения) $f(x, y)$ для двух переменных определяется как:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{16}$$

Чтобы выразить это уравнение в дискретной форме, мы можем использовать дискретные формы производных второго порядка. Дискретный Лапласиан для двух переменных можно определить как:

$$\nabla^2 f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y) \tag{17}$$

Это уравнение можно реализовать с помощью свертки с ядром слева. Это ядро изотропно для вращений с шагом 90° относительно осей x и y. Это означает, что его отклик не зависит от направления скачков яркости вдоль осей x и y. Чтобы сделать его изотропным и по диагоналям, мы можем добавить еще четыре составляющих и настроить центральную составляющую так, чтобы общая сумма всех элементов в ядре была равна нулю.

Маски Лапласиана часто используются для повышения резкости изображений.

Обнаружение линий Линии обнаруживаются образом, аналогичным обнаружению точек. Согласно нашим предыдущим наблюдениям, мы можем ожидать, что вторые производные приведут к более сильному отклику фильтра и будут давать более тонкие линии, чем первые производные. Таким образом, мы можем также использовать маски Лапласиана для обнаружения линий.

Крайнее левое изображение на слайде представляет собой двоичное изображение соединения проводов на электронной схеме. На следующем изображении показан его Лапласиан. Лапласиан содержит отрицательные значения, поэтому для его правильного отображения он был отмасштабирован: средний серый соответствует нулю, более темные оттенки серого представляют отрицательные значения, а более светлые оттенки – положительные. Эффект двойной линии отчетливо виден в увеличенной области. Сначала может показаться, что с отрицательными значениями можно справиться, просто взяв абсолютное значение Лапласиана. На следующем изображении показаны абсолютные значения Лапласиана. Как видите, такой подход удваивает толщину линий. Более подходящий подход – использовать только положительные значения Лапласиана. Изображение справа показывает, что этот подход приводит к более тонким линиям, которые обычно более полезны. Вы также можете заметить, что если линии шире, чем размер ядра Лапласа, они разделены нулевым «пространством». Это вполне ожидаемо. Например, когда ядро 3x3 центрируется на линии постоянной яркости шириной 5 пикселей, отклик будет нулевым. Таким образом, продемонстрированный подход полезен для обнаружения линий, которые тоньше размера детектора. Линии, которые не удовлетворяют этому условию, можно обрабатывать с помощью методов обнаружения краев, которые мы обсудим далее.

Но перед этим давайте рассмотрим еще один пример обнаружения линии.

Обнаружение направленных линий Ядро Лапласа изотропно, поэтому его отклик не зависит от направления (по отношению к четырем направлениям ядра 3x3: вертикальному, горизонтальному и двум диагоналям). Иногда нам интересны линии только одного конкретного направления. Как мы можем их обнаружить?

Рассмотрим ядра, представленные на экране. Предположим, что изображение с постоянным фоном и содержащее различные горизонтальные, вертикальные и диагональные линии фильтруется с помощью первого ядра. Максимальные отклики будут иметь место в тех областях изображения, где горизонтальная линия проходит через средний ряд ядра. Это легко проверить, нарисовав простой массив единиц с линией изменяющейся яркости (скажем, 5 с), горизонтально проходящий через массив. Подобный эксперимент показал бы, что второе справа ядро, показанное на экране, лучше всего реагирует на линии, ориентированные на $+45°$, третье ядро – на вертикальные линии и четвертое ядро – на линии, ориентированные на $-45°$. Предпочтительное направление каждого ядра умножается на больший коэффициент (т.е. 2), чем другие возможные направления. Сумма коэффициентов в каждом ядре равна нулю, что указывает на нулевой отклик в областях с постоянной интенсивностью.

Предположим, что нам нужно найти все линии толщиной в один пиксель, ориентированные в направлении $+45°$ на левом изображении во второй строке. Для этого мы используем второе ядро в верхнем ряду. Изображение в середине нижней строки – результат фильтрации с помощью этого ядра. Есть два основных сегмента изображения, ориентированных в направлении $+45°$: один в верхнем левом углу и один в правом нижнем углу. Отрезок прямой линии в правом нижнем углу светлее, чем отрезок в верхнем левом углу, потому что сегмент линии в правом нижнем углу исходного изображения имеет толщину в один пиксель, а сегмент в левом верхнем углу – нет. Ядро «настроено» на обнаружение линий толщиной в один пиксель в направлении $+45°$, поэтому мы ожидаем, что его отклик будет сильнее при обнаружении таких линий. Поскольку нас интересует самый сильный отклик, мы устанавливаем высокий порог T, равный 254 (максимальное значение минус один). На нижнем правом изображении белым цветом показаны точки, значения которых превысили порог: мы видим, что линия толщиной в один пиксель, ориентированная на $+45°$, правильно обнаруживается.

Обнаружение краев Теперь посмотрим на профили яркости краев. Идеальный край характеризуется перепадом между двумя уровнями яркости, происходящим на расстоянии в один пиксель. Левое изображение на слайде показывает разрез вертикального перепада и горизонтальный профиль яркости этого края. Перепады встречаются, например, в созданных компьютером изображениях, использующихся в таких областях, как моделирование движения твердых тел и анимация. Эти идеальные перепады могут появиться на расстоянии одного пикселя при условии, что дополнительная обработка (например, сглаживание), чтобы края выглядели «настоящими», не используется.

На практике края цифровых изображений размыты и зашумлены. В таких ситуациях края можно более точно смоделировать как имеющие возрастающий профиль яркости, например край на втором изображении на экране. Наклон склона обратно пропорционален степени размытия края. В такой модели у нас больше нет единой «точки края» вдоль профиля. Вместо этого точкой края теперь является любая точка склона, а сегментом края тогда будет считаться набор таких связанных точек.

Третий тип края – это так называемый край крыши, характеристики которого показаны на последнем изображении. Края крыши представляют собой линии, ширина которых определяется толщиной и резкостью линии.

Когда его основание крыши имеет ширину в один пиксель, край крыши представляет собой не что иное, как линию толщиной в один пиксель, проходящую через область изображения.

Изображения часто содержат все три типа краев.

Градиент До сих пор мы говорили об одномерных профилях интенсивности. Но изображения являются двумерными. Итак, инструмент выбора для определения силы и направления края в произвольной точке (x, y) изображения f – это градиент, определяемый как вектор

$$\nabla f(x,y) = \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} \\ \frac{\partial f(x,y)}{\partial y} \end{bmatrix} \tag{18}$$

Этот вектор обладает хорошо известным свойством: он направлен в сторону наибольшего изменения интенсивности f в точке (x, y). Таким образом, для вертикальных краев второй элемент этого вектора будет равен 0, показывая, что частная производная по y равна нулю. Для горизонтальных краев первый элемент этого вектора равен нулю, поскольку частная производная по x равна нулю.

Если вычислить градиент для всех допустимых значений x и y, то $f(x, y)$ становится векторным изображением, каждый элемент которого является вектором градиента. Амплитуда такого вектора градиента в точке (x, y) задается его евклидовой нормой. Это значение скорости изменения направления вектора градиента в точке (x, y). Обратите внимание, что массивы частных производных, градиента и его амплитуды имеют тот же размер, что и f, и создаются, когда x и y могут изменяться по всем местоположениям пикселей в f. Обычно массив величин называют градиентным изображением.

Направление вектора градиента в точке (x, y) задается арктангенсом отношения частных производных по y и x.

Дифференцирование и свертка Для получения градиента изображения необходимо вычислить частные производные $\partial f/\partial x$ and $\partial f/\partial y$ в каждом пикселе изображения. Частная производная функции двух переменных $f(x, y)$ по переменной x определяется как

$$\frac{\partial f}{\partial x} = \lim_{x \to \epsilon} \left(\frac{f(x+\epsilon, y)}{\epsilon} - \frac{f(x,y)}{\epsilon} \right) \tag{19}$$

Для дискретной функции мы можем получить оценку производной первого порядка через разности

$$\frac{\partial f}{\partial x} \approx f(x+1, y) - f(x,y) \tag{20}$$

Частная производная линейна и инвариантна к переносу. Это означает, что его можно представить в виде свертки. В частности, уравнение для $\partial f/\partial x$ может быть вычислено путем фильтрации $f(x, y)$ с помощью одномерного ядра, представленного на экране.

Аналогично, частная производная по y

$$\frac{\partial f}{\partial y} \approx f(x, y+1) - f(x, y) \tag{21}$$

может быть вычислена как результат фильтрации с вертикальным ядром.

Вычисление градиента изображения На экране представлены популярные операторы для вычисления градиента.

Когда нас интересует диагональное направление границ, нам нужны двумерные ядра. Одной из самых ранних попыток использования двумерного ядра с диагональным направлением были перекрестные операторы Робертса. Рассмотрим маску 3×3 в верхней части слайда. Операторы Робертса основаны на использовании диагональных разностей

$$\frac{\partial f}{\partial x} = (z_9 - z_5) \tag{22}$$

и

$$\frac{\partial f}{\partial y} = (z_8 - z_6) \tag{23}$$

Эти производные могут быть получены путем фильтрации изображения при помощи масок, показанных на слайде.

Ядра размера 2×2 концептуально просты, но они не так полезны для вычисления направлений границ, как ядра, симметричные относительно своих центров, самые маленькие из которых имеют размер 2×2. Такие ядра учитывают характер данных на противоположных сторонах от центральной точки и, таким образом, несут больше информации о направлении границы. Простейшие цифровые приближения частных производных с использованием ядер размером 3×3 – это операторы Прюитта, также показанные на слайде. В таком виде разница между третьей и первой строками области 3×3 аппроксимирует производную по x, а разница между третьим и первым столбцами аппроксимирует производную по y. Такие оценки производных более точны, чем оценки, полученные с помощью операторов Робертса.

Небольшое изменение двух предыдущих уравнений увеличивает вес центрального коэффициента в два раза. Использование коэффициента 2 в центре обеспечивает сглаживание изображения. Такие ядра называются операторами Собеля.

Ядра Прюитта проще реализовать, чем ядра Собеля, но разница во времени вычислений между ними небольшая, и, обычно, не является проблемой. Поскольку ядра Собеля обладают лучшими характеристиками шумоподавления (сглаживания), то они используются чаще.

Обратите внимание, что коэффициенты всех ядер на этом слайде суммируются до нуля, что дает нулевой отклик в областях с постоянной яркостью, как и ожидалось от производных операторов.

Пример Любая из пар ядер, которые мы рассмотрели на предыдущем слайде, сворачивается с изображением для получения компонентов градиента $G_x = \partial f/\partial x$ и $G_y = \partial f/\partial y$ для каждого пикселя. Два массива частных производных затем используются для оценки резкости и направления края. Чтобы получить величину градиента необходимо вычислить евклидову векторную норму, для чего требуется вычислить квадраты и квадратные корни, а это является ресурсо-затратным. Чтобы избежать этой проблемы, используется приближение величины градиента через абсолютные значения:

$$M(x, y) \approx |G_x| + |G_y| \tag{24}$$

Это уравнение более привлекательно с точки зрения вычислений, и оно также сохраняет относительные изменения уровней яркости.

Давайте посмотрим на пример. Верхнее левое изображение на экране – это исходное изображение со значениями яркости, нормированными до диапазона $[0, 1]$. Правое верхнее изображение представляет собой компонент градиента по x, а нижнее левое изображение – компонент градиента по y, оба изображения получены с использованием соответствующих ядер Собеля. И нижнее правое изображение – это градиентное изображение, полученное из суммы этих двух компонентов. Направленность горизонтальной и вертикальной составляющих градиента видна на изображениях частичных градиентов. Обратите внимание, например, насколько хорошо различимы черепица, горизонтальные стыки кирпича и горизонтальные сегменты окон на верхнем правом изображении по сравнению с другими краями. Однако, нанижнем левом изображении лучше видны вертикальные элементы фасада и окон.

В общей терминологии используется термин карта краев (edge map), когда речь идет об изображении, основными характеристиками которого являются края, например таком, как изображение значений градиента.

Поведение производных на перепадах яркости с шумом Краевые модели, которые мы использовали до сих пор, не имели шума. Однако естественные изображения часто зашумлены, и производные могут усилить этот шум. Давайте посмотрим, что происходит с производными профилей интенсивности с шумом.

Сегменты изображения на экране показывают крупные планы четырех краев склона яркости, которые переходят от черной области слева к белой области справа (заметим, что весь переход от черного к белому – это один край). Крайний левый сегмент изображения лишен шума. Остальные три изображения искажены аддитивным гауссовским шумом с нулевым средним и стандартными отклонениеми уровней интенсивности 0,1, 1,0 и 10,0, соответственно. График под каждым изображением представляет собой горизонтальный профиль яркости, проходящий через центр изображения. Все изображения имеют разрешение по яркости 8 бит, где 0 и 255 представляют, соответственно, черный и белый цвет.

Эти фрагменты изображений очень похожи. Даже самое шумное правое изображение явно воспринимается человеком как переход от черного к белому.

Теперь давайте посмотрим на производные для этих изображений. Первая производная фрагмента изображения без шума выглядит ожидаемо – она равна нулю в областях постоянной интенсивности, что соответствует двум черным полосам, показанным на изображении производной. Производные в точках склона постоянны и равны наклону склона. Эти постоянные значения на изображении производных показаны серым цветом.

По мере того как мы переходим к изображениям со все большим шумом, производные отличаются все больше от производных в отсутствии шума. Действительно, было бы сложно связать изображение производной и профиль последнего изображения на экране с первым изображением производной незашумленного края склона. В то же время на исходных изображениях в верхней части экрана шум практически не обнаруживается визуально. Этот пример показывает нам, что первые производные очень чувствительны к шуму.

А как же производные второго порядка? Они еще больше чувствительны к шуму. Вторая производная бесшумного изображения показана в левом нижнем углу слайда. Тонкие белые и черные вертикальные линии – это положительная и отрицательная составляющие второй производной. Серый на этих изображениях показывает ноль. Единственный график второй производной зашумленного изображения, который слегка похож на бесшумное изображение, соответствует шуму со стандартным отклонением 0,1. Остальные изображения и профили второй производной ясно демонстрируют, что обнаружить их положительные и отрицательные компоненты будет чрезвычайно сложно.

Влияние шума на производные Мы видели, что даже очень небольшой с визуальной точки зрения шум может оказать значительное влияние на две ключевые производные, используемые для обнаружения краев. Из-за шума значения соседних пикселей могут сильно отличаться, и это различие будет улавливаться масками на основе конечных разностей. Чем больше шум, тем больше отклик фильтров производных.

Что тут можно сделать? Может помочь сглаживание изображения. В применении к зашумленным изображениям, фильтры сглаживания изображения следует применять перед дифференцированием.

Сглаживание перед дифференцированием Давайте посмотрим на другой пример. На экране мы видим зашумленную функцию и ее первую производную. В данном случае невозможно обнаружить край.

Теперь давайте применим ядро сглаживания Гаусса. Сейчас на экране первая строка соответствует той же зашумленной функции, что и раньше. Во второй строке показан фильтр Гаусса, который мы планируем применить. Третья строка является результатом свертки зашумленной функции с гауссовым ядром. И последняя строка – это первая производная отфильтрованного сигнала. В отличие от производной исходного сигнала, эту производную можно использовать для обнаружения краев. Чтобы обнаружить край, нам достаточно нужно найти экстремум производной.

Использование свойств свертки Свертка и дифференцирование ассоциативны:

$$\frac{\mathrm{d}}{\mathrm{d}x}(f \star g) = f \star \frac{\mathrm{d}}{\mathrm{d}x}g \tag{25}$$

Это означает, что мы можем изменить порядок этих операций. И если мы предварительно вычислим производную фильтра Гаусса, а значит, мы можем применить этот фильтр непосредственно к зашумленному изображению вместо того, чтобы сначала сглаживать гауссовым ядром, а затем дифференцировать.

Производная Гауссова фильтра Функция производной Гаусса широко используется в анализе изображений. У нее много интересных свойств.

На этом рисунке, показанном на экране, представлен график двумерной гауссианы и ее производной первого порядка по x. Как обсуждалось ранее, производная первого порядка по x может быть вычислена посредством свертки с фильтром $[1, -1]$. На слайде показан результат применения такого фильтра к функции Гаусса. У производной функции появилось два «выступа» вместо одного «выступа» у исходной функции, причем в противоположных направлениях.

Сейчас на экране мы видим две частные производные – по x и по y. Если мы рассмотрим их как изображения, мы получим изображения в нижнем ряду. Производные Гаусса тесно связаны с фильтрами Габора, широко используемыми для анализа текстур изображений в традиционном компьютерном зрении.

Сглаживание и локализация краев Хотя сглаживание перед дифференцированием чрезвычайно полезно для подавления шума, в то же время оно размывает края. Так, более широкий фильтр производной Гаусса приводит к сглаживанию большей площади и, следовательно, лучше подавляет шум, но в то же время приводит к большему размытию на карте краев, а также смещает расположение краев. На экране мы видим сравнение трех различных по размеру фильтров производной Гаусса.

Лапласиан Гауссова-фильтра Подобно производной Гаусса, мы можем комбинировать сглаживание фильтром Гаусса и обнаружение краев с помощью производной второго порядка, или Лапласиана. $\nabla^2 G$ – Лапласиан Гауссова-фильтра, часто сокращаемый до LoG, является еще одним популярным оператором обнаружения краев.

$$\nabla^2 G(x, y) = \left(\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4}\right) e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{26}$$

На экране показан трехмерный график, изображение и поперечный разрез негатива функции LoG. На графике поперечного сечения вы можете заметить, что нулевые пересечения LoG происходят в точках $x^2 + y^2 = 2\sigma^2$, которые определяют окружность радиуса 2σ с центром в вершине функции Гаусса. Из-за своей формы функцию log иногда называют оператором мексиканской шляпы. Вы также можете увидеть на экране ядро 5x5, которое по форме похоже на log. Это приближение не уникальное. Его цель – уловить основную форму функции log: положительный центральный компонент, окруженный соседней отрицательной областью, значения которой уменьшаются в зависимости от расстояния до начала координат, и нулевую внешнюю область. Сумма коэффициентов должна равняться нулю, чтобы отклик ядра был нулевой в областях с постоянной интенсивностью.

Ядра фильтров произвольного размера (и в то же время с фиксированным значением σ) могут быть сгенерированы путем выборки функции LoG и масштабированием коэффициентов так, чтобы их сумма равнялась нулю.

Недостатки градиента при локализации краев Хотя градиент, градиент Гаусса и LoG можно использовать для улучшения краев изображений, они не являются идеальными детекторами края. Эти операторы приводят к смещению положений краев. Кроме того, они не дают связанных краев. На экране видно, что края обнаруживаются неидеально.

Свойства хорошего детектора краев Давайте подумаем, какими свойствами должен обладать хороший детектор краев.

1. Высокий уровень обнаружения, низкий уровень ошибок: все края должны быть найдены и не должно быть ложных срабатываний. Пример, показанный синим цветом на экране – это пример плохого детектора, когда некоторые точки края не обнаружены, а обнаруженные точки очень далеки от истинного края.

2. Хорошая локализация: краевые точки должны быть хорошо локализованы. Расположенные края должны быть как можно ближе к истинным краям. То есть расстояние между точкой, отмеченной детектором как край, и центром истинного края должно быть минимальным. Пример, показанный на слайде зеленым цветом – это пример плохой локализации. Обнаруженные краевые точки находятся не совсем там, где находится истинный край.

3. Единственная точка границы: Детектор должен возвращать только одну точку для каждой точки истинного края. То есть количество локальных максимумов вокруг истинного края должно быть минимальным. Это означает, что детектор не должен идентифицировать несколько краевых пикселей, если существует только одна краевая точка. Пример, показанный черным цветом, показывает случай, когда детектор дает слишком много откликов на крайние точки.

Детектор Кэнни – один из популярных детекторов края, который был разработан с учетом всех выше перечисленных свойств.

Детектор Кэнни Алгоритм детектора краев Кэнни работает следующим образом:

1. Применить ко входному изображению оператор производной Гаусса: это сгладит изображение, удалит шум и обнаружит неоднородности в интенсивности.

2. Вычислить значение и направления градиента.

3. Выделить локальные максимумы на изображении значений градиента. Градиентное изображение обычно содержит широкие гребни вокруг локальных максимумов. Цель этого шага – проредить эти гребни. Суть такого подхода состоит в том, чтобы найти ряд дискретных направлений краев (векторов градиента) и оставить только точки с наибольшей величиной для каждого направления.

4. Использовать двойной порог и анализ связности для обнаружения и связывания краев. Алгоритм Кэнни использует обрезание по порогу, который использует два порога: низкий порог, T_L, и высокий порог, T_H. Если пиксель проходит через высокий порог, он считается «сильным» краевым пикселем, а пиксели, прошедшие нижний порог, считаются «слабыми» краевыми пикселеми. Предполагается, что каждый сильный пиксель является допустимым краевым пикселем. Слабый краевой пиксель помечается как допустимый краевой пиксель, только если он подключен к другому сильному краевому пикселю.

Детектор края Кэнни по своим характеристикам превосходит те детекторы края, которые мы обсуждали до сих пор.

Пример Давайте посмотрим на результаты каждого шага работы детектора Кэнни, примененного к изображению Lena.

Первый шаг – применить оператор производной Гаусса и вычислить его величину.

Следующим шагом является применение порога.

И последний шаг – выделение локальных максимумов.

Приведем еще один пример для сравнения результатов обнаружения краев с использованием оператора Собеля и алгоритма детектора краев Кэнни. Обратите внимание, что результат детектора края Кэнни имеет более тонкие края.