

Основы обработки мультимедийных данных с онлайн курсом "Обработка изображений"

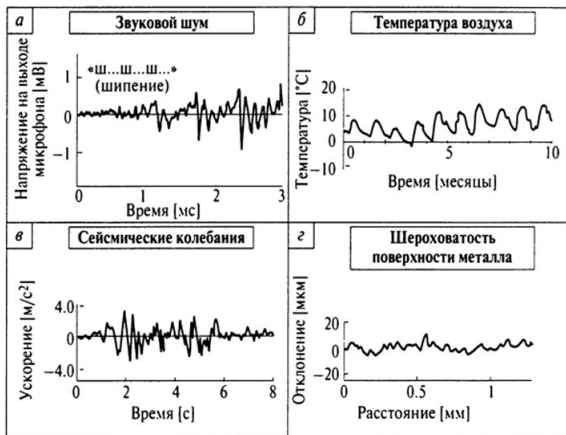
М.В. Хлопотов

Университет ИТМО

СПб, 2020

Определение

Физическая величина, которая содержит в себе определённую информацию называется *сигналом*.



Рассмотренные выше сигналы имеют одну независимую переменную. Однако существуют сигналы, имеющие более одной независимой переменной. Например, *сигнал изображения* - это функция яркости точки в зависимости от её координат. Сигнал изображения и его представление в цифровом виде:

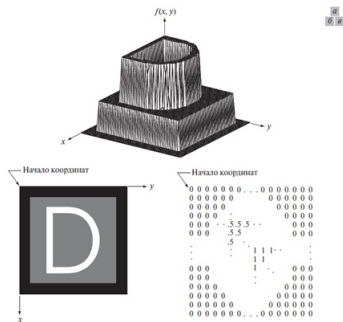


Рис. 2.18. (а) График изображения как объемной поверхности. (б) Отображение в форме массива видимых яркостей. (в) Изображение в виде числовой матрицы (значения 0, 0,5 и 1 представляют соответственно черное, серое и белое)

Определение

Мультимедиа - комплекс аппаратных и программных средств, позволяющих пользователю работать в диалоговом режиме с разнородными данными (графика, текст, звук, видео), организованными в виде единой информационной среды. В более узком смысле - это программное и аппаратное обеспечение, предназначенное для создания, хранения и передачи информации, представленной в различных цифровых форматах.

Одна из основных проблем передачи и хранения мультимедиа сигналов (речь, изображение, видео) состоит в том, что эти сигналы изначально являются аналоговыми.

Аналоговые сигналы - это непрерывные функции непрерывного аргумента, например, время. *Дискретный сигнал* может быть дискретным по множеству значений функции и(или) по множеству аргумента.

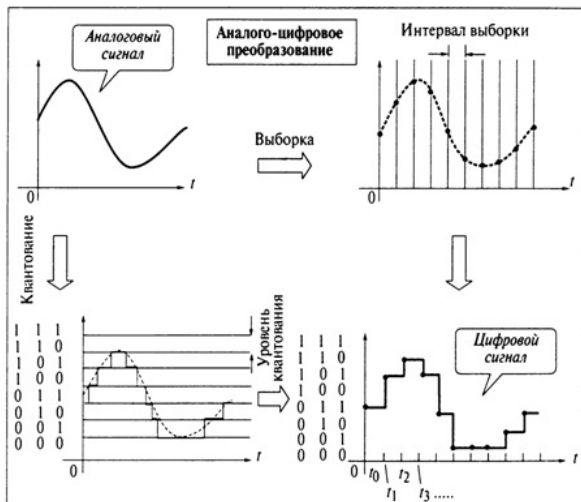
Аналого-цифровое преобразование (АЦП) включает в себя две операции:

- дискретизация, т.е. взятие отсчётов непрерывного сигнала;
- квантование непрерывного сигнала по уровню.

Вопрос 1

Что такое дискретизация для изображений?

Аналого-цифровое преобразование



Аналого-цифровое преобразование

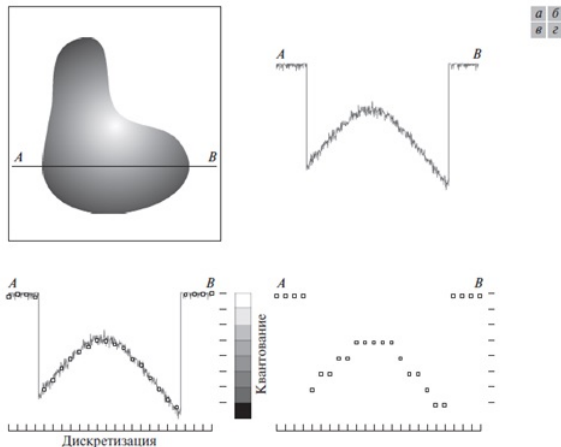


Рис. 2.16. Формирование цифрового изображения. (а) Непрерывное изображение. (б) Профиль вдоль линии сканирования между точками A и B на непрерывном изображении, который используется для иллюстрации понятий дискретизации и квантования. (в) Дискретизация и квантование. (г) Цифровое представление строки изображения

Аналого-цифровое преобразование

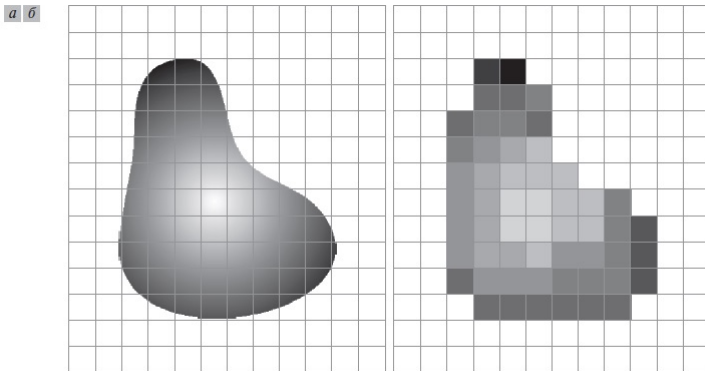


Рис. 2.17. (а) Проекция непрерывного изображения на матрицу чувствительных элементов. (б) Результат дискретизации и квантования изображения

Сжатие мультимедийных данных

Число битов, требуемое для хранения таких сигналов очень велико. Поэтому возникает проблема сжатия мультимедийных данных для того, чтобы передавать их быстрее и хранить эффективнее.

Вопрос 2

Сколько байт требуется для хранения несжатого цветного видео длиной полтора часа с расширением 1280x720?

Системы сжатия мультимедийных данных делят на системы сжатия без потерь и с потерями. В настоящем курсе рассматриваются теоретические аспекты сжатия с потерями: сжатие на основе преобразований (JPEG, CELP).

Кроме сжатия мультимедийных данных в курсе рассматривается их обработка и анализ изображений, видео и аудио.

Вопрос 3

Чем обработка данных отличается от анализа?

Многие методы обработки цифровых сигналов могут применяться к данным разного типа. В курсе наибольшее внимание будет уделено обработке и анализу изображений.

Обработка и анализ изображений

Типичные задачи *обработки изображений (digital image processing, DIP)*:

- Геометрические преобразования.
- Цветовая коррекция.
- Сравнение двух и более изображений.
- Комбинирование изображений.
- Интерполяция и сглаживание.
- Разделение изображения на области.
- Редактирование и ретуширование.
- Компенсация потери резкости.

Обработка и анализ изображений

Обработка - начальный этап *анализа изображений*. К числу задач анализа относят, например:

- Распознавание текста
- Обработка данных для выделения различных характеристик интересующего нас объекта
- Обработка изображений в медицине
- Идентификация личности (по лицу, радужке, дактилоскопическим данным)
- Определение перемещения объекта
- Наложение фильтров
- Обработка спутниковых снимков

Формирование цифровых изображений

Предположим, что непрерывное изображение преобразовано в двумерный массив $f(x, y)$ — матрицу из M строк и N столбцов, где (x, y) — дискретные координаты.

Для ясности обозначений и большего удобства мы будем использовать для этих координат целочисленные значения: $x = 0, 1, 2, \dots, M - 1$ и $y = 0, 1, 2, \dots, N - 1$, принимая за начало координат левый верхний угол изображения, где $(x, y) = (0, 0)$. Следующим значением координат вдоль первой строки изображения будет точка $(x, y) = (0, 1)$.

Формирование цифровых изображений

Для выполнения процесса оцифровки изображения необходимо принять решения относительно значений M и N , а также числа дискретных уровней (градаций) яркости L . Для M и N не устанавливается специальных ограничений помимо того, что они должны быть положительными целочисленными значениями. Однако значение L по соображениям удобства построения оборудования для хранения и квантования обычно выбирают равным целочисленной степени двойки: $L = 2^k$.

Вопрос 4

Чему обычно равно k ?

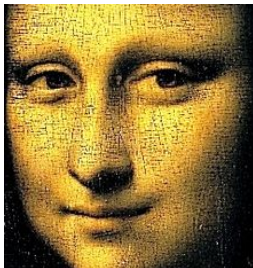
Формирование цифровых изображений

Дискретные уровни яркости расположены с постоянным шагом (т. е. используется равномерное квантование) и принимают целые значения в интервале $[0, L - 1]$. Иногда интервал значений яркости неформально называют динамическим диапазоном изображения (отношение максимально возможного значения измеренной яркости к минимальному уровню яркости, обнаруживаемому этой системой).

Расширение динамического диапазона путём комбинирования изображений с разной экспозицией (HDR) - интересная тема, которая находится за рамками курса.

Основные характеристики изображения

К числу основных характеристик цифровых изображений относятся рассмотренные выше M , N , L .



Вопрос 5

Какие ещё?

Пространственное и яркостное разрешение

Пространственное разрешение — это размер мельчайших различимых деталей на изображении. Количественно пространственное разрешение можно определять многими способами; к наиболее употребительным относятся число пар линий на единицу длины и число точек (пикселей) на единицу длины.

Число точек на единицу длины — мера разрешения изображения, которая повсеместно применяется в полиграфии и издательском деле. Эту характеристику обычно выражают в точках на дюйм (dpi).

Уменьшение пространственного разрешения

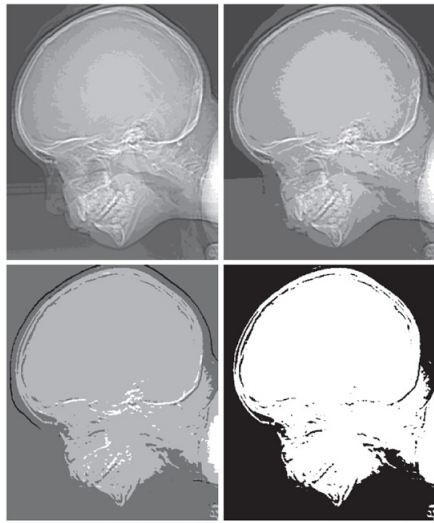


Пространственное и яркостное разрешение

Яркостное (или полутоновое) разрешение — мельчайшее различимое изменение яркости.

При выборе числа градаций яркости приходится в значительной степени учитывать особенности аппаратуры. По этой причине число градаций обычно выбирается равным степени 2. Наиболее частым решением является выбор 8-битового представления (256 градаций яркости), но в некоторых приложениях используется 16 бит, если необходимо иметь более точное представление полутонов. Квантование уровней яркости с точностью 32 бит используется редко.

Уменьшение яркостного разрешения



Методы обработки изображений

Методы обработки изображений классифицируют обычно по количеству пикселей участвующих в одном шаге преобразования:

- *поточечные методы* в процессе выполнения преобразуют значение в точке в значение независимо от соседних точек;
- *локальные (окрестностные) методы* для вычисления значения используют значения соседних точек в окрестности;
- *глобальные методы* определяют значение на основе всех значений исходного изображения.

Вопрос 6

К каким методам (поточечным, локальным или глобальным) относятся:

- уменьшение (увеличение) яркостного и пространственного разрешения?
- сжатие изображений?
- применение фильтров?
- ретуширование изображений?
- увеличение яркости?
- увеличение контрастности?

Лабораторная работа 1.1

Предварительная подготовка: Обработка изображений (ВШЭ)

Поточечные методы обработки

В цветовом пространстве greyscale требуется реализовать алгоритм скалярного равномерного квантования с переменной скоростью для постепенного уменьшения уровней яркости изображения. Сохранить изображение, в котором количество уровней яркости $\ll 256$, при этом визуально приемлемое качество.

В отчёт по лабораторной работе необходимо включать программный код, входные данные (изображение), выходные данные (обработанные изображения), выводы.

Лабораторная работа 1.2

Задача децимации (прореживания)

В технологии JPEG применяется формат YUV для промежуточного представления изображения. В этом формате два канала (2 и 3) хранят информацию о цвете, а один канал - о яркости. Требуется написать алгоритм прореживания (децимации) цветоразностных каналов. Необходимо реализовать отдельно кодирование и декодирование.

В отчёт по лабораторной работе необходимо включать программный код, входные данные (изображение), выходные данные (обработанные изображения), выводы.

Лабораторная работа 1.2

RGB to YUV

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

$$Cb = -0.1687 * R - 0.3313 * G + 0.5 * B + 128$$

$$Cr = 0.5 * R - 0.4187 * G - 0.0813 * B + 128$$

YUV to RGB

$$R = Y + 1.402 * (Cr - 128)$$

$$G = Y - 0.34414 * (Cb - 128) - 0.71414 * (Cr - 128)$$

$$B = Y + 1.772 * (Cb - 128)$$

Лабораторная работа 1.3

Энтропия (поговорим об этом на следующих занятиях):

$$H = -p_i \log p_i$$

Среднеквадратичная ошибка: $MSE = \frac{1}{n} \sum_{i=1}^n (P_i - Q_i)^2$

Энтропия и MSE

Для оценки работы алгоритмов в заданиях 1.1 и 1.2, необходимо посчитать энтропию и среднеквадратичную ошибку.

В отчёт по лабораторной работе необходимо включать программный код, результаты и выводы.