УНИВЕРСИТЕТ ИТМО

С.В. Шаветов

ОСНОВЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ: ЛАБОРАТОРНЫЙ ПРАКТИКУМ

Санкт-Петербург

2017

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

УНИВЕРСИТЕТ ИТМО

С.В. Шаветов

ОСНОВЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ: ЛАБОРАТОРНЫЙ ПРАКТИКУМ

РЕКОМЕНДОВАНО К ИСПОЛЬЗОВАНИЮ
В УНИВЕРСИТЕТЕ ИТМО
по направлению подготовки 15.03.06 Мехатроника и робототехника
в качестве учебно-методического пособия для реализации
основных профессиональных образовательных программ
высшего образования бакалавриата

Санкт-Петербург
2017

С.В. Шаветов. Основы технического зрения: лабораторный практикум. — СПб: Университет ИТМО, 2017. — 86 с.

Рецензенты:
С.М. Власов, канд. техн. наук, ассистент, Университет ИТМО.
В.В. Сыроквашин, канд. техн. наук, руководитель направления, ООО «Вексон».

Учебно-методическое пособие посвящено основам обработки и анализа цифровых изображений. Представлены шесть лабораторных работ с последовательно увеличивающейся сложностью, начиная от представления изображений как массива чисел до сегментации изображения на семантические области. Пособие предназначено для студентов, обучающихся по направлению подготовки бакалавров 15.03.06 Мехатроника и робототехника.

Пособие подготовлено при финансовой поддержке Министерства Образования и Науки Российской Федерации (Проект 2.8878.2017/8.9).

УНИВЕРСИТЕТ ИТМО

Университет ИТМО — ведущий вуз России в области информационных и фотонных технологий, один из немногих российских вузов, получивших в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО — участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Цель Университета ИТМО — становление исследовательского университета мирового уровня, предпринимательского по типу, ориентированного на интернационализацию всех направлений деятельности.

Содержание

Лабораторная работа №1
Гистограммы, профили и проекции

Цель работы

Освоение основных яркостных и геометрических характеристик изображений и их использование для анализа изображений.

Методические рекомендации

До начала работы студенты должны ознакомиться с основными функциями среды MATLAB по работе с гистограммами, профилями и проекциями изображений. Лабораторная работа рассчитана на 5 часов.

Теоретические сведения

Пиксель цифрового изображения характеризуется тремя параметрами: (x,y,I), где пара целочисленных значений (x,y) описывает геометрическое положение пикселя в плоскости изображения, а значение I характеризует его яркость (интенсивность) в точке плоскости. Таким образом, в изображении можно выделить яркостную и геометрическую составляющие. В общем случае данные составляющие не связаны между собой (например, изменение в освещенности сцены не изменит геометрических параметров объектов на сцене). Из-за этого проще исследовать отдельно яркостные свойства изображения и отдельно — геометрические. Такой подход понижает порядок исследуемого изображения в случае геометрических свойств с $n = 3$ (x,y,I) до $n = 2$ (x,y), а в случае яркостных свойств — до $n = 1$ (I). Яркостная составляющая изображения характеризуется одномерным массивом гистограммы, из которого можно вычислить контраст.

Гистограмма — это распределение частоты встречаемости пикселей одинаковой яркости на изображении.

Яркость — это среднее значение интенсивности сигнала.

Контраст — это интервал значений между минимальной и максимальной яркостями изображения.

Для сведéния геометрических составляющих изображения к одномерному массиву данных $n = 1$ используются такие характеристики, как *«профили»* и *«проекции»* изображения.

Профиль вдоль линии — это функция интенсивности изображения, распределенного вдоль данной линии *(прорезки)*.

Проекция на ось — это сумма интенсивностей пикселей изображения взятая в направлении перпендикулярном данной оси.

Гистограмма изображения

Для 8-битного полутонового изображения гистограмма яркости представляет собой одномерный целочисленный массив Hist из 256 элементов $[0 \dots 255]$. Элементом гистограммы $\mathrm{Hist}[i]$ является сумма пикселей изображения с яркостью i. По визуальному отображению гистограммы можно оценить необходимость изменения яркости и контрастности изображения, оценить площадь, занимаемую светлыми и темными элементами, определить местоположение на плоскости изображения отдельных объектов, соответствующих некоторым диапазонам яркости. Для цветного RGB-изображения необходимо построить три гистограммы по каждому цвету.

Листинг 1.1. Построение гистограмм изображения.

```
1 [numRows numCols Layers] = size(I);
2 imhist(I(:,:,1)); %red
3 imhist(I(:,:,2)); %green
4 imhist(I(:,:,3)); %blue
```

Если гистограмма неравномерна, то для улучшения изображения можно ее выровнять, причем выравнивание гистограммы в зависимости от решаемой задачи можно выполнять различным образом.

Арифметические операции

Простейшими способами выравнивания гистограммы являются *арифметические операции* с изображениями. Например, в случае, если большинство значений гистограммы находятся слева, то изображение является темным. Для увеличения детализации темных областей можно сдвинуть гистограмму правее, в более светлую область, например, на 50 градаций для каждого цвета. В среде MATLAB программная реализация имеет вид:

```
Inew(i,j) = I(i,j) + 50 / 255;
```

Гистограмма исходного изображения сдвигается в среднюю часть диапазона, которая является более приемлемой с точки зрения визуального восприятия. Данный подход обладает следующим недостатком: повышение интенсивностей темных областей приводит к сдвигу светлых к максимуму, что может привести к потере информации в светлых областях.

Растяжение динамического диапазона

Если интенсивности пикселей областей интереса находятся в узком динамическом диапазоне, то можно растянуть этот диапазон. Подобные преобразования выполняются согласно следующему выражению:

$$I_{new} = \left(\frac{I - I_{min}}{I_{max} - I_{min}} \right)^{\alpha}, \tag{1.1}$$

где I и I_{new} — массивы значений интенсивностей исходного и нового изображений соответственно; I_{min} и I_{max} — минимальное и максимальное значения интенсивностей исходного изображения соответственно; α — коэффициент нелинейности.

Данное выражение является нелинейным из-за коэффициента α. В случае, если $\alpha = 1$, применение формулы (1.1) к исходному изображению не даст желаемого эффекта, поскольку гистограммы цветовых компонент изображения занимают весь возможный диапазон. Нелинейные преобразования проводятся для каждой цветовой составляющей.

Листинг 1.2. Нелинейное растяжение динамического диапазона при $\alpha = 0{,}5$.

```
1 [numRows numCols Layers] = size(I);
2 for k=1:1:Layers;
3     Imin = min(min(I(:, :, k)));
4     Imax = max(max(I(:, :, k)));
5     for i=1:1:numRows;
6         for j=1:1:numCols;
7             Inew(i,j,k) = ...
8                 (((I(i,j,k) - Imin) / ...
9                     (Imax - Imin)))^0.5;
```

```
10              if Inew(i,j,k) > 1;
11                  Inew(i,j,k) = 1;
12              end
13              if Inew(i,j,k) < 0;
14                  Inew(i,j,k) = 0;
15              end
16          end
17      end
18  end
```

Равномерное преобразование

Осуществляется по следующей формуле:

$$I_{new} = (I_{max} - I_{min}) \cdot P(I) + I_{min}, \tag{1.2}$$

где I_{min}, I_{max} — минимальное и максимальное значения интенсивностей исходного изображения I; $P(I)$ — функция распределения вероятностей исходного изображения, которая аппроксимируется *кумулятивной гистограммой*:

$$P(I) \approx \sum_{m=0}^{i} \mathrm{Hist}(m). \tag{1.3}$$

Кумулятивная гистограмма исходного изображения в среде MATLAB может быть вычислена с помощью функции `cumsum()`:

```
CH = cumsum(H) ./ (numRows * numCols);
```

где `H` — гистограмма исходного изображения, `numRows` и `numCols` — число строк и столбцов исходного изображения соответственно.

Согласно формуле (1.2) можно вычислить значения интенсивностей пикселей результирующего изображения. В среде MATLAB программная реализация имеет вид:

```
Inew(i,j) = CH(ceil(255 * I(i,j) + eps));
```

Параметр `eps` необходим для защиты программы от присваивания индексам кумулятивной гистограммы нулевых значений.

Экспоненциальное преобразование

Осуществляется по следующей формуле:

$$I_{new} = I_{min} - \frac{1}{\alpha} \cdot \ln(1 - P(I)), \tag{1.4}$$

где α — постоянная, характеризующая крутизну преобразования. Согласно формуле (1.4) можно вычислить значения интенсивностей пикселей результирующего изображения. В среде MATLAB программная реализация имеет вид:

```
Inew(i,j) = Imin -(1 / alfa) * ...
    log10(1 - CH(ceil(255 *  I(i,j) + eps)));
```

Преобразование по закону Рэлея

Осуществляется по следующей формуле:

$$I_{new} = I_{min} + \left(2\alpha^2 \frac{1}{1 - P(I)}\right)^{1/2}, \tag{1.5}$$

где α — постоянная, характеризующая гистограмму распределения интенсивностей элементов результирующего изображения. В среде MATLAB программная реализация имеет вид:

```
Inew(i,j) = Imin + sqrt(2 * alfa^2 * ...
    log10(1 / (1 - CH(ceil(255 * I(i,j) + eps)))));
```

Преобразование по закону степени $2/3$

Осуществляется по следующей формуле:

$$I_{new} = (P(I))^{2/3}. \tag{1.6}$$

В среде MATLAB программная реализация имеет вид:

```
Inew(i,j) = (CH(ceil(255 * I(i,j) + eps)))^(2/3);
```

Гиперболическое преобразование

Осуществляется по следующей формуле:

$$I_{new} = \alpha^{(P(I))}, \tag{1.7}$$

где α — постоянная, относительно которой осуществляется преобразование и, как правило, равная минимальному значению интенсивности элементов исходного изображения $\alpha = I_{min}$.

В среде MATLAB программная реализация имеет вид:

```
Inew(i,j) = Imin^(CH(ceil(255 * I(i,j) + eps)));
```

Рассмотренные методы преобразования гистограмм могут применяться для устранения искажений при передаче уровней квантования, которым были подвергнуты изображения на этапе формирования, передачи или обработки данных. Кроме того, данные методы могут применяться не только ко всему изображению, но использоваться локально в *скользящем окне*, что позволит повысить детализированность отдельных областей.

В среде MATLAB реализовано несколько функций, автоматически выравнивающих гистограммы полутонового изображения:

1. `imadjust()` — повышает контрастность изображения, изменяя диапазон интенсивностей исходного изображения;

2. `histeq()` — эквализирует (выравнивает) гистограмму методом распределения значений интенсивностей элементов исходного изображения;

3. `adapthisteq()` — выполняет контрастно-ограниченное адаптивное выравнивание гистограммы методом анализа и эквализации гистограмм локальных окрестностей изображения.

Для обработки цветного изображения данные функции можно применять поочередно к каждому цвету.

Профиль изображения

Профилем изображения вдоль некоторой линии называется функция интенсивности изображения, распределенного вдоль данной линии *(прорезки)*. Простейшим случаем профиля изображения является профиль строки:

$$\text{Profile}\, i(x) = I(x,i), \tag{1.8}$$

где i — номер строки изображения I.

Профиль столбца изображения:

$$\text{Profile}\, j(y) = I(j,y), \tag{1.9}$$

где j — номер столбца изображения I.

В общем случае профиль можно рассматривать вдоль любой прямой, ломаной или кривой линии, пересекающей изображение. После формирования массива профиля изображения проводится его анализ стандартными средствами. Анализ позволяет автоматически выделять особые точки функции профиля, соответствующие контурам изображения, пересекаемым данной линией. Например, на рис. 1.1 представлен профиль изображения штрих-кода, взятого вдоль оси Ox. Данный профиль содержит всю необходимую информацию для считывания штрих-кода, поскольку позволяет определить последовательность чередования «толстых» и «тонких» штрихов и пробелов различной ширины. В среде MATLAB профиль изображения можно найти при помощи функции `improfile()`.

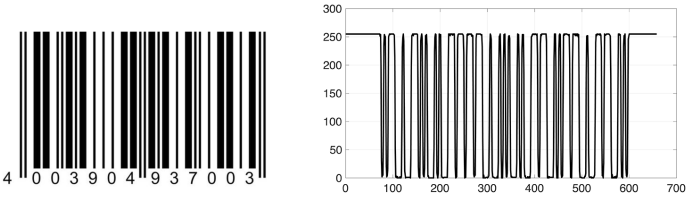


Рис. 1.1 — Слева — штрих-код,
справа — его профиль вдоль оси Ox.

Листинг 1.3. Поиск профиля штрих-кода вдоль оси Ox.

```
I = imread('code.jpg');
[numRows, numCols, Layers] = size(I);
x = [1 numCols];
y = [ceil(numRows/2) ceil(numRows/2)];
figure
improfile(I,x,y),grid on;
```

Для интерактивного указания линии (ломаной) вдоль которой следует построить профиль необходимо использовать функцию `improfile` без параметров.

Листинг 1.4. Интерактивное задание линии профиля.

```
1 I = imread('code.jpg');
2 imshow(I);
3 improfile
```

Проекция изображения

Проекцией изображения на некоторую ось называется сумма интенсивностей пикселей изображения в направлении, перпендикулярном данной оси. Простейшим случаем проекции двумерного изображения являются вертикальная проекция на ось Ox, представляющая собой сумму интенсивностей пикселей *по столбцам* изображения:

$$\mathrm{Proj}\, X(y) = \sum_{y=0}^{\dim Y - 1} I(x,y), \tag{1.10}$$

и горизонтальная проекция на ось Oy, представляющая собой сумму интенсивностей пикселей *по строкам* изображения:

$$\mathrm{Proj}\, Y(x) = \sum_{x=0}^{\dim X - 1} I(x,y). \tag{1.11}$$

Запишем выражение для проекции на произвольную ось. Допустим, что направление оси задано единичным вектором с координатами (e_x,e_y). Тогда проекция изображения на ось Oe определяется следующим выражением:

$$\mathrm{Proj}\, E(t) = \sum_{xe_x + ye_y = t} I(x,y). \tag{1.12}$$

Анализ массива проекции позволяет выделять характерные точки функции проекции, которые соответствуют контурам объектов

на изображении. Например, если на изображении имеются контрастные объекты, то в проекции будут видны перепады или экстремумы функции, соответствующие положению каждого из объектов.

На рис. 1.2 показан пример проекции на ось Oy машиночитаемого документа. Видно, что две машиночитаемые текстовые строки порождают два существенных экстремума функции проекции. Подобные проекции могут быть использованы в алгоритмах обнаружения и сегментации текстовых строк в системах распознавания текста.

Листинг 1.5. Определение положения текста.

```
for i=1:1:numRows
    Proj(i,1) = (round(sum(I(i,:,1))) + ...
    round(sum(I(i,:,2))) + ...
    round(sum(I(i,:,3)))) / (256 * 3);
end
```

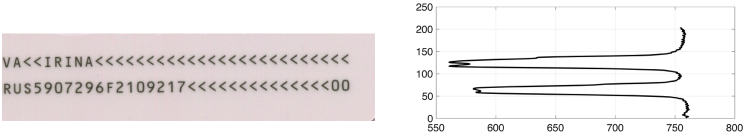


Рис. 1.2 — Слева — фрагмент заграничного паспорта, справа — его проекция на ось Oy.

Порядок выполнения работы

1. *Гистограммы.* Выбрать произвольное слабоконтрастное изображение. Выполнить выравнивание гистограммы и растяжение контраста, использовать рассмотренные преобразования и встроенные функции пакета MATLAB. Сравнить полученные результаты.

2. *Профили.* Выбрать произвольное изображение, содержащее монотонные области и выделяющиеся объекты. Произвести

построение проекций изображения на вертикальную и горизонтальную оси. Определить границы областей объектов.

3. *Проекции.* Выбрать произвольное изображение, содержащие штрих-код. Выполнить построение профиля изображения вдоль штрих-кода.

Содержание отчета

1. Цель работы.
2. Теоретическое обоснование применяемых методов и функций геометрических преобразований.
3. Ход выполнения работы:
   (a) Исходные изображения;
   (b) Листинги программных реализаций;
   (c) Комментарии;
   (d) Результирующие изображения.
4. Выводы о проделанной работе.

Вопросы к защите лабораторной работы

1. Что такое контрастность изображения и как её можно изменить?
2. Чем эффективно использование профилей и проекций изображения?
3. Каким образом можно найти объект на равномерном фоне?

Лабораторная работа №2
Геометрические преобразования изображений

Цель работы

Освоение основных видов отображений и использование геометрических преобразований для решения задач пространственной коррекции изображений.

Методические рекомендации

До начала работы студенты должны ознакомиться с основными функциями среды MATLAB по работе с геометрическими преобразованиями изображений. Лабораторная работа рассчитана на 5 часов.

Теоретические сведения

Геометрические преобразования изображений подразумевают пространственное изменение местоположения множества пикселей с целочисленными координатами (x,y) в другое множество с координатами (x',y'), причем интенсивность пикселей сохраняется. В двумерных плоских геометрических преобразованиях, как правило, используется евклидово пространство $\mathbf{P}^2$ с ортонормированной декартовой системой координат. В этом случае пикселю изображения соответствует пара декартовых координат, которые интерпретируются в виде двумерного вектора, представленного отрезком из точки (0,0) до точки $X_i = (x_i,y_i)$. Двумерные преобразования на плоскости можно представить в виде движения точек, соответствующих множеству пикселей.

Для общности с дальнейшими преобразованиями будем использовать *однородные координаты*, обладающие тем свойством, что определяемый ими объект не меняется при умножении всех координат на одно и то же ненулевое число. Из-за этого свойства необходимое количество координат для представления точек всегда на одну

больше, чем размерность пространства $\mathbf{P}^n$, в котором эти координаты используются. Например, для представления точки $X = (x,y)$ на плоскости в двумерном пространстве $\mathbf{P}^2$ необходимо три координаты $X = (x,y,w)$. Проиллюстрируем это следующим примером:

$$\begin{bmatrix} x' \\ y' \\ w \end{bmatrix} = w \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} x' & y' & w \end{bmatrix} = \begin{bmatrix} x & y & 1 \end{bmatrix} w, \tag{2.1}$$

где w — скалярный произвольный множитель, $x = \dfrac{x'}{w}$, $y = \dfrac{y'}{w}$.

При помощи троек однородных координат и матриц третьего порядка можно описать любое линейное преобразование плоскости. Таким образом, геометрические преобразования являются матричными преобразованиями, и множества координат пикселей преобразованного и исходного изображений связаны следующим матричным соотношением либо в строчном виде $X' = XT$, либо в столбцовом $X' = T^{\mathrm{T}}X$. Распишем эти матричные соотношения:

$$\begin{bmatrix} x' & y' & w' \end{bmatrix} = \begin{bmatrix} x & y & w \end{bmatrix} \cdot \begin{bmatrix} A & D & G \\ B & E & H \\ C & F & I \end{bmatrix} \tag{2.2}$$

$$\Leftrightarrow \begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & I \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ w \end{bmatrix}. \tag{2.3}$$

Точка с декартовым координатами (x,y) в однородных координатах запишется как $(x,y,1)$:

$$\begin{bmatrix} x' & y' & 1 \end{bmatrix} = \begin{bmatrix} x & y & 1 \end{bmatrix} \cdot \begin{bmatrix} A & D & 0 \\ B & E & 0 \\ C & F & 1 \end{bmatrix} \tag{2.4}$$

$$\Leftrightarrow \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A & B & C \\ D & E & F \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{2.5}$$

Формулу (2.4) можно представить в виде следующей системы уравнений:

$$\begin{cases} x' = Ax + By + C, \\ y' = Dx + Ey + F. \end{cases} \tag{2.6}$$

Линейные преобразования

Конформное отображение — это отображение, при котором сохраняется форма бесконечно малых фигур и углы между кривыми в точках их пересечения. Основными линейными конформными преобразованиями являются евклидовы преобразования. К ним относятся сдвиг, отражение, однородное масштабирование и поворот. Конформные преобразования являются подмножеством аффинных преобразований.

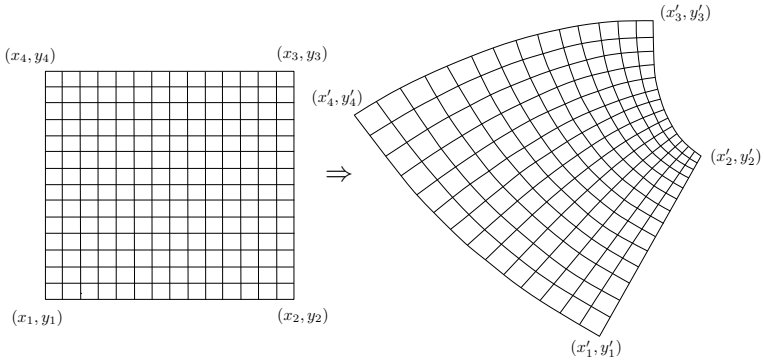


Рис. 2.1 — Конформное отображение: перпендикулярность линий в местах их пересечений сохраняется.

Сдвиг изображения

Система уравнений (2.6) и матрица преобразования координат T в случае *сдвига* изображения $A = E = 1$, $B = D = 0$ примут вид:

$$\begin{cases} x' = x + C, \\ y' = y + D, \end{cases} \tag{2.7}$$

$$\Leftrightarrow \begin{bmatrix} x' & y' & 1 \end{bmatrix} = \begin{bmatrix} x & y & 1 \end{bmatrix} T \Rightarrow T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ C & F & 1 \end{bmatrix}, \qquad (2.8)$$

где C и F — сдвиг по осям Ox и Oy соответственно.

Листинг 2.1. Сдвиг изображения на 50 и 100 пикселей по осям Ox и Oy соответственно. Размер исходного рисунка `roadSign.jpg` 300×300 пикселей.

```
I = imread('roadSign.jpg');
T = [1 0 0; 0 1 0; 50 100 1];
tform = affine2d(T);
I_shift = imwarp(I, tform, ...
    'Interp', 'nearest', ...
    'OutputView', imref2d(size(I), ...
    [1 size(I,2)], [1 size(I,1)]) ...
);
```

Функция `affine2d()` создает матрицу преобразования, которая применяется к изображению `I` при помощи функции `imwarp()`. Дополнительные параметры задают одинаковые координаты для исходного и преобразованного изображений, а также метод интерполяции для неопределенных пикселей.

Отражение изображения

Система уравнений (2.6) и матрица преобразования координат T в случае *отражения* изображения вдоль оси Ox при $A = 1$, $E = -1$, $B = C = D = F = 0$ примут вид:

$$\begin{cases} x' = x \\ y' = -y \end{cases} \Rightarrow T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (2.9)$$

Листинг 2.2. Отражение относительно оси Ox.

```
T = [1 0 0; 0 -1 0; 0 0 1];
tform = affine2d(T);
I_reflect = imwarp(I, tform);
```

Однородное масштабирование изображения

Система уравнений (2.6) и матрица преобразования координат T в случае *масштабирования* изображения $A = \alpha$, $E = \beta$, $B = C = D = F = 0$ примут вид:

$$\begin{cases} x' = \alpha x, \alpha > 0 \\ y' = \beta y, \beta > 0 \end{cases} \Rightarrow T = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.10}$$

если $\alpha < 1$ и $\beta < 1$, то изображение уменьшается, если $\alpha > 1$ и $\beta > 1$ — увеличивается. Если $\alpha \neq \beta$, то пропорции будут не одинаковыми по ширине и высоте. В общем случае данное отображение будет являться аффинным, а не конформным.

Листинг 2.3. Увеличение исходного изображения в два раза.

```
1 T = [2 0 0; 0 2 0; 0 0 1];
2 tform = affine2d(T);
3 I_scale = imwarp(I, tform);
```

Поворот изображения

Система уравнений (2.6) и матрица преобразования координат T в случае *поворота* изображения по часовой стрелке $A = \cos\varphi$, $B = -\sin\varphi$, $D = \sin\varphi$, $E = \cos\varphi$, $C = F = 0$ примут вид:

$$\begin{cases} x' = x\cos\varphi - y\sin\varphi \\ y' = x\sin\varphi + y\cos\varphi \end{cases} \Rightarrow T = \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.11}$$

Листинг 2.4. Поворот изображения на $\varphi = 17°$.

```
1 phi = 17*pi/180;
2 T = [cos(phi) sin(phi) 0;
3     -sin(phi) cos(phi) 0;
4     0 0 1];
5 tform = affine2d(T);
6 I_rotate = imwarp(I, tform);
```

Рис. 2.2 — Конформные преобразования, верхний ряд слева направо: исходное изображение, сдвиг, отражение, вращение; нижний ряд: однородное масштабирование.

Если $\varphi = 90°$, то $\cos\varphi = 0$ и $\sin\varphi = 1$ и выражение (2.11) примет вид:

$$\begin{cases} x' = -y \\ y' = x \end{cases} \Rightarrow T = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.12}$$

Аффинное отображение — это отображение, при котором параллельные прямые переходят в параллельные прямые, пересекающиеся в пересекающиеся, скрещивающиеся в скрещивающиеся; сохраняются отношения длин отрезков, лежащих на одной прямой (или на параллельных прямых), и отношения площадей фигур. Базовыми преобразованиями являются конформные преобразования, скос и неоднородное масштабирование. Произвольное аффинное преобразование можно получить при помощи последовательного произведения матриц базовых преобразований. В непрерывной геометрии любое аффинное преобразование имеет обратное аффинное преобразование, а произведение прямого и обратного дает единичное преобразование, которое оставляет все точки на месте. Аффинные преобразования являются подмножеством проекционных преобразований.

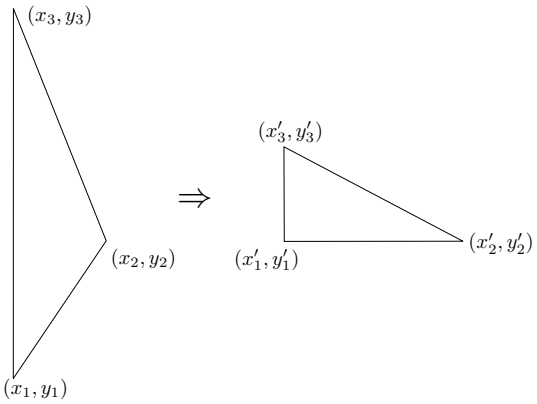


Рис. 2.3 — Аффинное отображение.

Скос изображения

Система уравнений (2.6) и матрица преобразования координат T в случае *скоса* изображения $A = E = 1$, $B = s$, $C = D = F = 0$ примут вид:

$$\begin{cases} x' = x + sy, \\ y' = y, \end{cases} \Rightarrow T = \begin{bmatrix} 1 & 0 & 0 \\ s & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.13}$$

Листинг 2.5. Скос изображения, $s = 0.3$.

```
1 T = [1 0 0; 0.3 1 0; 0 0 1];
2 tform = affine2d(T);
3 I_bevel = imwarp(I, tform);
```

Кусочно-линейные преобразования

Кусочно-линейное отображение — это отображение, при котором изображение разбивается на части, а затем к каждой из этих частей применяются различные линейные преобразования.

Листинг 2.6. Пример кусочно-линейного отображения — левая половина изображения остается без изменений, а правая растягивается в два раза вдоль оси Ox.

```
1 imid = round(size(I,2) / 2);
```

```
2 I_left = I(:, 1:imid, :);
3 stretch = 2;
4 I_right = I(:, (imid + 1:end), :);
5 T = [stretch 0 0; 0 1 0; 0 0 1];
6 tform = affine2d(T);
7 I_scale = imwarp(I_right, tform);
8 I_piecewiselinear = [I_left I_scale];
```

Рис. 2.4 — Слева — скос, справа — кусочно-линейное растяжение.

Нелинейные преобразования

При рассмотрении геометрических преобразований предполагается, что изображения получены при помощи идеальной модели камеры. В действительности формирование изображений сопровождается различными нелинейными искажениями, см. рис. 2.5. Для их коррекции используются различные нелинейные функции.

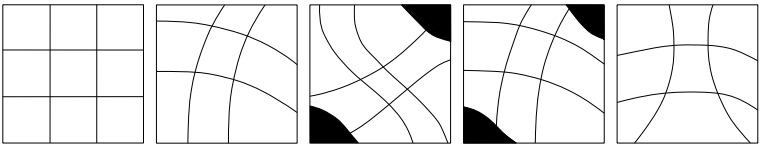


Рис. 2.5 — Примеры нелинейных искажений.

Проекционное преобразование

Проекционное отображение — это отображение, при котором прямые линии остаются прямыми линиями, однако геометрия фигуры может быть нарушена, т.к. данное отображение в общем случае не сохраняет параллельности линий. Свойством, сохраняющимся при проективном преобразовании, является *коллинеарность* точек: три точки, лежащие на одной прямой (коллинеарные), после

преобразования остаются на одной прямой. Проекционное отображение может быть как *параллельным* (изменяется масштаб), так и *проективным* (изменяется геометрия фигуры).

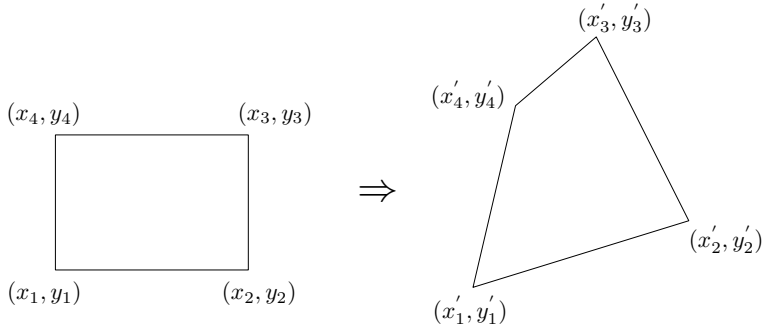


Рис. 2.6 — Проекционное проективное отображение: прямые остались прямыми, но параллельные отобразились в скрещивающиеся.

В случае проективного отображения точки трехмерной сцены $\mathbf{P}^3$ проецируются на двумерную плоскость $\mathbf{P}^2$ изображения. Такое преобразование $\mathbf{P}^3 \to \mathbf{P}^2$ отображает евклидову точку сцены $P = (x,y,z)$ (в однородных координатах (x',y',z',w')) в точку изображения $X = (x,y)$ (в однородных координатах (x',y',w')). Для нахождения декартовых координат точек из однородных координат воспользуемся следующими соотношениями: $P = (\frac{x'}{w'}, \frac{y'}{w'}, \frac{z'}{w'})$ — координаты вектора $\mathbf{P}$ и $X = (\frac{x'}{w'}, \frac{y'}{w'})$ — координаты вектора $\mathbf{X}$. Подставляя в (2.2) $w = 1$ для векора $\mathbf{X}$ получим систему уравнений:

$$\begin{cases} x' = \dfrac{Ax + By + C}{Gx + Hy + I}, \\ y' = \dfrac{Dx + Ey + F}{Gx + Hy + I}, \end{cases} \quad \Rightarrow T = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & 1 \end{bmatrix}. \tag{2.14}$$

Из-за нормирования координат на w' в общем случае проекционное отображение является нелинейным.

Листинг 2.7. Пример проективного отображения при $A = 1.1$, $B = 0.35$, $C = F = 0$, $D = 0.2$, $E = 1.1$, $G = 0.00075$, $H = 0.0005$, $I = 1$.

```
1 T = [1.1 0.35 0; 0.2 1.1 0; 0.00075 0.0005 1];
2 tform = projective2d(T);
3 I_projective = imwarp(I, tform);
```

Полиномиальное преобразование

Полиномиальное отображение — это отображение исходного изображения с помощью полиномов. В данном случае матрица преобразования координат T будет содержать коэффициенты полиномов соответствующих порядков для координат x и y. Например, в случае полиномиального преобразования *второго порядка* система уравнений 2.6 примет вид:

$$\begin{cases} x' = a_1 + a_2x + a_3y + a_4x^2 + a_5xy + a_6y^2, \\ y' = b_1 + b_2x + b_3y + b_4x^2 + b_5xy + b_6y^2, \end{cases} \tag{2.15}$$

где x, y — координаты точек в одной системе координат; x', y' — координаты этих точек в другой системе координат; $a_1 \dots a_6$, $b_1 \dots b_6$ — коэффициенты преобразования.

Листинг 2.8. Пример полиномиального отображения с заданной матрицей преобразования T и в котором не используются предопределенные функции MATLAB. Результирующее преобразованное изображение не будет содержать интерполированных пикселей.

```
1 [numRows, numCols, Layers] = size(I);
2 T = [0 0; 1 0; 0 1; 0.00001 0;
3     0.002 0; 0.001 0];
4 for k=1:1:Layers
5     for y=1:1:numCols
6         for x=1:1:numRows
7             xnew = round(T(1,1)+T(2,1)*x+...
8                 T(3,1)*y+T(4,1)*x^2+...
9                 T(5,1)*x*y+T(6,1)*y^2);
10            ynew = round(T(1,2)+T(2,2)*x+...
```

```
11                  T(3,2)*y+T(4,2)*x^2+...
12                  T(5,2)*x*y+T(6,2)*y^2);
13              I_polynomial(xnew,ynew,k) = ...
14                  I(x,y,k);
15          end
16      end
17 end
```

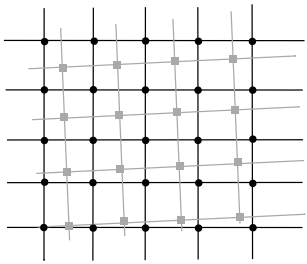


Рис. 2.7 — Смещение координат пикселей в преобразованном изображении.

Как правило, при нелинейном искажении коэффициенты $a_1 \dots a_6$, $b_1 \dots b_6$ и матрица преобразования координат T *неизвестны*. В случае, когда помимо искаженного изображения доступно исходное, можно вычислить коэффициенты преобразования, найдя пары соответствующих точек на исходном и преобразованном изображениях. Число минимально необходимых пар точек вычисляется по формуле:

$$t_{min} = \frac{(n+1)(n+2)}{2}, \tag{2.16}$$

где n — порядок преобразования.

Если имеет место полиномиальное искажение *второго порядка*, то согласно (2.16) необходимо знать координаты хотя бы *шести пар* соответствующих точек до и после трансформации:

$(x_1, y_1) \dots (x_6, y_6)$ и $(x'_1, y'_1) \dots (x'_6, y'_6)$. Согласно (2.15) запишем уравнения для вычисления коэффициентов в матричном виде:

$$\begin{bmatrix} a_1 \\ \dots \\ a_6 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1 y_1 & y_1^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_6 & y_6 & x_6^2 & x_6 y_6 & y_6^2 \end{bmatrix}^{-1} \begin{bmatrix} x'_1 \\ \dots \\ x'_6 \end{bmatrix}, \tag{2.17}$$

$$\begin{bmatrix} b_1 \\ \dots \\ b_6 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1 y_1 & y_1^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_6 & y_6 & x_6^2 & x_6 y_6 & y_6^2 \end{bmatrix}^{-1} \begin{bmatrix} y'_1 \\ \dots \\ y'_6 \end{bmatrix}. \tag{2.18}$$

Для интерактивного указания одинаковых точек на исходном и преобразованном изображениях можно воспользоваться функцией MATLAB `cpselect('source.jpg','transformed.jpg')`, которая возвращает два массива `movingPoints` и `fixedPoints` с координатами преобразованного и исходного изображений соответственно.

Синусоидальное искажение

В качестве еще одного из примеров нелинейного преобразования можно рассмотреть гармоническое искажение изображения.

Листинг 2.9. Пример синусоидального искажения изображения.

```
1 [xi,yi] = meshgrid(1:ncols,1:nrows);
2 imid = round(size(I,2)/2);
3 u = xi + 20*sin(2*pi*yi/90);
4 v = yi;
5 tmap_B = cat(3,u,v);
6 resamp = makeresampler('linear','fill');
7 I_sinusoid = tformarray(I,[],resamp,...
8     [2 1],[1 2],[],tmap_B,.3);
```

Коррекция дисторсии

При формировании изображения оптической системой на нем может возникнуть *дисторсия*. *Дисторсия* — это оптическое искажение, выражающееся в искривлении прямых линий. Световые лу-

Рис. 2.8 — Слева — полиномиальное искажение, в центре — проективное, справа — синусоидальное.

чи, проходящие через центр линзы сходятся в точке, расположенной дальше от линзы, чем лучи, проходящие через ее края. Прямые линии исккривляются за исключением тех, которые лежат в одной плоскости с оптической осью. Например, изображение квадрата, центр которого пересекает оптическая ось, имеет вид подушки (*подушкообразная дисторсия*) при положительной дисторсии и вид бочки (*бочкообразная дисторсия*) при отрицательной дисторсии (рис. 2.9).

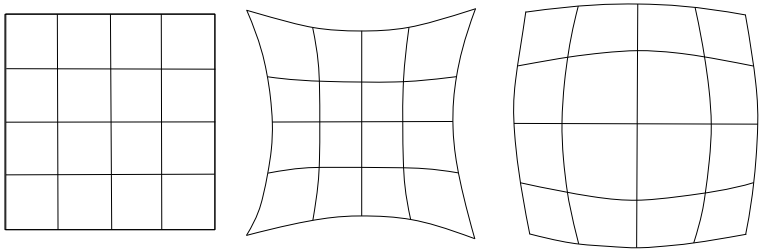


Рис. 2.9 — Примеры дисторсий. Слева — исходное изображение, по центру — подушкообразная дисторсия, справа — бочкообразная.

Пусть $\mathbf{r} = (x,y)$ — вектор, задающий две координаты в плоскости, расположенной перпендикулярно оптической оси. Для идеального изображения все лучи, вышедшие из этой точки и прошедшие

через оптическую систему, попадут в точку изображения с координатами $\mathbf{R}$, которые определяются по формуле:

$$\mathbf{R} = b_0\mathbf{r}, \tag{2.19}$$

где b_0 — коэффициент линейного увеличения.

Если присутствует дисторсия более высокого порядка (для осесимметричных оптических систем дисторсия может быть только нечетных порядков: третьего, пятого, седьмого и т.д.), то в выражение (2.19) необходимо добавить соответствующие слагаемые:

$$\mathbf{R} = b_0\mathbf{r} + F_3 r^2\mathbf{r} + F_5 r^4\mathbf{r} + \ldots, \tag{2.20}$$

где r — длина вектора $\mathbf{r}$; $F_i, i = 3{,}5,\ldots,n$ — коэффициенты дисторсии n-го порядка, которые вносят наибольший вклад в искажение формы изображения. При дисторсии третьего порядка, если коэффициент F_3 имеет тот же знак, что и b_0: $\mathrm{sign}(F_3) = \mathrm{sign}(b_0)$, возникает подушкообразное искажение, в противном случае — бочкообразное.

Для коррекции дисторсии используется подход, описанный выше для проективного отображения. Используется изображение регулярной сетки и его искаженное изображение, находятся пары точек на этих изображениях и вычисляются коэффициенты корректирующего преобразования.

Бочкообразная дисторсия

Листинг 2.10. Пример наложения бочкообразной дисторсии пятого порядка на исходное изображение.

```
 1 [xi,yi] = meshgrid(1:ncols,1:nrows);
 2 imid = round(size(I,2)/2);
 3 xt = xi(:) - imid;
 4 yt = yi(:) - imid;
 5 [theta,r] = cart2pol(xt,yt);
 6 F3 = .000001;
 7 F5 = .0000012;
 8 R = r + F3 * r.^2 + F5 * r.^4;
 9 [ut,vt] = pol2cart(theta, R);
10 u = reshape(ut, size(xi)) + imid;
11 v = reshape(vt, size(yi)) + imid;
```

```
12 tmap_B = cat(3, u, v);
13 resamp = makeresampler('linear','fill');
14 I_barrel = tformarray(I,[],resamp,...
15     [2 1],[1 2],[],tmap_B,.3);
```

Для формирования сетки изображения используется функция `meshgrid()`. Затем на строках 2-4 происходит получение отцентрированного набора координат всех пикселей. После этого все координаты функцией `cart2pol()` переводятся из декартовой системы координат в полярную, для более удобного применения формулы (2.20) через радиус-векторы каждой точки. Затем координаты пикселей преобразуются при помощи функций `pol2cart()` и `reshape()` обратно в декартовы координаты. Из итоговой сетки формируется трехмерная матрица преобразования `tmap_B` для функции `tformarray`, которая производит само преобразование исходного изображения по заданной сетке.

Подушкообразная дисторсия

Листинг 2.11. Пример наложения подушкообразной дисторсии третьего порядка на исходное изображение.

```
1 F3 = -0.003;
2 R = r + F3 * r.^2;
```

Рис. 2.10 — Слева — исходное изображение, в центре — бочкообразная дисторсия, справа — подушкообразная дисторсия.

«Сшивка» изображений

Геометрические преобразования можно использовать, например, для построения мозаики из нескольких изображений. Мозаика («сшивка», «склейка») — это объединение двух или более изображений в единое целое, причем системы координат склеиваемых

изображений могут отличаться из-за разного ракурса съемки, изменения положения камеры или движения самого объекта. Однако необходимо, чтобы оба изображения имели области перекрытия, т.е. на них присутствовали одинаковые объекты.

Основной задачей обработки таких изображений является приведение их в общую систему координат. В качестве общей системы координат можно использовать систему первого изображения, тогда требуется найти преобразование координат всех пикселей второго изображения (x,y) в общую систему координат (x',y'). Если имеет место полиномиальное искажение, то для пересчета координат можно воспользоваться системой уравнений (2.15). В случае аффиного отображения система (2.15) примет вид:

$$\begin{cases} x' = a_1 + a_2 x + a_3 y, \\ y' = b_1 + b_2 x + b_3. \end{cases} \tag{2.21}$$

Поэтому необходимо найти лишь по 3 коэффициента преобразования по каждой координате:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix}, \tag{2.22}$$

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} y'_1 \\ y'_2 \\ y'_3 \end{bmatrix}. \tag{2.23}$$

Для этого на обоих изображениях следует выбрать соответствующие пары точек (три пары в случае аффинного искажения и не менее шести пар в случае полиномиального искажения). Интерактивно выполнить эту операцию можно при помощи функции MATLAB `cpselect()`, либо воспользоваться специальными алгоритмами. Например, можно определить эти точки на основе коэффициента корреляции, который в MATLAB вычисляется с помощью функции `corr2()`.

Рассмотрим простой случай «склейки» двух неискаженных изображений, имеющих одинаковую ширину. Необходимо склеить их по вертикали, т.е. добавить к первому второе снизу. Однако,

граница склейки неизвестна. Эту задачу можно реализовать при помощи корреляционного подхода.

Листинг 2.12. Пример «склейки» изображений. Считывание верхней и нижней картинок соответственно:

```
topPart = imread('itmoTop.jpg');
botPart = imread('itmoBot.jpg');
```

Для простоты определения коэффициента корреляции преобразуем цветные изображения в полутоновые:

```
topPartHT = im2double(rgb2gray(topPart));
botPartHT = im2double(rgb2gray(botPart));
```

Определим размеры полутоновых изображений и инициализируем промежуточные массивы для вычислений:

```
[numRows, numCols, Layers] = size(topPartHT);
[numRowsBot, numColsBot] = size(botPartHT);
botPartCorrHT = zeros(intersecPart,numCols);
topPartCorrHT = zeros(intersecPart,numCols);
correlationArray = [];
```

Определим верхнюю строку нижнего изображения для вычисления коффициента корреляции со строками верхнего:

```
intersecPart = 5;
for j = 1:1:numCols
    for i = 1:1:intersecPart
        botPartCorrHT(i,j) = botPartHT(i,j);
    end
end
```

Сравним полученную строку со строками верхнего изображения и вычислим коэффициент корреляции:

```
for j = 0:1:numRows-intersecPart
    for i = 1:1:intersecPart
        topPartCorrHT(i,:) = topPartHT(i+j,:);
    end
    correlationCoefficient = ...
        corr2(topPartCorrHT, botPartCorrHT);
    correlationArray = ...
```

```
23          [correlationArray
24          correlationCoefficient];
25      correlationCoefficient = 0;
26  end
27  [M, I] = max(correlationArray);
```

Построим цветное «склеенное» изображение `result_img` с границей склейки на основе полученного индекса, соответствующего номеру строки с максимальным коэффициентом корреляции:

```
28  numRowsBotCorr = numRowsBot + I - 1;
29  for k = 1:1:Layers
30      for j = 1:1:numCols
31          for i = 1:I-1
32              result_img(i,j,k) = ..
33                  topPart(i,j,k);
34          end
35          for i = I:1:numRowsBotCorr
36              result_img(i,j,k) = ...
37                  botPart(i-I+1,j,k);
38          end
39      end
40  end
```

Рис. 2.11 — Слева — `itmoTop.jpg`, в центре — `itmoBot.jpg`, справа — `result_img`.

Порядок выполнения работы

1. *Простейшие геометрические преобразования.* Выбрать произвольное изображение. Выполнить над ним линейные и нелинейные преобразования (конформные, аффинные и проективные отображения).

2. *Коррекция дисторсии.* Выбрать произвольное изображение либо с подушкообразной, либо с бочкообразной дисторсией. Выполнить коррекцию изображения.

3. *«Склейка» изображений.* Выбрать два изображения (снимки с фотокамеры, фрагменты сканированного изображения и пр.) на которых имеется область пересечения. Выполнить коррекцию второго изображения для его перевода в систему координат первого; затем выполнить автоматическую «склейку» из двух изображений в одно.

Содержание отчета

1. Цель работы.

2. Теоретическое обоснование применяемых методов и функций геометрических преобразований.

3. Ход выполнения работы:

   (a) Исходные изображения;

   (b) Листинги программных реализаций;

   (c) Комментарии;

   (d) Результирующие изображения.

4. Выводы о проделанной работе.

Вопросы к защите лабораторной работы

1. Каким образом можно выполнить поворот изображения, не используя матрицу поворота?

2. Какое минимальное количество соответствующих пар точек необходимо задать на исходном и искаженном изображениях, если порядок преобразования $n = 4$?

3. После геометрического преобразования изображения могут появиться пиксели с неопределенными значениями интенсивности. С чем это связано и как решается данная проблема?

Лабораторная работа №3
Фильтрация и выделение контуров

Цель работы

Освоение основных способов фильтрации изображений от шумов и выделения контуров.

Методические рекомендации

До начала работы студенты должны ознакомиться с основными функциями среды MATLAB фильтрации изображений. Иметь представление о низкочастной и высокочастотной фильтрации. Лабораторная работа рассчитана на 5 часов.

Теоретические сведения

Типы шумов

Цифровые изображения, полученные различными оптико-электронными приборами, могут содержать в себе разнообразные искажения, обусловленные разного рода помехами, которые принято называть *шумом*. Шум на изображении затрудняет его обработку автоматическими средствами и, поскольку шум может иметь различную природу, для его успешного подавления необходимо определить адекватную математическую модель. Рассмотрим наиболее распространенные модели шумов. В среде MATLAB шум может быть наложен на изображение с помощью функции `imnoise()`.

Импульсный шум

При импульсном шуме сигнал искажается выбросами с очень большими отрицательными или положительными значениями малой длительностью и может возникать, например, из-за ошибок декодирования. Такой шум приводит к появлению на изображении белых («соль») или черных («перец») точек, поэтому зачастую называется *точечным* шумом. Для его описания следует принять во внимание тот факт, что появление шумового выброса в каждом пикселе $I(x,y)$ не зависит ни от качества исходного изображения, ни от наличия шума в других точках и имеет вероятность появления

Рис. 3.1 — Исходное полутоновое изображение.

p, причем значение интенсивности пикселя $I(x,y)$ будет изменено на значение $d \in [0,255]$:

$$I(x,y) = \begin{cases} d, \text{ с вероятностью } p, \\ s_{x,y}, \text{ с вероятностью } (1-p), \end{cases} \tag{3.1}$$

где $s_{x,y}$ — интенсивность пикселя исходного изображения, I — зашумленное изображение, если $d = 0$ — шум типа «перец», если $d = 255$ — шум типа «соль». В среде MATLAB задается параметром `'salt & pepper'` функции `imnoise()`: `imnoise(I,'salt & pepper')`.

Аддитивный шум

Аддитивный шум описывается следующим выражением:

$$I_{new}(x,y) = I(x,y) + \eta(x,y), \tag{3.2}$$

где I_{new} — зашумленное изображение, I — исходное изображение, η — не зависящий от сигнала аддитивный шум с гауссовым или любым другим распределением функции плотности вероятности.

Мультипликативный шум

Мультипликативный шум описывается следующим выражением:

$$I_{new}(x,y) = I(x,y) \cdot \eta(x,y), \tag{3.3}$$

где I_{new} — зашумленное изображение, I — исходное изображение, η — не зависящий от сигнала мультипликативный шум, умножающий

зарегистрированный сигнал. В качестве примера можно привести зернистость фотопленки, ультразвуковые изображения и т.д. Частным случаем мультипликативного шума является *спекл*-шум, который появляется на изображениях, полученных устройствами с когерентным формированием изображений, например, медицинскими сканерами или радарами. На таких изображениях можно отчетливо наблюдать светлые пятна, крапинки (спеклы), которые разделены темными участками изображения. В среде MATLAB спекл-шум накладывается на изображение `I` функцией `imnoise(I,'speckle')`.

Гауссов (нормальный) шум

Гауссов шум на изображении может возникать в следствие недостатка освещенности сцены, высокой температуры и т.д. Модель шума широко распространена в задачах низкочастотной фильтрации изображений. Функция распределения плотности вероятности $p(z)$ случайной величины z описывается следующим выражением:

$$p(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(z-\mu)^2}{2\sigma^2}}, \tag{3.4}$$

где z — интенсивность изображения (например, для полутонового изображения $z \in [0{,}255]$), μ — среднее (математическое ожидание) случайной величины z, σ — среднеквадратичное отклонение, дисперсия σ^2 определяет мощность вносимого шума. Примерно 67% значенией случайной величины z сосредоточено в диапазоне $[(\mu-\sigma),(\mu+\sigma)]$ и около 96% в диапазоне $[(\mu-2\sigma),(\mu+2\sigma)]$. В среде MATLAB шум может быть задан с помощью функции `imnoise(I,'gaussian')` или `imnoise(I,'localvar')` в случае нулевого математического ожидания.

Шум квантования

Зависит от выбранного шага квантования и самого сигнала. Шум квантования может приводить, например, к появлению ложных контуров вокруг объектов или убирать слабо контрастные детали на изображении. Такой шум не устраняется. Приближенно шум квантования можно описать распределением Пуассона и задать функцией `imnoise(I,'poisson')`.

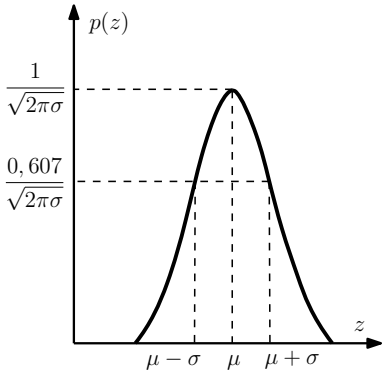


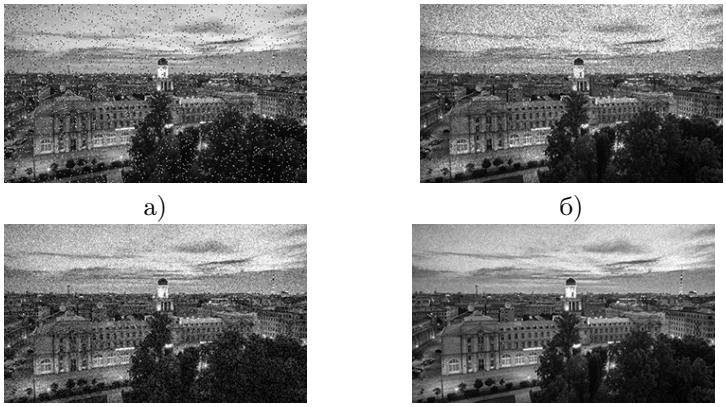
Рис. 3.2 — Функция распределения плотности вероятности $p(z)$.

Фильтрация изображений

Рассмотрим основные методы фильтрации изображений. Если для вычисления значения интенсивности каждого пикселя учитываются значения соседних пикселей в некоторой окрестности, то такое преобразование называется *локальным*, а сама окрестность — *окном*, представляющим собой некоторую матрицу, называемую *маской*, *фильтром*, *ядром фильтра*, а сами значения элементов матрицы называются *коэффициентами*. Центр маски совмещается с анализируемым пикселем, а коэффициенты маски умножаются на значения интенсивностей пикселей, накрытых маской. Как правило, маска имеет квадратную форму размера 3×3, 5×5 и т.п. Фильтрация изображения I, имеющего размеры $M \times N$, с помощью маски размера $m \times n$ описывается формулой:

$$I_{new}(x,y) = \sum_s \sum_t w(s,t) I(x+s, y+t), \tag{3.5}$$

где s и t — координаты элементов маски относительно ее центра (в центре $s = t = 0$). Такого рода преобразования называются *линейными*. После вычисления нового значения интенсивности пикселя $I_{new}(x,y)$ окно w, в котором описана маска фильтра, сдвигается и вычисляется интенсивность следующего пикселя, поэтому подобное преобразование называется *фильтрацией в скользящем окне*.



а) б)

в) г)

Рис. 3.3 — Результат наложения: а) шума типа «соль» и «перец», б) спекл-шума, в) нормального шума, г) шума Пуассона.

В среде MATLAB фильтрация изображения может быть осуществлена при помощи функции `filter2(mask,I)`, где матрица `mask` задает маску фильтра. Маска может задаваться либо вручную, либо с помощью функции `fspecial()`.

Низкочастотная фильтрация

Низкочастотные пространственные фильтры ослабляют высокочастотные компоненты (области с сильным изменением интенсивностей) и оставляют низкочастотные компоненты изображения без изменений. Используются для снижения уровня шума и удаления высокочастотных компонент, что позволяет повысить точность исследования содержания низкочастотных компонент. В результате применения низкочастотных фильтров получим сглаженное или размытое изображение. Главными отличительными особенностями являются:

1. неотрицательные коэффициенты маски;
2. сумма всех коэффициентов равна единице.

Примеры ядер низкочастотных фильтров:

$$w = \frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}, w = \frac{1}{10}\begin{bmatrix}1 & 1 & 1\\ 1 & 20 & 1\\ 1 & 1 & 1\end{bmatrix}. \tag{3.6}$$

Рассмотрим основные виды низкочастотных сглаживающих фильтров.

Арифметический усредняющий фильтр

Данный фильтр усредняет значение интенсивности пикселя по окрестности с использованием маски с одинаковыми коэффициентами, например, для маски размером 3×3 коэффициенты равны $1/9$, при 5×5 — $1/25$. Благодаря такому нормированию значение результата фильтрации будет приведено к диапазону интенсивностей исходного изображения. Графически двумерная функция, описывающая маску фильтра, похожа на параллелепипед, поэтому в англоязычной литературе используется название *box*-фильтр. Арифметическое усреднение достигается при использовании следующей формулы:

$$I_{new}(x,y) = \frac{1}{m \cdot n}\sum_{i=0}^{m}\sum_{j=0}^{n} I(i,j), \tag{3.7}$$

где $I_{new}(x,y)$ — значение интенсивности пикселя отфильтрованного изображения, $I(i,j)$ — значение интенсивностей пикселей исходного изображения в маске, m и n — ширина и высота маски фильтра соответственно. Данный алгоритм эффективен для слабо зашумленных изображений. В среде MATLAB фильтрация изображения `I` с маской размера 3×3 может быть осуществлена при помощи функции `filter2(fspecial('average',3),I)`.

Геометрический усредняющий фильтр

Геометрическое усреднение рассчитывается при помощи формулы:

$$I_{new}(x,y) = \left[\prod_{i=0}^{m}\prod_{j=0}^{n} I(i,j)\right]^{\frac{1}{m \cdot n}}. \tag{3.8}$$

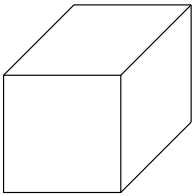


Рис. 3.4 — Графическое представление *box*-фильтра.

Эффект от применения данного фильтра аналогичен предыдущему методу, однако отдельные объекты исходного изображения искажаются меньше. Фильтр может использоваться для подавления высокочастотного аддитивного шума с лучшими статистическими характеристиками по сравнению с арифметическим усредняющим фильтром.

Гармонический усредняющий фильтр

Фильтр базируется на выражении:

$$I_{new}(x,y) = \frac{m \cdot n}{\sum_{i=0}^{m} \sum_{j=0}^{n} \frac{1}{I(i,j)}}. \tag{3.9}$$

Фильтр хорошо подавляет шумы типа «соль» и не работает с шумами типа «перец».

Контргармонический усредняющий фильтр

Фильтр базируется на выражении:

$$I_{new}(x,y) = \frac{\sum_{i=0}^{m} \sum_{j=0}^{n} I(i,j)^{Q+1}}{m \cdot n \cdot \sum_{i=0}^{m} \sum_{j=0}^{n} I(i,j)^{Q}}, \tag{3.10}$$

где Q — порядок фильтра. Контргармонический фильтр является обобщением усредняющих фильтров и при $Q > 0$ подавляет шумы типа «перец», а при $Q < 0$ — шумы типа «соль», однако одновременное удаление белых и черных точек невозможно. При $Q = 0$ фильтр превращается в арифметический, а при $Q = -1$ — в гармонический.

Фильтр Гаусса

Пиксели в скользящем окне, расположенные ближе к анализируемому пикселю, должны оказывать большее влияние на результат фильтрации, чем крайние. Поэтому коэффициенты весов маски можно описать колоколообразной функцией Гаусса (3.4). При фильтрации изображений используется двумерный фильтр Гаусса:

$$G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-y^2}{2\sigma^2}}. \tag{3.11}$$

Чем больше параметр σ, тем сильнее размывается изображение. Как правило, радиус фильтра $r = 3\sigma$. В таком случае размер маски $2r+1 \times 2r+1$ и размер матрицы $6\sigma+1 \times 6\sigma+1$. За пределами данной окрестности значения функции Гаусса будут пренебрежимо малы. В среде MATLAB фильтрация изображения фильтром Гаусса может быть осуществлена при помощи функции `imgaussfilt(I)`.

Нелинейная фильтрация

Низкочастотные фильтры линейны и оптимальны в случае, когда имеет место нормальное распределение помех на цифровом изображении. Линейные фильтры локально усредняют импульсные помехи, сглаживая изображения. Для устранения импульсных помех лучше использовать нелинейные, например, *медианные* фильтры.

Медианная фильтрация

В классическом медианном фильтре используется маска с единичными коэффициентами. Произвольная форма окна может задаваться при помощи нулевых коэффициентов. Значения интенсивностей пикселей в окне представляются в виде вектора-столбца и сортируются по возрастанию. Отфильтрованному пикселю присвается медианное (среднее) в ряду значение интенсивности. Номер медианного элемента после сортировки может быть вычислен по формуле $n = \frac{N+1}{2}$, где N — число пикселей, участвующих в сортировке. В MATLAB может быть реализован с использованием функции `medfilt2(I)`.

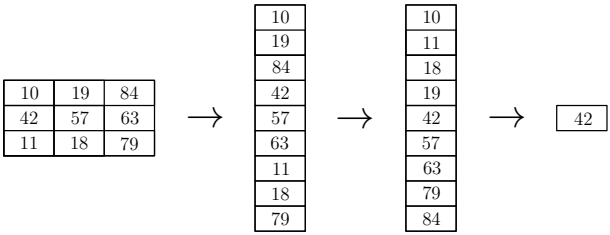


Рис. 3.5 — Иллюстрация работы медианной фильтрации в окне 3×3.

Взвешенная медианная фильтрация

В данной модификации медианной фильтрации в маске используются весовые коэффициенты (числа $2, 3$ и т.д.), чтобы отразить бóльшее влияние на результат фильтрации пикселей, расположенных ближе к фильтруемому элементу. Медианная фильтрация качественно удаляет имульсные шумы, а также на полутоновых изображениях не вносит новых значений интенсивностей. При увеличении размера окна увеличивается шумоподавляющая способность фильтра, но начинают искажаться очертания объектов. В MATLAB может быть реализован с использованием функции `medfilt2(I, [m n])`, где второй аргумент `[m n]` функции `medfilt2()` задает маску фильтра размера $m \times n$.

Адаптивная медианная фильтрация

В данной модификации фильтра скользящее окно размера $s \times s$ адаптивно увеличивается в зависимости от результата фильтрации. Обозначим через z_{min}, z_{max}, z_{med} минимальное, максимальное и медианное значения интенсивностей в окне, $z_{i,j}$ — значение интенсивности пикселя с координатами (i,j), s_{max} — максимально допустимый размер окна. Алгоритм адаптивной медианной фильтрации состоит из следующих шагов:

1. Вычисление значений z_{min}, z_{max}, z_{med}, $A_1 = z_{med} - z_{min}$, $A_2 = z_{med} - z_{max}$ пикселя (i,j) в заданном окне.

   (a) Если $A_1 > 0$ и $A_2 < 0$, перейти на шаг 2. В противном случае увеличить размер окна.

(b) Если текущий размер окна $s \leqslant s_{max}$, повторить шаг 1. В противном случае результат фильтрации равен $z_{i,j}$.

2. Вычисление значений $B_1 = z_{i,j} - z_{min}$, $B_2 = z_{i,j} - z_{max}$.

   (a) Если $B_1 > 0$ и $B_2 < 0$, результат фильтрации равен $z_{i,j}$. В противном случае результат фильтрации равен z_{med}.

3. Изменить координаты (i,j).

   (a) Если не достигнут предел изображения, перейти на шаг 1. В противном случае фильтрация окончена.

Основной идеей является увеличение размера окна до тех пор, пока алгоритм не найдет медианное значение, не являющееся импульсным шумом, или пока не достигнет максимального размера окна. В последнем случае алгоритм вернет величину $z_{i,j}$.

Ранговая фильтрация

Обобщением медианной фильтрации является *ранговый фильтр* порядка r, выбирающий из полученного вектора-столбца элементов маски пиксель с номером $r \in [1,N]$, который и будет являться результатом фильтрации.

1. Если число пикселей в окне N нечетное и $r = \frac{N+1}{2}$, тогда ранговый фильтр является *медианным*. В случае окна 3×3 в MATLAB можно воспользоваться функцией `ordfilt2(I,5,ones(3,3))`.

2. Если $r = 1$, фильтр выбирает наименьшее значение интенсивности и называется *min-фильтром*. В MATLAB задается функцией `ordfilt2(I,1,ones(3,3))`.

3. Если $r = N$, фильтр выбирает максимальное значение интенсивности и называется *max-фильтром*. В MATLAB задается функцией `ordfilt2(I,9,ones(3,3))`.

Ранг иногда записывается в процентах, например для *min-фильтра* ранг равен 0%, медианного — 50%, *max-фильтра* — 100%.

Винеровская фильтрация

Использует пиксельно-адаптивный метод Винера, основанный на статистических данных, оцененных из локальной окрестности каждого пикселя.

$$\mu = \frac{1}{n \cdot m} = \sum_{i=0}^{m} \sum_{j=0}^{n} I(i,j), \tag{3.12}$$

$$\sigma^2 = \frac{1}{n \cdot m} = \sum_{i=0}^{m} \sum_{j=0}^{n} I^2(i,j) - \mu^2, \tag{3.13}$$

$$I_{new}(x,y) = \mu + \frac{\sigma^2 - v^2}{\sigma^2} \left(I(x,y) - \mu \right), \tag{3.14}$$

где μ — среднее в окрестности, σ^2 — дисперсия, v^2 — дисперсия шума. В MATLAB реализуется при помощи функции `wiener2(I,[m n])`.

Высокочастотная фильтрация

Высокочастотные пространственные фильтры усиливают высокочастотные компоненты (области с сильным изменением интенсивностей) и ослабляют низкочастотные составляющие изображения. Используются для выделения перепадов интенсивностей и определения границ (контуров) на изображениях. Для этого в MATLAB может быть использована функция `edge()`. В результате применения высокочастотных фильтров повышается резкость изображения.

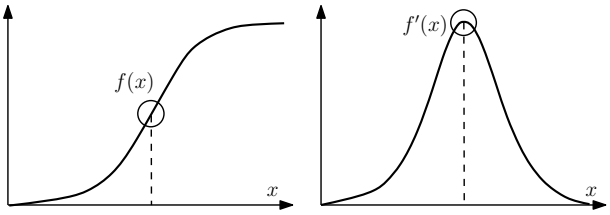


Рис. 3.6 — Функция интенсивности и ее первая производная, максимум производной соответствует краю.

Высокочастотные фильтры аппроксимируют вычисление производных по направлению, при этом приращение аргумента Δx принимается равным 1 или 2. Главными отличительными особенностями являются:

1. коэффициенты маски фильтра могут принимать отрицательные значения;

2. сумма всех коэффициентов равна нулю.

Фильтр Робертса

Фильтр Робертса работает с минимально допустимой для вычисления производной маской размерности 2×2, поэтому является быстрым и довольно эффективным. Возможные варианты масок для нахождения градиента по осям Ox и Oy:

$$G_x = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, G_y = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \tag{3.15}$$

либо

$$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, G_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{3.16}$$

В результате применения дифференциального оператора Робертса получим оценку градиента по направлениям G_x и G_y. Модуль градиента всех краевых детекторов может быть вычислен по формуле $G = \sqrt{G_x^2 + G_y^2} = |G_x| + |G_y|$, а направление градиента — по формуле $\arctan\left(\frac{G_y}{G_x}\right)$. В MATLAB с использованием дифференциального оператора Робертса границы можно выделить при помощи функции `edge(I, 'Roberts')`.

Фильтр Превитта

В данном подходе используются две ортогональные маски размером 3×3, позволяющие более точно вычислить производные по осям Ox и Oy:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}. \tag{3.17}$$

В MATLAB границы фильтром Превитта можно выделить при помощи функции `edge(I, 'Prewitt')`.

Фильтр Собела

Данный подход аналогичен фильтру Робертса, однако используются разные веса в масках. Типичный пример фильтра Робертса:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \tag{3.18}$$

В MATLAB границы фильтром Собела можно выделить при помощи функции `edge(I, 'Sobel')`.

Фильтр Лапласа

Фильтр Лапласа использует аппроксимацию вторых производных по осям Ox и Oy в отличие от предыдущих подходов, использующих первую производную:

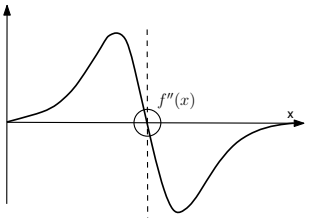


Рис. 3.7 — Вторая производная функции яркости меняет знак (проходит через ноль в точке, соответствующей краю).

$$L(I(x,y)) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}. \tag{3.19}$$

Формула 3.19 может быть аппроксимирована следующей маской:

$$w = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{3.20}$$

Алгоритм Кэнни

Одним из самых распространенных и эффективных алгоритмов выделения контуров на изображении является *алгоритм Кэнни*. Данный алгоритм позволяет не только определять краевые пиксели, но и связные граничные линии. Алгоритм состоит из следующих шагов:

1. Устранение мелких деталей путем сглаживания исходного изображения с помощью фильтра Гаусса.

2. Использование дифференциального оператора Собела для определения значений модуля градиента всех пикселей изображения, причем результат вычисления округляется с шагом $45°$.

3. Анализ значений модулей градиента пикселей, расположенных ортогонально исследуемому. Если значение модуля градиента исследуемого пикселя больше, чем у ортогональных соседних пикселей, то он является *краевым*, в противном случае — *немаксимумом*.

4. Выполнение двойной пороговой фильтрации краевых пикселей, отобранных на предыдущем шаге:

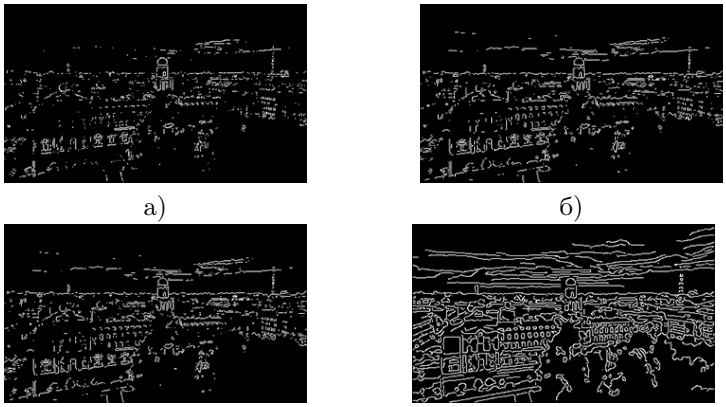
   (a) Если значение модуля градиента выше порога t_2, то наличие края в пикселе является достоверным.

   (b) Если значение модуля градиента ниже порога t_1, то пиксель однозначно не является краевым.

   (c) Если значение модуля градиента лежит в интервале $[t_1, t_2]$, то такой пиксель считается *неоднозначным*.

5. Подавление всех неоднозначных пикселей, не связанных с достоверными пикселями по 8-связности.

В MATLAB алгоритмом Кэнни можно выделить границы при помощи функции `edge(I, 'Canny')`.



а) б) в) г)

Рис. 3.8 — Результат выделения границ: а) фильтром Робертса, б) фильтром Превитта, в) фильтром Собела, г) алгоритмом Кэнни.

Порядок выполнения работы

1. *Типы шумов.* Выбрать произвольное изображение. Получить искаженные различными шумами изображения с помощью функции `imnoise()` с отличными от значений по умолчанию параметрами.

2. *Низкочастотная фильтрация.* Обработать полученные в предыдущем пункте искаженные изображения фильтром Гаусса и контргармоническим усредняющим фильтром с различными значениями параметра Q.

3. *Нелинейная фильтрация.* Обработать полученные в первом пункте искаженные изображения медианной, взвешенной медианной, ранговой и Винеровской фильтрациями при различных размерах маски и ее коэффициентов. Реализовать адаптивную медианную фильтрацию.

4. *Высокочастотная фильтрация.* Выбрать исходное изображение. Выделить границы фильтрами Робертса, Превитта, Собела, Лапласа, алгоритмом Кэнни.

Содержание отчета

1. Цель работы.
2. Теоретическое обоснование применяемых методов и функций фильтрации изображений.
3. Ход выполнения работы:
   (a) Исходные изображения;
   (b) Листинги программных реализаций;
   (c) Комментарии;
   (d) Результирующие изображения.
4. Выводы о проделанной работе.

Вопросы к защите лабораторной работы

1. В чем заключаются основные недостатки адаптивных методов фильтрации изображений?
2. При каких значениях параметра Q контргармонический фильтр будет работать как арифметический, а при каких — как гармонический?
3. Какими операторами можно выделить границы на изображении?
4. Для чего на первом шаге выделения контуров, как правило, выполняется низкочастотная фильтрация?

Лабораторная работа №4
Сегментация изображений

Цель работы

Освоение основных способов сегментации изображений на семантические области.

Методические рекомендации

До начала работы студенты должны ознакомиться с основными функциями среды MATLAB по преобразованию цветовых пространств изображений и способами определения порогов. Лабораторная работа рассчитана на 5 часов.

Теоретические сведения

Бинаризация изображений

Простейшим способом сегментации изображения на два класса (фоновые пиксели и пиксели объекта) является *бинаризация*. Бинаризацию можно выполнить по порогу или по двойному порогу. В первом случае:

$$I_{new}(x,y) = \begin{cases} 0, I(x,y) \leqslant t, \\ 1, I(x,y) > t, \end{cases} \tag{4.1}$$

где I — исходное изображение, I_{new} — бинаризованное изображение, t — порог бинаризации. Бинаризация данным методом в среде MATLAB может быть выполнена с использованием функций `im2bw()` (устаревшая) или `imbinarize()`.

Листинг 4.1. Бинаризация.

```
I = imread('pic.jpg');
L = 255;
t = 127 / L; %norm to 0...1
Inew = im2bw(I, t);
```

Бинаризация по двойному порогу (диапазонная бинаризация):

$$I_{new}(x,y) = \begin{cases} 0, I(x,y) \leqslant t_1, \\ 1, t_1 < I(x,y) \leqslant t_2, \\ 0, I(x,y) > t_2, \end{cases} \tag{4.2}$$

где I — исходное изображение, I_{new} — бинаризованное изображение, t_1 и t_2 — верхний и нижний пороги бинаризации. Бинаризация данным методом в среде MATLAB может быть выполнена с использование функции `roicolor()`. Для преобразования полноцветного изображения в полутоновое можно предварительно воспользоваться функцией `rgb2gray()`.

Листинг 4.2. Бинаризация по двойному порогу.

```
1 I = imread('pic.jpg');
2 t1 = 110;
3 t2 = 200;
4 Igray = rgb2gray(I);
5 Inew = roicolor(Igray, t1, t2);
```

Пороги бинаризации t, t_1 и t_2 могут быть либо заданы вручную, либо вычислены с помощью специальных алгоритмов. В случае автоматического вычисления порога можно воспользоваться следующими алгоритмами.

1. Поиск максимального I_{max} и минимального I_{min} значений интенсивности исходного полутонового изображения и нахождение их среднего арифметического. Среднее арифметическое будет являться глобальным порогом бинаризации t:

$$t = \frac{I_{max} - I_{min}}{2}. \tag{4.3}$$

2. Поиск оптимального порога t на основе модуля градиента яркости каждого пикселя. Для этого сначала вычисляется модуль градиента в каждой точке (x,y):

$$G(x,y) = \max\left\{|I(x+1,y) - I(x-1,y)|, |I(x,y+1) - I(x,y-1)|\right\}, \tag{4.4}$$

затем вычисляется оптимальный порог t:

$$t = \frac{\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} I(x,y)G(x,y)}{\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} G(x,y)}. \tag{4.5}$$

3. Вычисление оптимального порога t статистическим методом Отсу (Оцу, англ. Otsu), разделяющим все пиксели на два класса 1 и 2, минимизируя дисперсию внутри каждого класса $\sigma_1^2(t)$ и $\sigma_2^2(t)$ и максимизируя дисперсию между классами.

Алгоритм вычисления порога методом Отсу:

1. Вычисление гистограммы интенсивностей изображения и вероятности $p_i = \dfrac{n_i}{N}$ для каждого уровня интенсивности, где n_i — число пикселей с уровнем интенсивности i, N — число пискселей в изображении.

2. Задание начального порога $t = 0$ и порога $k \in (0, L)$, разделяющего все пиксели на два класса, где L — максимальное значение интенсивности изображения. В цикле для каждого значения порога от $k = 1$ до $k = L - 1$:

   (a) Вычисление вероятностей двух классов $\omega_j(0)$ и средних арифметических $\mu_j(0)$, где $j = \overline{1,2}$:

$$\omega_1(k) = \sum_{s=0}^{k} p_s, \tag{4.6}$$

$$\omega_2(k) = \sum_{s=k+1}^{L} p_s = 1 - \omega_1(k), \tag{4.7}$$

$$\mu_1(k) = \sum_{s=0}^{k} \frac{s \cdot p_s}{\omega_1}, \tag{4.8}$$

$$\mu_2(k) = \sum_{s=k+1}^{L} \frac{s \cdot p_s}{\omega_2}. \tag{4.9}$$

   (b) Вычисление межклассовой дисперсии $\sigma_b^2(k)$:

$$\sigma_b^2(k) = \omega_1(k)\omega_2(k)(\mu_1(k) - \mu_2(k))^2. \tag{4.10}$$

(c) Если вычисленное значение $\sigma_b^2(k)$ больше текущего значения t, то присвоить порогу значение межклассовой дисперсии $t = \sigma_b^2(k)$.

3. Оптимальный порог t соответствует максимуму $\sigma_b^2(k)$.

В среде MATLAB порог t методом Отсу может быть вычислен с использованием функции `graythresh()`:

Листинг 4.3. Бинаризация методом Отсу.

```
1 I = imread('pic.jpg');
2 t = graythresh(I);
3 Inew = im2bw(I, t);
```

либо с использованием функции `otsuthresh()` на основе гистограммы изображения:

Листинг 4.4. Бинаризация методом Отсу на основе гистограммы.

```
1 I = imread('pic.jpg');
2 Igray = rgb2gray(I);
3 [counts,x] = imhist(Igray);
4 t = otsuthresh(counts);
5 Inew = imbinarize(Igray, t);
```

4. Адаптивные методы, работающие не со всем изображением, а лишь с его фрагментами. Такие подходы зачастую используются при работе с изображениями, на которых представлены неоднородно освещенные объекты. В среде MATLAB порог t адаптивным методом может быть вычислен при помощи функции `adaptthresh()`:

Листинг 4.5. Бинаризация адаптивным методом.

```
1 I = imread('pic.jpg');
2 Igray = rgb2gray(I);
3 t = adaptthresh(Igray);
4 Inew = imbinarize(Igray, t);
```

Помимо рассмотренных методов существуют и многие другие, например методы Бернсена, Эйквела, Ниблэка, Яновица и Брукштейна и др.

Сегментация изображений

Рассмотрим несколько основных методов сегментации изображений.

На основе принципа Вебера

Алгоритм предназначен для сегментации полутоновых изображений. *Принцип Вебера* подразумевает, что человеческий глаз плохо воспринимает разницу уровней серого между $I(n)$ и $I(n) + W(I(n))$, где $W(I(n))$ — функция Вебера, n — номер класса, I — кусочно-нелинейная функция градаций серого. Функция Вебера может быть вычислена по формуле:

$$W(I) = \begin{cases} 20 - \dfrac{12I}{88}, 0 \leqslant I \leqslant 88, \\ 0{,}002(I - 88)^2, 88 < I \leqslant 138, \\ \dfrac{7(I - 138)}{117} + 13, 138 < I \leqslant 255. \end{cases} \tag{4.11}$$

Можно объединить уровни серого из диапазона $[I(n),I(n) + W(I(n))]$ заменив их одним значением интенсивности.

Алгоритм сегментации состоит из следующих шагов:

1. Инициализация начальных условий: номер первого класса $n = 1$, уровень серого $I(n) = 0$.

2. Вычисление значения $W(I(n))$ по формуле Вебера и присваивание значения $I(n)$ всем пикселям, интенсивность которых находится в диапазоне $[I(n),I(n) + W(I(n))]$.

3. Поиск пикселей с интенсивностью выше $G = I(n) + W(I(n)) + 1$. Если найдены, то увеличение номера класса $n = n + 1$, присваивание $I(n) = G$ и переход на второй шаг. В противном случае закончить работу. Изображение будет сегментировано на n классов интенсивностью $W(I(n))$.

Сегментация RGB-изображений по цвету кожи

Общим приницпом данного подхода является определение критерия близости интенсивности пикселей к оттенку кожи. Аналитически описать *оттенок кожи* довольно затруднительно, поскольку

его описание базируется на человеческом восприятии цвета, меняется при изменении освещения, отличается у разных народностей, и т.д.

Существует несколько аналитических описаний для изображений в цветовом пространстве RGB, позволяющих отнести пиксель к классу «кожа» при выполнении условий:

$$\begin{cases} R > 95, \\ G > 40, \\ B < 20, \\ \max R,G,B - \min R,G,B > 15, \\ |R - G| > 15, \\ R > G, \\ R > B, \end{cases} \tag{4.12}$$

или

$$\begin{cases} R > 220, \\ G > 210, \\ B > 170, \\ |R - G| \leqslant 15, \\ G > B, \\ R > B, \end{cases} \tag{4.13}$$

или

$$\begin{cases} r = \frac{R}{R+G+B}, \\ g = \frac{G}{R+G+B}, \\ b = \frac{B}{R+G+B}, \\ \frac{r}{g} > 1{,}185, \\ \frac{rb}{(r+g+b)^2} > 0{,}107, \\ \frac{rg}{(r+g+b)^2} > 0{,}112. \end{cases} \tag{4.14}$$

На основе цветового пространства CIE Lab

В цветовом пространстве **Lab** значение светлоты отделено от значения хроматической составляющей цвета (тон, насыщенность).

Светлота задается координатой $\boldsymbol{L}$, которая может находиться в диапазоне от 0 (темный) до 100 (светлый). Хроматическая составляющая цвета задается двумя декартовыми координатами $\boldsymbol{a}$ (означает положение цвета в диапазоне от *зеленого* (-128) до *красного* (127)) и $\boldsymbol{b}$ (означает положение цвета в диапазоне от *синего* (-128) до *желтого* (127)). Бинарное изображение получается при нулевых значениях координат $\boldsymbol{a}$ и $\boldsymbol{b}$. Идея алгоритма состоит в разбиении цветного изображения на сегменты доминирующих цветов.

В качестве исходных данных выберем следующее цветное изображение:

Рис. 4.1 — Исходное цветное изображение.

В первую очередь, чтобы уменьшить влияние освещенности на результат сегментации, преобразуем полноцветное изображение из цветового пространства **RGB** в пространство **Lab**. Для этого в среде MATLAB используется функция `rgb2lab()`.

Листинг 4.6. Сегментация на основе цветового пространства **Lab**.

```
1 I = imread('pic2.jpg');
2 Ilab = rgb2lab(I);
3 L = Ilab(:,:,1);
4 a = Ilab(:,:,2);
5 b = Ilab(:,:,3);
```

На следующем шаге необходимо определить количество цветов, на которые будет сегментировано изображение, и задать области, содержащие пиксели примерно одного цвета. Области можно задать для каждого цвета интерактивно в виде многоугольников при помощи функции `roipoly()`:

```
6 numColors = 3;
7 sampleAreas = false([size(I, 1)
```

```
 8     size(I, 2) numColors]);
 9 for i=1:1:numColors
10     [BW, xi, yi] = roipoly(I);
11     sampleAreas(:,:,i) = roipoly(I, xi, yi);
12 end
```

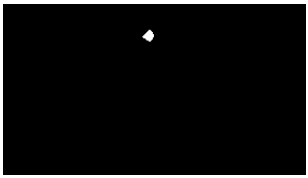


Рис. 4.2 — Пример выделенной красной области из четырех точек.

После этого требуется определить цветовые метки для каждого из сегментов путем расчета среднего значения цветности в каждой выделенной области. Средние значения можно вычислить при помощи функции `mean2()`:

```
13 colorMarks = zeros([numColors, 2]);
14 for i=1:1:numColors
15     colorMarks(i,1) = ...
16         mean2(a(sampleAreas(:,:,i)));
17     colorMarks(i,2) = ...
18         mean2(b(sampleAreas(:,:,i)));
19 end
```

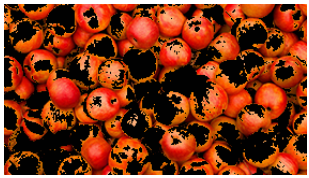
Затем используем принцип ближайшей окрестности для классификации пикселей путем вычисления евклидовых метрик между пикселями и метками: чем меньше расстояние до метки, тем лучше пиксель соответствует данному сегменту. Евклидова метрика по двум цветовым координатам расчитывается по формуле: $\sqrt{(a(x,y) - a_{mark})^2 + (b(x,y) - b_{mark})^2}^2$. Для поиска минимального расстояния будем использовать функцию `min()`. Приведем листинг для поиска меток сегментов `label` для каждого пикселя:

```
20 distance = zeros([size(a), numColors]);
21 for i=1:1:numColors
22     distance(:,:,i) = ...
```

```
23          ((a-colorMarks(i,1)).^2 + ...
24          (b-colorMarks(i,2)).^2).^0.5;
25      colorLabels = i;
26  end
27  [~, label] = min(distance, [], 3);
28  label = colorLabels(label);
```

Таким образом, матрица `label` рамерности равном исходному изображению будет содержать идентификаторы классов для каждого пикселя. Для сегментации изображения на фрагменты `segmentedFrames` используем следующий листинг:

```
29  rgbLabel = repmat(label, [1 1 3]);
30  segmentedFrames = ...
31      zeros([size(I), numColors], 'uint8');
32  for i=1:1:numColors
33      color = I;
34      color(rgbLabel ~= colorLabels(i)) = 0;
35      segmentedFrames(:,:,:,i) = color;
36  end
```



а)

б)

Рис. 4.3 — Сегментированные области: а) красные, б) желтые.

Отметим распредление сегментированных пикселей на координатной плоскости (a,b):

```
37  figure
38  for i=1:1:numColors
39      plot(a(label == i), b(label==i), ...
40          '.', 'MarkerEdgeColor', ...
41          plotColors{i}, ...
42          'MarkerFaceColor', plotColors{i});
```

```
43      hold on
44  end
```

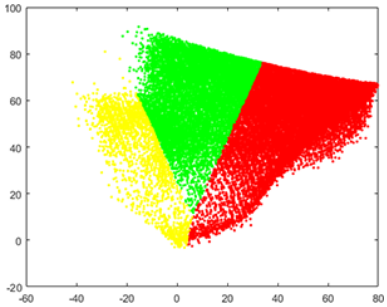


Рис. 4.4 — Распределение сегментированных пикселей на координатной плоскости (a,b).

На основе кластеризации методом k-средних

Идея метода заключается в определении центров k-кластеров и отнесении к каждому кластеру пикселей, наиболее близко относящихся к этим центрам. Все пиксели рассматриваются как векторы $x_i, i = \overline{1,p}$. Алгоритм сегментации состоит из следующих шагов:

1. Определение случайным образом k векторов $m_j, j = \overline{1,k}$, которые объявляются начальными центрами кластеров.

2. Обновление значений средних векторов m_j путем вычисления расстояний от каждого вектора x_i до каждого m_j и их классификации по критерию минимальности расстояния от вектора до кластера, пересчет средних значений m_j по всем кластерам.

3. Повторение шагов 2 и 3 до тех пора, пока центры кластеров не перестанут изменяться.

Реализация метода очень похожа на предыдущий подход и содержит ряд аналогичных действий (используем исходное изображение рис. 4.1). Будем работать в цветовом пространстве **Lab**, поэтому первым шагом перейдем из пространства **RGB** в **Lab**:

Листинг 4.7. Сегментация на основе кластеризации методом k-средних.

```
I = imread('pic2.jpg');
Ilab = rgb2lab(I);
L = Ilab(:,:,1);
a = Ilab(:,:,2);
b = Ilab(:,:,3);
```

Рассмотрим координатную плоскость (a,b). Для этого сформируем трехмерный массив `ab`, а затем функцией `reshape()` превратим его в двумерный вектор, содержащий все пиксели изображения:

```
ab(:,:,1) = a;
ab(:,:,2) = b;
nrows = size(I, 1);
ncols = size(I, 2);
ab = reshape(ab, nrows * ncols, 2);
```

Кластеризация методом k-средних в среде MATLAB осуществляется функцией `kmeans()`. Аналогично предыдущему способу разобьем изображение на три области соответствующих цветов. Используем евклидову метрику (параметр `'distance'` со значением `'sqEuclidean'` и для повышения точности повторим процедуру кластеризации три раза (параметр `'Replicates'` со значением `3`)):

```
k = 3;
[ids, centers] = kmeans(ab, k, 'distance',...
    'sqEuclidean', 'Replicates', 3);
label = reshape(ids, nrows, ncols);
```

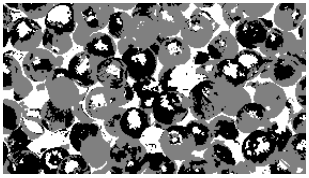
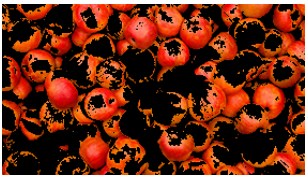
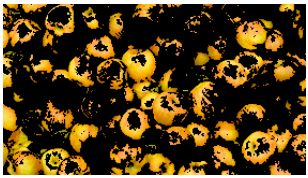


Рис. 4.5 — Метки классов.

Матрица `label` размера равном исходному изображению будет содержать идентификаторы классов для каждого пикселя. Для сегментации изображения на фрагменты `segmentedFrames` используем следующий листинг:



а)



б)

Рис. 4.6 — Сегментированные области: а) красные, б) желтые.

```
15 segmentedFrames = cell(1, 3);
16 rgbLabel = repmat(label, [1 1 3]);
17 for i = 1:1:k
18     color = I;
19     color(rgbLabel ~= i) = 0;
20     segmentedFrames{i} = color;
21     figure, imshow(segmentedFrames{i});
22 end
```

Массив L содержит значение о светлоте изображения. Используя эти данные можно, например, сегментированные красные области разделить на светло-красные и темно-красные сегменты.

Текстурная сегментация

При текстурной сегментации для описания текстуры применяются три основных метода: статистический, структурный и спектральный. В лабораторной работе будем рассматривать статистический подход, который описывает текстуру сегмента как гладкую, грубую или зернистую. Характеристики соответствующих текстурам параметров приведены в табл. 4.1. Рассмотрим пример изображения, представленного на рис. 4.7, на котором имеется два типа текстур. Их разделение в общем случае невозможно выполнить с использованием только лишь простой бинаризации.

Будем рассматривать интенсивность изображения I как случайную величину z, которой соответствует вероятность распределения $p(z)$, вычисляемая из гистограммы изображения. *Централь-*

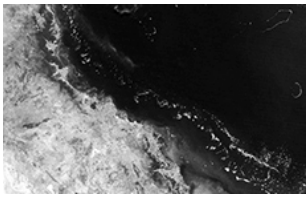


Рис. 4.7 — Исходное полутоновое изображение.

ным моментом порядка n случайной величины z называется параметр $\mu_n(z)$, вычисляемый по формуле:

$$\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i), \tag{4.15}$$

где L — число уровней интенсивностей, m — среднее значение случайной величины z:

$$m = \sum_{i=0}^{L-1} z_i p(z_i). \tag{4.16}$$

Из 4.15 следует, что $\mu_0 = 1$ и $\mu_1 = 0$. Для описания текстуры важна *дисперсия* случайной величины, равная второму моменту $\sigma^2(z) = \mu_2(z)$ и являющаяся мерой яркостного контраста, которую можно использовать для вычисления признаков *гладкости*. Введем меру относительной гладкости R:

$$R = 1 - \frac{1}{1 + \sigma^2(z)}, \tag{4.17}$$

которая равна нулю для областей с постоянной интенсивностью (нулевой дисперсией) и приближается к единице для больших значений дисперсий $\sigma^2(z)$. Для полутоновых изображений с интервалом интенсивностей [0,255] необходимо нормировать дисперсию до интервала [0,1], поскольку для исходного диапазона значения дисперсий будут слишком велики. Нормирование осуществляется де-

Таблица 4.1 — Значения параметров текстур.

Текстура	m	s	$R \in [0,1]$
Гладкая	82,64	11,79	0,0020
Грубая	143,56	74,63	0,0079
Периодическая	99,72	33,73	0,0170

Текстура	$\mu_3(z)$	U	E
Гладкая	−0,105	0,026	5,434
Грубая	−0,151	0,005	7,783
Периодическая	0,750	0,013	6,674

лением дисперсии $\sigma^2(z)$ на $(L-1)^2$. В качестве характеристики текстуры также зачастую используется *стандартное отклонение*:

$$s = \sigma(z). \tag{4.18}$$

Третий момент является *характеристикой симметрии гистограммы*. Для оценки текстурных особенностей используется функция *энтропии* E, определяющая разброс интенсивностей соседних пикселей:

$$E = -\sum_{i=0}^{L-1} p(z) \log_2 p(z_i). \tag{4.19}$$

Еще одной важной характеристикой, описывающей текстуру, является *мера однородности* U, оценивающая равномерность гистограммы:

$$U = \sum_{i=0}^{L-1} p^2(z_i). \tag{4.20}$$

После вычисления какого-либо признака или набора признаков необходимо построить бинарную маску, на основе которой и будет производится сегментация изображения. Например, можно использовать энтропию E в окрестности каждого пикселя (x,y). Для этого в среде MATLAB используется функция `entropyfilt()`, по умолчанию у которой используется окрестность размера 9×9. Для нор-

мирования функции энтропии в диапазоне от 0 до 1 используем функцию `mat2gray()`, а для построения маски бинаризуем полученный нормированный массив `Eim` методом Отсу.

Листинг 4.8. Текстурная сегментация.

```
I = imread('pic3.jpg');
E = entropyfilt(I);
Eim = mat2gray(E);
BW1 = imbinarize(Eim,graythresh(Eim));
```

а) б)

Рис. 4.8 — а) Энтропия исходного изображения, б) бинаризованное изображение.

После этого используем морфологические фильтры (будут рассмотрены подробнее в лабораторной работе №6) сначала для удаления связных областей, содержащих менее заданного количества пикселей (функция `bwareaopen()`), а затем для удаления внутренних *дефектов формы* или «дырок» (функция `imclose()` со структурным элементом размера 9×9). Оставшиеся крупные «дырки» заполним при помощи функции `imfill()`. Таким образом, получим маску:

```
BWao = bwareaopen(BW1,2000);
nhood = true(9);
closeBWao = imclose(BWao,nhood);
Mask1 = imfill(closeBWao,'holes');
```

Применив полученную маску к исходному изображению выделим сегменты воды и суши.

Границу между текстурами рассчитаем с использованием функции определения периметра `bwperim()`:

а) б)

Рис. 4.9 — а) Результат выполнения функции `bwareaopen()`; б) результат выполнения функции `imclose()`.

а) б)

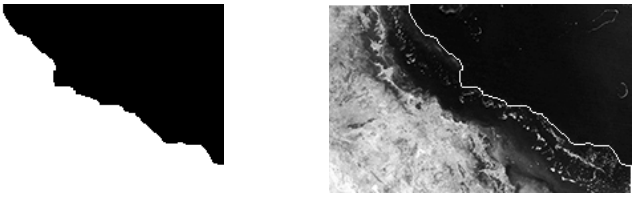
Рис. 4.10 — а) Текстура суши, б) текстура воды.

```
 9 boundary = bwperim(Mask1);
10 segmentResults = I;
11 segmentResults(boundary) = 255;
```

Аналогичный подход можно применить для построения маски относительно суши:

```
12 I2 = I;
13 I2(Mask1) = 0;
14 E2 = entropyfilt(I2);
15 E2im = mat2gray(E2);
16 BW2 = imbinarize(E2im, graythresh(E2im));
17 Mask2 = bwareaopen(BW2,2000);
18 boundary = bwperim(Mask2);
19 segmentResults = I;
20 segmentResults(boundary) = 255;
```

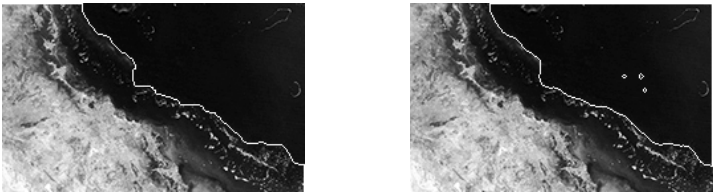
Найдем текстуры суши и воды:



а) б)

Рис. 4.11 — а) Результат выполнения функции `imfill()`, б) выделенная граница функцией `bwperim()`.

```
21 texture1 = I;
22 texture1(~Mask2) = 0;
23 texture2 = I;
24 texture2(Mask2) = 0;
```



а) б)

Рис. 4.12 — а) Результат сегментации относительно воды, б) результат сегментации относительно суши.

Порядок выполнения работы

1. *Бинаризация.* Выбрать произвольное изображение. Выполнить бинаризацию изображения при помощи рассмотренных методов. В зависимости от изображения использовать бинаризацию по верхнему или нижнему порогу.

2. *Сегментация 1.* Выбрать произвольное изображение, содержащее лицо(-а). Выполнить сегментацию изображения либо по принципу Вебера, либо на основе цвета кожи (на выбор).

3. *Сегментация 2.* Выбрать произвольное изображение, содержащее ограниченное количество цветных объектов. Выполнить сегментацию изображения в пространстве **CIE Lab** либо по методу ближайших соседей, либо по методу k-средних (на выбор).

4. *Сегментация 3.* Выбрать произвольное изображение, содержащее две разнородные текстуры. Выполнить текстурную сегментацию изображения, оценить не менее трех параметров выделенных текстур, определить к какому классу относятся текстуры.

Содержание отчета

1. Цель работы.

2. Теоретическое обоснование применяемых методов и функций сегментации изображений.

3. Ход выполнения работы:

   (a) Исходные изображения;

   (b) Листинги программных реализаций;

   (c) Комментарии;

   (d) Результирующие изображения.

4. Выводы о проделанной работе.

Вопросы к защите лабораторной работы

1. В каких случаях целесообразно использовать сегментацию по принципу Вебера?

2. Какие значения имеют цветовые координаты $\boldsymbol{a}$ и $\boldsymbol{b}$ цветового пространства **CIE Lab** в полутновом изображении?

3. Зачем производить сегментацию в цветовом пространстве **CIE Lab**, а не в исходном **RGB**?

4. Что такое *цветовое пространство* и *цветовой охват*?

Лабораторная работа №5
Преобразование Хафа

Цель работы

Освоение преобразования для поиска геометрических примитивов.

Методические рекомендации

До начала работы студенты должны ознакомиться с функциями среды MATLAB для работы с преобразованием Хафа. Знать о подходе «голосования» точек. Лабораторная работа рассчитана на 5 часов.

Теоретические сведения

Идея преобразования Хафа (англ. Hough, возможные варианты перевода Хох, Хо) заключается в поиске общих *геометрических мест точек* (ГМТ). Например, данный подход используется при построении треугольника по трем заданным сторонам, когда сначала откладывается одна сторона треугольника, после этого концы отрезка рассматриваются как центры окружностей радиусами равными длинам второго и третьего отрезков. Место пересечения двух окружностей является общим ГМТ, откуда и проводятся отрезки до концов первого отрезка. Иными словами можно сказать, что было проведено *голосование* двух точек в пользу вероятного расположения третьей вершины треугольника. В результате «голосования» «победила» точка, набравшая два «голоса» (точки на окружностях набрали по одному голосу, а вне их — по нулю).

Обобщим данную идею для работы с реальными данными, когда на изображении имеется большое количество особых характеристических точек, участвующих в голосовании. Допустим, необходимо найти в бинарном точечном множестве окружность известного радиуса R, причем в данном множестве могут присутствовать и ложные точки, не лежащие на искомой окружности. Набор центров возможных окружностей искомого радиуса вокруг каждой характеристической точки образует окружность радиуса R, см. рис. 5.2.

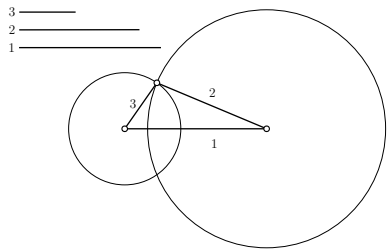


Рис. 5.1 — Построение треугольника по трем заданным сторонам.

Таким образом, точка, соответствующая максимальному пересечению числа окружностей, и будет являться центром окружности искомого радиуса.

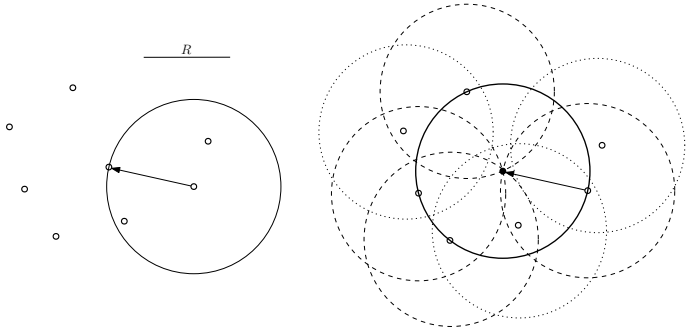


Рис. 5.2 — Обнаружение окружности известного радиуса в точечном множестве.

Классическое преобразование Хафа, базирующееся на рассмотренной идее голосования точек, изначально было предназначено для выделения прямых на бинарных изображениях. В преобразовании Хафа для поиска геометрических примитивов используется пространство параметров. Самым распространенным параметрическим уравнением прямых является:

$$y = kx + b, \tag{5.1}$$

$$x\cos\Theta + y\sin\Theta = \rho, \tag{5.2}$$

где ρ — радиус-вектор, проведенный из начала координат до прямой; Θ — угол наклона радиус-вектора.

Пусть в декартовой системе координат прямая задана уравнением (5.1), из которого легко вычислить радиус-вектор ρ и угол Θ (5.2). Тогда в пространстве параметров Хафа прямая будет представлена точкой с координатами (ρ_0,Θ_0), см. рис. 5.3.

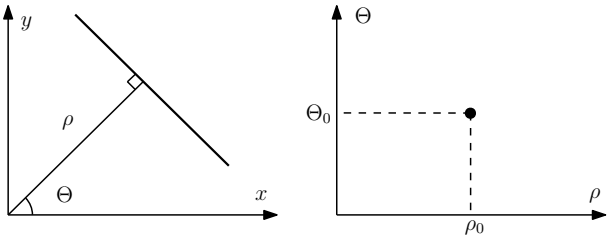


Рис. 5.3 — Представление прямой в пространстве Хафа.

Подход преобразования Хафа заключается в том, что для каждой точки пространства параметров суммируется количество голосов, поданных за нее, поэтому в дискретном виде пространство Хафа называется *аккумулятором* и представляет собой некоторую матрицу $A(\rho,\Theta)$, хранящую информацию о голосовании. Через каждую точку в декартовой системе координат можно провести бесконечное число прямых, совокупность которых породит в пространстве параметров синусоидальную функцию отклика. Таким образом, любые две синусоидальные функции отклика в пространстве параметров пересекутся в точке (ρ,Θ) только в том случае, если порождающие их точки в исходном пространстве лежат на прямой, см. рис. 5.4. Исходя из этого можно сделать вывод, что для того, чтобы найти прямые в исходном пространстве, необходимо найти все локальные максимумы аккумулятора.

Рассмотренный алгоритм поиска прямых может быть таким же образом использован для поиска любой другой кривой, описываемой в пространстве некоторой функцией с определенным числом параметров $F = (a_1,a_2,...,a_n,x,y)$, что повлияет лишь на размерность пространства параметров. Воспользуемся преобразованием

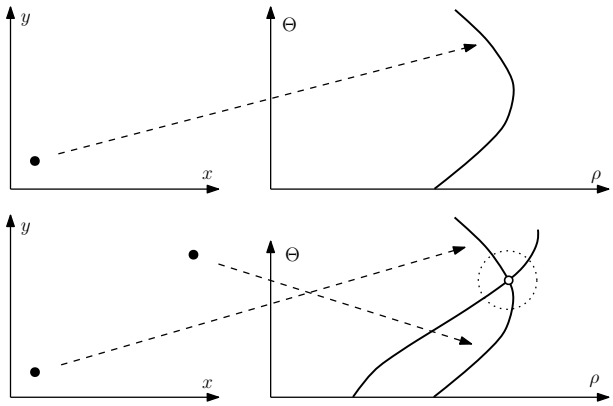


Рис. 5.4 — Процедура голосования.

Хафа для поиска окружностей заданного радиуса R. Известно, что окружность на плоскости описывается формулой $(x - x_0)^2 + (y - y_0)^2 = R^2$. Набор центров всех возможных окружностей радиуса R, проходящих через характеристическую точку, образует окружность радиуса R вокруг этой точки, поэтому функция отклика в преобразовании Хафа для поиска окружностей представляет окружность такого же размера с центром в голосующей точке. Тогда аналогично предыдущему случаю необходимо найти локальные максимумы аккумуляторной функции $A(x,y)$ в пространстве парамаетров (x,y), которые и будут являться центрами искомых окружностей.

Преобразование Хафа инвариантно к сдвигу, масштабированию и повороту. Учитывая, что при проективных преобразованиях трехмерного пространства прямые линии всегда переходят только в прямые линии (в вырожденном случае — в точки), преобразование Хафа позволяет обнаруживать линии инвариантно не только к аффинным преобразованиям плоскости, но и к группе проективных преобразований в пространстве.

Пусть задано некоторое изображение. Выделим контуры алгоритмом Кэнни и выполним преобразование Хафа функцией `hough()`.

Листинг 5.1. Поиск прямых преобразованием Хафа.

```
I = imread('pic.png');
Iedge = edge(I, 'Canny');
[H,Theta,rho] = hough(Iedge);
figure, imshow(imadjust(mat2gray(H)),[],...
    'YData',rho,'XData',Theta,...
    'InitialMagnification','fit');
xlabel('\rho'), ylabel('\Theta')
axis on, axis normal, hold on
```

Вычислим пики функцией `houghpeaks()` в пространстве Хафа и нанесем их на полученное изображение функций откликов:

```
peaks = houghpeaks(H,100,'threshold',...
    ceil(0.5*max(H(:))));
x = Theta(peaks(:,2));
y = rho(peaks(:,1));
plot(x,y,'s','color','white');
```

Определим на основе пиков прямые функцией `houghlines()` и нанесем их на исходное изображение:

```
lines = houghlines(Iedge,Theta,rho,peaks,...
    'FillGap',5,'MinLength',10);
figure, imshow(I), hold on
for k = 1:length(lines)
   xy = [lines(k).point1; lines(k).point2];
   plot(xy(:,1),xy(:,2),'LineWidth',2,...
        'Color','green');
end
```

Для поиска окружностей преобразованием Хафа можно воспользоваться функцией `imfindcircles(I,R)`.

Порядок выполнения работы

1. *Поиск прямых.* Выбрать три произвольных изображения, содержащие прямые. Осуществить поиск прямых с помощью преобразования Хафа как для исходного изображения, так и для изображения, полученного с помощью использования какого-либо дифференциального оператора. Отразить найденные линии на исходном изображении. Отметить точки на-

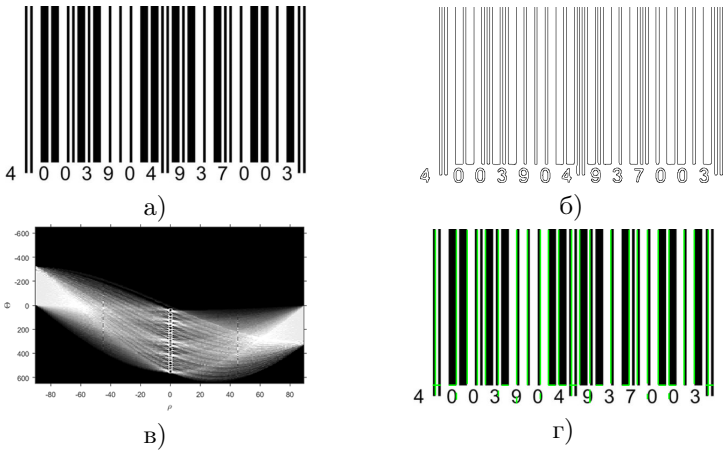


Рис. 5.5 — а) Исходное изображение, б) обработанное алгоритмом Кэнни, в) пространство параметров, г) выделенные линии.

чала и окончания линий. Определить длины самой короткой и самой длинной прямых, вычислить количество найденных прямых.

2. *Поиск окружностей.* Выбрать три произвольных изображения, содержащие окружности. Осуществить поиск окружностей как определенного радиуса, так и из заданного диапазона с помощью преобразования Хафа как для исходного изображения, так и для изображения, полученного с помощью использования какого-либо дифференциального оператора. Отразить найденные окружности на исходном изображении.

Содержание отчета

1. Цель работы.
2. Теоретическое обоснование применяемого преобразования для поиска геометрических примитивов.
3. Ход выполнения работы:

(a) Исходные изображения;

(b) Листинги программных реализаций;

(c) Комментарии;

(d) Результирующие изображения.

4. Выводы о проделанной работе.

Вопросы к защите лабораторной работы

1. Какая идея лежит в основе преобразования Хафа?
2. Можно ли использовать преобразование Хафа для поиска произвольных контуров, которые невозможно описать аналитически?
3. Что такое *рекуррентное* и *обобщенное* преобразования Хафа?
4. Какие бывают способы параметризации в преобразовании Хафа?

Лабораторная работа №6
Морфологический анализ изображений

Цель работы

Освоение принципов математической морфологии в области обработки и анализа изображений.

Методические рекомендации

До начала работы студенты должны ознакомиться с функциями среды MATLAB для работы с бинарной морфологией. Знать основные операции и положения бинарной морфологии. Лабораторная работа рассчитана на 5 часов.

Теоретические сведения

Термин *морфология* дословно переводится как «наука о форме». Морфологический анализ используется во многих областях знаний, в т.ч. и в обработке изображений. Морфологический анализ является относительно универсальным подходом, поэтому стоит обособленно от всех рассмотренных ранее методов. С использованием математической морфологии при обработке изображений можно осуществлять фильтрацию шумов, сегментацию объектов, выделение контуров, реализовать поиск заданного объекта на изображении, вычислить «скелет» образа и т.д. Рассмотрим основные операции бинарной морфологии над изображением A структурным элементом B:

1. Дилатация (расширение, наращивание): $A \oplus B$, в MATLAB выполняется функцией `imdilate(A,B)`, расширяет бинарный образ A структурным элементом B;

2. Эрозия (сжатие, сужение): $A \ominus B$, в MATLAB выполняется функцией `imerode(A,B)`, сужает бинарный образ A структурным элементом B;

3. Открытие (отмыкание, размыкание, раскрытие): $(A \ominus B) \oplus B$, в MATLAB выполняется функцией `imopen(A,B)`, удаляет внешние дефекты бинарного образа A структурным элементом B;

4. Закрытие (замыкание): $(A \oplus B) \ominus B$, в MATLAB выполняется функцией `imclose(A,B)`, удаляет внутренние дефекты бинарного образа A структурным элементом B,

где $\oplus$ и $\ominus$ — сложение и вычитание Минковского, соответственно. В среде MATLAB имеются следующие полезные функции, часто используемые совместно с морфологическими операциями:

1. `strel()` — задает структурный элемент;

2. `bwmorph(A,operation,n)` — применяет морфологическую фильтрацию `operation` к образу `A` `n` раз;

3. `bwhitmiss(A,B,C)` — выполняет операцию «hit & miss», определяющую пиксели, окрестности которых совпадают по форме со структурным элементом `B` и не совпадают со структурным элементом `C`;

4. `bwareaopen(A,dim)` — удаляет объекты на изображении `A`, содержащие менее `dim` пикселей;

5. `bwselect()` — выбирает определенные объекты на изображении;

6. `bwarea()` — вычисляет площадь объектов;

7. `bwlabel()` — выполняет маркировку связных объектов бинарного изорбражения;

8. `bweuler()` — вычисляет число Эйлера бинарного изображения.

Примеры использования морфологических операций

Выделение границ объектов

Границы объектов могут быть выделены с использованием следующего подхода:

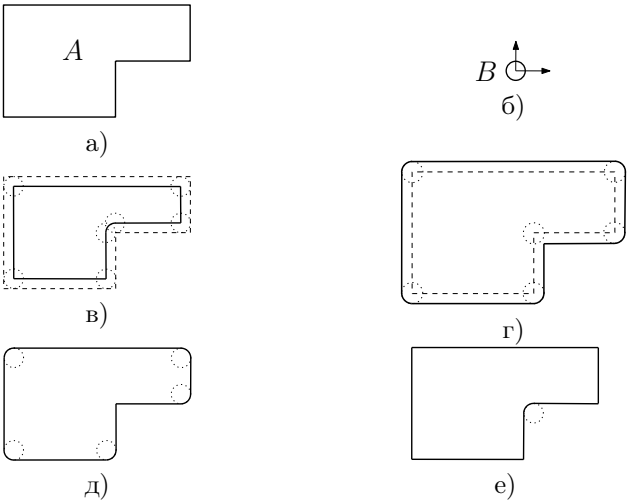


Рис. 6.1 — Результат выполения операций бинарной морфологии: а) бинарный образ A, б) дисковый структурный элемент B, в) операция дилатации, г) операция эрозии, д) операция открытия, е) операция закрытия.

1. $C = A - (A \ominus B)$ — формирование внутреннего контура;
2. $C = (A \oplus B) - A$ — формирование внешнего контура.

Разделение «склеенных» объектов

Одним из примеров использования морфологических операций над изображением может быть задача разделения склеившихся на изображении объектов. Задача может быть решена с достаточной степенью точности при помощи последовательного выполнения нескольких раз фильтра сжатия, а затем максимально возможного расширения полученного результата. Пересечение исходного изображения с обработанным позволит разделить склеенные объекты.

Листинг 6.1. Разделение «склееных» объектов.

```
1 I = imread('pic.jpg');
2 t = graythresh(I);
```

```
3 Inew = im2bw(I,t);
4 Inew = ~Inew;
5 BW2 = bwmorph(Inew,'erode',7);
6 BW2 = bwmorph(BW2,'thicken',Inf);
7 Inew = ~(Inew & BW2);
```

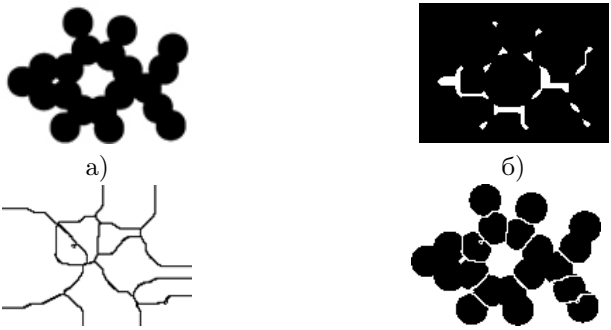


Рис. 6.2 — Разделение «склеенных» объектов: а) исходное изображение, б) эрозия бинарного изображения, в) расширение объектов, г) результат разделения.

Сегментация методом управляемого водораздела

Рассмотрим еще один из примеров применения математической морфологии к задаче сегментации изображения. В подходе сегментации по водоразделам изображение рассматривается как карта высот, на котором интенсивности пикселей описывают высоты относительно некоторого уровня. На такую «высотную местность» «льет дождь», образуя множество *водосборных бассейнов*. Постепенно вода из переполненных бассейнов переливается, и бассейны объединяются в более крупные, см. рис. 6.3. Места объединения бассейнов отмечаются как линии водораздела. Если «дождь» остановить рано, тогда изображение будет сегментировано на мелкие области, а если поздно — на крупные. В таком подходе все пиксели подразделяются на три типа:

1. *локальные минимумы*;

2. находящиеся *на склоне* (с которых вода скатывается в один и тот же локальный минимум);

3. *локальные максимумы* (с которых вода скатывается более чем в один минимум).

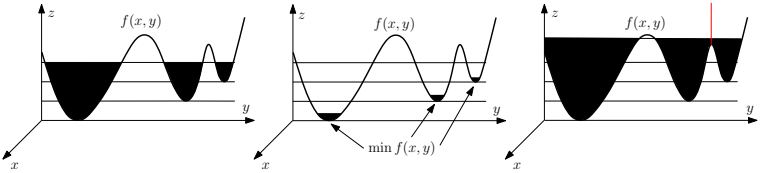


Рис. 6.3 — Водораздел областей.

При реализации данного метода необходимо определить водосборные бассены и линии водораздела путем обработки локальных областей и вычисления их характеристик. Алгоритм сегментации состоит из следующих шагов:

1. Вычисление функции сегментации. Как правило, для этого используется градиентное представление изображения.

2. Вычисление маркеров переднего плана на основании связности пикселей каждого объекта.

3. Вычисление маркеров фона, представляющих пиксели, не являющиеся объектами.

4. Модифицирование функции сегментации с учетом взаиморасположения маркеров переднего плана и фона.

В результате работы алгоритма будет получена маска, где пиксели одинаковых сегментов будут помечены одинаковыми метками и будут образовывать связную область. Пусть задано изображение рис. 6.4.

Выполним морфологическую фильтрацию изображения с использованием базовых морфологических операций и морфологической реконструкции `imreconstruct()`. Функция `imcomplement()`

Рис. 6.4 — Исходное изображение.

вычисляет изображение-дополнение к изображению-аргументу и представляет собой инвертированное изображение.

Листинг 6.2. Сегментации методом водораздела.

```
1 rgb = imread('pic.jpg');
2 A = rgb2gray(rgb);
3 B = strel('disk',6);
4 C = imerode(A,B);
5 Cr = imreconstruct(C,A);
6 Crd = imdilate(Cr,B);
7 Crdr = imreconstruct(imcomplement(Crd), ...
8     imcomplement(Cr));
9 Crdr = imcomplement(Crdr);
```

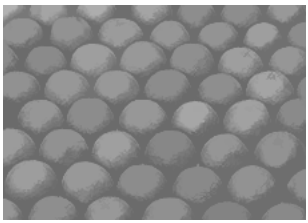
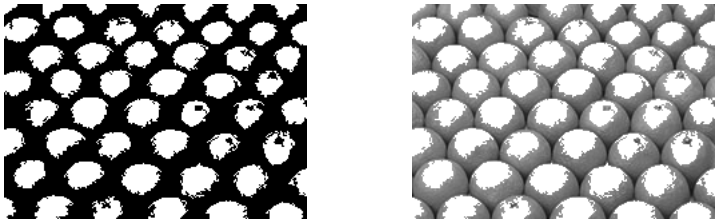


Рис. 6.5 — Отфильтрованное изображение.

После этого определим локальные максимумы функцией `imregionalmax()` для определения маркеров переднего плана (*foreground markers* или `fgm`) и, для наглядности, наложим маркеры на исходное изображение:

```
10 fgm = imregionalmax(Crdr);
11 A2 = A;
12 A2(fgm) = 255;
```



а) б)

Рис. 6.6 — Маркеры переднего плана: а) вычисленные, б) наложенные на исходное изображение.

Как видно из представленных изображений, маркеры сильно изрезаны. Для сглаживания маркеров переднего плана воспользуемся следующей последовательностью морфологических операций:

```
13 B2 = strel(ones(5,5));
14 fgm = imclose(fgm,B2);
15 fgm = imerode(fgm,B2);
16 fgm = bwareaopen(fgm,20);
17 A3 = A;
18 A3(fgm) = 255;
```

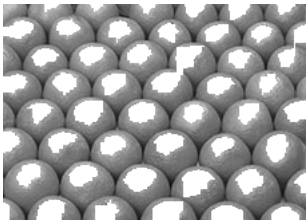
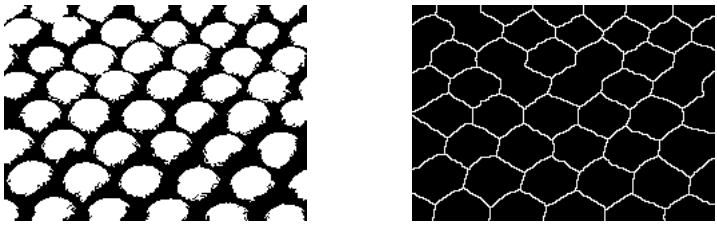


Рис. 6.7 — Отфильтрованные маркеры переднего плана.

На следующем шаге требуется определить маркеры фона (*background markers* или `bgm`). Для этого бинаризуем отфильтро-

ванное изображение `Crdr`, найдем евклидово расстояние от каждого черного до ближайшего белого пикселя функцией `bwdist()` и вычислим матрицу `L`, содержащую метки сегментов, полученных методом водораздела с использованием функции `watershed()`:

```
19 bw = imbinarize(Crdr);
20 D = bwdist(bw);
21 L = watershed(D);
22 bgm = L == 0;
```



а) б)

Рис. 6.8 — а) бинаризованное изображение, б) маркеры фона.

Затем требуется модифицировать функцию сегментации. Для этого можно воспользоваться функцией `imimposemin()`, которая точно определяет локальные минимумы изображения. Благодаря этому, функция корректирует значения градиентов на изображении и уточняет расположение маркеров переднего плана и фона. Перед этим нужно определить градиентное представление изображения, которое и будет скорректировано:

```
23 hy = fspecial('sobel');
24 hx = hy';
25 Ay = imfilter(double(A), hy, 'replicate');
26 Ax = imfilter(double(A), hx, 'replicate');
27 grad = sqrt(Ax.^2 + Ay.^2);
28 grad = imimposemin(grad, bgm | fgm);
```

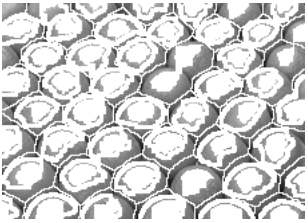
После этого выполняется операция сегментации уточненного градиентного представления изображения на основе водораздела и визуализируется результат работы алгоритма:

а)

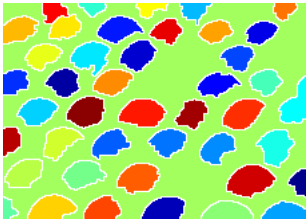
б)

Рис. 6.9 — Градиентное представление изображения: а) исходное, б) модифицированное.

```
29 L = watershed(grad);
30 A4 = A;
31 A4(imdilate(L == 0, ...
32      ones(3,3)) | bgm | fgm) = 255;
33 Lrgb = label2rgb(L, 'jet', 'w', 'shuffle');
```



а)



б)

Рис. 6.10 — Маркеры и границы, наложенные на: а) исходное изображение, б) представленное в цветах rgb.

Порядок выполнения работы

1. *Базовые морфологические операции.* Выбрать произвольное изображение, содержащее дефекты формы (внутренние «дырки» или внешние «выступы») объектов. Используя базовые морфологические операции полностью убрать или минимизировать дефекты.

2. *Разделение объектов.* Выбрать произвольное бинарное изображение, содержащее перекрывающиеся объекты. Использовать операции бинарной морфологии для разделения объектов. Выделить контуры объектов.

3. *Сегментация.* Выбрать произвольное изображение, содержащее небольшое число локальных минимумов. Выполнить сегментацию изображения по водоразделам.

Содержание отчета

1. Цель работы.

2. Теоретическое обоснование математической морфологии для анализа изображений.

3. Ход выполнения работы:

   (a) Исходные изображения;

   (b) Листинги программных реализаций;

   (c) Комментарии;

   (d) Результирующие изображения.

4. Выводы о проделанной работе.

Вопросы к защите лабораторной работы

1. Включает ли результат открытия в себя результат закрытия?

2. Какой морфологический фильтр необходимо применить, чтобы убрать у объекта выступы?

3. Каким образом с помощью морфологических операций можно найти контур объекта?

4. Что такое *морфология*?

Список литературы

[1] *Журавель И.М.* Краткий курс теории обработки изображений: [Электронный ресурс]. URL: http://matlab.exponenta.ru/imageprocess/book2/index.php. (Дата обращения: 19.12.2017).

[2] *MATLAB Documentation.* Computer Vision System Toolbox: [Электронный ресурс]. URL: https://www.mathworks.com/help/vision/index.html. (Дата обращения: 19.12.2017).

[3] *Визильтер Ю.В.* Обработка и анализ изображений в задачах машинного зрения: Курс лекций и практических занятий / Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В., Ососков М.В., Моржин А.В. — М.: Физматкнига, 2010. — 672 с. — ISBN 978-5-89155-201-2.

[4] *Визильтер Ю.В.* Обработка и анализ цифровых изображений с примерами на LabVIEW IMAQ Vision / Визильтер Ю.В., Желтов С.Ю., Князь В.А., Ходарев А.Н., Моржин А.В. — М.: ДМК Пресс, 2007. — 464 с. — ISBN 5-94074-404-4.

[5] *Грузман И.С.* Цифровая обработка изображений в информационных системах: Учебное пособие / Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И., Спектор А.А. — Новосибирск: Изд-во НГТУ, 2002. — 352 с.

[6] *Старовойтов В.В.* Цифровые изображения: от получения до обработки / Старовойтов В.В., Голуб Ю.И. — Минск: ОИПИ НАН Беларуси, 2014. — 202 с. — ISBN 978-985-6744-80-1.

УНИВЕРСИТЕТ ИТМО

Миссия университета — генерация передовых знаний, внедрение инновационных разработок и подготовка элитных кадров, способных действовать в условиях быстро меняющегося мира и обеспечивать опережающее развитие науки, технологий и других областей для содействия решению актуальных задач.

КАФЕДРА ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ ПРОМЫШЛЕННОЙ РОБОТОТЕХНИКИ

Кафедра Интеллектуальных Технологий Промышленной Робототехники (ИТПРТ), в настоящее время входящая в состав мегафакультета Компьютерных технологий и управления факультета Систем управления и робототехники, была основана в 2012 году как базовая кафедра ООО «Тепловое Оборудование», выпускающей продукцию под брендом Thermex. В настоящее время кафедрой реализуются основные профессиональные образовательные программы высшего образования бакалавриата «Интеллектуальные технологии в робототехнике» и магистратуры «Индустриальная робототехника» по направлениям подготовки 15.03.06 и 15.04.06 «Мехатроника и робототехника» соответственно. Основной упор образовательных программ сделан на подготовку кадров в области современного цифрового производства и индустриальных кибер-физических систем (CPS). Кадры, подготавливаемые кафедрой, получают не только теоретические знания, но и практические навыки при прохождении производственной практики на современном предприятии Thermex, которое спроектировали и построили ведущие европейские инженеры из Италии, Швейцарии, Германии и России.

Шаветов Сергей Васильевич

Основы технического зрения: лабораторный практикум

Учебно-методическое пособие

В авторской редакции
Редакционно-издательский отдел Университета ИТМО
Зав. РИО Н.Ф. Гусарова
Подписано к печати
Заказ №
ТиражОтпечатано на ризографе

Редакционно-издательский отдел
Университета ИТМО
197101, Санкт-Петербург, Кронверкский пр., 49