

ОБРАБОТКА И РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ

Леонид Моисеевич Местецкий
профессор

кафедра математических методов
прогнозирования ВМК МГУ
кафедра интеллектуальных систем МФТИ

ОПЕРАЦИИ НАД ИЗОБРАЖЕНИЯМИ

- Точечные
- Пространственные
- Алгебраические
- Геометрические
- Межкадровые

Алгебраические операции

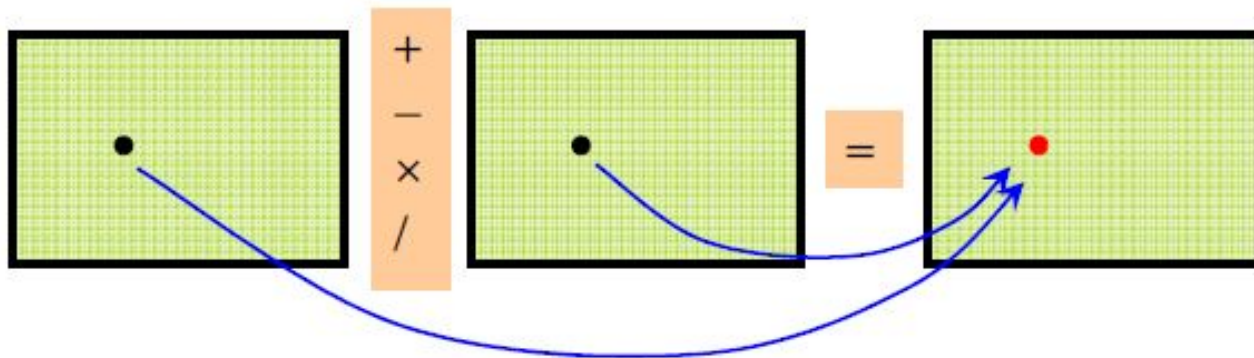
Составляют новое изображение из
поточечных сумм, разностей,
произведений и
частных двух исходных изображений.

Сумма: $C(x, y) = A(x, y) + B(x, y)$

Разность: $C(x, y) = A(x, y) - B(x, y)$

Произведение: $C(x, y) = A(x, y) \times B(x, y)$

Частное: $C(x, y) = A(x, y) / B(x, y)$



Сложение изображений

Используется для:

- осреднения множественных изображений

одной и той же сцены с целью уменьшения

влияния случайного шума;

- для выделения содержимого одного образа

над другим, создания эффекта двойной экспозиции

Уменьшение шума

Осреднения множественных изображений одной и той же сцены



$D_i(x, y) = S(x, y) + N_i(x, y)$, $i = 1, \dots, m$ - множество образов

$S(x, y)$ - образ, $N_i(x, y)$ - шум.

Шум некоррелированный и имеет нулевое среднее значение:

$$E[N_i(x, y)] = 0$$

Отношение сигнал/шум

Безразмерная величина, равная отношению мощности полезного сигнала к мощности шума

SNR – signal-to-noise ratio

$$SNR = \frac{P_{\text{сигнал}}}{P_{\text{шум}}} = \frac{A_{\text{сигнал}}^2}{A_{\text{шум}}^2}$$

P – средняя мощность

A – среднеквадратичное значение амплитуды

Изменение отношения

Отношение мощностей сигнал/шум вырастает в \sqrt{m} раз:

$$P(x, y) = \frac{S^2(x, y)}{E[N^2(x, y)]}$$

При осреднении m образов

$$\bar{D}(x, y) = \frac{1}{m} \sum_{i=1}^m (S(x, y) + N_i(x, y))$$

имеем

$$\begin{aligned} \bar{P}(x, y) &= \frac{S^2(x, y)}{E \left[\frac{1}{m} \sum_{i=1}^m (S(x, y) + N_i(x, y)) - S(x, y) \right]^2} = \\ &= \frac{m^2 S^2}{E \left[\sum_{i=1}^m N_i^2 \right] + E \left[\sum_{i, j=1; i \neq j}^m N_i \cdot N_j \right]} = \frac{m^2 S^2}{\sum_{i=1}^m E[N_i^2]} = m \cdot P(x, y) \end{aligned}$$

Результат осреднения

$S(x, y)$ - образ,

$N_i(x, y)$ - шум,

$\bar{D}(x, y) = \frac{1}{m} \sum_{i=1}^m (S(x, y) + N_i(x, y))$ - осреднённый образ,

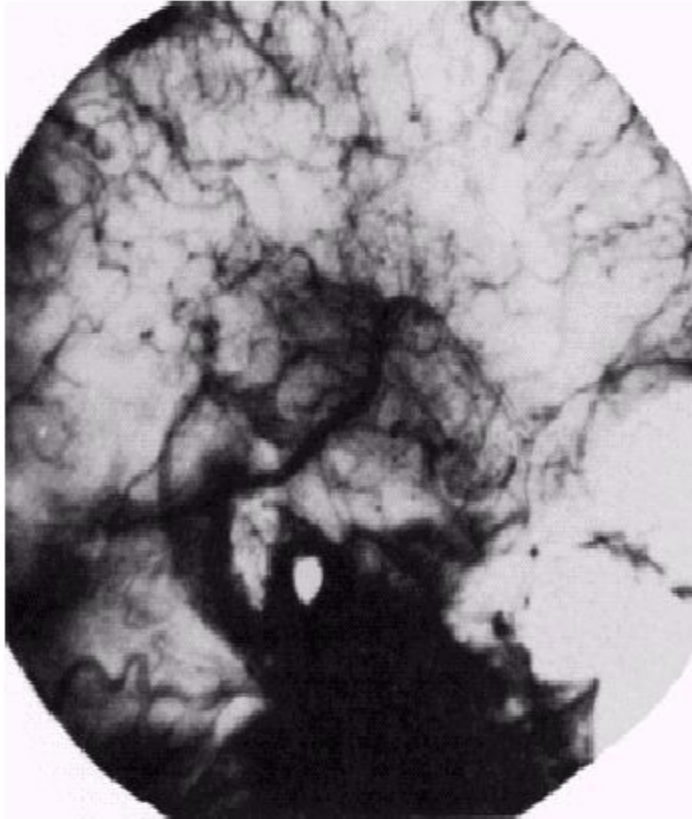
$\sqrt{\bar{P}(x, y)} = \sqrt{m} \cdot \sqrt{P(x, y)}$ - отношение «сигнал/шум» возрастает в \sqrt{m} раз.

Вычитание изображений

Используется для

- удаления нежелательного образа из изображения (медленно меняющийся фон, периодический шум, другие аддитивные помехи);
- определение изменений между двумя изображениями одной и той же сцены.

Выделение объекта на фоне

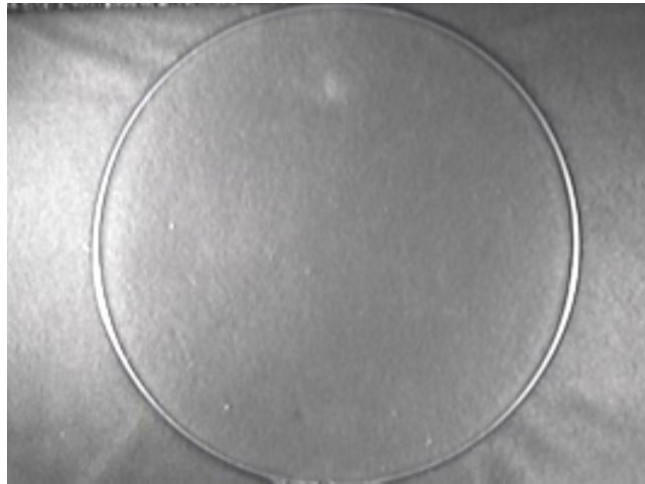


Исходное
изображение



Изображение с вводом
контрастного вещества и
вычитанием исходного

Сегментация изображения



Фон $B(x,y)$



Образ $A(x,y)$

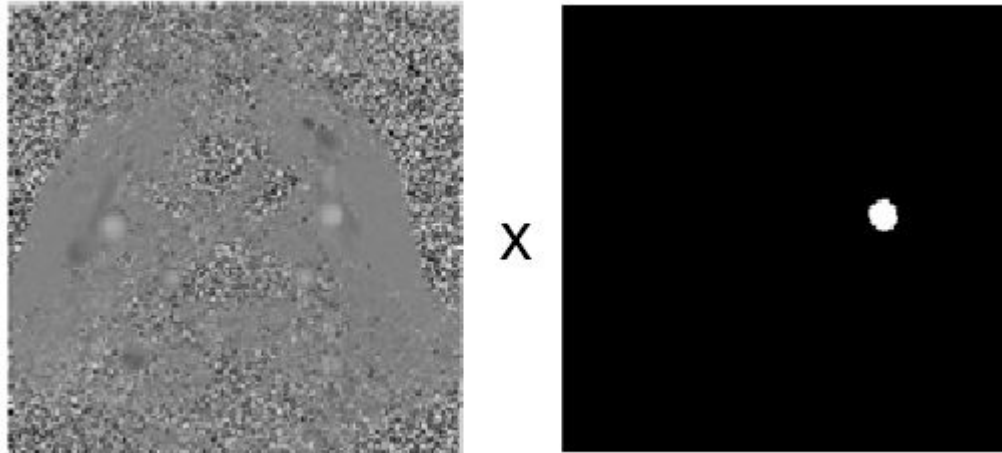


Вычитание $A(x,y)-B(x,y)$



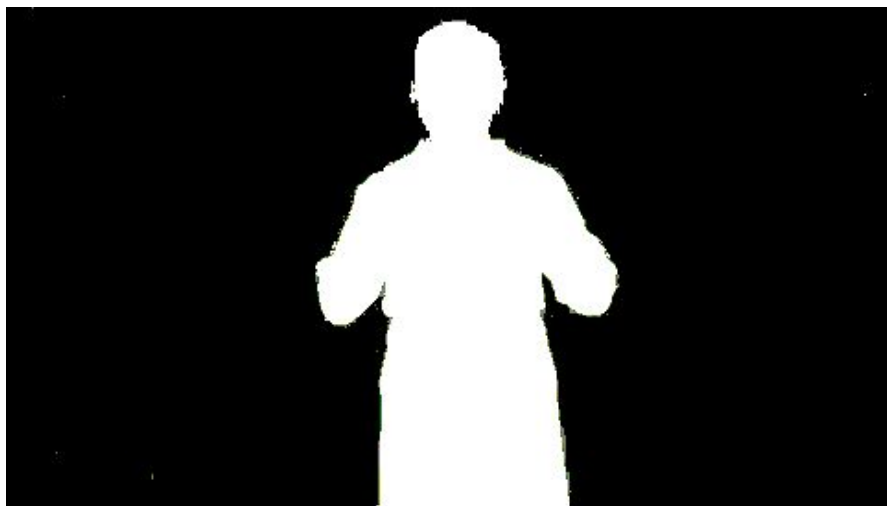
Умножение $k(A(x,y)-B(x,y))$

Умножение изображений



- Используется для выделения с помощью маски областей, которые не должны учитываться при вычислениях.
- Магнитно-резонансная томограмма шеи человека (слева) умножается на маску (справа), в результате чего выделяются пиксели, соответствующие одной из артерий.
- Осреднение яркости пикселей внутри выделенной области и умножение на площадь области даёт возможность оценить средний кровоток через артерию.

Построение маски и выделение объект



Применение умножения для синтеза сцены



Деление изображений

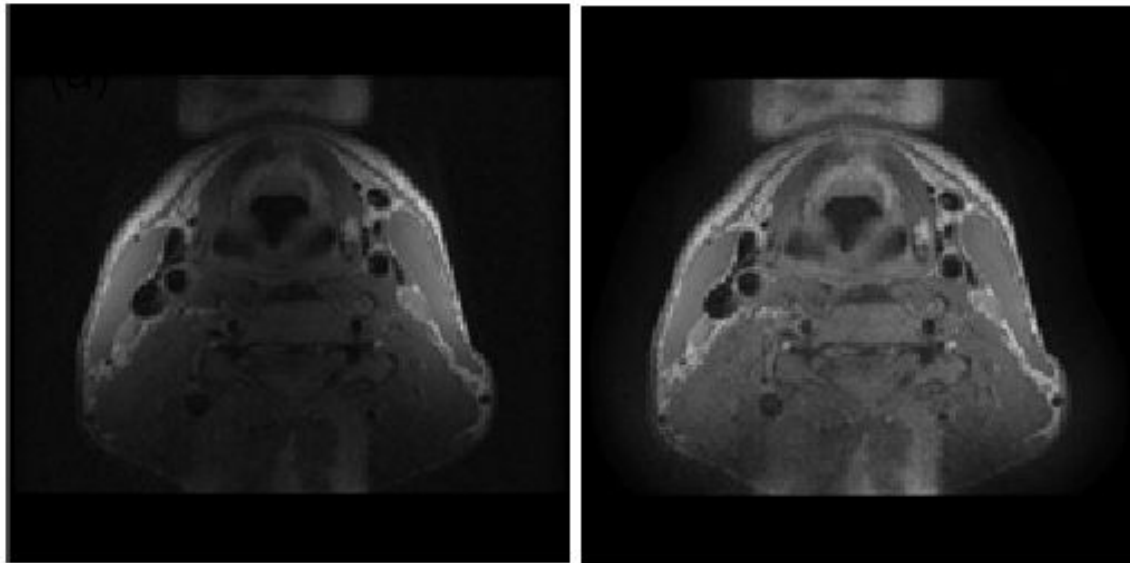
Формирование изображения часто можно рассматривать, как мультипликативный процесс:

$$A(x, y) = I(x, y) \cdot R(x, y)$$

$A(x, y)$ - наблюдаемое изображение

$I(x, y)$ - низкочастотная помеха (неравномерное освещение)

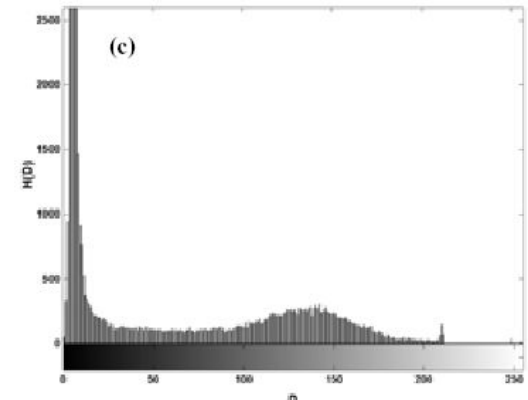
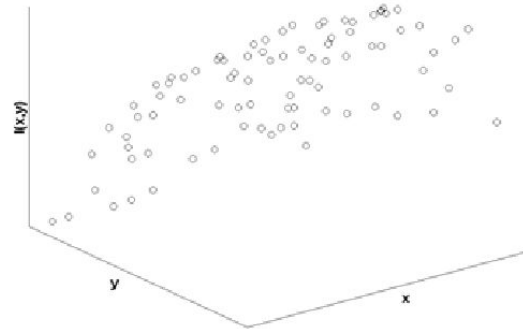
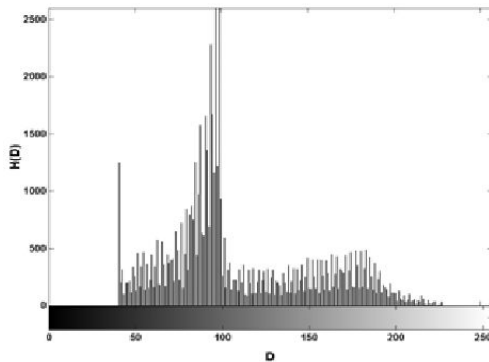
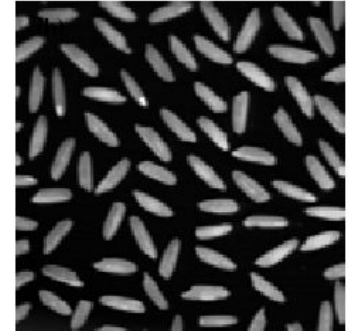
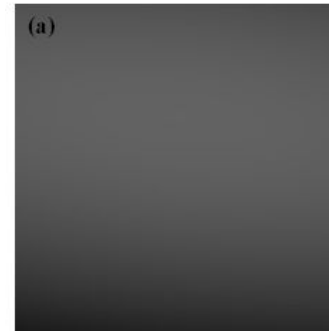
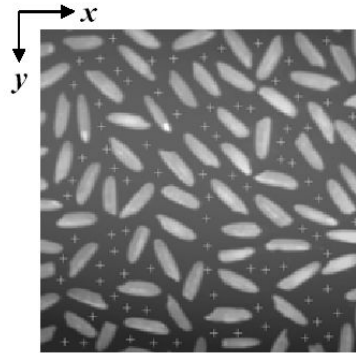
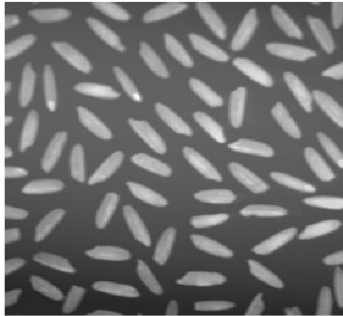
$R(x, y)$ - нужная компонента изображения



1. Оценить компоненту $I(x, y)$
2. Разделить $A(x, y)$ на $I(x, y)$

На левом рисунке изображение кровотока через шею, полученное с помощью магнитно-резонансного томографа. На правом – улучшенное изображение с помощью операции деления.

Модель фона



Изображение и его гистограмма яркости.

Фоновые пиксели и распределение их яркости

Модель фоновой яркости и результат деления изображений

Модель фоновой яркости

$$B(x, y) = a_0 + a_1x + a_2y + a_3x^2 + a_4y^2 + a_5xy$$

$(x_i, y_i), i = 1, \dots, k$ - выбранные фоновые пиксели,

$c_i = I(x_i, y_i)$ - яркость выбранных фоновых пикселей.

Аппроксимация по методу наименьших
квадратов

$$C(a_0, \dots, a_5) = \sum_{i=1}^k (c_i - B(x_i, y_i))^2$$

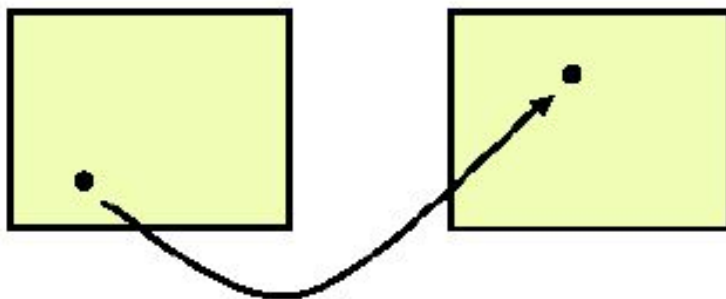
$\frac{\partial C}{\partial a_i} = 0, i = 0, \dots, 5$ - система линейных уравнений

Геометрические операции



Геометрические операции меняют пространственные отношения между объектами в изображении.

Геометрические операции



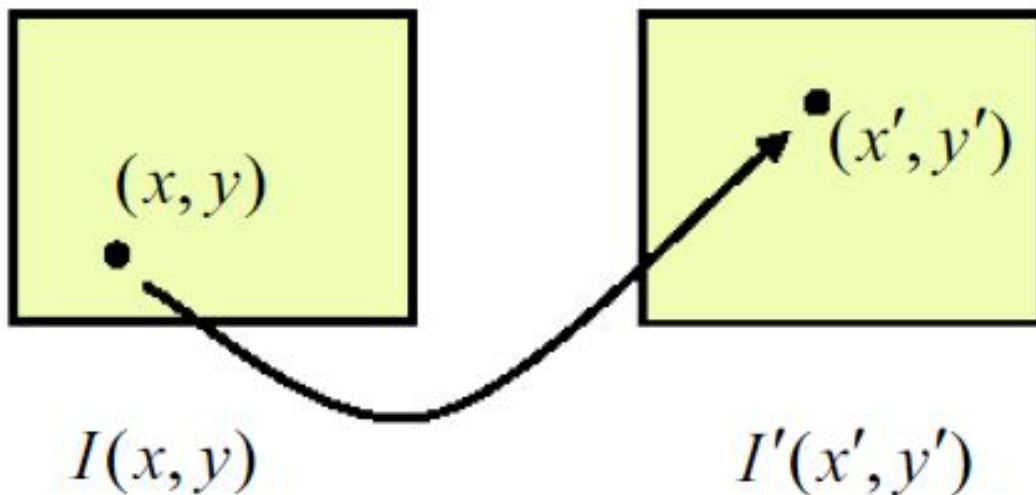
1. Результат зависит только от координат пикселя
2. Результат не зависит от окружающих пикселей
3. Пример: $I(x,y) = I(x+a, y+b)$

Геометрические операции

$$x \rightarrow f_x(x, y) = x'$$

$$y \rightarrow f_y(x, y) = y'$$

$$I(x, y) = I'(f_x(x, y), f_y(x, y))$$



Прямое отображение

$$x \rightarrow f_x(x, y) = x'$$

$$y \rightarrow f_y(x, y) = y'$$



Основные проблемы:

пропуск или перекрытие пикселей образа

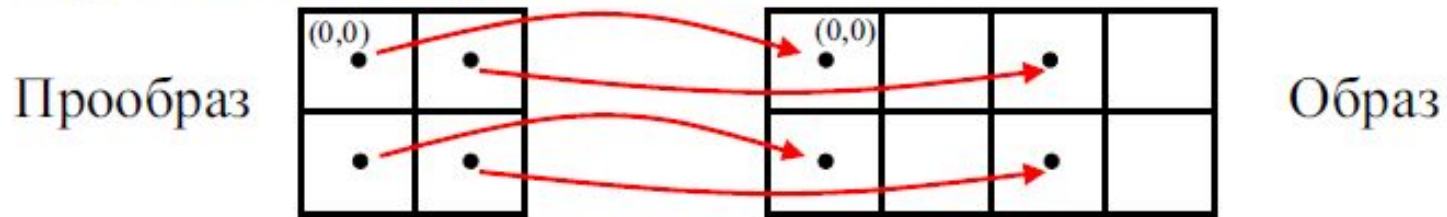
Обратное отображение

$$x' \rightarrow f_x^{-1}(x', y') = x$$
$$y' \rightarrow f_y^{-1}(x', y') = y$$

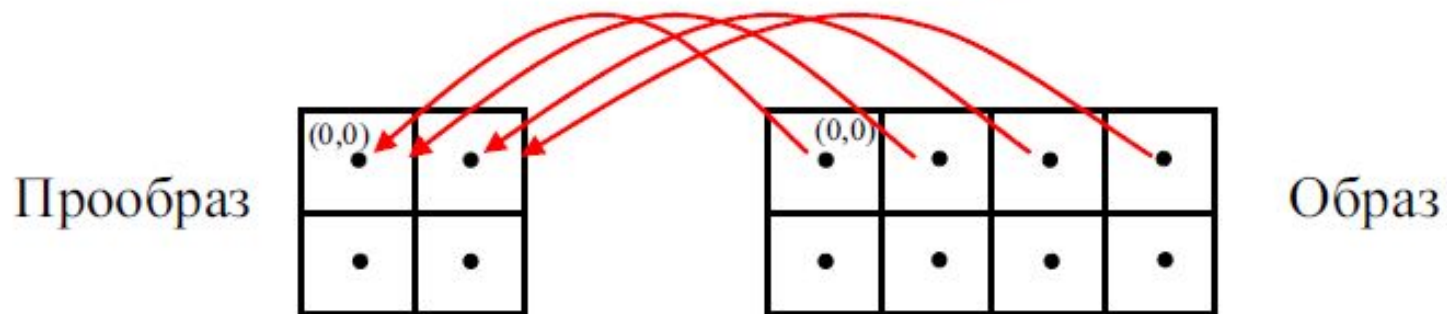


Пример: растяжение по оси X

Прямое отображение: $x' = 2x$ $y' = y$



Обратное отображение: $x = \frac{x'}{2}$ $y = y'$

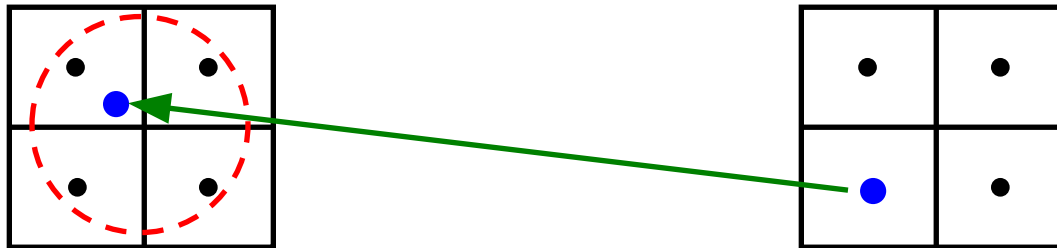


Интерполяция

Что делать, когда значения (x, y) не целые числа?

$$x = f_x^{-1}(x', y')$$

$$y = f_y^{-1}(x', y')$$



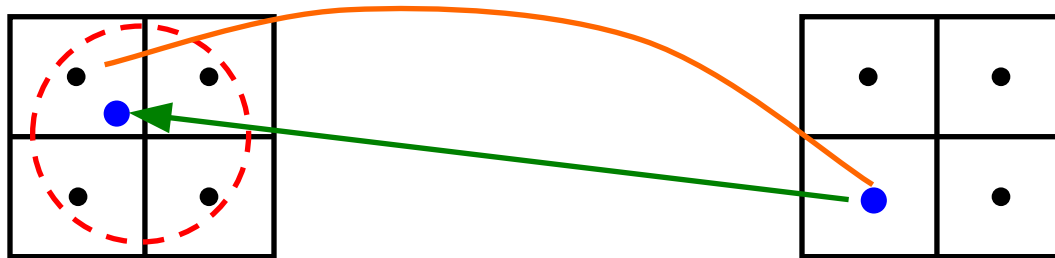
Интерполяция – генерация нового пикселя на основе анализа окружающих пикселей

Интерполяция по ближайшему соседу

Яркость определяется по цвету ближайшего пикселя к прообразу

Достоинство – быстро работает

Недостаток – нарушение непрерывности цветовых переходов

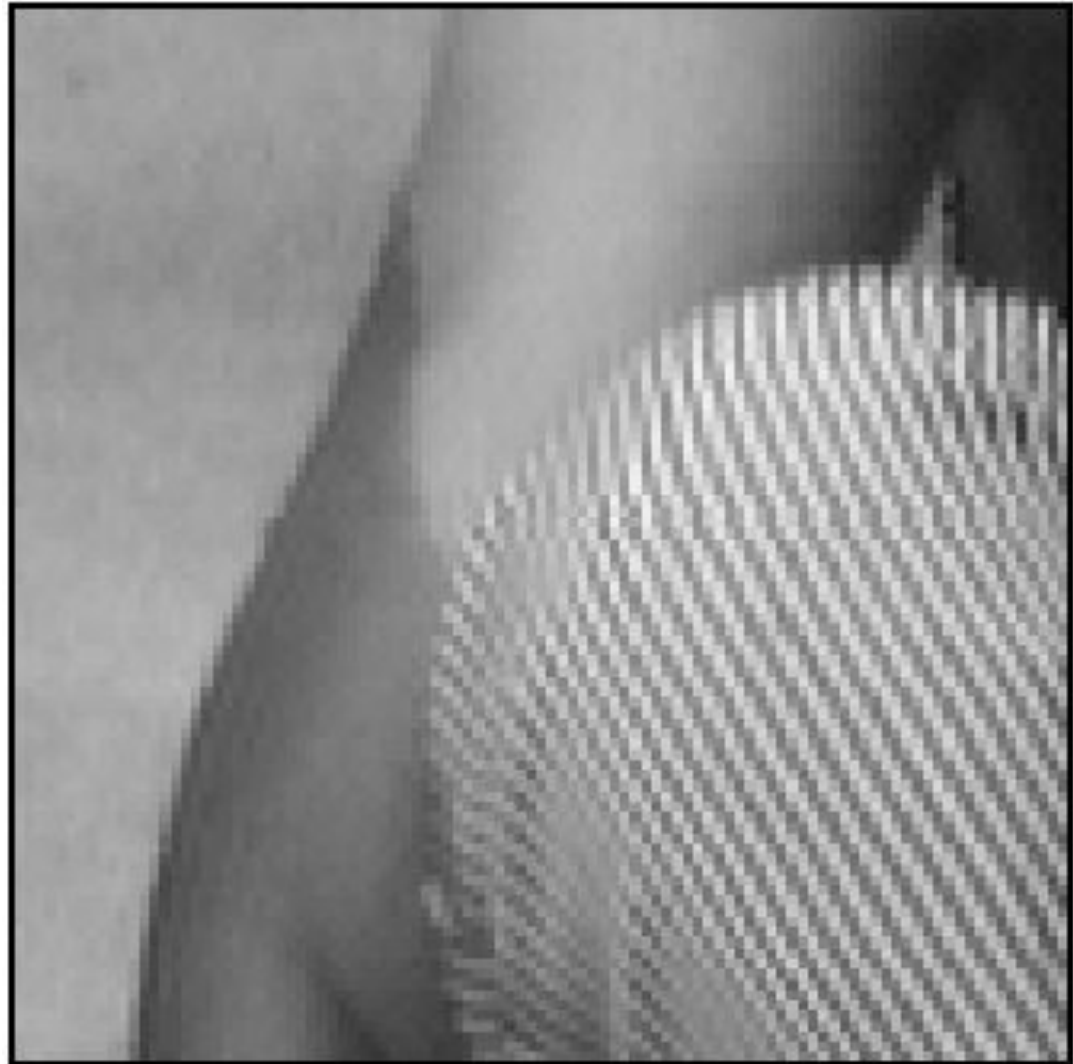
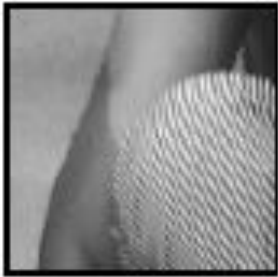


$$I'(x', y') = I(\text{round}[f_x^{-1}(x', y')], \text{round}[f_y^{-1}(x', y')])$$

Пример «поворот»

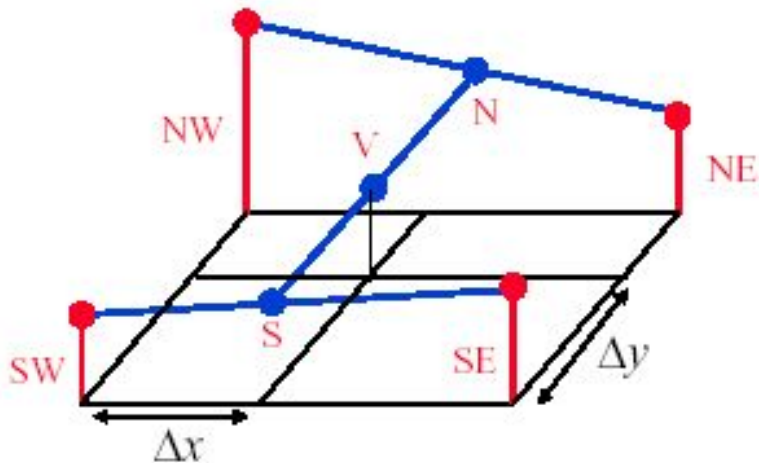
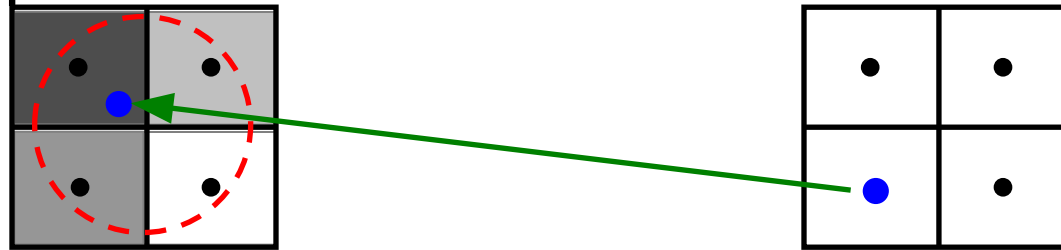


Пример «масштабирование»



Билинейная интерполяция

Яркость определяется путём интерполяции по цвету между четырьмя ближайшими к прообразу пикселями



$$S = SW \cdot (1 - \Delta x) + SE \cdot \Delta x$$
$$N = NW \cdot (1 - \Delta x) + NE \cdot \Delta x$$
$$V = S \cdot (1 - \Delta y) + N \cdot \Delta y$$

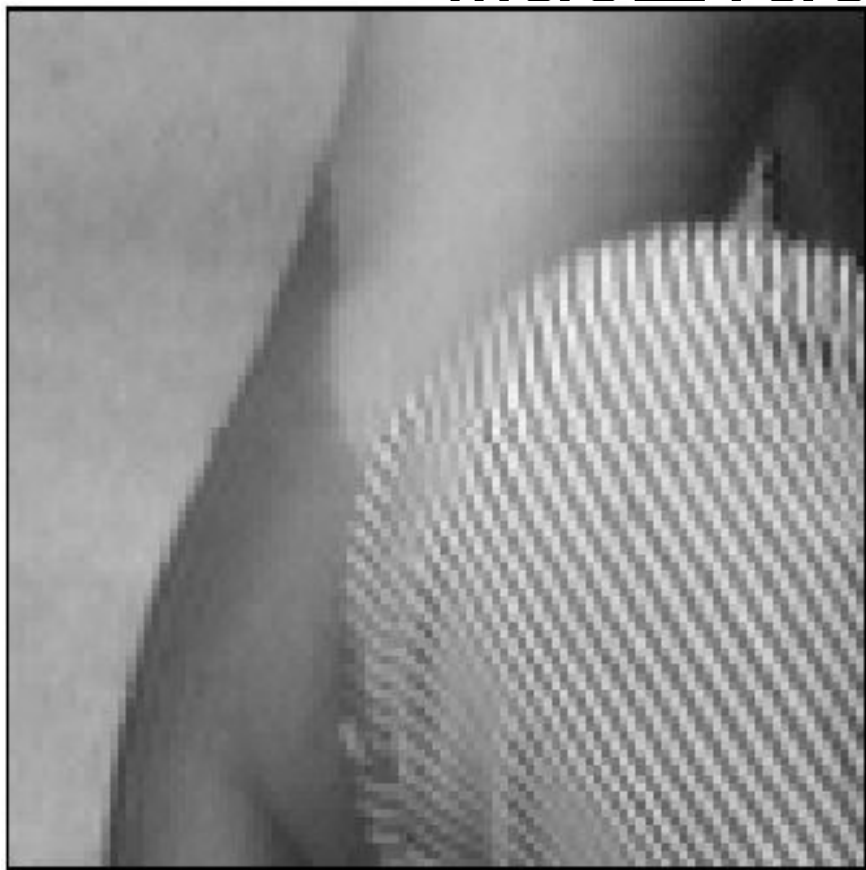
Сравнение двух вариантов интерполяции «поворот»



Интерполяция по
ближайшему соседу

Билинейная
интерполяция

Сравнение двух вариантов интерполяции «масштабирование»

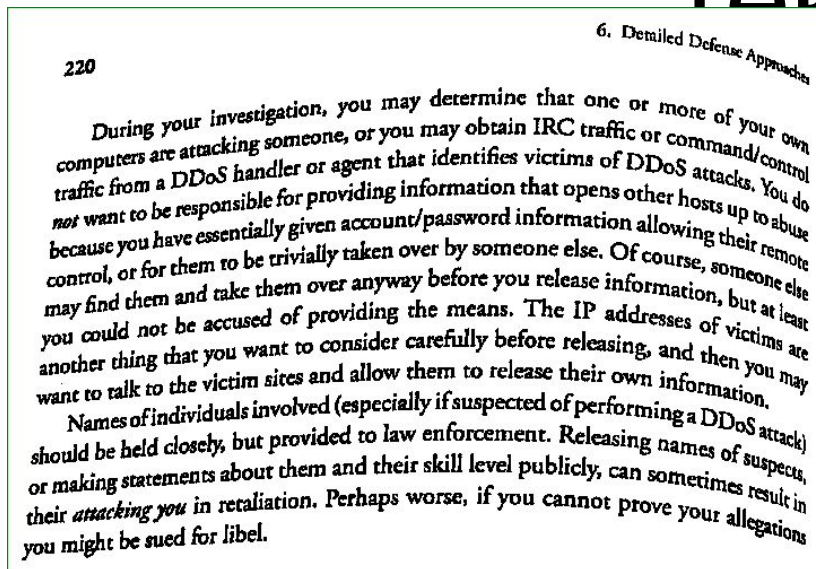


Интерполяция по
ближайшему соседу

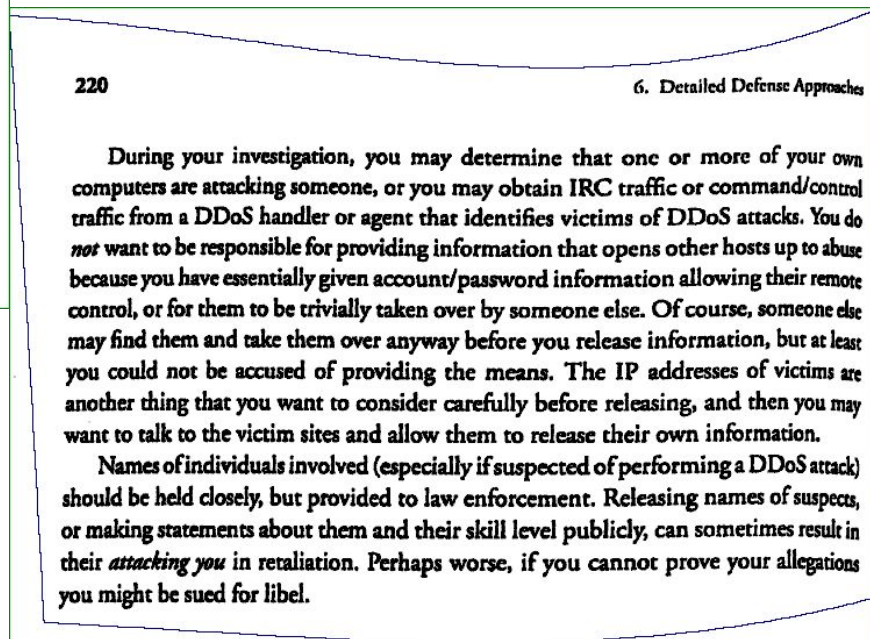


Билинейная
интерполяция

Устранение деформации текста

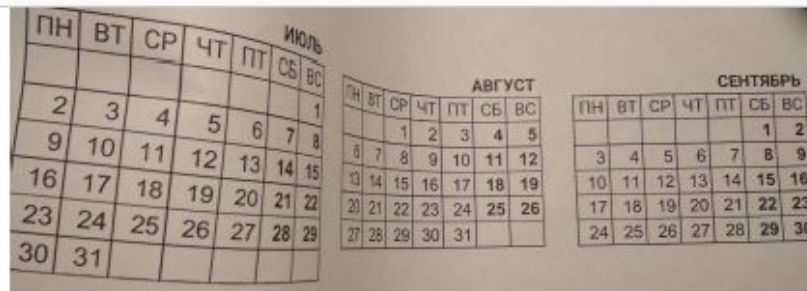


Исходное
изображение



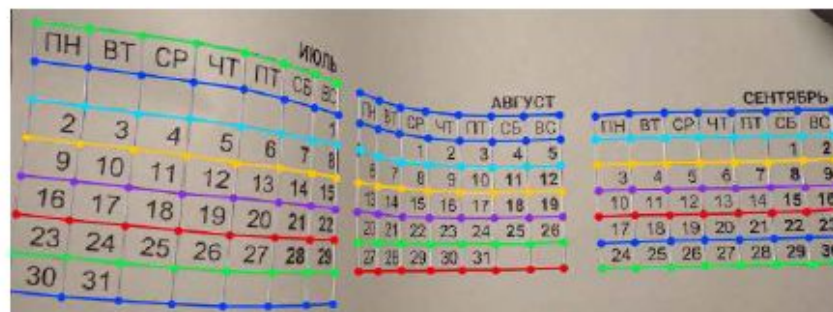
Распрямленное изображение
со сглаживанием (серое)

Устранение деформации таблицы



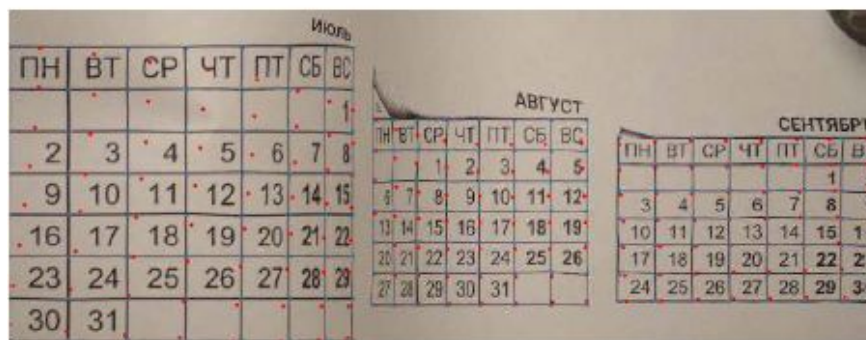
The image shows a calendar for July, August, and September that is severely distorted. The text is slanted and the grid lines are curved, making it difficult to read. The days of the week are labeled in Cyrillic: ПН (Monday), ВТ (Tuesday), СР (Wednesday), ЧТ (Thursday), ПТ (Friday), СБ (Saturday), and ВС (Sunday). The dates are arranged in a grid, but the lines are not straight.

Исходное искажённое изображение



The same distorted calendar is shown, but with horizontal lines drawn across it. These lines are color-coded: blue for the first row, yellow for the second, purple for the third, red for the fourth, and green for the fifth. The lines help to identify the horizontal structure of the calendar grid.

Детектирование горизонтальных линий таблицы



The calendar is now perfectly aligned and straight. The text is clear and the grid lines are straight. The days of the week and dates are easily readable. The color-coded lines from the previous step are still visible, showing how they were used to detect the horizontal structure.

Результат выравнивания таблицы