

Сравнение изображений:  
глобальные признаки

# Содержание

<b>1</b>	<b>Свойства изображений и их сравнение: традиционный подход к компьютерному зрению</b>	<b>2</b>
1.1	Зачем сравнивать изображения? . . . . .	2
1.2	Как сравнивать изображения? . . . . .	4
1.2.1	Глубокое обучение vs. традиционное компьютерное зрение	5
1.3	Признаки изображений . . . . .	6
1.4	Признаки по содержанию: цвет, текстура, форма . . . . .	7
1.5	Цвет . . . . .	7
1.5.1	Использование гистограмм . . . . .	7
1.5.2	Часто используемые функции для сравнения гистограмм	8
1.6	Сравнение цветов . . . . .	10
1.6.1	Гистограммы . . . . .	10
1.7	Текстуры . . . . .	14
1.7.1	Фильтры Габора . . . . .	15
1.7.2	Фильтры ICA . . . . .	16

# 1 Свойства изображений и их сравнение: традиционный подход к компьютерному зрению

## 1.1 Зачем сравнивать изображения?

В предыдущих лекциях мы обсуждали низкоуровневые методы обработки изображений, которые могут быть очень полезны для улучшения изображений. В таких случаях как на входе, так на выходе были изображения, но изображения, которые мы получали на выходе, были в некотором смысле лучше, чем на входе: менее зашумленными или с более высокой контрастностью. Рассмотренные нами методы упрощают анализ изображений и извлечение семантической информации как для человека, так и для компьютера. Они являются основой для других более сложных алгоритмов анализа изображений и распознавания изображений, которые мы обсудим далее.

Во вводной части курса мы рассмотрели различные задачи и области применения анализа и распознавания изображений: поиск изображений, классификация, сегментация, обнаружение и распознавание объектов и многие другие. Многие из этих задач основаны на понятии сходства изображений. Определение сходства может меняться от задачи к задаче, но умение сравнивать два изображения для решения конкретной задачи имеет первостепенное значение во всех областях применения.

**Поиск изображений** Для решения задачи поиска изображений мы считаем два изображения похожими, если они имеют как визуальное, так и семантическое сходство. Но это не обязательно должны быть изображения одинаковых объектов. Посмотрим на два изображения, представленные на экране: они похожи, поскольку на них изображены сельские горные районы в летнее время. В таком случае мы обращаем внимание на схожесть изображений в целом.

**Классификация сцен** Для задачи классификации сцен семантическое сходство играет более важную роль, чем визуальное сходство. Если мы хотим классифицировать сцены с точки зрения места, где они происходят - в помещении или на улице, то можно ожидать, что все изображения на этом слайде будут рассматриваться как похожие, потому что на всех этих изображениях представлены сцены на открытом воздухе. Также, нам бы хотелось, чтобы зимний пейзаж с большим количеством белого снега имел большую степень сходства с зеленым травянистым летним пейзажем, чем с чистой белой комнатой.

**Определение объектов и сегментация** Но если наша задача - обнаружение объектов, и нас интересуют автобусы, трамваи и автомобили, то изображения, рассмотренные на предыдущем слайде, попадут в разные классы. Мы ожидаем, что верхнее и нижнее изображения во втором столбце будут считаться похожими, также как и два изображения в третьем столбце, потому что на этих парах изображений представлены объекты одного типа. В задаче обнаружения объектов мы обычно ищем сходство между фрагментами изображения, отображающими интересующие нас объекты, а фон не принимаем во внимание.

Задача семантической сегментации аналогична задаче обнаружения объекта, но требует гораздо более точной локализации объекта. Вместо примерного местоположения объекта цель сегментации изображения - предоставить метки для каждого пикселя - независимо от того, принадлежит он объекту или нет. Для этого нам может понадобиться измерить сходство не только между любым заданным фрагментом анализируемого изображения и некоторым золотым шаблоном интересующего нас объекта, но также между различными фрагментами анализируемого изображения - чтобы понять, принадлежат два соседних фрагмента к тому же объекту или же нет.

**Аннотация изображений** Для аннотации изображений нужно выбрать теги как для объектов, так и для фона в изображениях. Например, на многих изображениях есть небо, на некоторых - дорога. И в таком случае нужно будет сравнить фрагменты изображений с некоторыми представлениями конкретных тегов. Эта задача решается путем многоклассовой классификации - каждое изображение может иметь несколько меток, и мы сравниваем каждое изображение, которое нам дано, с представлениями каждой метки в словаре имеющихся меток.

**Поиск неполных дубликатов, выравнивание изображений, трехмерная реконструкция, отслеживание движения** Во всех задачах, о которых мы говорили до сих пор, два изображения могут считаться похожими, даже если на них представлена не одна и та же сцена или не одинаковые объекты. Если наша задача состоит в обнаружении трамвая, то представленные на экране изображения трамваев будут принадлежать к одному классу и должны считаться похожими. Но для некоторых других задач, таких как поиск неполных дубликатов, выравнивание изображений, 3D-реконструкция, отслеживание движения, нам необходимо иметь возможность определять, одна и та же сцена (один и тот же объект) представлена ли на двух изображениях или нет. В задачах такого рода изображения зеленых трамваев не следует считать одинаковыми, тогда как два изображения желтого пикапа и два снимка местности должны считаться одинаковыми. Нам также необхо-

димо научиться определять позиции совпадающих точек для решения задач такого рода.

## 1.2 Как сравнивать изображения?

Теперь, когда мы уже поняли, что нам нужно сравнивать сходство изображений, возникает следующий вопрос: как это сделать? Компьютеру необходимо числовое представление изображений и метрика подобия, определенная в некотором пространстве представления. Числовое представление изображения обычно называют вектором признаков. Самый простой вектор признаков, который вы можете использовать, - это непосредственно интенсивности необработанных пикселей. Можно сравнивать изображения попиксельно, используя базовые векторы признаков, но такой метод хорошо работает только для определения точных копий изображений. Метрики попиксельного сходства не способны уловить семантическое сходство между изображениями. Даже два изображения одной и той же сцены из одного и того же положения дадут нам разные значения пикселей из-за потенциальных различий в условиях освещения, экспозиции или фокусировки. Вот почему представление изображения в виде необработанного набора значений интенсивности пикселей обычно является лишь отправной точкой для построения более сложных признаков, которые затем используются для сравнения изображений.

Результатом выделения признаков является вектор признаков. Вектор признаков - это просто вектор, который содержит информацию, описывающую важные характеристики изображения. Векторы признаков обычно являются более компактным представлением изображения, чем исходные значения пикселей, но в то же время они более точно описывают изображение и сохраняют ключевую семантическую информацию, представляющую интерес.

Когда мы получаем новое представление изображений в виде векторов признаков, мы можем сравнивать эти векторы, используя меры сходства, подходящие для решения конкретных задач и определенные в соответствующих пространствах признаков.

**Пространства признаков** Итак, вектор признаков - это абстракция изображения, используемая для описания и количественной оценки содержимого изображения. Векторы обычно состоят из вещественных, целых или двоичных значений. Проще говоря, вектор признаков - это список чисел, используемых для представления изображения. При обработке изображений могут использоваться различные признаки. Традиционные алгоритмы обработки изображений и компьютерного зрения часто полагаются на очень сложные наборы признаков. И выбор признаков для конкретной системы и задачи компьютерного зрения может сильно зависеть от конкретной проблемы.

Выбор информативных и независимых отличительных признаков является важным шагом для создания эффективных алгоритмов распознавания, классификации и регрессии. Обычно используются числовые функции, иногда также используются структурные элементы, такие как строки и графики.

Векторное пространство, связанное с векторами признаков и функцией подобия, определенной для этого пространства, часто называется пространством признаков. Другими словами, пространство признаков - это просто пространство всех возможных векторов признаков с заданной функцией, которую можно использовать для вычисления сходства (или расстояния) между любыми заданными двумя векторами признаков.

В некоторых задачах недостаточно выделить только один тип признаков для получения соответствующей информации из изображения. В таких случаях извлекаются два или более различных вектора признаков. И каждое из этих пространств признаков может иметь собственную связанную с ним меру сходства.

Векторы признаков могут быть объединены в один длинный объединенный вектор признаков, а затем мы можем определить функцию подобия для этого нового пространства признаков более высокой размерности.

**Комбинирование векторов признаков** Существует простой способ определить функцию подобия для комбинации различных векторов признаков - использовать линейную комбинацию функций подобия, определенных для каждого из исходных пространств признаков. Если у нас есть функция подобия  $d_1$ , определенная для признаков  $X$ , функция подобия  $d_2$  определенная для признаков  $Y$  и т. д., мы можем определить функцию подобия для объединенного пространства признаков как взвешенную сумму всех  $d_i$ . Таким образом, при измерении сходства между двумя изображениями мы будем учитывать различные признаки, а также можем определить важность каждого признака для решения конкретной задачи, что позволит нам соответствующим образом скорректировать веса во взвешенной сумме.

### 1.2.1 Глубокое обучение vs. традиционное компьютерное зрение

Алгоритм извлечения признаков может быть разработан и запрограммирован вручную. Этот этап также называется построением признаков. Построение признаков - сочетание искусства и науки, на этом шаге требуется множество экспериментов, которые основываются на интуиции и знаниях специалиста в предметной области. Основные достижения в области традиционного компьютерного зрения связаны с применением все более сложных методов извлечения признаков. За последние несколько лет во многих сферах применения традиционных методов компьютерного зрения на смену ручному

построению признаков пришли автоматические методы обучения признаков. Глубокое обучение позволило компьютеру не только использовать для обучения признаки, выделенные вручную, но обучиться на других признаках. Глубокое обучение включает в себя как трудоемкое построение признаков, так и автоматизированное принятие решений. Сегодня во многих задачах можно подавать необработанные значения пикселей в алгоритм машинного обучения и получать более точные прогнозы, чем когда-либо прежде. Мы поговорим о глубоком обучении в области компьютерного зрения в следующих лекциях. Сейчас мы остановимся на традиционных алгоритмах. Несмотря на то, что в настоящий момент глубокое обучение является наиболее распространенным решением для некоторых задач распознавания изображений, в других задачах, таких как выравнивание изображений и 3D-реконструкция, исследователи по-прежнему часто полагаются на традиционные методы построения признаков, а именно, на признаки, выделенные вручную.

### 1.3 Признаки изображений

**Текстовые и визуальные признаки** Теперь давайте посмотрим на различные виды признаков изображений. Есть разные способы их категоризации. Во-первых, признаки изображений можно разделить на текстовые и визуальные. Текстовые признаки состоят из текстовых аннотаций, созданных пользователем, и метаданных изображений, таких как геотеги, дата создания, настройки камеры и т.д. Они не связаны с пикселями изображения, но тем не менее могут быть очень полезны для различных задач и часто используются в сочетании с визуальными признаками.

Визуальные признаки - это признаки, полученные из необработанных значений пикселей. Такие признаки базируются на содержании. Они могут характеризовать распределение цвета в изображении, текстуру изображения, форму объектов, взаимное расположение объектов, шаблоны цвета или текстуры.

**Глобальные и локальные** Визуальные признаки можно разделить на глобальные и локальные. Глобальные признаки описывают изображение в целом. Обычно глобальные признаки вычисляются на основе значений всех пикселей изображения, например, цветовая гистограмма изображения или средняя интенсивность изображения. Такие признаки могут использоваться для задач поиска изображений и классификации сцен, но в настоящее время они в основном заменены глубоким обучением.

Локальные признаки описывают фрагмент изображения. Многие локальные признаки вычисляются для областей, расположенных поблизости от так называемых ключевых точек или точек, представляющих интерес. Поиск ключевых точек и локальные признаки по-прежнему широко исполь-

зуются для решения многих задач, таких как выравнивание изображений, 3D-реконструкция, отслеживание движения. Всего несколько лет назад подобные построенные вручную локальные признаки успешно использовались также в обнаружении и сегментации объектов, но в настоящее время для решения этих задач машинное и глубокое обучение заменили ручное построение признаков. Мы перейдем к глубокому обучению немного позже, а теперь давайте поговорим в целом о признаках, основанных на содержании, а также изучим пару наиболее известных методов выделения локальных признаков, которые все еще используются в настоящее время в ряде задач компьютерного зрения.

## **1.4 Признаки по содержанию: цвет, текстура, форма**

Существует множество различных классификаций признаков по содержанию. Цвет, текстура и форма - одни из самых интуитивно понятных для человека свойств изображения. В компьютерном зрении существует множество подходов к описанию и представлению этих важных свойств. Векторы признаков иногда также называют дескрипторами. Дескрипторы цвета, текстуры и формы описывают цвет, текстуру и форму всего изображения или какой-либо части изображения. Информация о пространственном расположении частей изображения также бывает важной для решения некоторых задач, в таком случае помимо информации о цвете, текстуре и форме изображения могут использоваться в том числе и пространственные характеристики, предоставляющие информацию о пространственном расположении фрагментов изображения.

## **1.5 Цвет**

### **1.5.1 Использование гистограмм**

Цвет, несомненно, является наиболее важным признаком визуального восприятия изображения человеком. Именно поэтому важную роль в анализе изображений играют цветовые дескрипторы. Естественный способ сравнить цветовые характеристики двух изображений - это сравнить распределение цветов на этих изображениях. Таким образом, многие характеристики цвета являются дескрипторами распределения вероятностей; а функции расстояния для распределений вероятности, являются общепринятыми метриками, используемыми для сравнения характеристик цвета.

Наиболее очевидно представить распределение цвета в виде гистограммы. Гистограммы часто используются в качестве признаков изображения. Цветовая гистограмма - это представление распределения цветов в изображении. Она получается путем подсчета и нормирования количества пиксе-



лей в изображении, которые попадают в каждый из возможных цветов в данном цветовом пространстве. Таким образом, когда цветовые гистограммы используются в качестве цветовых признаков, изображения будут представлены векторами длины, равной общему количеству цветов в выбранном цветовом пространстве, где каждый элемент вектора признаков представляет собой долю пикселей определенного цвета в данном изображении. Затем два изображения можно легко сравнить, вычислив расстояние между гистограммами цвета, которые их описывают. Для этого могут использоваться различные метрики, перечислим несколько из них:  $L_1$  или просто сумма поэлементных разностей двух векторов,  $L_2$  - евклидова метрика,  $L_\infty$ , расстояние хи-квадрат, расстояние Вассерштейна (в англоязычной литературе можно часто встретить название этой функции Earth Mover Distance или EMD).

Поскольку мы говорим о распределениях, стандартные статистические метрики также иногда используются для более компактного представления цветовых распределений изображения, а именно три первые момента: среднее значение, дисперсия и асимметрия. Стандартные метрики могут также использоваться для измерения расстояния между векторами признаков.

### 1.5.2 Часто используемые функции для сравнения гистограмм

Мы упомянули множество различных функций, которые можно использовать для сравнения двух гистограмм. На практике чаще всего используются пересечение гистограмм и хи-квадрат. Чтобы сравнить гистограммы цветов двух изображений разного размера, важно сначала нормировать гистограммы.

Пересечение гистограмм - это самый простой и быстрый способ вычисления метрики расстояний. Оно определяется как

$$histint(H_1, H_2) = 1 - \sum_{i=1}^K \min(H_1(i), H_2(i)) \quad (1)$$

где  $H_1$  и  $H_2$  - это две гистограммы с  $K$  интервалами. Эта функция расстояния привлекательна тем, что может обрабатывать частичные совпадения, когда области двух гистограмм (сумма по всем столбцам) различны. Когда гистограммы нормированы (области двух гистограмм равны), пересечение гистограмм эквивалентно нормированному расстоянию  $L_1$ .

Другой часто используемой функцией является расстояние хи-квадрат. Оно более трудоемко с точки зрения вычислений, но работает зачастую лучше. Это расстояние определяется как

$$\chi^2(H_1, H_2) = \frac{1}{2} \sum_{i=1}^K \frac{(H_1(i) - H_2(i))^2}{H_1(i) + H_2(i)} \quad (2)$$

Основное преимущество расстояния хи-квадрат заключается в том, что оно рассчитывается на основе относительных значений. Таким образом, небольшая разница между двумя небольшими значениями будет иметь большее влияние, чем точно такая же небольшая разница между двумя большими значениями.

**Квантование пространства при построении гистограмм** Выбор функции расстояния для гистограмм - не единственное, что влияет на результаты. Еще один важный аспект - квантование. Хотя мы имеем дело с цифровыми изображениями и, следовательно, все значения уже дискретны, часто требуется более грубое квантование цветового пространства. Это уменьшит размер векторов признаков, а также может привести к лучшим конечным результатам, если функция расстояния сравнивает только ячейки (например, как в тех случаях, что мы обсуждали ранее).

Выбор правильной схемы квантования может быть непростым. Если используются функции, сравнивающие расстояния между ячейками, то слишком мелкозернистое квантование приведет к большим расстояниям даже между изображениями с похожими, но не совсем одинаковыми цветами. Слишком грубое квантование приведет к тому, что разные цвета будут отнесены к одной и той же ячейке, и, таким образом, мы получим небольшое расстояние между изображениями, которые на самом деле отличаются по цвету.

**Гистограммы в многомерных пространствах** Другой сложный аспект вопроса о квантовании и использовании гистограмм связан с многомерными пространствами. Цветовые пространства обычно трехмерны: координаты R, G, B в пространстве RGB, H, S, V - в HSV и т.д. В многомерном пространстве мы можем строить совместные или маргинальные гистограммы. Совместные гистограммы данных в N-мерном пространстве являются N-мерными. Мы просто помещаем точки данных в N-мерные ячейки. Совместные гистограммы часто имеют много пустых ячеек, особенно при мелкозернистом квантовании. Грубое квантование помогает уменьшить общее количество ячеек, и, таким образом, будет меньше пустых ячеек, но грубое квантование означает потерю разрешения, кроме того, в одной ячейке окажутся не слишком похожие значения.

Иногда вместо этого в многомерных пространствах используются маргинальные гистограммы. Размерность маргинальных гистограмм меньше, чем N-мерных гистограмм, поскольку одно или несколько измерений в них опущены. На экране вы видите двумерное пространство и пример совместной гистограммы слева и двух маргинальных гистограмм справа. Совместные и маргинальные гистограммы очень похожи на совместные и маргинальные

распределения. В маргинальных гистограммах мы получим больше точек данных в ячейке по сравнению с совместной гистограммой при одинаковых размерах ячеек для каждого признака.

**Квантование пространства при помощи кластеризации** Альтернативный способ квантования многомерного пространства - это кластеризация признаков. Количество ячеек определяется количеством кластеров. Этот метод допускает неоднородное квантование пространства признаков данных. Центры кластеров обычно определяются с помощью набора доступных примеров изображений. Затем, когда необходимо обработать новое ранее не встречавшееся изображение, его признаки распределяются по ближайшим ячейкам.

На экране представлено двумерное пространство признаков. Мы можем квантовать это пространство на две ячейки, используя алгоритм кластеризации с двумя кластерами. Точки  $X$  на изображении, показанном на экране, - это центры двух кластеров.

Квантование с помощью кластеризации - хорошая альтернатива для пространств признаков с большой размерностью. Оно не часто используется для простых цветовых гистограмм, потому что цветовые пространства обычно имеют только три измерения. Но позже мы рассмотрим по крайней мере один пример, когда этот метод будет использоваться для других пространств признаков.

## 1.6 Сравнение цветов

### 1.6.1 Гистограммы

Гистограммы с функциями расстояния между ячейками - мощный и часто используемый инструмент для сравнения цветов двух изображений. Но он не идеален. Даже с использованием интеллектуальных методов квантования цветового пространства этот метод все равно не принимает во внимание сходство между различными цветовыми ячейками. Человек ожидает, что расстояние между гистограммами  $H_1$  и  $H_2$  будет меньше, чем расстояние между гистограммами  $H_1$  и  $H_3$  представленными на экране, потому что, хотя  $H_1$  и  $H_2$  не имеют пересекающихся интервалов, общее цветовое распределение воспринимается человеком как аналогичное, т.к. есть похожие цвета. При использовании любой функции расстояния между ячейками расстояние между  $H_1$  и  $H_2$  будет больше, чем расстояние между  $H_1$  и  $H_3$ , потому что  $H_1$  и  $H_3$  имеют одну общую ячейку.

Один из возможных способов решить эту проблему - использовать кумулятивные гистограммы и по-прежнему использовать любую функцию расстояния между ячейками на ваш выбор. Кумулятивная гистограмма - это

гистограмма, в которой каждый интервал подсчитывает значения для этого интервала плюс значения всех предыдущих интервалов. Кумулятивная гистограмма, соответствующая гистограмме  $H_1$  из нашего предыдущего примера, будет выглядеть следующим образом. Первый интервал будет таким же как и в исходной гистограмме, второй интервал содержит сумму первого и второго интервалов исходной гистограммы и так далее. Поскольку второй, третий и четвертый интервалы в исходной гистограмме пусты, первые четыре интервала совокупной гистограммы будут иметь одинаковое значение. Затем в пятый интервал нам нужно добавить содержимое предыдущего, т.е. четвертого интервала, и пятого. И так далее. Цвета на экране сохранены, поэтому легко проследить создание кумулятивной гистограммы.

Итак, кумулятивная гистограмма, соответствующая исходной гистограмме  $H_1$ , выглядит так, как показано на экране. Теперь построим кумулятивные гистограммы, соответствующие  $H_2$  и  $H_3$ . Они окрашены в разные цвета, чтобы было легче сравнивать. Как вы видите, простое расстояние между ячейками, например,  $L_1$ , теперь дает результат, который лучше соответствует интуитивно ожидаемому результату. Расстояние между кумулятивными гистограммами  $H_1$  и  $H_2$  меньше, чем между  $H_1$  и  $H_3$ .

Важно отметить, что этот подход будет работать только в том случае, если соседние ячейки гистограммы соответствуют подобным цветам.

Можно также использовать расстояние Вассерштейна (Earth Mover's Distance или EMD), чтобы сравнить сходство между ячейками. Оно также известно как метрика Канторовича – Рубинштейна. Интуитивно понятно, что если каждую гистограмму рассматривать как некоторое количество земли, распределенной по ячейкам, метрикой является минимальная «стоимость» превращения одной кучи в другую, которая, как предполагается, равна количеству земли, которую необходимо переместить, умноженной на расстояние, на которое она должна быть перемещена. Итак, возвращаясь к нашему исходному примеру, все три гистограммы,  $H_1$ ,  $H_2$  и  $H_3$  имеют три единицы «земли» в трех разных ячейках. Чтобы преобразовать  $H_1$  в  $H_2$ , нам нужно, чтобы содержимое всех трех ячеек было сдвинуто на шаг вправо. Таким образом, стоимость переезда составляет 3 шага. Чтобы преобразовать  $H_1$  в  $H_3$ , красная ячейка остается на месте, но две ячейки - светло-зеленую и голубую - нужно переместить на большее расстояние: 3 шага для светло-зеленой и 6 шагов для голубой. Итак, всего 9 шагов. Расстояние Вассерштейна между  $H_1$  и  $H_2$  меньше, чем между  $H_1$  и  $H_3$ , что ближе к восприятию человеком.

Как и кумулятивная гистограмма, расстояние Вассерштейна «работает» как ожидается только в том случае, если интервалы соседних столбцов гистограммы соответствуют похожим значениям (похожим цветам в случае цветовых гистограмм).

Другой вариант - использовать иную функцию расстояния, которая мо-

жет измерять сходство между ячейками. Примером такой функции расстояния является расстояние в квадратичной форме, показанное на экране. Это расстояние было предложено в статье Ниблака и др. (1993) для поиска изображений на основе цвета. Это обобщенное евклидово расстояние с матрицей  $A$ , которая включает в себя информацию о сходстве между ячейками. Она содержит коэффициенты подобия для всех ячеек разбитого на части пространства.

**Пространственное расположение** Еще одна проблема с цветовыми гистограммами заключается в том, что они не учитывают пространственное расположение цветов. Они представляют собой распределение цвета без какой-либо пространственной информации. Например, цветовые гистограммы показанных на экране трех синтетических изображений совершенно одинаковы, т.к. все три изображения имеют одинаковое количество белых и черных пикселей.

Рассмотрим еще один пример трех разных изображений с одинаковой цветовой гистограммой, которые, однако, сильно различаются и не будут восприниматься человеком как похожие.

**Пространственное расположение цветов** На заре компьютерного зрения люди иногда разбивали изображения на определенные или нечеткие блоки и строили цветовые гистограммы для каждого блока независимо. Этот прием добавляет пространственную информацию к цветовым характеристикам, но он не адаптивен - позиции блоков фиксированы для всех изображений.

Вот почему одного цвета недостаточно для сравнения изображений: в сочетании с цветом часто используются текстура и другие признаки.

**Гистограммы и квантование - выводы** Подводя итог, цветовые гистограммы - это очень простой, но в то же время наглядный метод описания цветов. Метод цветовых гистограмм был разработан одним из первых, и по-прежнему является один из наиболее важных методов описания. Несмотря на то, что в задачах распознавания изображений метод цветовых гистограмм был в основном заменен более сложными методами, он может быть достаточно полезным для определенных наборов изображений. Гистограммы до сих пор широко используются для описания изображения.

Для правильной работы гистограмм важен выбор квантования. С меньшим количеством интервалов вероятность того, что две гистограммы будут иметь ненулевые значения для одних и тех же интервалов, выше, даже когда точек данных не так много, поэтому метрики поинтервального сравнения будут работать лучше. Но это также означает более грубое представление

данных, когда разные значения могут быть помещены в один и тот же интервал.

При большом количестве интервалов мы получаем более детальное представление данных, но становится сложнее сопоставить гистограммы - при небольшом количестве точек данных многие интервалы будут пустыми, а две гистограммы могут иметь нулевое пересечение, поэтому их будет невозможно сравнить используя поинтервальные метрики сравнения. Расстояние между двумя изображениями с похожими, но не идентичными цветами будет большим. Гистограммы с большим количеством интервалов очень разреженные и, следовательно, чувствительны к шуму.

Квантование также может быть равномерным и независимым от данных, например, квантование по сетке, когда мы разделяем все пространство на ячейки. Для многомерного пространства это будет декартовым произведением однородных ячеек в каждом измерении. Это легко реализовать, но это приведет к очень большому количеству ячеек в случае многомерного пространства. И поскольку это разделение не зависит от данных, многие ячейки могут оказаться пустыми.

Другой способ разбить пространство - это использовать кластеризацию. Этот подход основан на данных и поэтому разбивает пространство на неравномерные ячейки в соответствии с распределением данных. Это более затратно с точки зрения вычислений: чтобы найти центры ячеек, нам нужно сгруппировать репрезентативное подмножество данных. Но в то же время для пространств большой размерности такой метод может привести к уменьшению количества ячеек при достаточной степени детализации.

**Сравнение гистограмм: выводы** Для вычисления расстояния между двумя гистограммами можно использовать разные функции. Большинство используемых в настоящее время функций расстояния - это функции расстояния между ячейками, которые не учитывают сходство между ячейками. Наиболее часто используемые метрики - это пересечение гистограмм, которое вычисляется быстро, и расстояние хи-квадрат, которое вычислить сложнее, но этот метод зачастую работает лучше, чем простое пересечение гистограммы или  $L_1$ .

В случае, когда соседние ячейки гистограммы представляют аналогичные значения, для сравнения ячеек можно использовать кумулятивные гистограммы со стандартными функциями расстояния между ячейками или переключиться на расстояние движителя земли.

Квадратичное расстояние - это более общая функция расстояния, которую можно использовать для сравнения ячеек. Оно основано на произвольном подобии ячеек, коэффициенты подобия указаны в матрице  $A$ .

## 1.7 Текстуры

Несмотря на то, что признаки цвета несут важную информацию о содержанием изображения, часто этого недостаточно. Еще одна важная для зрительного восприятия человеком характеристика изображения - это текстура. Она дает нам информацию о структурном расположении поверхностей и объектов на изображении. Мы не будем обсуждать текстуры так же подробно, как признаки цвета, но обсудим пару признаков текстуры, которые чаще всего использовались ранее.

Что такое текстура? Любой человек способен распознать текстуру; однако дать формальное определение этому понятию сложно. Описание текстуры интуитивно определяет такие свойства, как гладкость, грубость и периодичность. Давайте посмотрим на три примера на экране. Белые квадраты на трех изображениях на экране показывают (слева направо) гладкие, грубые и периодические текстуры. Описание текстуры дает количественную оценку воспринимаемой текстуры изображения. Текстура изображения содержит информацию о пространственном расположении цвета или интенсивности всего изображения или определенной его области. Текстуру невозможно определить для отдельного пикселя, она зависит от распределения интенсивности в изображении.

Анализ текстуры играет важную роль для сравнения изображений, дополняя цветовые характеристики. Существует множество методов представления текстуры и сравнения изображений на основе текстуры. Текстурные признаки можно разделить на статистические, геометрические, модельные и спектральные.

Статистические признаки описывают распределение интенсивности в изображении с помощью различных статистических параметров. Признаки этой категории являются одними из первых текстурных признаков, использовавшихся в компьютерном зрении. Один из простейших подходов к описанию текстуры - использование статистических моментов гистограммы интенсивности изображения или его области. Другие статистические признаки текстуры включают в себя признаки, основанные на матрицах смежности, а также гистограммы текстур, построенные на признаках Тамуры.

Геометрические признаки описывают текстуру через элементы текстуры или примитивы. В этот класс признаков входят методы представления текстуры путем использования диаграмм Вороного и структурных методов. Последние описывают текстуру через определение примитивов текстуры и правил размещения этих примитивов относительно друг друга. Структурные методы не подходят для описания нерегулярных текстур.

Модельные методы текстурного анализа основаны на построении модели, которую можно использовать не только для описания, но и для создания

текстуры. К этой категории относятся признаки на основе марковских случайных полей и фракталов.

Спектральные характеристики описывают текстуру в частотной области. Они часто основаны на разложении сигнала на взвешенную сумму базисных функций и использовании весов как элементов вектора признаков. Среди спектральных подходов особого внимания заслуживают фильтры Габора. Фильтры Габора широко используются во множестве приложений для обработки изображений. Эти фильтры позволяют аппроксимировать характеристики определенных клеток зрительной коры некоторых млекопитающих. Фильтры Габора обладают оптимальными свойствами локализации как в пространственной, так и в частотной области и, таким образом, хорошо подходят для задач сегментации текстуры.

### 1.7.1 Фильтры Габора

Фильтры Габора - это линейные фильтры, используемые во многих приложениях обработки изображений для обнаружения границ, анализа текстуры, выделения признаков и т.д. С помощью фильтров Габора можно проанализировать, есть ли в изображении какое-либо конкретное частотное содержание в определенном направлении. Фильтры Габора - это особые классы полосовых фильтров, то есть они пропускают определенную «полосу» частот и отклоняют другие. Многие ученые, работающие в области компьютерного зрения утверждают, что частотные и пространственные представления фильтров Габора аналогичны представлениям зрительной системы человека.

В пространственной области двумерный фильтр Габора представляет собой гауссову функцию ядра, модулированную синусоидальной волной определенной частоты и ориентации. Один из таких двумерных фильтров Габора показан на экране: двумерное ядро Гаусса, модулированное синусоидой, ориентированной на 30° от оси X. Синусоиду иногда называют несущей, а двумерную функцию гауссовой формы называют огибающей. Фильтрация изображения с помощью такого фильтра обнаружит наличие краев в определенном направлении. На практике обычно используется набор фильтров Габора с разными коэффициентами сжатия (растяжения) и разных направлений. Этот набор фильтров часто называют банком фильтров.

Например, сейчас на экране показан набор фильтров Габора из 40 фильтров для 5 коэффициентов сжатия и 8 направлений. Проще говоря, фильтры Габора обнаруживают градиенты изображения определенной направленности. Свертка участков изображения с множеством различных частот и направлений позволяет извлечь информацию о локальной текстуре и закодировать ее в низкоразмерный вектор признаков.

Фильтры сворачиваются с исходным изображением, в результате получа-



ется так называемое пространство Габора. Затем можно вычислить значения энергии отклика для каждого фильтра, и сформировать из значений энергии вектор признаков текстуры для данного изображения.

### 1.7.2 Фильеры ICA

Хотя фильтры Габора считаются хорошим приближением к структуре работы определенных клеток зрительной коры головного мозга человека, они не могут изменяться в зависимости от данных. Одним из примеров фильтров, который принимает в расчет изначальные данные (появившихся до широкого распространения глубокого обучения), являются фильтры ICA, где ICA означает анализ независимых компонентов. Эти фильтры также можно использовать для анализа текстуры. Они построены с использованием независимого компонентного анализа (independent component analysis). Репрезентативный набор изображений используется для «изучения» «независимых компонентов». В своей простейшей форме ICA - это алгоритм поиска линейного преобразования, минимизирующего статистическую зависимость между компонентами входного вектора. Когда ICA применяется к естественным изображениям, этот метод создает наборы визуальных фильтров, которые работают как простые клетки первичной зрительной коры. Они пространственно локализованы, ориентированы и избирательны к структуре в различных пространственных масштабах. Эти фильтры являются примитивами, которые представляют собой пространственные шаблоны, встречающиеся в разных сценах. Каждый фрагмент изображения может быть представлен как комбинация наборов таких примитивов. На экране показаны примеры фильтров ICA, извлеченные авторами из фрагментов естественных изображений, найденных в Интернете.

**Фильтры ICA - модель ответа** После того, как фильтры получены, каждое изображение может быть охарактеризовано с помощью  $N$  общих фильтров как набор из  $N$  ответов. Энергетические отклики изображения  $I(x, y)$  при выбранным наборе фильтров оцениваются как квадраты свертки изображения с каждым фильтром. Операция возведения в квадрат, которая соответствует энергии ответа, необходима для устранения внутренней неоднозначности в отношении знака сигналов, которые оцениваются с помощью ICA. Вместо возведения в квадрат можно также использовать абсолютное значение.

Итак, после фильтрации с помощью  $N$  фильтров у нас есть  $N$  ответов. Теперь мы можем выбрать разные дескрипторы (или сигнатуры): использовать только среднее значение ответов, среднее значение и дисперсию или гистограмму. Гистограммы содержат больше информации о распределении,

поэтому это представление и было выбрано авторами.

Текстуры двух изображений теперь можно сравнивать путем сравнения распределения ответов этих изображений на фильтры ИСА. Авторы фильтров ИСА предложили использовать расхождение Кульбака-Лейблера для сравнения распределений ответов. Это еще одна часто используемая мера, используемая для сравнения распределения вероятностей. Иногда ее также называют относительной энтропией. Она определяется как:

$$KL(f_1, f_2) = - \int_{-\infty}^{\infty} f_1(x) \log \left( \frac{f_1(x)}{f_2(x)} \right) dx \quad (3)$$

Благодаря вогнутости логарифмической функции эта мера положительна, когда  $f_1$  и  $f_2$  различны, и равна нулю, если  $f_1$  равно  $f_2$ . Но она не симметрична и не удовлетворяет неравенству треугольника. Первый недостаток можно решить с помощью использования симметричного отображения этой меры:

$$KL_s(f_1, f_2) = KL(f_1, f_2) + KL(f_2, f_1) \quad (4)$$

Если в качестве сигнатур изображений используются гистограммы, то различие между изображениями, представленными гистограммами  $H_1$  и  $H_2$ , вычисляется как:

$$\begin{aligned} KL_s(H_1, H_2) &= \sum_{b=1}^B H_1(b) \log \left( \frac{H_1(b)}{H_2(b)} \right) + \sum_{b=1}^B H_2(b) \log \left( \frac{H_2(b)}{H_1(b)} \right) \\ &= \sum_{b=1}^B ((H_1(b) - H_2(b)) \log \left( \frac{H_1(b)}{H_2(b)} \right)) \end{aligned} \quad (5)$$

Здесь  $B$  - это общее количество интервалов в гистограммах: обе гистограммы  $H_1$  и  $H_2$  имеют одинаковое количество интервалов.

Но каждое изображение характеризуется не одной гистограммой, а набором из  $N$  гистограмм. Каждый ответ можно рассматривать как плотность независимой переменной, а набор из  $N$  ответов как многомерное распределение. Расхождение Кульбака-Лейблера между двумя многомерными распределениями с независимыми компонентами представляет собой сумму расхождений Кульбака-Лейблера между каждыми двумя компонентами. Таким образом, для каждого изображения, представленного гистограммами  $N$ , несходство между изображениями можно вычислить как:

$$KL_I(I_1, I_2) = \sum_{i=1}^N KL_s(H_{1,i}, H_{2,i}) \quad (6)$$

Итак, для каждого фильтра  $i = 1 \dots N$  мы вычисляем расхождение Кульбака-Лейблера между ответами изображения  $I_1$  и изображения  $I_2$  на этот фильтр, а затем суммируем все эти различия для каждого фильтра.

Мы рассмотрели несколько примеров глобальных признаков, используемых для анализа для цвета и текстуры. Существует большое количество разных подходов к описанию цвета и текстуры. Также было предложено несколько дескрипторов формы. Но в последние годы с развитием глубокого обучения все меньше и меньше подобных признаков используется на практике.