

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

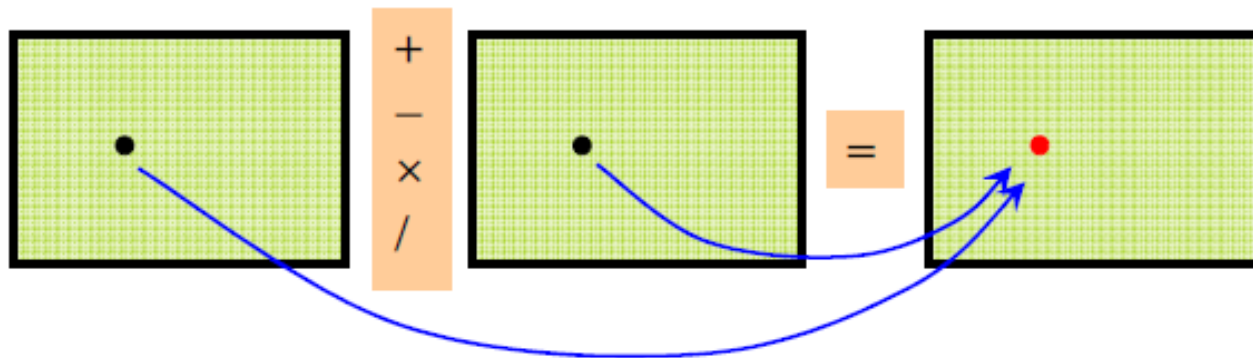
Составляют новое изображение из поточечных сумм, разностей, произведений и частных двух исходных изображений.

Сумма: $C(x, y) = A(x, y) + B(x, y)$

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

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

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



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*

- средняя мощность
- 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,

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

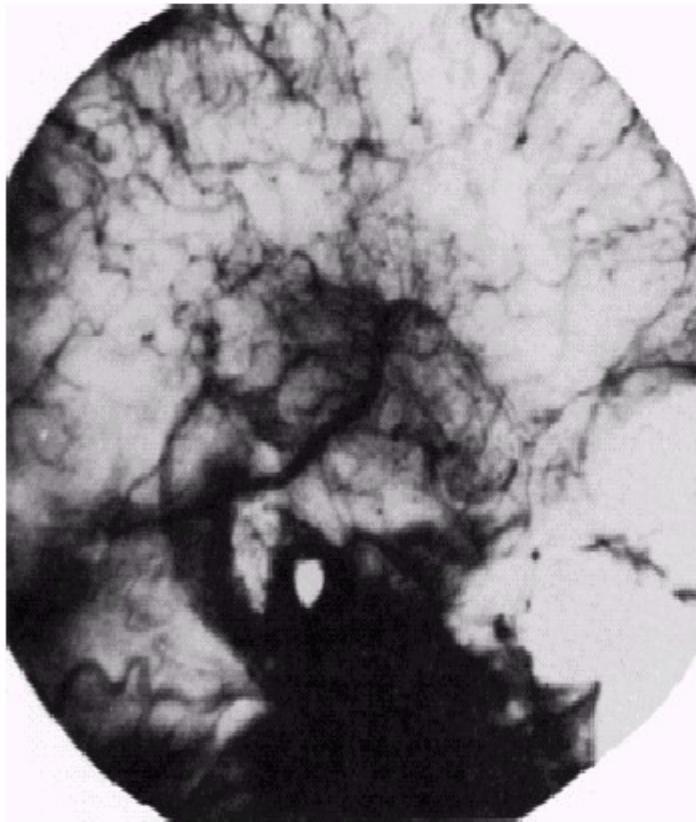
$\sqrt{\overline{P}(x, y)} = \sqrt{m} \cdot \sqrt{P(x, y)}$ - отношение «сигнал/шум» возрастает в \sqrt{m} раз.

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

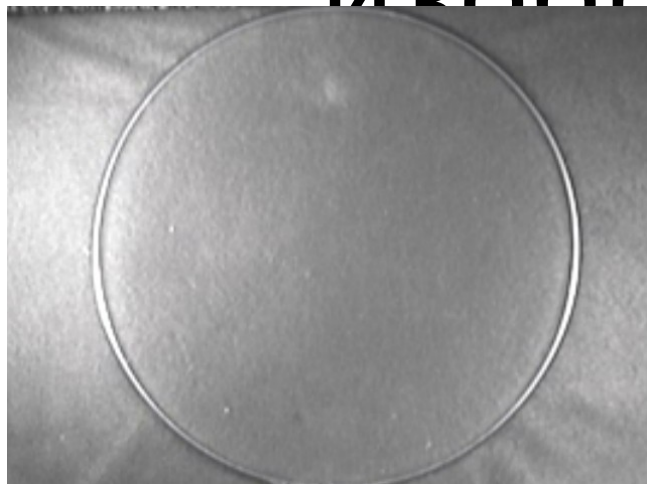


Исходное
изображение



Изображение с
вводом контрастного
вещества и
вычитанием

Сегментация изображения



Фон $B(x,y)$



Образ $A(x,y)$

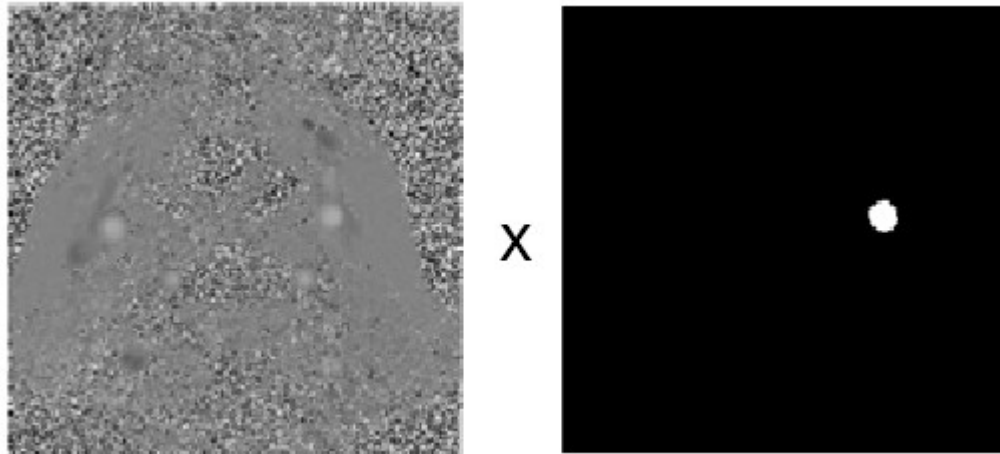


Вычитание $A(x,y)-B(x,y)$



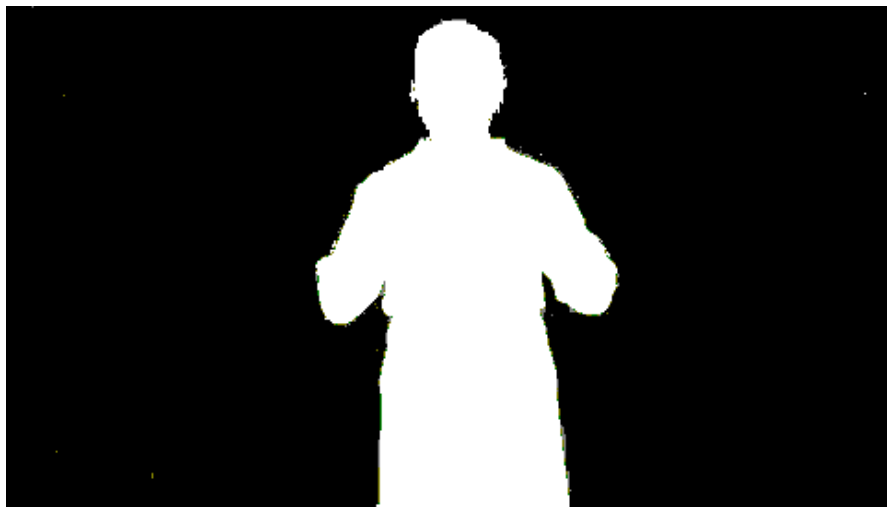
Умножение $k(A(x,y)-B(x,y))$

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

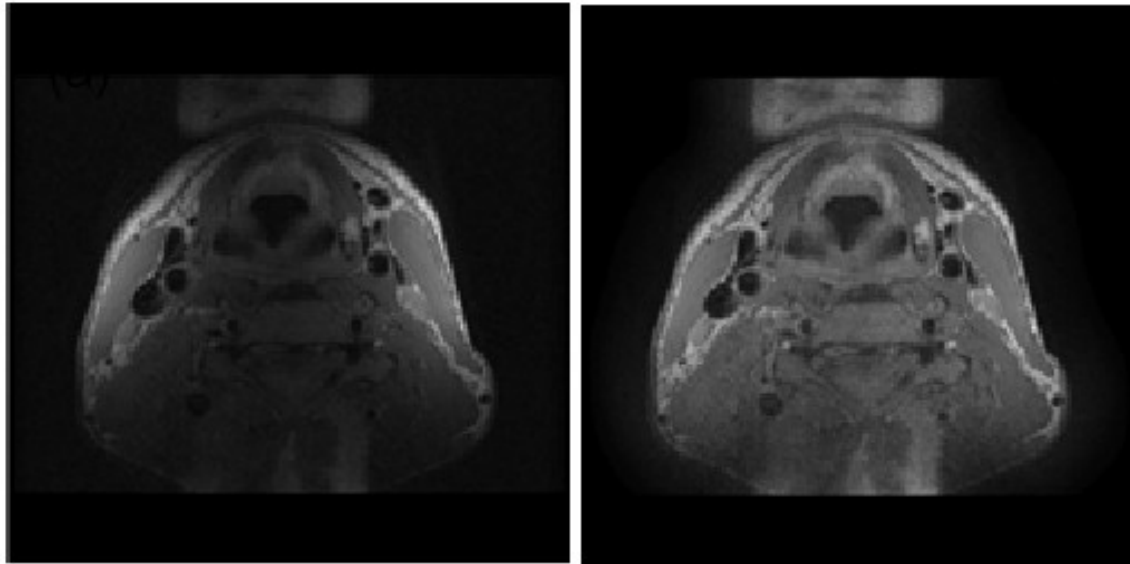
Формирование изображения часто можно рассматривать, как мультипликативный процесс:

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

$A(x, y)$ - наблюдаемое изображение

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

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

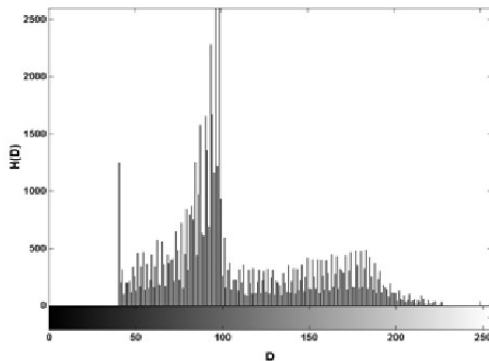
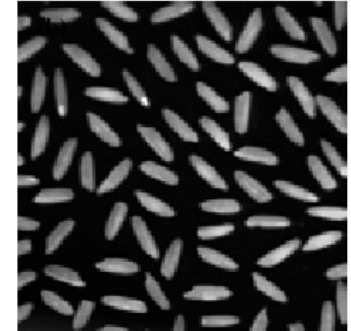
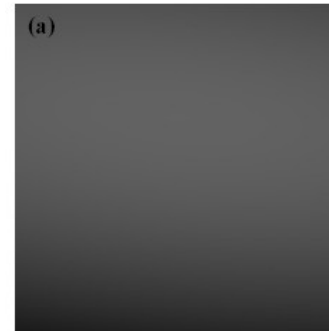
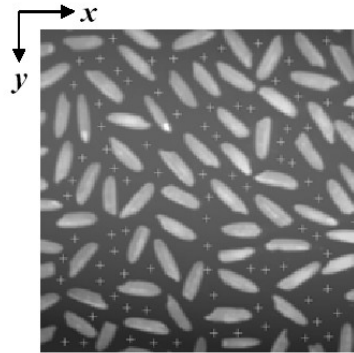
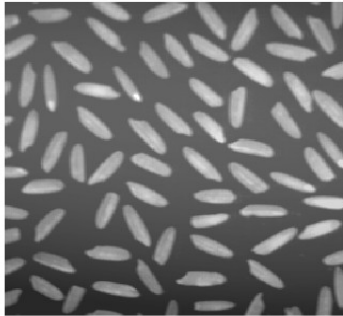


1. Оценить компоненту $I(x, y)$

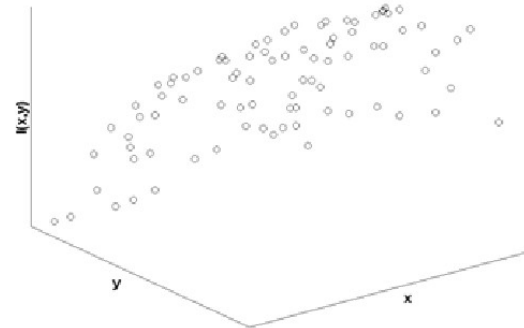
2. Разделить $A(x, y)$ на $I(x, y)$

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

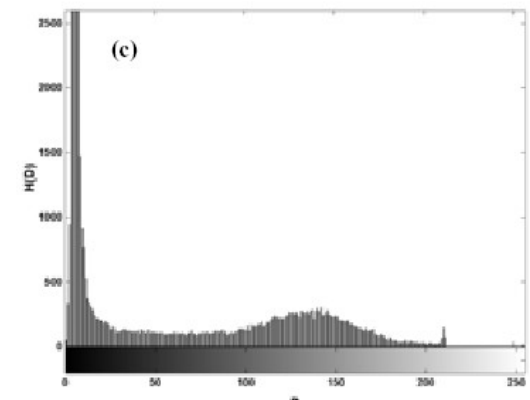
Модель фона



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

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

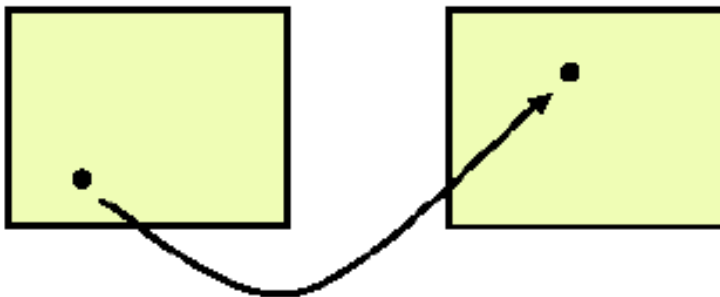
$\frac{\partial C}{\partial a_i} = 0, i = 0, \dots, 5$ - система линейных уравнений

Геометрические операции



Геометрические операции меняют пространственные отношения между объектами в изображении.

Геометрические операции



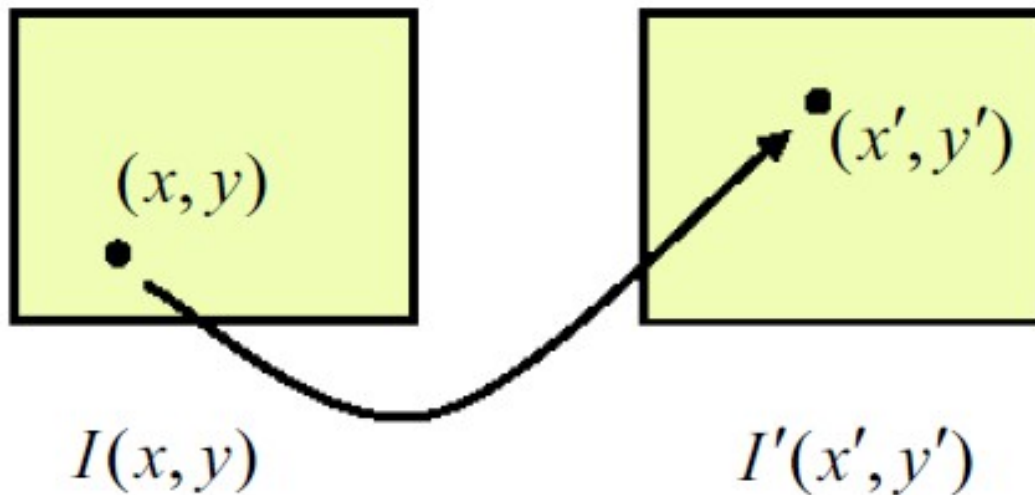
1. Результат зависит только от координат пикселя
2. Результат не зависит от окружающих пикселей
3. Пример: $I(x,y) = I(x+a, y+b)$

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

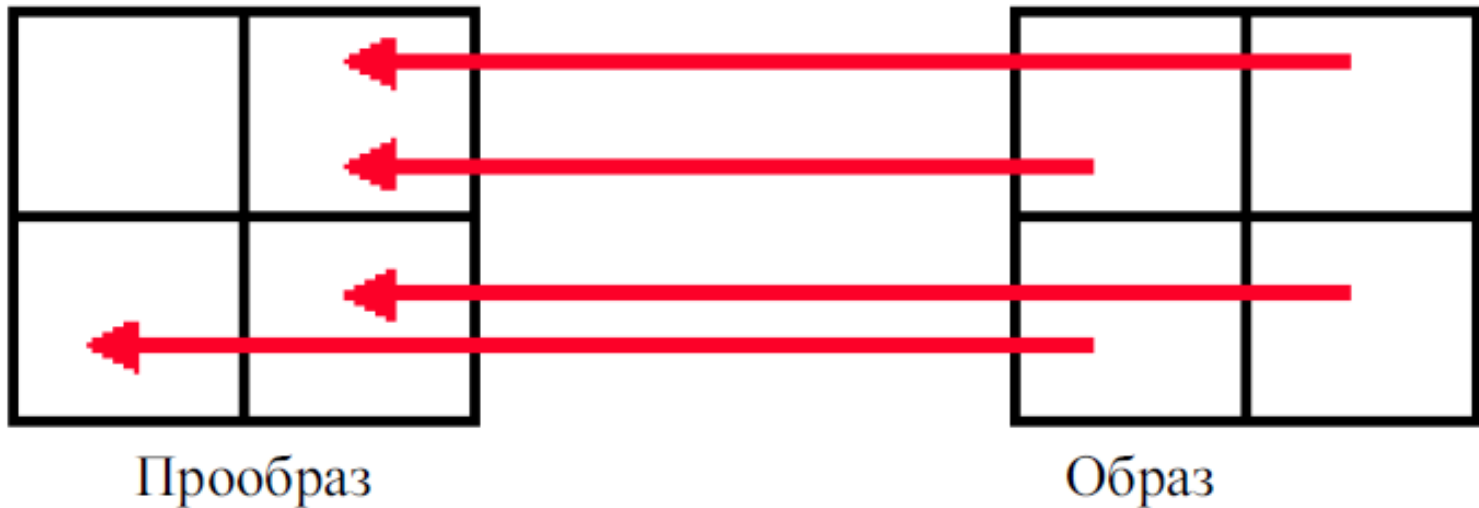
$$x \rightarrow f_x(x, y) = x'$$
$$y \rightarrow f_y(x, y) = y'$$



Основные проблемы:
пропуск или перекрытие пикселей
образа

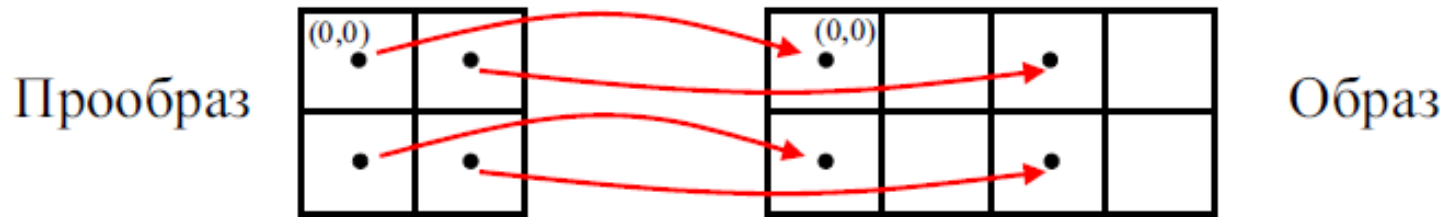
Обратное отображение

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

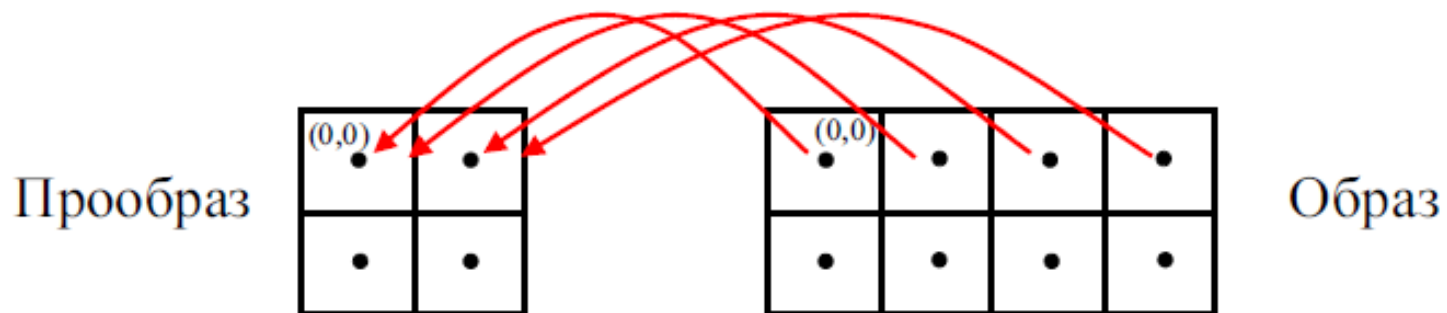


Пример: растяжение по оси X

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



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

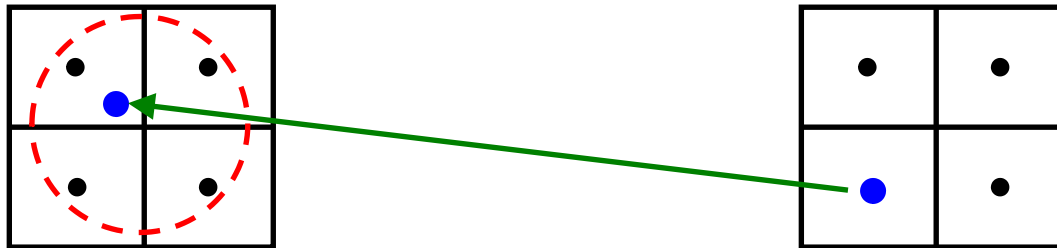


Интерполяция

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

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

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



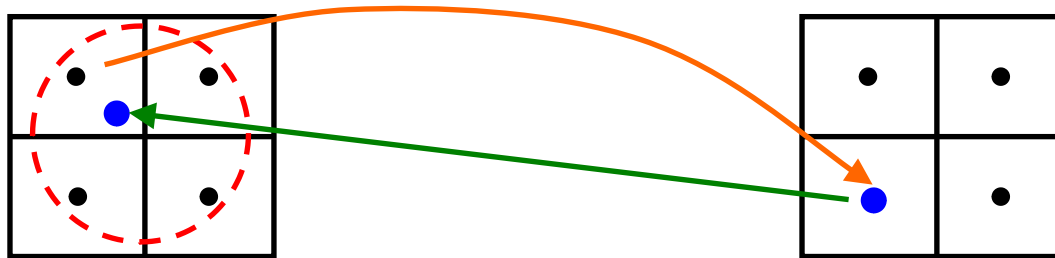
Интерполяция – генерация нового пикселя на основе анализа окружающих пикселей

Интерполяция по ближайшему соседу

Яркость определяется по цвету ближайшего пикселя к прообразу

Достоинство – быстро работает

Недостаток – нарушение непрерывности цветовых переходов

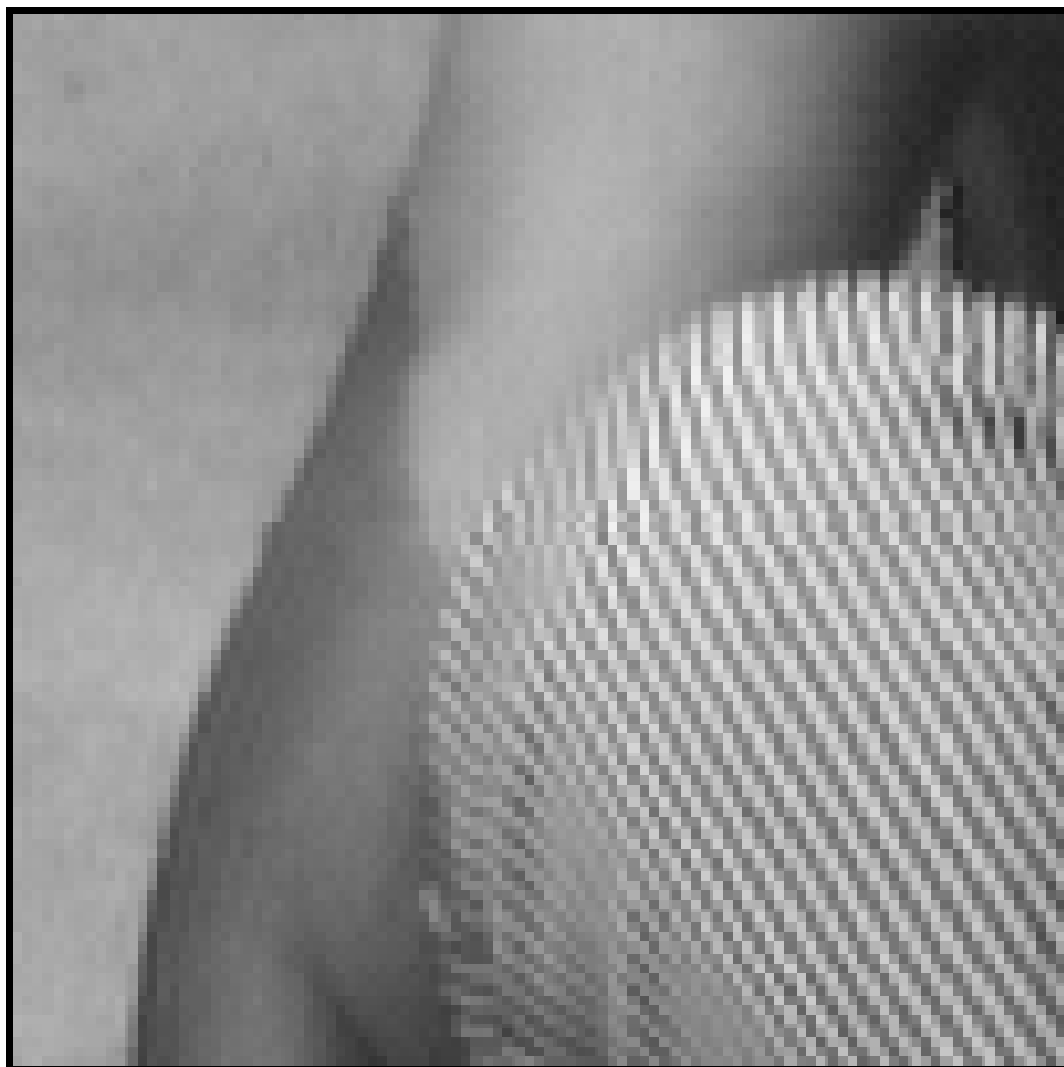
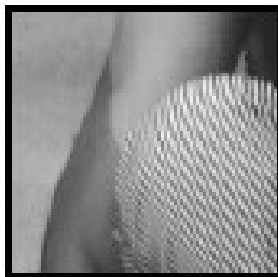


$$I'(x', y') = I(\text{round}[f_x^{-1}(x', y')], \text{round}[f_y^{-1}(x', y')])$$

Пример «поворот»

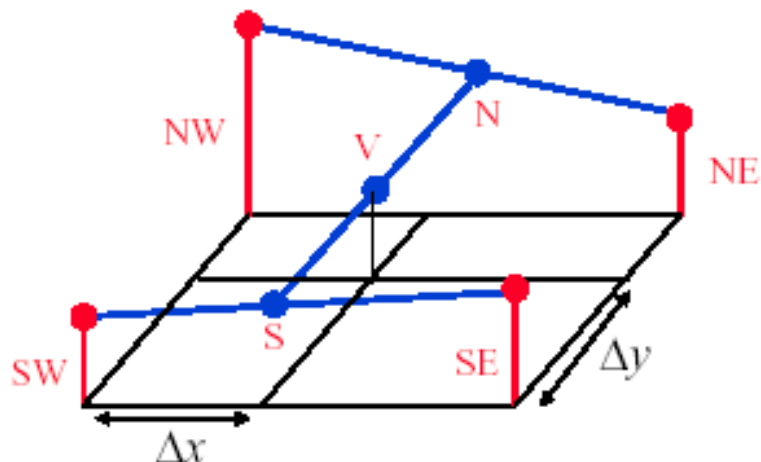
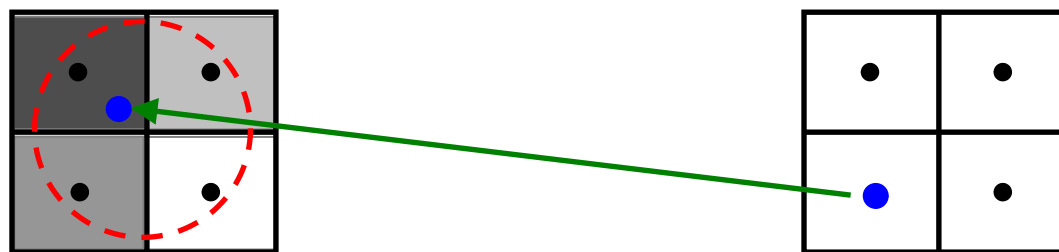


Пример «масштабирование»



Билинейная интерполяция

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



$$\begin{aligned} S &= SW \cdot (1 - \Delta x) + SE \cdot \Delta x \\ N &= NW \cdot (1 - \Delta x) + NE \cdot \Delta x \\ V &= S \cdot (1 - \Delta y) + N \cdot \Delta y \end{aligned}$$

Сравнение двух вариантов интерполяции «поворот»

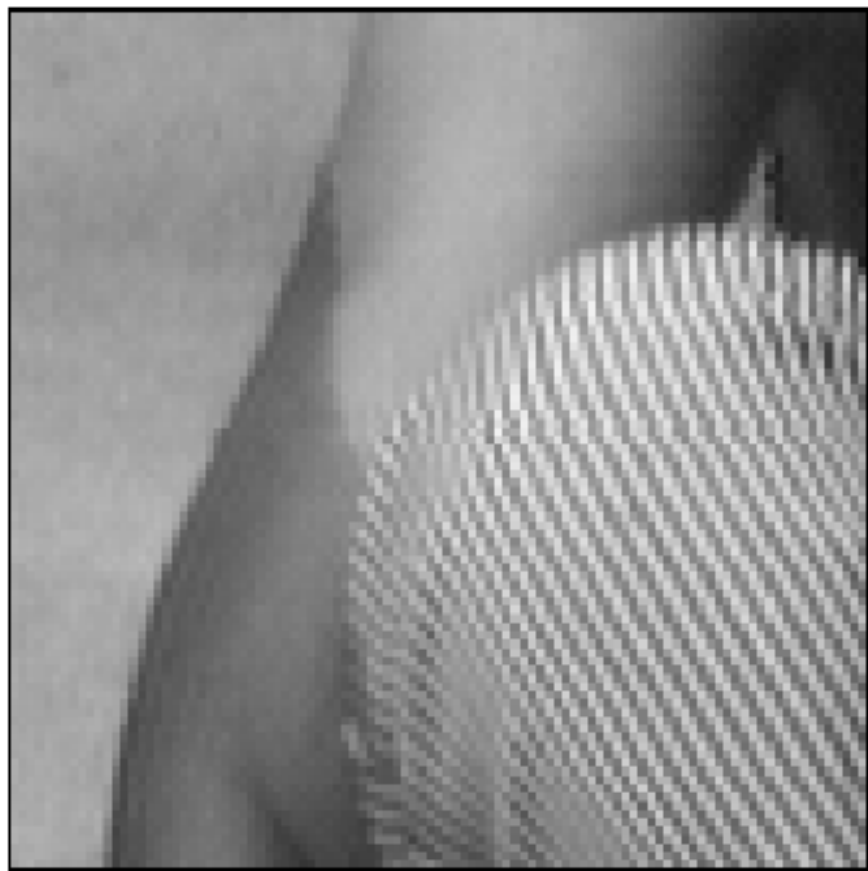


Интерполяция по
ближайшему соседу



Билинейная интерполяция

Сравнение двух вариантов интерполяции «масштабирование»



Интерполяция по
ближайшему соседу



Билинейная интерполяция

Устранение деформации текста

220

6. Detailed Defense Approaches

During your investigation, you may determine that one or more of your own computers are attacking someone, or you may obtain IRC traffic or command/control traffic from a DDoS handler or agent that identifies victims of DDoS attacks. You do *not* want to be responsible for providing information that opens other hosts up to abuse because you have essentially given account/password information allowing their remote control, or for them to be trivially taken over by someone else. Of course, someone else may find them and take them over anyway before you release information, but at least you could not be accused of providing the means. The IP addresses of victims are another thing that you want to consider carefully before releasing, and then you may want to talk to the victim sites and allow them to release their own information.

Names of individuals involved (especially if suspected of performing a DDoS attack) should be held closely, but provided to law enforcement. Releasing names of suspects, or making statements about them and their skill level publicly, can sometimes result in their *attacking you* in retaliation. Perhaps worse, if you cannot prove your allegations you might be sued for libel.

Исходное
изображение

220

6. Detailed Defense Approaches

During your investigation, you may determine that one or more of your own computers are attacking someone, or you may obtain IRC traffic or command/control traffic from a DDoS handler or agent that identifies victims of DDoS attacks. You do *not* want to be responsible for providing information that opens other hosts up to abuse because you have essentially given account/password information allowing their remote control, or for them to be trivially taken over by someone else. Of course, someone else may find them and take them over anyway before you release information, but at least you could not be accused of providing the means. The IP addresses of victims are another thing that you want to consider carefully before releasing, and then you may want to talk to the victim sites and allow them to release their own information.

Names of individuals involved (especially if suspected of performing a DDoS attack) should be held closely, but provided to law enforcement. Releasing names of suspects, or making statements about them and their skill level publicly, can sometimes result in their *attacking you* in retaliation. Perhaps worse, if you cannot prove your allegations you might be sued for libel.

Распрямленное
изображение со
сглаживанием (серое)

Устранение деформации таблиц

The image shows three calendar tables for the months of July, August, and September. The tables are distorted due to perspective, with the rows and columns not being perfectly horizontal and vertical. The days of the week are abbreviated in Cyrillic: ПН (Monday), ВТ (Tuesday), СР (Wednesday), ЧТ (Thursday), ПТ (Friday), СБ (Saturday), and ВС (Sunday). The dates are numbered from 1 to 31.

Исходное искажённое изображение

This image shows the same calendar tables as the previous one, but with colored lines overlaid to detect horizontal lines. Blue lines are drawn along the rows of the tables, indicating the detected horizontal structure. The lines are slightly curved to follow the perspective of the original image.

Детектирование горизонтальных линий таблицы

This image shows the result of aligning the calendar tables. The tables are now perfectly horizontal and vertical, with the rows and columns straight. The days of the week and dates are clearly visible and aligned. The alignment process has corrected the perspective distortion seen in the original image.

Результат выравнивания таблицы