



Введение в курс.  
Представление и обработка  
изображений

2023  
Обнинск

# Содержание курса

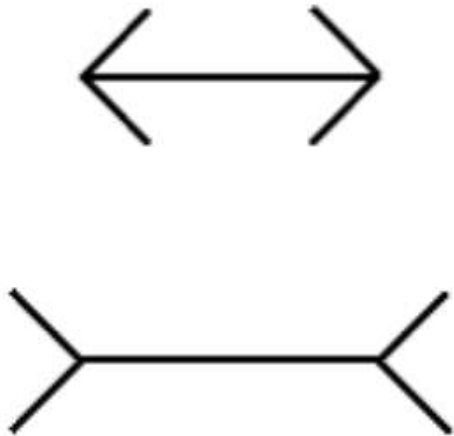
- Обработка изображений и видео с введением в OpenCV
- Теоретическое знакомство с основными задачами CV
- Решение задач CV с использованием Google Colab



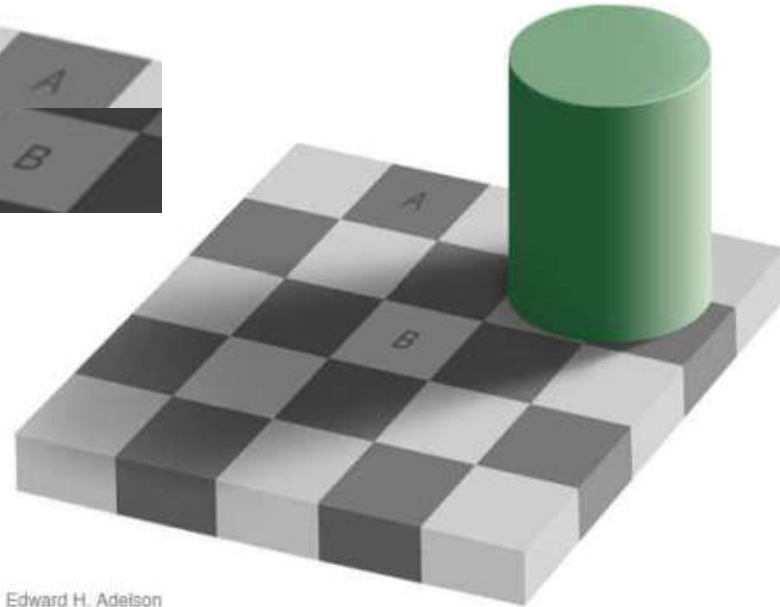
# Что такое компьютерное зрение?



# Что такое компьютерное зрение?



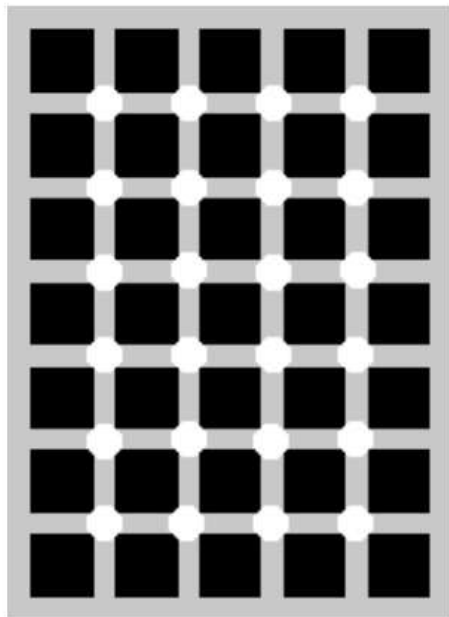
(a)



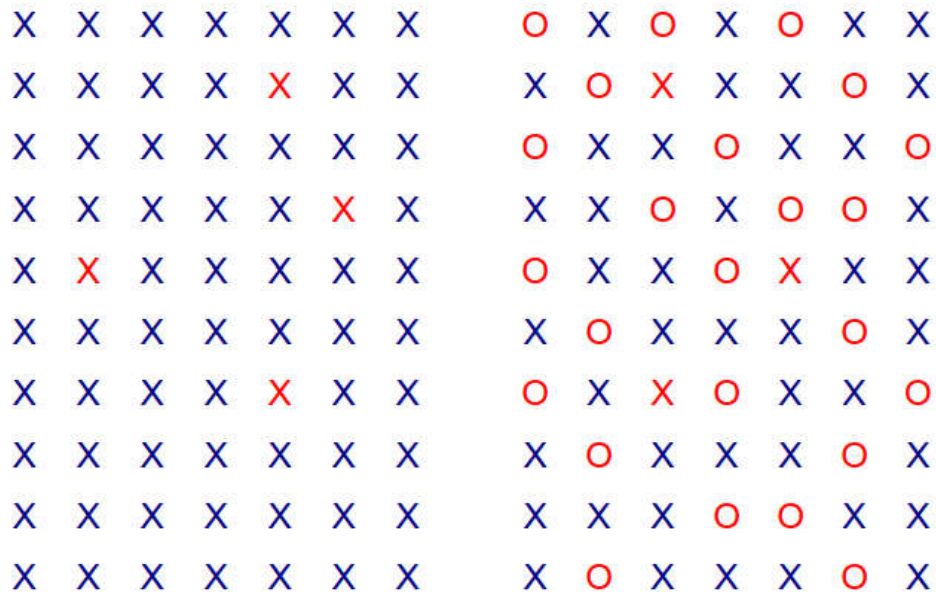
Edward H. Adelson

(b)

# Что такое компьютерное зрение?

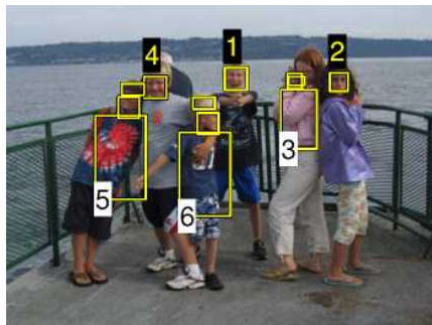


(c)



(d)

# Что такое компьютерное зрение?



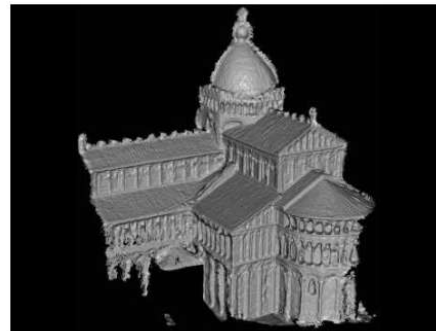
(a)



(b)



(c)



(d)

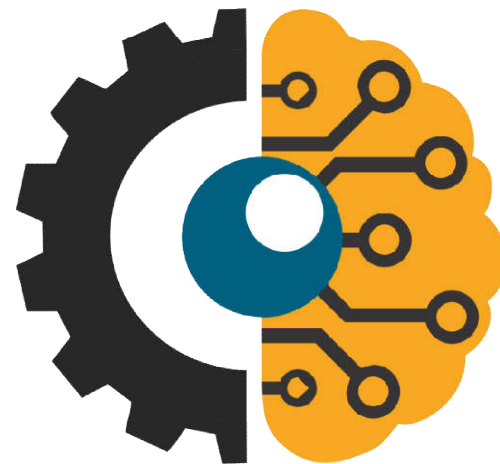
# Что такое компьютерное зрение?

Компьютерное зрение (*Computer Vision, CV*) – это область науки, которая включает в себя набор методов, наделяющих компьютер способностью «видеть» и понимать увиденное.



187	183	174	168	160	162	129	181	172	161	165	166
185	182	163	74	75	62	93	17	110	210	180	164
180	180	60	14	54	6	10	33	48	106	169	181
206	109	6	234	131	111	132	204	166	15	66	180
194	68	137	261	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	36	74	206
188	68	179	209	186	216	211	158	129	76	20	163
189	97	166	86	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	262	236	231	148	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	198	227	210	127	102	36	101	265	224
190	214	173	66	103	143	36	90	2	109	249	215
187	196	236	75	1	81	47	0	4	217	266	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	177	123	200	176	18	36	218

187	183	174	168	160	162	129	181	172	161	165	166
185	182	163	74	75	62	93	17	110	210	180	164
180	180	60	14	54	6	10	33	48	106	169	181
206	109	6	234	131	111	132	204	166	15	66	180
194	68	137	261	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	36	74	206
188	68	179	209	186	216	211	158	129	76	20	163
189	97	166	86	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	262	236	231	148	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	198	227	210	127	102	36	101	265	224
190	214	173	66	103	143	36	90	2	109	249	215
187	196	236	75	1	81	47	0	4	217	266	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	177	123	200	176	18	36	218



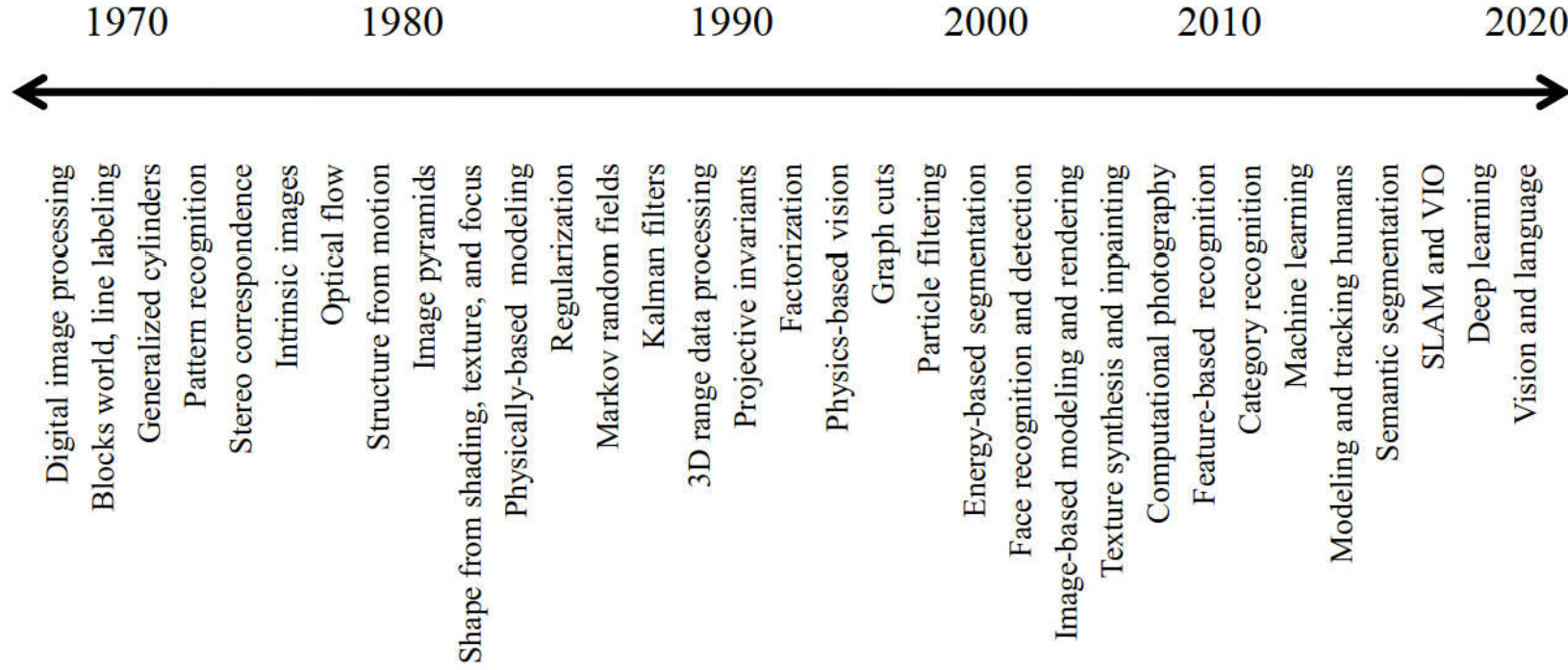
# История развития CV

- Лоуренс (Ларри) Робертс – автор одной из первых статей о компьютерном зрении
- С конца 1990-х наблюдается рост количества приложений CV, где применяют (применимы) камеры

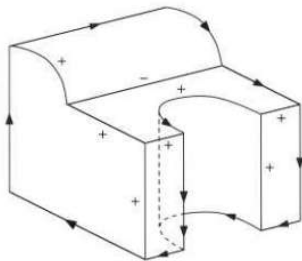




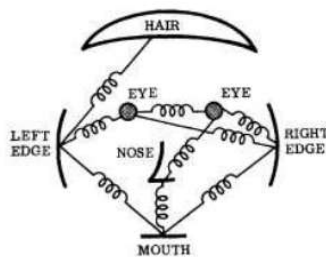
# История развития CV



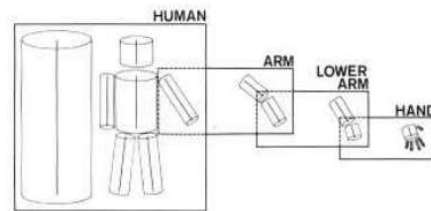
# История развития CV. 1970-е



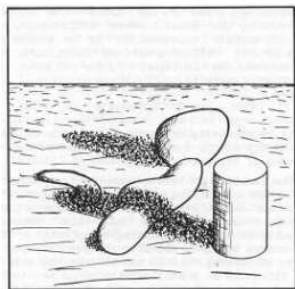
(a)



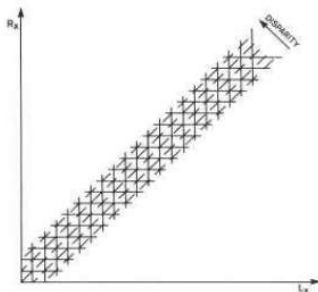
(b)



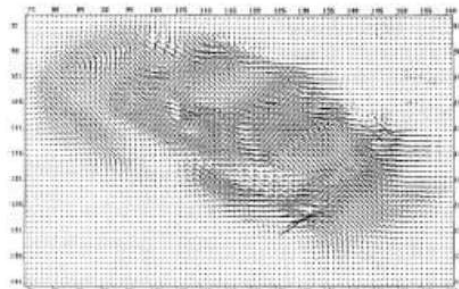
(c)



(d)



(e)



(f)

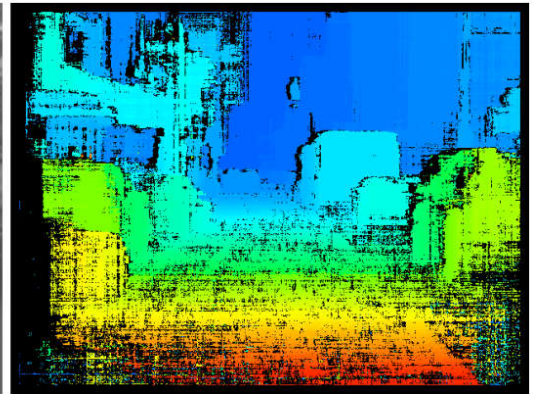
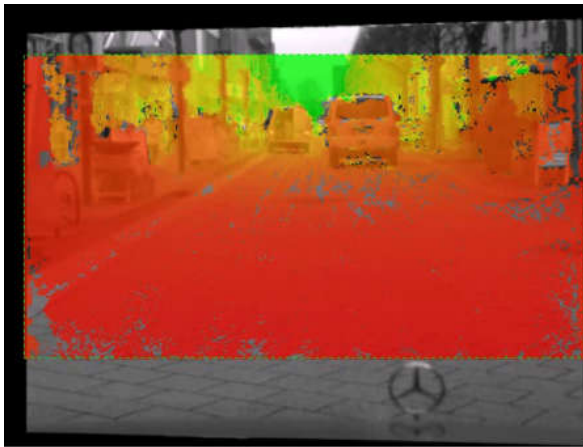
# Стереоанализ



Расстояние до цели	1 м	10 м	100 м	1 км	10 км	100 км
Время отклика	6.7 нс	67 нс	0.67 мкс	6.7 мкс	67 мкс	0.67 мс

Принцип действия лидара

# Вычислительное стереозрение



# История развития CV

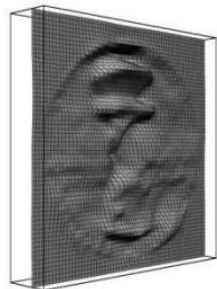
Три уровня описания системы обработки визуальной информации (Дэвид Марр):

- Теория вычислений
- Представления и алгоритмы
- Аппаратная реализация

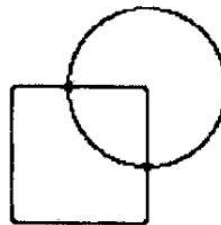
# История развития CV. 1980-е



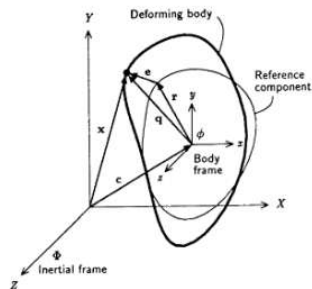
(a)



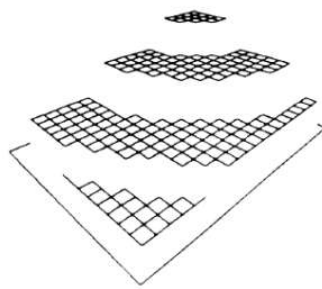
(b)



(c)



(d)



(e)



(f)

# История развития CV. 1990-е



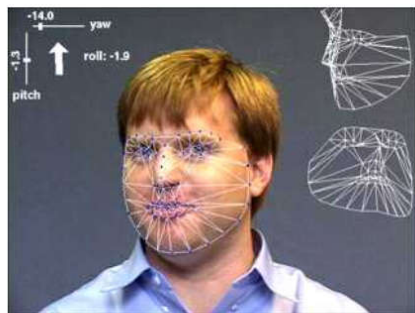
(a)



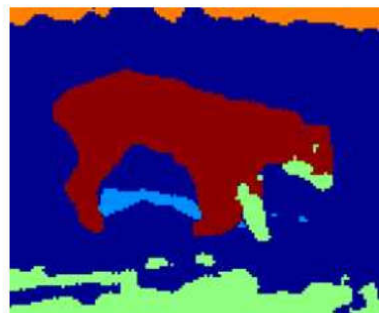
(b)



(c)



(d)

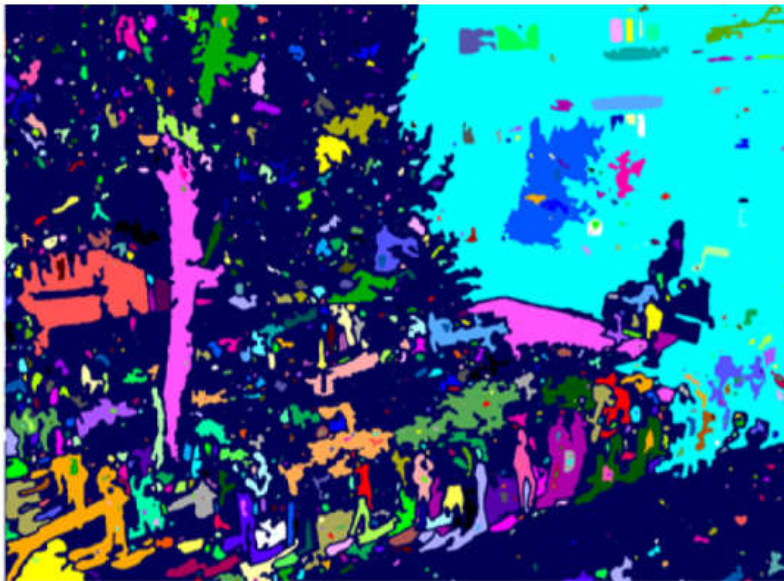


(e)



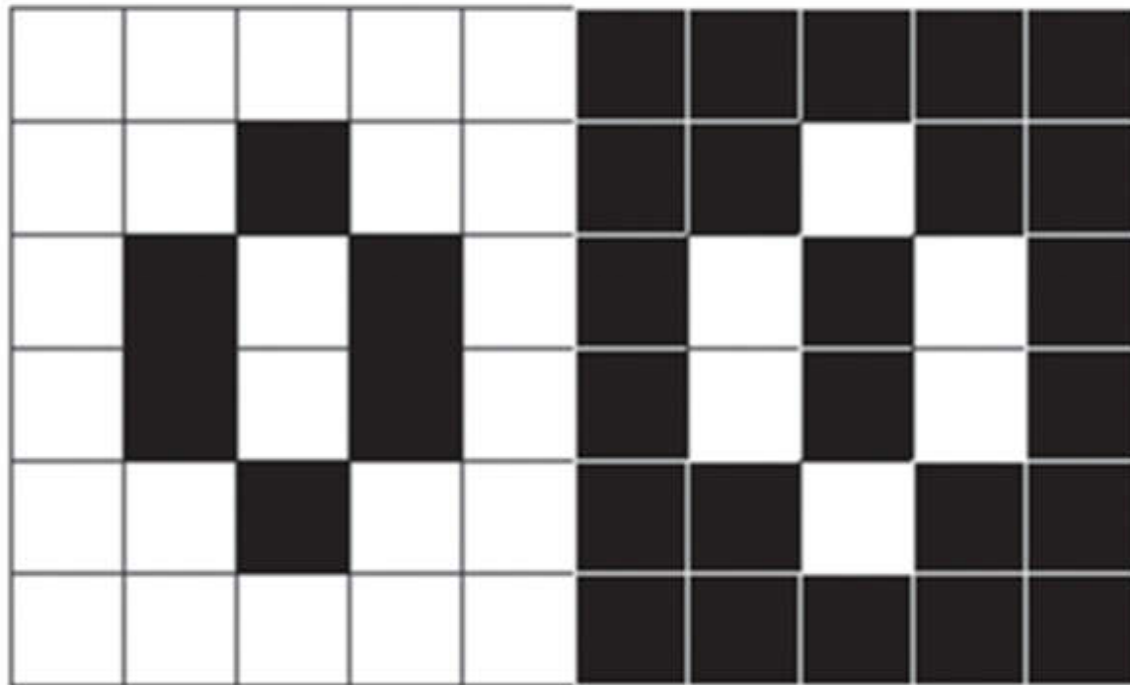
(f)

# Сегментация изображений

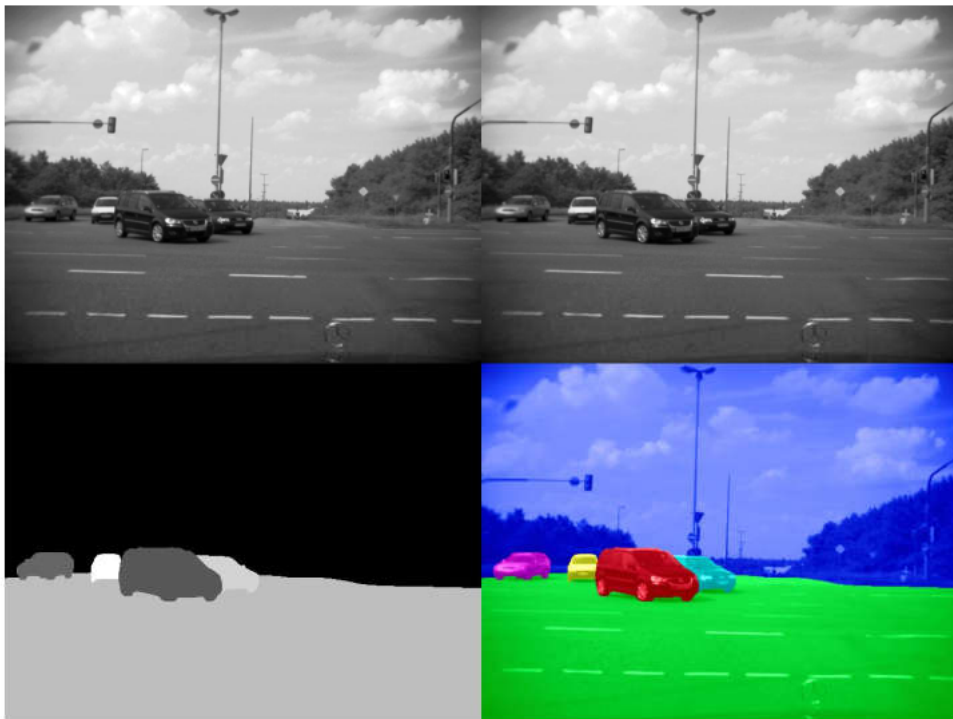




# Связность пикселей



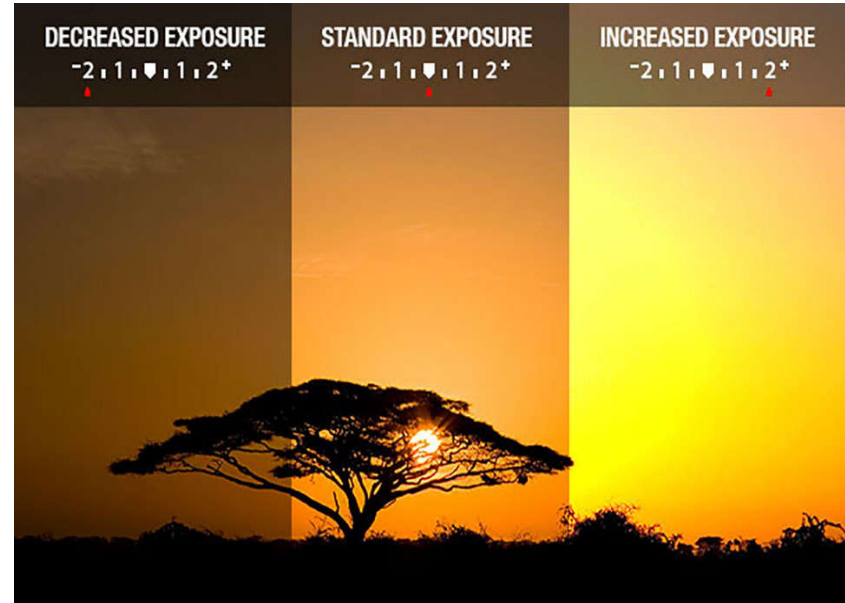
# Семантическая сегментация



# История развития CV. 2000-е. Вычислительная фотография

Может применяться для:

1. удаления шума;
2. в живописи (при отсутствующих пикселях);
3. создания особых эффектов;
4. улучшения характеристик изображения и т.д.



# Художественные фильтры



Фильтр, эмулирующий стиль картин Винсента Ван Гога

# Эмуляция пуантилизма



Фильтр, эмулирующий пуантилизм

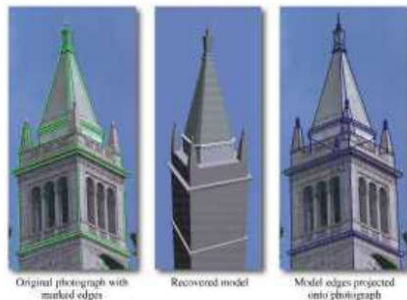
# Добавление тумана на фото



# История развития CV. 2000-е



(a)



(b)



(c)



(d)



(e)



(f)







# Что такое изображение?

0 – черный цвет

255 – белый цвет

1-254 – уровни серого

Каждое число отображается  
в виде заштрихованного  
маленького квадрата,  
называемого пикселем.





# Что такое изображение?



Пример улучшения контрастности

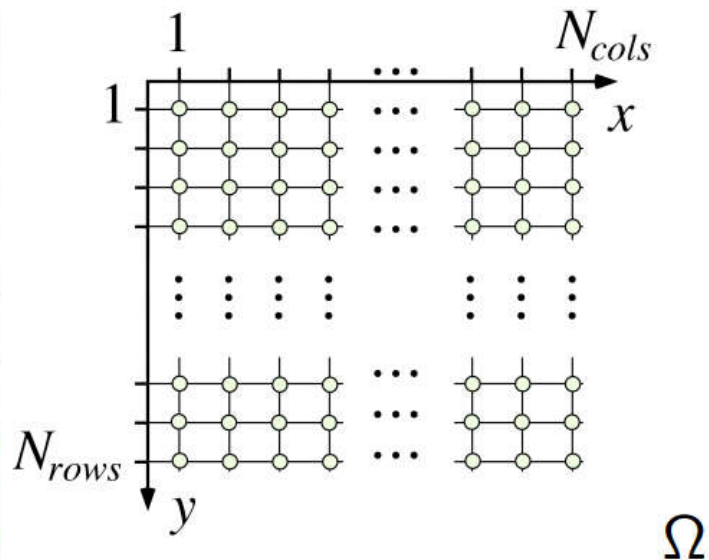
# Изображение в пространственной области

Цифровое изображение определяется посредством дискретизации непрерывных аналоговых данных в пространственной области. Оно состоит из прямоугольного массива пикселей  $(x, y, u)$ , каждый из которых является комбинацией местоположения  $(x, y)$  и значения  $u$ , представляющего отсчёт в точке  $(x, y)$ . Точки  $(x, y)$  образуют регулярную сетку. Формально изображение  $I$  определено на прямоугольном множестве  $\Omega$ , которое называется носителем ( $\mathbb{Z}$  – множество всех целых чисел):

$$\Omega = \{(x, y) : 1 \leq x \leq N_{cols} \wedge 1 \leq y \leq N_{rows}\} \subset \mathbb{Z}^2$$

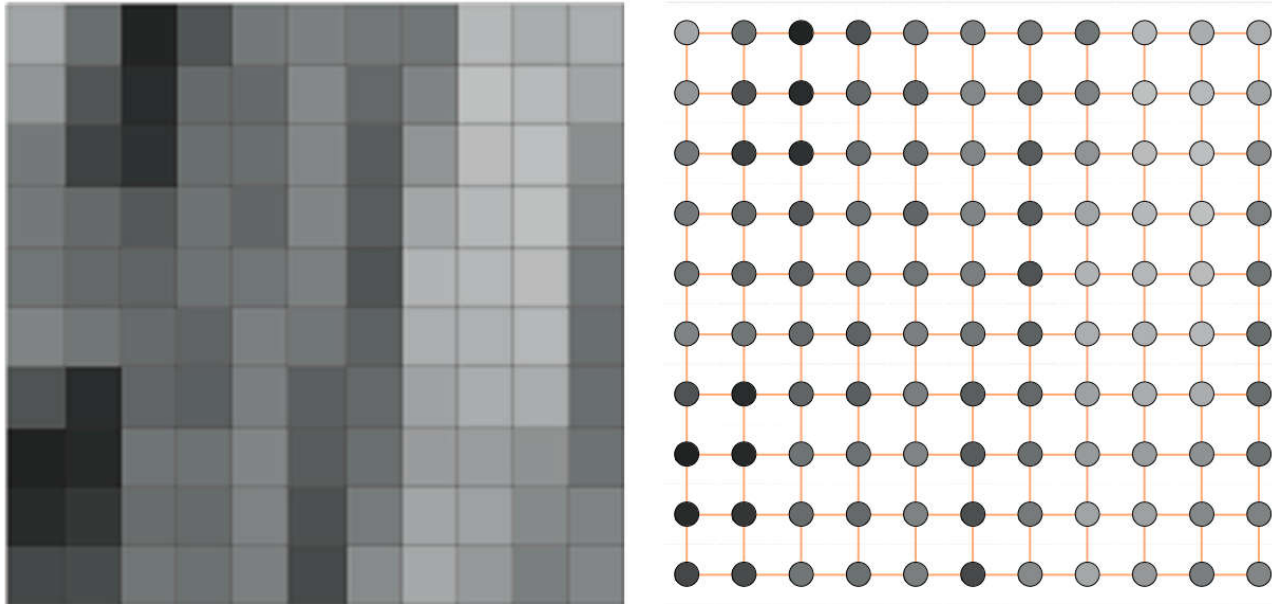
и содержит узлы сетки, или местоположения пикселей.

# Система координат изображения



Левосторонняя система координат. Большой палец указывает положительное направление оси  $x$ , указательный палец – положительное направление оси  $y$ .

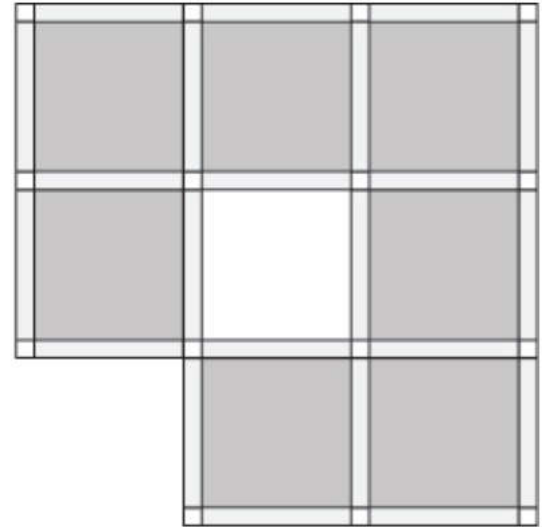
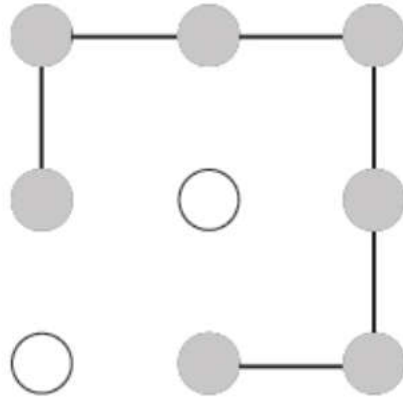
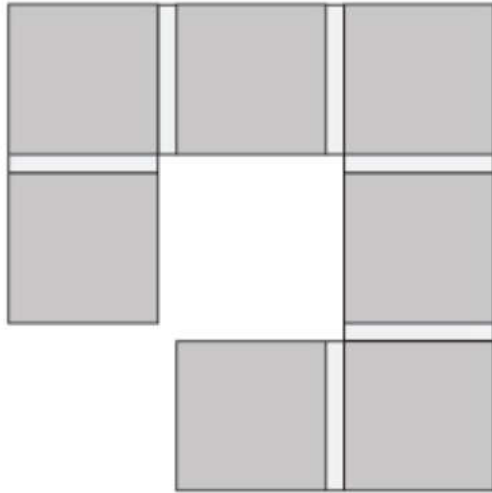
# Пиксели



Слева – ячейки сетки.

Справа – метки в узлах сетки, являющихся центрами ячеек сетки.

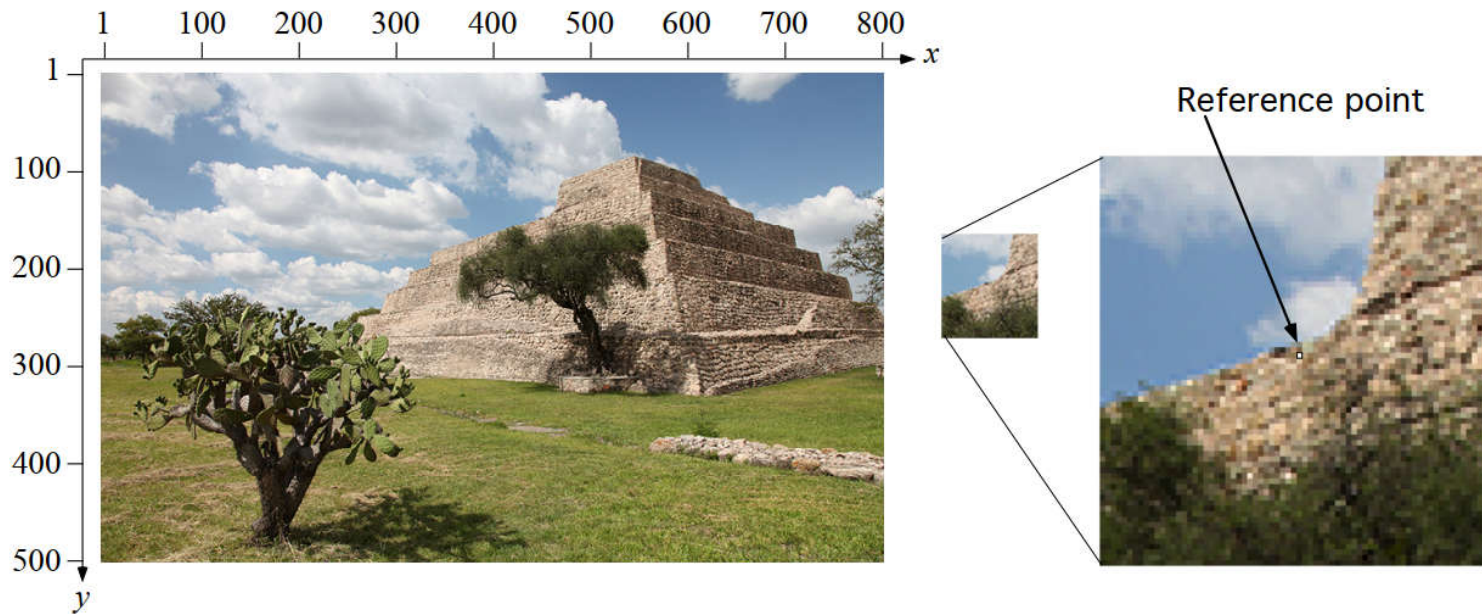
# Смежность



Смежность определяет связность и связанные наборы пикселей



# Окна изображения



Обработка изображений часто является локальной

# Значения изображения

Значениями скалярного изображения являются целые числа:

$$u \in \{0, 1, \dots, 2^a - 1\}$$

Например, *полутонное* изображение – это скалярное изображение, где скалярные значения представляют уровни серого, при этом 0 – черный и  $2^a - 1$  – белый; все остальные уровни серого линейно интерполируются между черным и белым

0

...

128

...

255



# Значения изображения

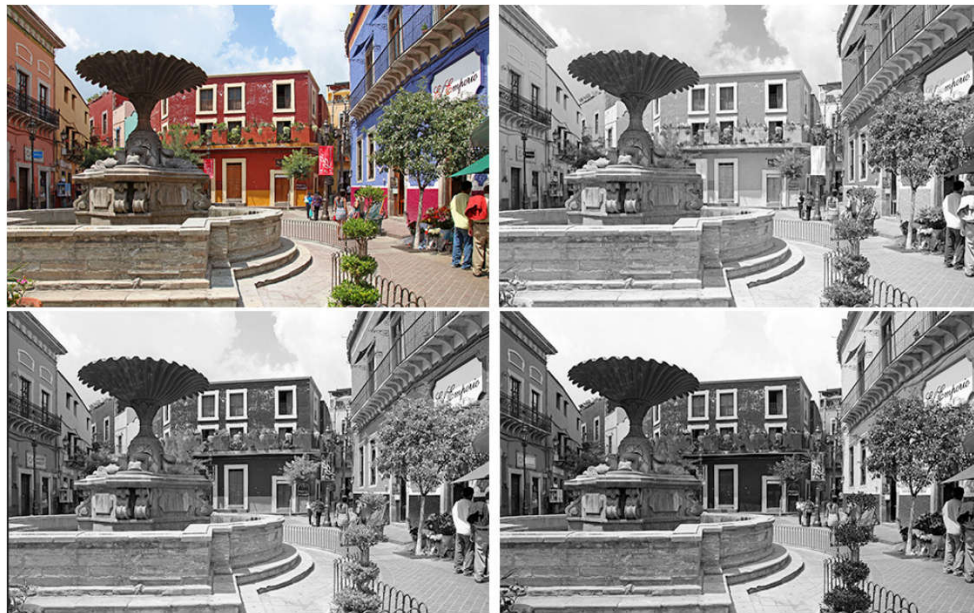
*Бинарное* изображение имеет два значения, традиционно 0 – чёрный и 1 – белый (чёрные объекты на белом фоне).

В *векторном* изображении имеется более одного канала. Значениями изображения являются векторы  $[u_1, \dots, u_{N_{channels}}]$

Например, изображения *RGB* имеют три канала, по одному для красной, зелёной и синей компонент. Каждый канал представляет собой полутоновое изображение.



# Пример изображения RGB



Цветное изображение RGB и три его канала для красного (вверху справа), зеленого (внизу слева) и синего (внизу справа)

# Основные статистики изображения. Среднее

Для скалярного изображения  $I$  (размера  $N_{cols} \times N_{rows}$ ) *среднее* (*средний уровень яркости*) определяется следующим образом:

$$\begin{aligned}\mu_I &= \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=1}^{N_{cols}} \sum_{y=1}^{N_{rows}} I(x, y) \\ &= \frac{1}{|\Omega|} \sum_{(x,y) \in \Omega} I(x, y)\end{aligned}$$

$|\Omega| = N_{cols} \times N_{rows}$  – мощность носителя  $\Omega$ , содержащего все пиксели

# Основные статистики изображения. Дисперсия и стандартное отклонение

*Дисперсией* скалярного изображения  $I$  называется величина:

$$\sigma_I^2 = \frac{1}{|\Omega|} \sum_{(x,y) \in \Omega} [I(x,y) - \mu_I]^2$$

Квадратный корень из нее называется *стандартным отклонением* изображения  $I$ . Применима также хорошо известная формула статистики:

$$\sigma_I^2 = \left[ \frac{1}{|\Omega|} \sum_{(x,y) \in \Omega} I(x,y)^2 \right] - \mu_I^2$$

# Гистограммы



*Слева сверху: исходное изображение. Справа сверху: более яркий вариант.  
Слева внизу: более тёмный вариант. Справа внизу: после выравнивания гистограммы.*

# Гистограммы

Пусть имеется скалярное изображение  $I$  с пикселями  $(i, j, u)$ . *Абсолютные частоты* определяются путём подсчёта того, сколько раз значение  $u$  встречается в носителе  $\Omega$ , содержащем все пиксели:

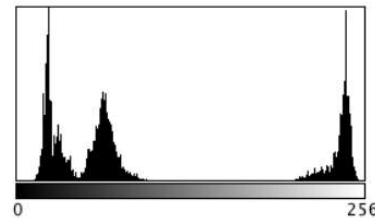
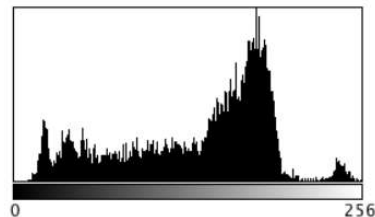
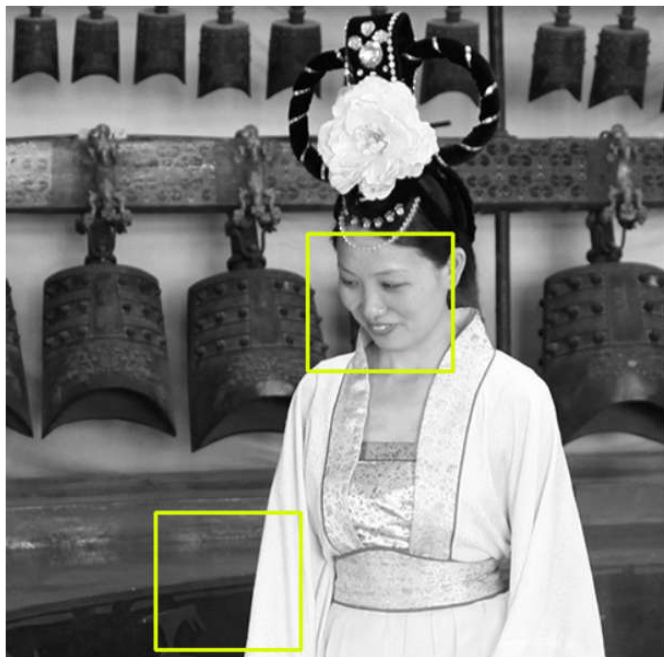
$$H_I(u) = |\{(x, y) \in \Omega : I(x, y) = u\}|$$

Относительные частоты – значения между 0 и 1 – можно сравнить с функцией плотности вероятности распределения дискретной случайной величины:

$$h_I(u) = \frac{H_I(u)}{|\Omega|}$$

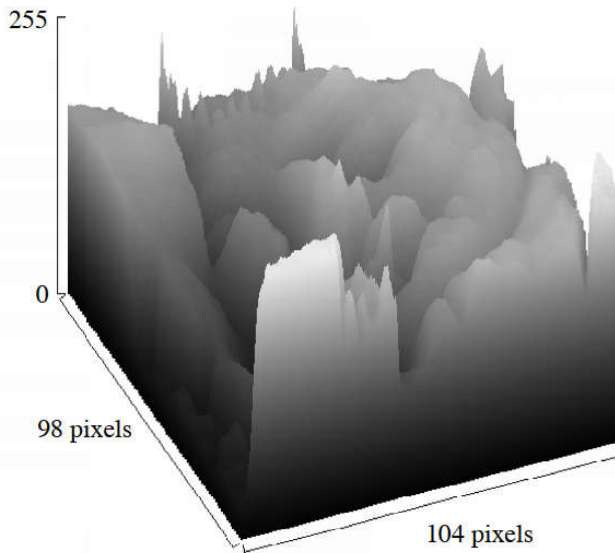
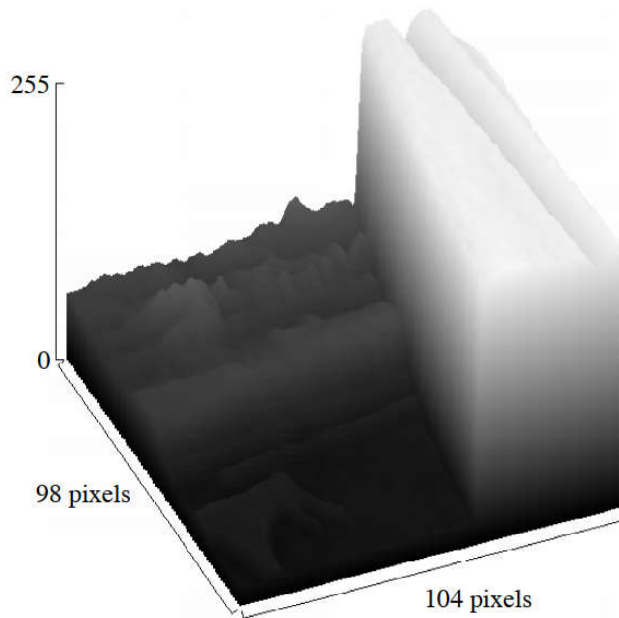


# Статистики значений в окне



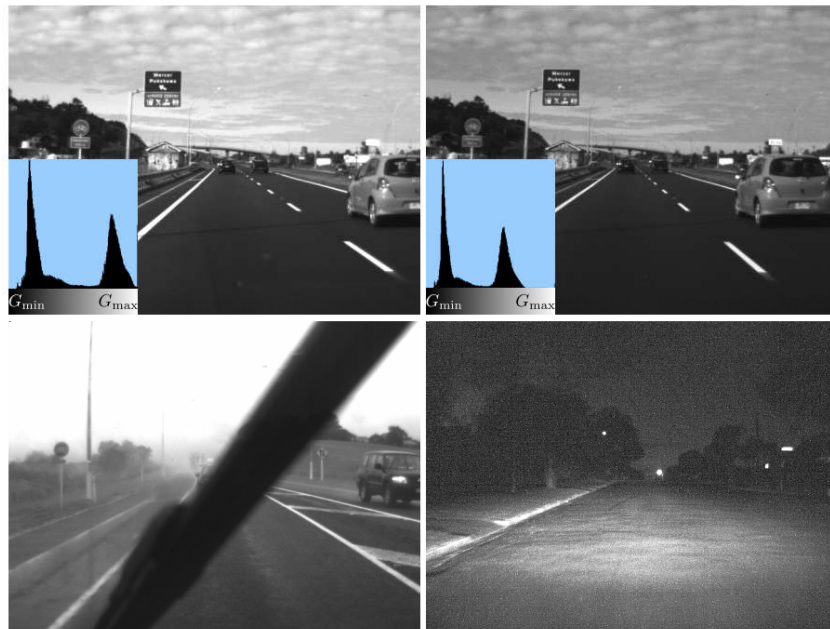
Пример двух окон размера 104x98 в изображении и соответствующих гистограмм

# Трёхмерное представление уровней яркости



Слева: «крутой переход от тёмного к яркому». Справа: «незначительное» изменение

# Обработка изображений



Вверху: два изображения, снятые в один момент времени с разной яркостью  
Слева внизу: размытие из-за дождя. Справа внизу: шум на сцене, снятой ночью

# Выравнивание гистограммы

Скалярное изображение  $I$  преобразуется таким образом, чтобы в новом изображении  $I_{new}$  все уровни яркости встречались с одинаковой частотой. Цель – выполнение условия:

$$H_{I_{new}}(u) = \frac{N_{cols} N_{rows}}{G_{max} + 1} = \frac{|\Omega|}{G_{max} + 1} = const$$

для всех  $u \in \{0, 1, \dots, G_{max}\}$ . В реальности используется:

$$g(u) = c_I(u) \cdot G_{max}$$

# Выравнивание гистограмм



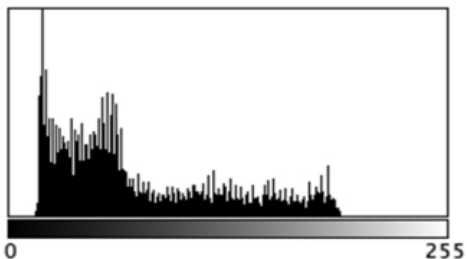
Слева: входное изображение и его гистограмма  
Справа: то же изображение после выравнивания гистограмм

# Линейное масштабирование

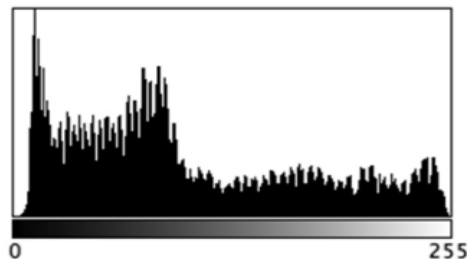
Предположим, что в гистограмме изображения  $I$  положительные значения сосредоточены в ограниченном интервале. Цель – сделать так, чтобы все значения, встречающиеся в  $I$ , были линейно распределены в интервале от 0 до  $G_{\max}$ .

$$a = -u_{\min} \quad \text{and} \quad b = \frac{G_{\max}}{u_{\max} - u_{\min}}$$
$$g(u) = b(u + a)$$

# Линейное масштабирование



Count: 232960  
Mean: 75.418  
StdDev: 49.903  
Min: 12  
Max: 195  
Mode: 19 (6693)



Count: 232960  
Mean: 101.575  
StdDev: 69.874  
Min: 0  
Max: 255  
Mode: 12 (3234)

# Локальное среднее и локальный максимум

Рассмотрим скользящее окно  $W_p$  размера  $(2k+1) \times (2k+1)$  с центром в качестве начальной точки  $p$ .

Локальное среднее вычисляется так ( $p = (x, y)$ ):

$$\mu_{W_p(I)} = \frac{1}{(2k+1)^2} \cdot \sum_{i=-k}^{+k} \sum_{j=-k}^{+k} I(x+i, y+j)$$

Для вычисления локального максимума справедливо ( $p = (x, y)$ ):

$$J(p) = \max\{I(x+i, y+j) : -k \leq i \leq k \wedge -k \leq j \leq k\}$$



# Локальные операторы



1	0	-1
2	0	-2
1	0	-1

 /1

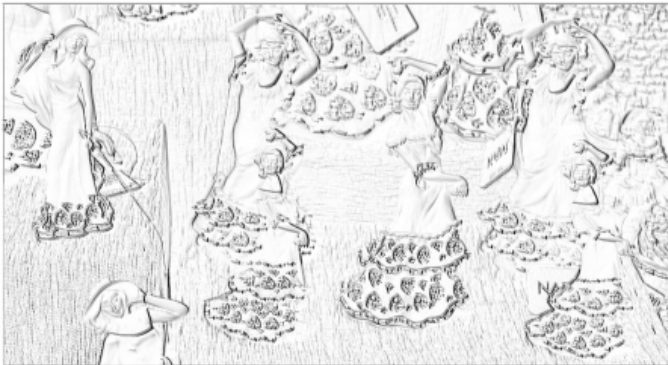
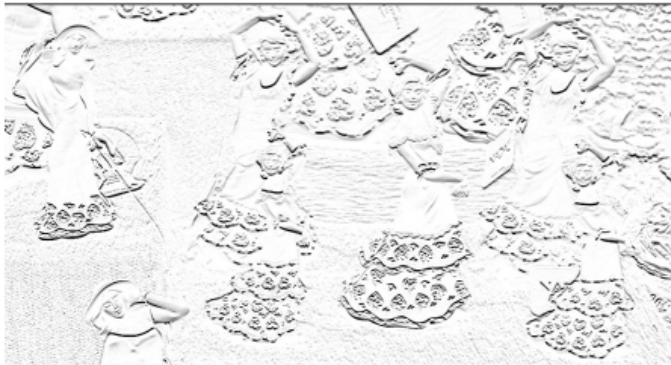
Слева вверху: исходное изображение. Справа вверху: локальный максимум для  $k=3$ .  
Слева внизу: локальный минимум для  $k=5$ . Справа внизу: локальный оператор с ядром фильтра, показанным справа

# Удаление шума



Слева сверху: исходное изображение с шумом. Справа сверху: прямоугольный фильтр  $3 \times 3$ .  
Слева внизу: сигма-фильтр  $3 \times 3$  с параметром  $\sigma=30$ . Посередине внизу: медианный фильтр  $3 \times 3$ .  
Справа внизу: фильтр Гаусса с  $\sigma=1$ .

# Детекторы границ. Оператор Собеля



# Детекторы границ. Лапласиан



0	1	0
1	-4	1
0	1	0

 /1

1	1	1
1	-8	1
1	1	1

 /1

-1	2	-1
2	-4	2
-1	2	-1

 /1

Слева: Исходное изображение I.  
Справа: Абсолютные значения лапласиана I (перевернутое изображение).