

# Обработка изображений

иколаев Д. П.

2 марта 2016 г.

### **Аннотация**

организационные моменты на сдаче: 1 вопрос по выбору студента - идеаль 2 вопроса по выбору преподавателя задача в зависимости от сложности можно выступить в течение семестра с докладом по части курса +1/2 балла сделать методичку? по теме опасные вопросы - вопросы про цвет, быстрое преобразование хафа хорошо гуглится - сжатие изображений задачи забываются заранее. в разные дни можно отвечать 1 и тот же вопрос

учебник - гонсалес/вудс, йане??, красивеньких "цифровая обработка изображений" про цвет - кандидатская диссертация димы

# Глава 1

сделать

## Глава 2

# lecture1

### 2.1 введение

Что такое изображение, к примеру фотографическое изображение. не протяженные во времени объекты. в основном двумерные объекты. Изображение это некоторый результат измерения, обычно с пространственными значениями и несущее какую-либо информацию.

через полиморфическую функцию. Измеряем потоки света, а хотим понять отражательные свойства и объекты трехмерной сцены расположенной перед фотоаппаратом. Можно работать с акустическими, рентгеновскими изображениями.

Постановка задачи обработки изображений в широком смысле В узком смысле

- анализ изображений - на выходе параметры изображения
- обработка - на выходе изображения
- синтез изображения - на входе параметры на выходе изображения

Что еще делаем с изображением - сжатие изображения - анализ + синтез изображений

Обработка изображений:

- восстановление изображения(super resolution, фильтрация)
- редактирование изображений(закрывать дырку на изображении (свести к минимуму градиент на краях...), правдоподобно вклеить объект в изображение)

анализ изображений: распознавание, выделение углов и т.д

где используется: везде: медицина, транспорт (штрафы, платные дороги), военные, финансовая безопасность, маркетинг, физика (геофизика, астрономия), нефть (томография)

Математический аппарат(требуется):

- дискран (алгоритмы на дискретных структурах)
- численные методы
- линейная алгебра
- урматы, дифуры
- матан, чуть-чуть
- статистика
- проективная геометрия
- оптимизация

Изображение - работаем с растром и говорим, что это изображение пусть существует сколько-то мерная таблица, в узлах которой - математической объекты есть несколько осей координат и задана евклидова система кроме растра существует векторное изображение (объектное). в растре может храниться:

- битики (черно-белые изображения, жесткие маски)
- скаляр (уровень сигнала, серые, скалярные изображения)
- вектор (цветное изображение, векторное)

Обсудим случайно выбранную задачу обработки изображений. Отличие обработки изображений от обработки сигналов. в обработке сигналов очень хорошо работают линейные методы (фурье, вейвлеты), в обработке изображений - очень плохо. Почему? в обработке сигналов - сигналы акустические, электромагнитные. непрерывные сигланы (как правило) про изображение так сказать нельзя ибо изображение это проекция и при проекции 3х мерного мира на 2ух мерный существуют заслонения, разрывы глубины, разрывы свойств объектов на разных планах.

необходима частотная разделимость сигнала и шума - тогда линейные методы хорошо работают

любые линейные фильтры будут сглаживать границы и размывать, что неприемлемо в обработке изображений

модель изображения - кусочно-гладкая функция с разрывами образующими граф, ребрами которого являются кусочно гладкие разрывы границы могут начинаться из какой-то точки и ни к чему не присоединяться

рассмотрим линейный метод фильтрации, окно размера  $m$  сглаживание - преобразование фурье, обработка спектра (свертка) - обратное преобразование. не устраивает сложность  $n^2$  дзять среднее по окну  $n \log(m)$

Хотим за  $O(n)$  способы 1. построить интегральное изображение 2. зная предыдущую сумму вычисляем следующую за 2 действия ( $O(2n)$ ) 2.1 хотим фильтр с сохранением границ - менять размер окна в зависимости от

изображения 2.2 основная теорема математического анализа - формула Ньютона-Лейбница площадь под изображением равна разнице первообразных. 2.3 формула посмотреть у иры (или) 2.3 что считаем?? 2.4 получилось за  $O(2n)$

Всегда следим за сложностью.

Представление чисел. Пусть сигнал 8 бит.  $n=2^{*30}$ ,  $m=200$  подсчитали  $255*2^{*30}=2^{*38}$ , влезает в  $2^{*64}$  в 64 бита однако достаточно 16 что будет, если 16 бит то в какой-то момент произойдет переполнение и старший бит пропадет получим лесенку (картинка у иры) в одно окно попадет максимум один разрыв

пока в одно окно попадает максимум одно переполнение - все хорошо.

проблема потери точности: осторожно с флоатами. когда мы накапливаем что-то большое во float то вначале все приводится к одному показателю, а потом складывается если меньшее число не попадает в мантиссу больше то большее число не изменится

осторожно с округлением вниз - можно потерять сигнал

проблема дрейфа значений - когда прибавляем, вычитаем - складываются погрешности и будут уплывать значения к правой части картинки

## 2.2 регистрация изображений

Изображение непрозрачных объектов в видимом диапазоне

Камера простейшая модель камеры - камера обскуры. маленькая дырка перед экраном в темной комнате практически идеальна по параметрам дисторсии обертки светового пучка очень маленькая использовалась художниками для получения реалистичных картин ящик с дырочкой (картинка у иры) заднее фокусное расстояние - расстояние между экраном и дыркой глубина резкости - насколько можно уходить из фокусной плоскости передние, задние фокусные расстояния, фокусная плоскость - расфокусировка - изменение резкости - изменение переднего фокусного расстояния изменение зума - заднее фокусное расстояние

чем хорошая камера хуже камеры обскуры вьюетирование - изображение темнее к краям - элемент площади светочувствительной матрицы находящийся скраю, отвечает за меньший телесный угол, чем находящийся в центре матрицы

калибровка для вьюетирования - запоминание коэффициентов усиления для каждого пикселя по отдельности размечаем вьюетирование на возрастание уровня шумов

дефокусировка - (для камеры обскуры - не все лучи проходят через одну точку и в каждой точке мы наблюдаем лучи еще и для соседних пикселей - можно считать, что была вычислена свертка со сглаживающим ядром)

свертка - вызывает дефокусировку ядро свертки - функция рассеяния точки или аппаратная функция

$$f(x) * g(x) = \int f(x) * g(x - a) dx$$

можно ли дефокусировку откалибровать? b да и нет по идее обратная свертка существует но это ядро - сглаживающее а вообще говорим, что нет

это не калибровка - это серьезная фильтрация с массой побочных эффектов  
 смаз - проявляется только в разговорах о времени - на изображении может быть движущийся объект, необходимо время для интегрирования яркости (время выдержки). если за время выдержки местоположение объекта изменилось - изображение получится смазанным или движущаяся камера смаз можно записать как свертку идеальной картинке с функцией ... смаз и дефокусировка - нелинейная свертка смаз из-за вращения камеры нельзя записать как свертку

хроматические aberrации - у линзы разная плотность для разных длин волн и как следствие фокусное расстояние тоже разное и разные длины волн накладываются друг на друга получаем двоение границ если глубины резкости хватает и мы не говорим, что для разных длин волн дефокус разный разные цветовые каналы имеют разный размер и это можно учесть и скалибровать

разные каналы будут проэективно сводимы

если оптическая ось не перпендикулярна матрице - это часть тангенциальной дисторсии проэективное преобразование? 8 параметров

$$x = \frac{ax + by + c}{gx + hy + i}$$

$$y = \frac{dx + ey + f}{gx + hy + i}$$

аффинное - 6 параметров преобразование

Пуассоновский шум увеличение шума, уменьшение шума но увеличение смаза

Апертура

Байеровская мозайка

Гамма коррекция

## Глава 3

## лекция2

### 3.1 цветность, цветовое пространство

в глазу колбочки (для дня) и палочки (для ночи) в сумерках работают 4 канала

в камере тоже светочувствительные элементы. рассмотрим 1 элемент, он считает падающие фотоны, с определенными видами волн рис 1 - блокнот рис 2 - функция светочувствительности элемента (спектральная характеристика детектора)

фотоны между собой не взаимодействуют, с красителем взаимодействуют до насыщения

неизвестная линейная функция о  $a_i$  интеграл

$F(\lambda)$  - спектральная характеристика, неотрицательная, не образует линейное пространство - выпуклый конус

при линейном преобразовании  $F \rightarrow a$

$\Rightarrow a$  - выпуклый трехмерный конус

множество  $a$  - не покрывает положительный октант, ибо не может одновременно одинаково ярко видеть все цвета (одной колбочкой)

всевозможные достижимые точки цветового пространства называют цветовым конусом, и характеризуются чувствительностью -  $\chi$

сечение плоскостью проходящей через начало координат - цветовой треугольник

все возможные цветовые треугольники образованные одним конусом - проективно подобны цветовой треугольник человека - рис3 - двуугольник

цветим лазером от инфокрасного до ультрафиолетового, это образует петлю рис 4 - с одной точкой излома. проецируем на плоскость цветности (считаем касательные)

касательные в нуле - разные - переходят в разные точки. рис 5

фиолетовые цвета - не спектральные. так как не существует пурпурного лазера



цветность - то, что не меняется, когда меняется яркость. яркость - мощностная характеристика входного излучения без изменения длины волны  
цветность двумерна

не зависящая от длины волны характеристика окраски - светлота не зависящая от длины волны характеристика излучения - яркость

рис 6 хотим извлечь информацию об отражающей поверхности главная задача цветового зрения - определение характеристик окрасок объектов

способность биологического зрения решать поставленную задачу - феномен цветовой константности

что такое окраска?

каждую точку излучающей поверхности характеризуем спектральной индикатриссой излучения  $B(r, \lambda)$

$r$  - направление, размерность 2

$B(r, \lambda) = B(r) * S(\lambda)$  - не работает для витражных стекол

с окраской - спектральная индикатриса рассеяния (характеристика отражающей способности) чтобы охарактеризовать каждую точку объекта нужна функция  $G(r_{out}, r_{in}, \lambda)$  -  $G$ ? спектральная двулучевая функция отражающей способности

у нас нет нелинейных объектов, только линейное рассеяние - поэтому одна  $\lambda$

можем вынести за скобку спектральные характеристики? не фига

доказательство - белый блик. он бы не существовал

для диэлектрических глянцевых объектов можно записать данное разложение  $G = G_1() * \phi(\lambda) + G_2 * \phi(\lambda)$

одна компонента диффузная, другая - зеркальная. у зеркальной  $\phi_2$  - константа. не всегда белый

$G_2$  - матовое  $G_1$  - глянцевое - ламбертова - переизлучает в половине телесных углов одинаково - не так

все неправда. зеркальная компонента - то, что возникает на поверхности, диффузная возникает в толще.

металлы и хорошо матированные диэлектрики имеют один член разложения, большинство гладких диэлектриков - 2. бывает больше - 3.

вводим понятие цветового тела - подмножество цветового пространства, которое достижимо при наблюдении всех сцен при зафиксированном источнике

цветовой конус параметризован чувствительностью глаза

цветовое тело выпукло, не конус, центральносимметрично (относительно точки, соответствующей нейтрально серой окраске  $(1/2, 1/2)$ ) рис 7

теорема максимова - по форме цветового тела, можно восстановить спектральные характеристики источника света. - не применимо на практике - все точки сцены освещаются в разной мощностью

метомирии излучений и окрасок

метомирия излучений - это свойство разных излучений давать одну и ту же реакцию

метамирия окрасок - это разные красители, с разным спектром излучения могут оказаться неразличимыми для определенного сенсора и при определенном излучении

$$\int_0^\infty s_1(\lambda)\phi_1(\lambda)\chi_1(\lambda)d\lambda = \int_0^\infty s_2(\lambda)\phi_2(\lambda)\chi_2(\lambda)d\lambda$$

## 3.2 цветовые пространства

$$a_c = \int_0^\infty F(\lambda)\chi_c(\lambda)$$

$$a_h = \int_0^\infty F(\lambda)\chi_h(\lambda)$$

зная  $a_c, a_h, \chi_c, \chi_h$  - нельзя восстановить F

но если есть ограничения на F - то можно в среднем минимизировать ошибку в пространстве  $\lambda$

человек привыкает ко всему, даже к перевернутому миру восприятие фрагментально

CIE Lab, XYZ - цветовые пространства

XYZ - цветовое пространство наблюдателя, зафиксированно. изменили базис, сдвинули, сделали измеримым - получили Lab

использование Lab - не обосновано в нем яркостная компонента цвета введена в отдельную координату - правильно

закон Гугера Ламберта Бера:

рис 8

$$T(\lambda, l) = \exp(l * t(\lambda))$$

принтеры, мониторы и т. д. спектральные характеристики можно складывать монитор просто замещает свои базисные спектры

с принтерами все не так несколько механизмов 1 - если на белую канву наносить пятна красителя не пересекая их - субтрактивная цветовая модель из максимума спектров (белой бумаги) вычитаем несколько спектров 2 - льем краситель непрерывно, цветность зависит от толщины красителя спектр ведет себя как рис 9 заслонение - краска на краску. не увидим следов первой краски.

только в Lab говорить о расстоянии

базис в цветовых пространствах задан кривыми чувствительности человека,

RGB - не опирается на человека, скорее на камеры

цилиндрические и конические системы координат - HSV, HSI, HSB вводятся насыщенность и цветовой тон рис 10 проецирование на плоскость цветности центральное или параллельное - от этого зависит цилиндричность или коничность системы

почему эти системы важны? легко говорить в терминах цветового тона, яркости, насыщенности

наиболее инвариантная компонента по отношению к окраске - цветовой тон

когда мы говорим об окраске - наиболее безопасно говорить о цветовом тоне

каналы. сколько должен иметь монитор для воспроизведения всех, желаемых нами цветов? меньше 3ех нельзя. могли бы изображать только отрезок на цветовом треугольнике

старые мониторы воспроизводили рис 11 лучше рис 11.б сделать при 3ех каналах невозможно 5 каналов - рис 11.в на боках треугольника - есть проблемы

рис 12 - спектр чувствительности человеческого глаза красный и зеленый очень скорелированы должны: между красным и зеленым - намного больше цветов, чем между синим и зеленым

сколько красок нужно принтеру - бесконечно много. 7 красок - можно рассчитывать на неплохую цветопередача

смотрим на красную рубашку, в цветовом пространстве видим - скорее размазанный отрезок (из точки 0 до 1, если есть блик)

цветовое распределение однородно окрашенного объекта лежит в пространствах меньшей разметности (очень редко - 0)

k-ближайщих в цветовом пространстве по этому даст убогие результаты

## Глава 4

### лекция 3

.....1 час лекции....

движения

$\delta x, \delta y$  - сдвиг - для определения преобразования необходимо 2 точки

$\delta x, \delta y, \phi$  - + поворот - необходимо 3 точки

$\delta x, \delta y, \phi, s$  - + масштабирование - необходимо 4 точки

аффинная группа преобразований - 6 параметров

проективная - 8 параметров

для любых потриколлинеарных единичный квадрат проективным преобразованием в странный четырехугольник - разорванный по идее проективное преобразование сохраняет выпуклость

проективное преобразование может преобразовать бесконечно удаленную прямую в собственную, проходящую через изображение и наоборот

перспективное преобразование - проективное, реализуемое камерой, 6 параметров, группа

#### 4.1 свертки

$f$  свернута с  $g$  по  $y$  это  $\int_{-\infty}^{\infty} f(x) * g(y-x) dx$

считать свертки от картинки.

дискретная свертка - вместо интеграла - сумма

$$\sum_i f_i * g_{j-i}$$

сложность  $O(N^2)$  в асимптотике  $O(N*M)$

теорема о свертке

$$f * g = F^{-1}(F(f) * F(g))$$

сложность  $O(N * \log(N))$

свертка с прямоугольником - за  $O(n)$

свертка с треугольником новый сдвиг вычисляется как сумма предыдущего и прямоугольных сверток с разным знаком  $O(n)$

свертка с параболой - разность двух парабол - две треугольных свертки - вначале один смотри вниз, потом вверх  $O(n*k)$

сложное задание - подсчитать свертку гауссианы аппроксимированной 3 звеньями параболы.

самые нужные - бокс, и гаусовская свертки

композиция гаусовских сверток - гаусовская свертка, ассоциативны, нулевой элемент

гаусовское ядро дает изотропное сглаживание и фильтр сепарабелен

размер ядра гауссиановой свертки должно быть пропорционален  $\sigma$  обычно  $4 - 6 * \sigma$

алгоритм Дерише, Ван Влиит?- Диана на следующей лекции

смотрели на фильтры как на FIR фильтры - с ограниченным носителем функции рассеяния точки можем считать ядро равным нулю, всюду, за исключением конечной области

можем смотреть на фильтры как на IIR

на входе сигнал, на выходе свертка его с экспоненциальное спадающей штукой, обрезанной с одной стороны

делаем:  $y_i = (1 - \alpha) * y_{i-1} + \alpha * x_i$

## 4.2 градиент

производная - свертка свертка, которая ничего не делает, - дельта-функция свертка с производной дельта-функции - производная.

дискретный аналог дельта-функции - везде 0, где-то 1 - 0000001000000

аналог производной дельта-функции - везде ноль, перед 1 в дельта-функции - -1, после 1 000000000-10+1000000, а лучше 0000000000-1+1000000

картинку сдвинута на полпикселя относительно исходной - помнитть

дискретная свертка для второй производной 0000000000-12-10000000000000

в двумерный мир: рис.13.1 -1 -1 0+1 +1

крест Робертса

+1 -1 - половина оператора

градиент повернут на 45 градусов, и все съехало на полпикселя

есть еще треугольник . рис13.2 только из этого оператора можно собрать лапласиан (div(grad))) - сумма вторых производных по x и Y

-1 -1 4-1 - оператор для лапласианна -1

куча методов вычисление градиента на картинке - сглаживание прямого креста

аналог градиента для цветного изображений считать для каждой компоненты отдельно? а что потом

использование псевдоградиента Di Zenzo

вы построим директор, вектор без знака

рис. 14

запишем уравнение

рис 15

первое собственное значение - максимальная скорость псевдоградиента, второе - минимальная первый собственный вектор - направление псевдоградиента, второй - нормаль к нему

модуль псевдоградиента считается разными способами:  
 $\sqrt{\lambda_+}$ ,  $\sqrt{\lambda_+ - \lambda_-}$ ,  $\sqrt{\lambda_+} - \sqrt{\lambda_-}$   
с помощью данного оператора хорошо искать границы на изображении  
говорили о псевдоградиенте  
теперь о псевдолапласиане  
лапласиан - это оператор второй производной  
однако, вместо лапласиана используется обычно DOG - разница гаусси-  
анн  
на самом деле нас интересует (картинка - сглаженная картинка) снизу  
ограничиваем нулем

### 4.3 морфология

оконные максимумы и оконные минимумы  
эрозии, делатации, размыкания, замыкания  
не вводят новые значения в изображения  
любые ранговые фильтры так работают  
эрозия и делатация - сепарабельны. для прямоугольного структурного  
элемента  
алгоритм Ван Херка-Дила-Вермана  
возьмем два дополнительных массива рис. 16  
 $O(3N)$   
структурный элемент.  
возьмем 9 пикселей  
закрасим часть пикселей черным, остальные белым  
ввели  $\omega_+$  - белые,  $\omega_-$  - черные  
рассчитаем  $\max(\min(\omega_+) - \max(\omega_-), 0)$   
это подсчет крестообразных элементов на изображении  
что делать, если пики - сигнал, а не шум.  
нашли сигнал без пиков, вычли его из изображения

### 4.4 преобразования расстояния

есть два множества, черное и белое. и для каждого элемента мы хотим найти  
расстояние для ближайшего черного, или белого  
расстояние может быть разным (евклидовым ( $L_2$ ), манхетонским ( $L_1$ ),  
 $L_{inf}$  и т. д.) рис. 17  
как считать distance transform/ он, подсчитанный для однобитного из-  
ображения, дает возможность подсчитать элементарный морфологический  
фильтр, с окном, являющимся порогом данного преобразования  
если расстояние от элемента до ближайшего черного больше окна -  
оставляем белым, иначе становим черным  
сегментация с помощью дистансе трансформ

## Глава 5

### лекция4

коноваленко хрень сопоставление точек ааааа сопоставление объектов бееее