

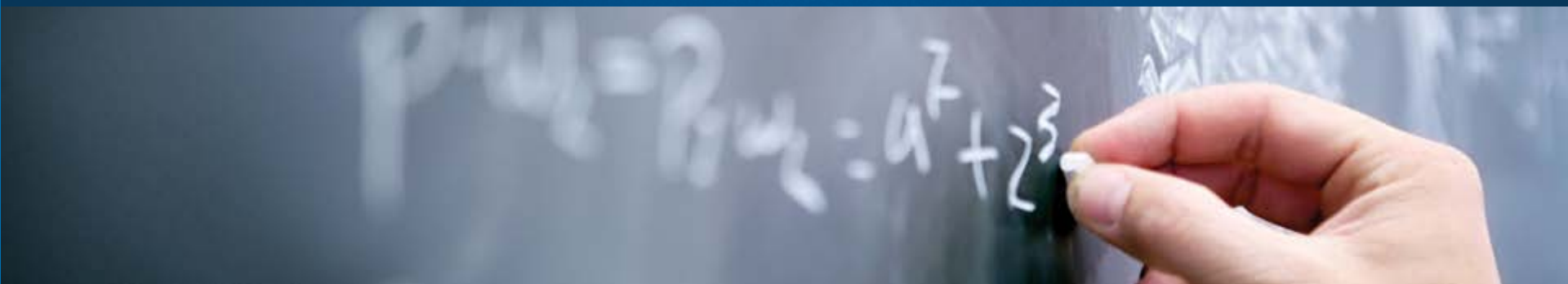
# Анализ изображений и видео

## Лекция 4: Построение признаков и сравнение изображений: глобальные признаки

Наталья Васильева

[nvassilieva@hp.com](mailto:nvassilieva@hp.com)

HP Labs Russia

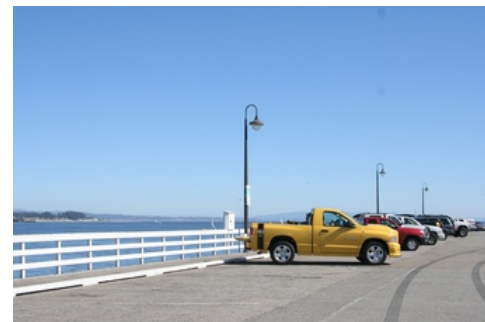


# Вопросы:

1. К исходному изображению применяется усредняющая маска для подавления шума, а затем маска лапласиана для улучшения мелких деталей. Изменится ли результат, если поменять очередность этих операций?
2. С помощью какой операции можно полностью удалить из изображения изолированные темные или светлые небольшие области? Каким должен быть размер маски, если известно, что размер таких областей не превосходит  $n$  пикселей?
3. Покажите, что второй проход операции эквализации гистограммы даст точно тот же результат, что и после первого прохода.
4. Совместное использование процедур фильтрации с усилением высоких частот и эквализации гистограммы является эффективным методом повышения резкости и улучшения контраста. Влияет ли порядок на окончательный результат? Почему?



# Зачем сравнивать изображения?



- Поиск
- Классификация
  - indoor/outdoor
  - landscape/city/forest/desert/...
- Аннотирование
- Image parsing

# Зачем сравнивать изображения?

Поиск изображений



# Зачем сравнивать изображения?

Классификация, кластеризация

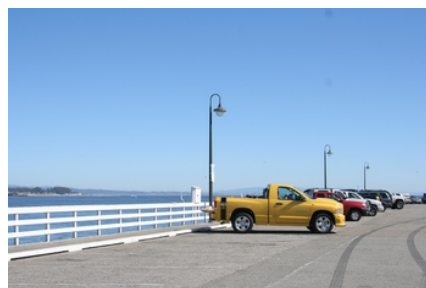


Обучающее множество:  
изображения и метки класса

Обучение классификатора:  
сопоставление особенностей  
изображений меткам классов

Обучение

Модель  
классификатора



Тестовое изображение

Сопоставление тестового  
изображения модели  
классификатора

Тестирование

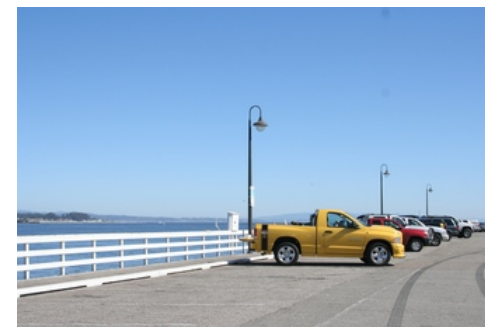
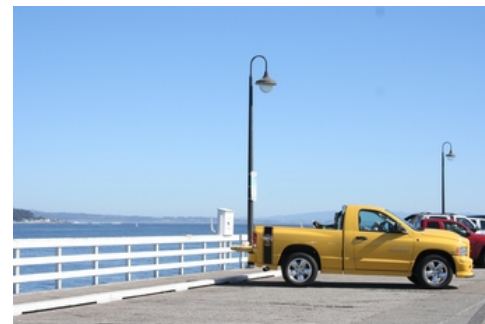
Предсказанные  
классификатором  
классы:  
– open outdoor





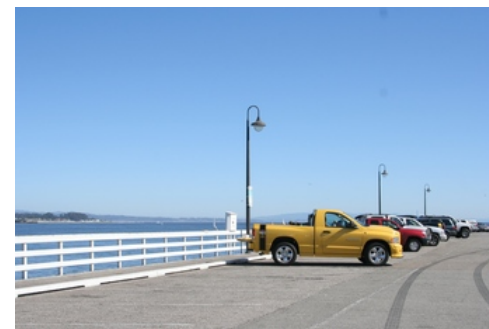
# Зачем сравнивать изображения?

Обнаружение объектов



# Зачем сравнивать изображения?

## Аннотирование



- Небо, гора, дом, трамвай, пикап, набережная, трава, солнечно, день, дорога,...

# Как сравнивать?

- Представить каждое изображение в виде набора признаков
  - компактность
  - описательность
- Сравнить наборы признаков между собой => сравнивать изображения





# Признаки изображений



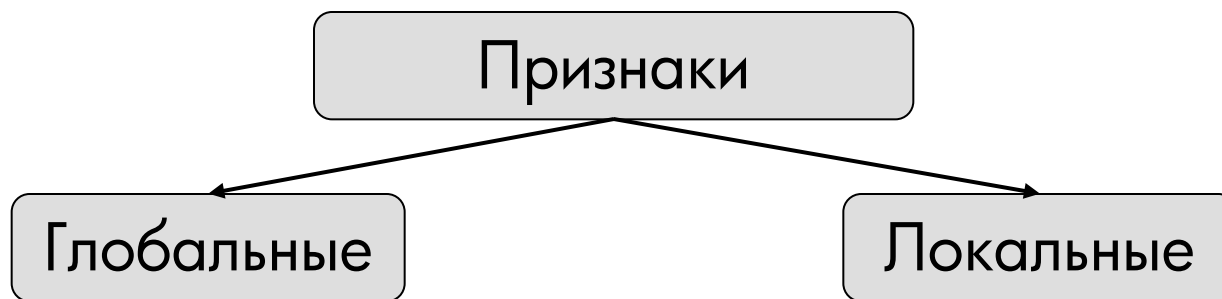
## Аннотации и метаданные:

- тэги, аннотации;
- дата создания;
- геотэги;
- названия файла;
- параметры съемки (выдержка, диафрагма, наличие вспышки...).

## Признаки, полученные по значениям пикселей:

- цветовые;
- текстурные;
- формы;
- пространственного расположения.

# Признаки изображений



Описывают **картинку целиком**:

- средняя яркость;
- среднее значение по красному каналу;
- ...

Обычно, признаки вычисляются по всем пикселям изображения

Описывают **часть** картинки:

- средняя яркость верхней левой четверти;
- среднее значение по красному каналу в окрестности центра изображения;
- ...

Сегментация, поиск точек интереса, построение признаков по окрестностям точек интереса

# Пространства признаков

**Вектор признаков** – набор числовых параметров, отражающих особенности изображения

**Пространство признаков** – пространство векторов признаков с заданной функцией подобия (расстояния) для их сравнения.

Image A



Image B

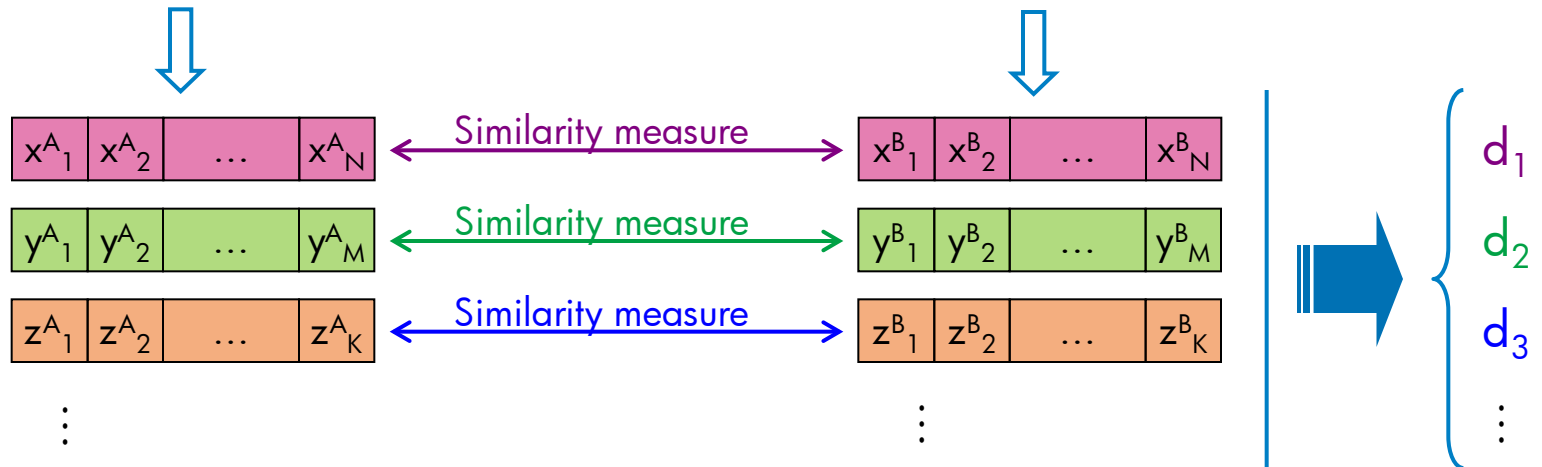


# Комбинирование признаков

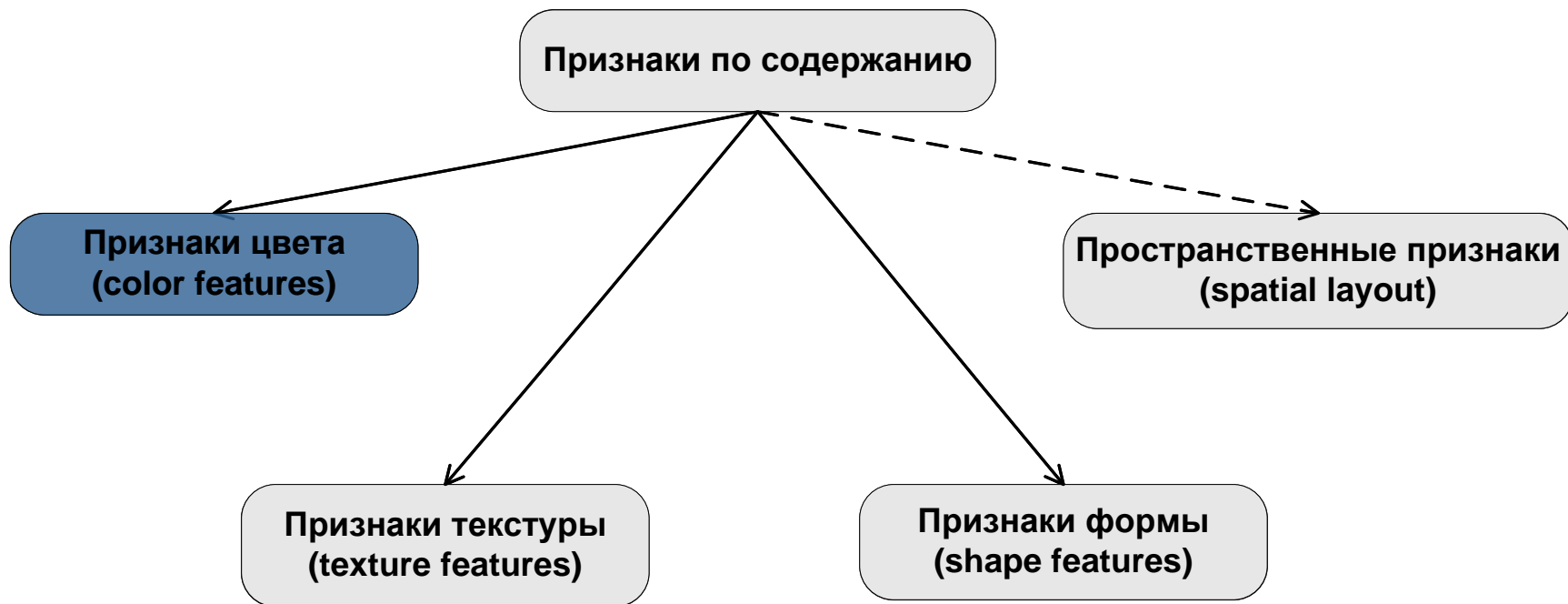
Image A



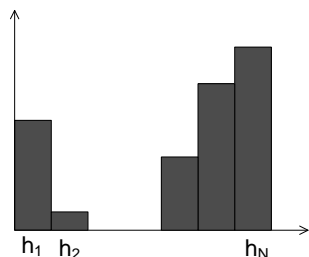
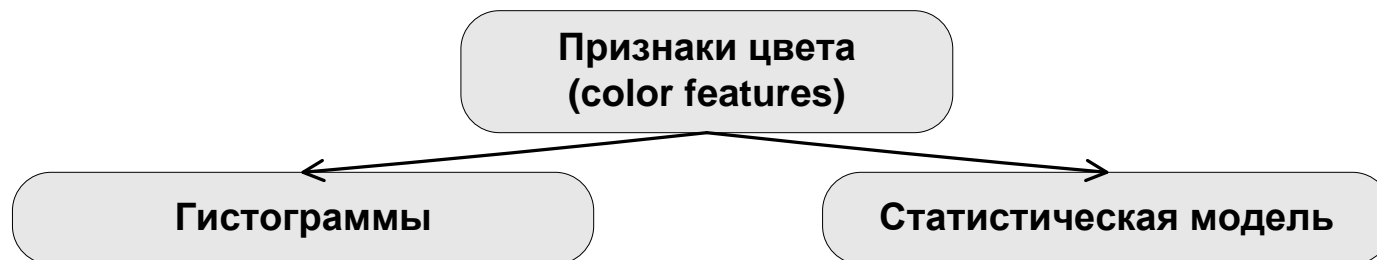
Image B



$$D = \sum_i c_i d_i$$







$$F(I) = (h_1^I, h_2^I, \dots, h_N^I)$$

Метрики:  $L_1, L_2, L_\infty, \chi^2,$   
EMD (earth mover's distance), ...

Мат. ожидание, дисперсия, 3-ий момент: для каждого цветового канала

$$F(I) = (E_1^I, E_2^I, E_3^I, \\ \sigma_1^I, \sigma_2^I, \sigma_3^I, \\ s_1^I, s_2^I, s_3^I)$$

Метрики:  $\sim L_1$

Stricker M., Orengo M. Similarity of Color Images. Proceedings of the SPIE Conference, vol. 2420, p. 381-392, 1995

# Популярные функции расстояния для гистограмм

- Пересечение гистограмм (Histogram intersection)
  - Гистограммы должны быть нормализованы
  - Эквивалента  $L_1$

$$\text{histint}(h_i, h_j) = 1 - \sum_{m=1}^K \min(h_i(m), h_j(m))$$

- Chi-squared Histogram matching distance

$$\chi^2(h_i, h_j) = \frac{1}{2} \sum_{m=1}^K \frac{[h_i(m) - h_j(m)]^2}{h_i(m) + h_j(m)}$$

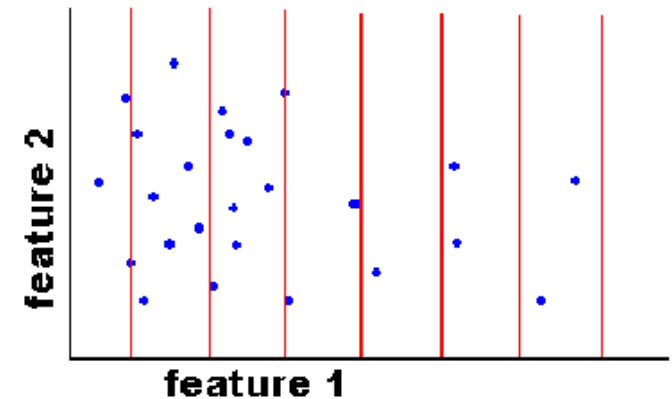
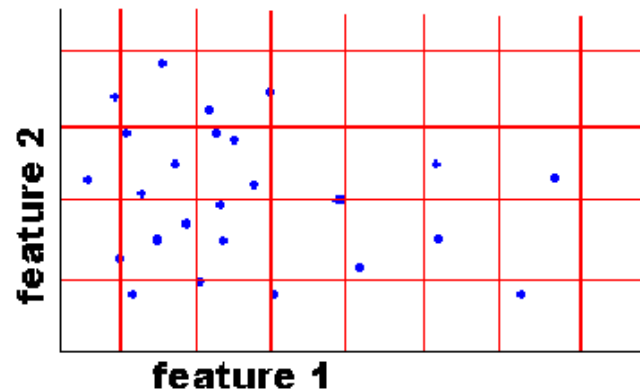
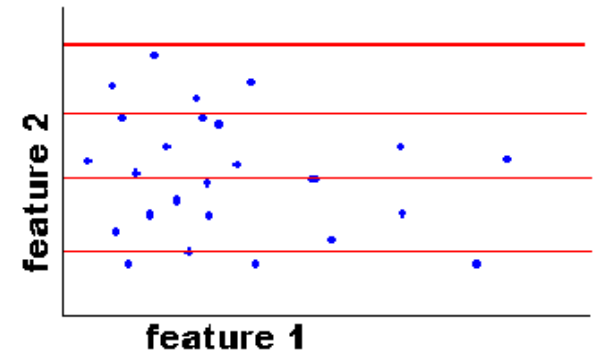
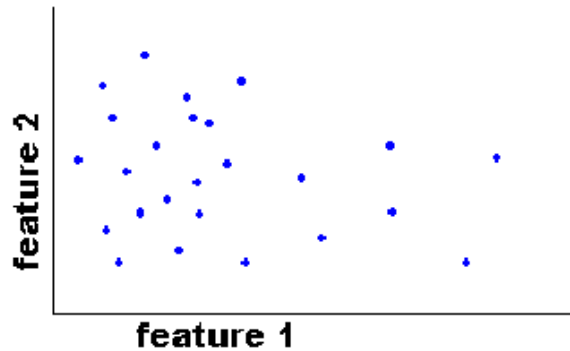


# Квантование пространства при построении гистограмм



- Схема квантования влияет на размер вектора признаков
- Если метрика не учитывает подобие цветов:
  - Много промежутков – расстояние между близкими по цвету может быть большим
  - Мало промежутков – расстояние между далекими по цвету может быть маленьким

# Квантование в случае многомерных признаков



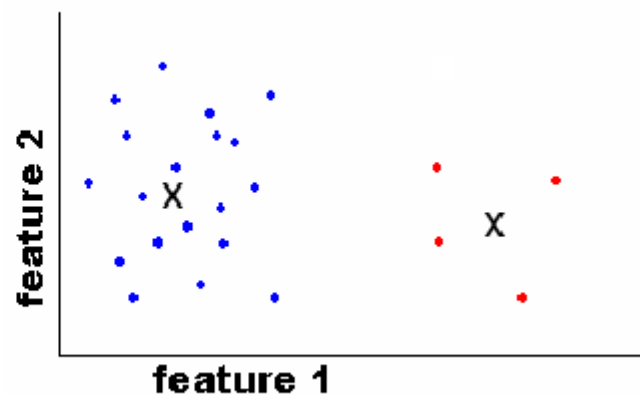
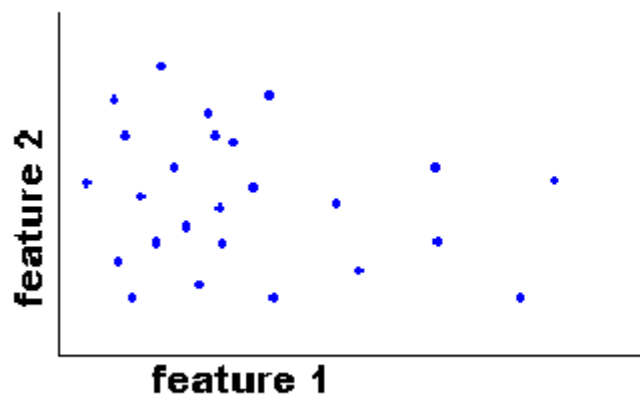
- Joint histogram

- Requires lots of data
- Loss of resolution to avoid empty bins

## Marginal histogram

- Requires independent features
- More data/bin than joint histogram

# Квантование пространства при помощи кластеризации



- Построение кластеров по признакам обучающего множества
- Определение центров кластеров
- Для пикселя тестового изображения – принадлежность промежутку квантования по ближайшему центру



# Выбор схемы квантования цветового пространства

Проверка существования пороговых значений шага квантования, таких что выбор меньших шагов не приводит к повышению качества результатов поиска

Число диапазонов	HSI*, nHSI*, HCL, nHCL	CIELab, nCIELab	nRGB
36	$\langle 6, 2, 3 \rangle$		
41	$\langle 6, 2, 3 \rangle_{th}$		
64			$\langle 4, 4, 4 \rangle$
72	$\langle 12, 2, 3 \rangle$	$\langle 2, 6, 6 \rangle$	
96	$\langle 6, 4, 4 \rangle$		
108	$\langle 18, 2, 3 \rangle$		
144		$\langle 4, 6, 6 \rangle$	
216	$\langle 6, 6, 6 \rangle$	$\langle 6, 6, 6 \rangle$	$\langle 6, 6, 6 \rangle$
288	$\langle 12, 4, 6 \rangle$	$\langle 2, 12, 12 \rangle$	
512			$\langle 8, 8, 8 \rangle$
576	$\langle 12, 6, 8 \rangle$		
586	$\langle 12, 6, 8 \rangle_{th}$		
864		$\langle 6, 12, 12 \rangle$	
1152		$\langle 8, 12, 12 \rangle$	
1728	$\langle 12, 12, 12 \rangle$	$\langle 12, 12, 12 \rangle$	$\langle 12, 12, 12 \rangle$
3060	$\langle 17, 12, 15 \rangle$		
3077	$\langle 17, 12, 15 \rangle_{th}$		
5832	$\langle 18, 18, 18 \rangle$	$\langle 18, 18, 18 \rangle$	$\langle 18, 18, 18 \rangle$

Пространства и схемы квантования, участвовавшие в экспериментах

- Равномерное квантование

$$ColorSpace\langle K \times L \times M \rangle = \{\Delta_{k,l,m}\}_{k=1..K, l=1..L, m=1..M},$$

$$\Delta_{k,l,m} = \{(x, y, z) | x \in (x_k, x_{k+1}], y \in (y_l, y_{l+1}], z \in (z_m, z_{m+1}]\},$$

$$x_j = j\delta_x, \quad y_j = j\delta_y, \quad z_j = j\delta_z.$$

- Равномерное квантование с граничными условиями пространств семейства HSV

$$ColorSpace\langle K \times L \times M \rangle_{th} = \Delta^B \cup \Delta^W \cup$$

$$\{\Delta_m^G\}_{m=1..M} \cup \{\Delta_{k,l,m}\}_{k=1..K, l=1..L, m=1..M},$$

$$\Delta^B = \{(x, y, z) | z \leq I^B\}, \quad \Delta^W = \{(x, y, z) | z \geq I^W\},$$

$$\Delta_m^G = \{(x, y, z) | y \leq S^G, z \in (z_m, z_{m+1}]\},$$

$$\Delta_{k,l,m} = \{(x, y, z) | x \in (x_k, x_{k+1}], y \in (y_l, y_{l+1}], z \in (z_m, z_{m+1}]\},$$

$$x_j = j\delta_x, \quad y_j = S^G + j\delta_y, \quad z_j = I^B + j\delta_z.$$



# Histograms: Implementation issues

- Quantization
  - Grids: fast but applicable only with few dimensions
  - Clustering: slower but can quantize data in higher dimensions



Few Bins

Need less data

Coarser representation

Many Bins

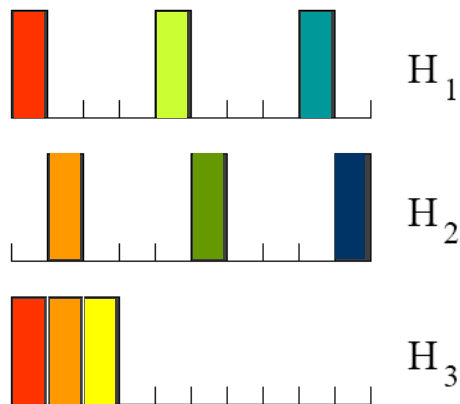
Need more data

Finer representation

- Matching
  - Histogram intersection or Euclidean may be faster
  - Chi-squared often works better
  - Earth mover's distance is good for when nearby bins represent similar values

# Цветовые гистограммы – недостатки

## 1. Не учитывается подобие цветов:



$$d(H_1, H_2) > d(H_1, H_3)$$

- Кумулятивные гистограммы

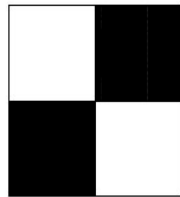
- $$d(H_1, H_2) = \sqrt{(H_1 - H_2) \cdot A \cdot (H_1 - H_2)^T}$$

A – матрица с коэффициентами  
«подобия» цветов

Niblack W., Barber R., et al. The QBIC project: Querying images by content using color, texture and shape. In IS&T/SPIE International Symposium on Electronic Imaging: Science & Technology, Conference 1908, Storage and Retrieval for Image and Video Databases, Feb. 1993

# Цветовые гистограммы – недостатки

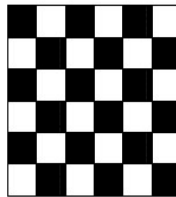
2. Не учитывается пространственное расположение цветов:



A



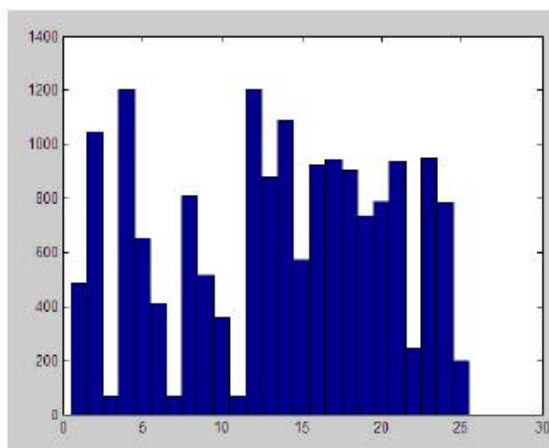
B



C

$$H_A = H_B = H_C$$

# But what about layout?

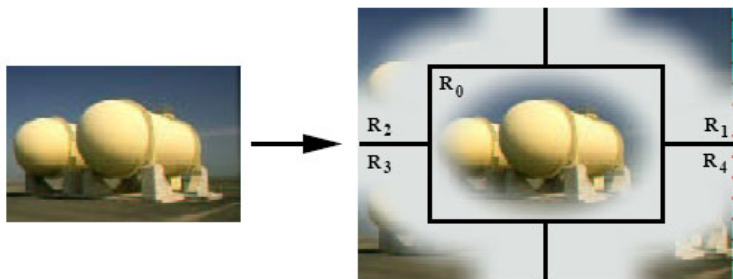


All of these images have the same color histogram



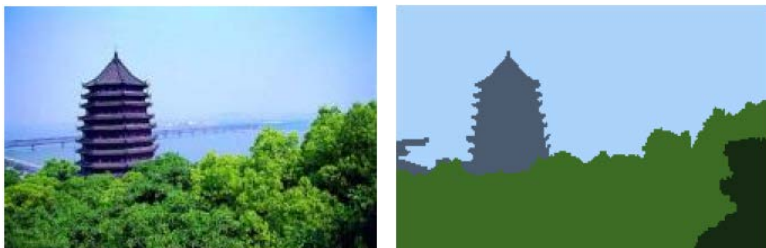
# Пространственное расположение цветов

- Разбиение изображения на фиксированные блоки
- «Нечеткие области»



Stricker M., Dimai A. Spectral Covariance and Fuzzy Regions for Image Indexing. Machine Vision and Applications, vol. 10., p. 66-73, 1997

- Сегментация

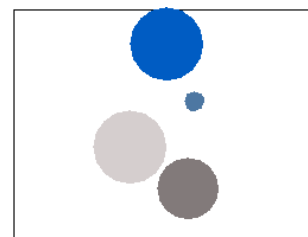


# Цветовая гистограмма с информацией о пространственном расположении цветов

## Вектор признаков

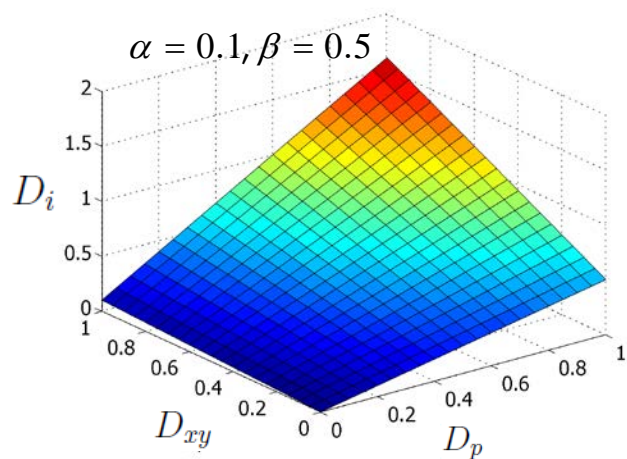
$$HistSP(I) = \{c_i \mid c_i = (p_i, x_i, y_i)\}_{i=1..N}$$

$N$  – число цветовых диапазонов,  
 $p_i$  – доля пикселей  $i$ -го диапазона,  
 $(x_i, y_i)$  – нормированные координаты центра масс пикселей  $i$ -го диапазона.



- $c_6 = (0.09, 59, 40)$
- $c_{12} = (0.31, 50, 15)$
- $c_{26} = (0.26, 57, 78)$
- $c_{27} = (0.31, 38, 60)$

## Функция подобия



$$D_{HistSP}(I^{(1)}, I^{(2)}) = \sum_{i=1}^N \underbrace{\left( (D_p(\mathbf{c}_i^{(1)}, \mathbf{c}_i^{(2)}) + \alpha) (D_{xy}(\mathbf{c}_i^{(1)}, \mathbf{c}_i^{(2)}) + \beta) - \alpha\beta \right)}_{D_i}$$

где

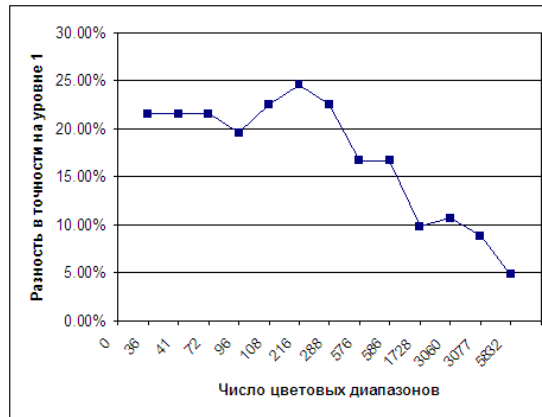
$$D_p(\mathbf{c}_i^{(1)}, \mathbf{c}_i^{(2)}) = |p_i^{(1)} - p_i^{(2)}|,$$

$$D_{xy}(\mathbf{c}_i^{(1)}, \mathbf{c}_i^{(2)}) = \begin{cases} \frac{\sqrt{(x_i^{(1)} - x_i^{(2)})^2 + (y_i^{(1)} - y_i^{(2)})^2}}{\max_{xy}}, & p_i^{(1)} > 0 \wedge p_i^{(2)} > 0 \\ 1, & p_i^{(1)} = 0 \vee p_i^{(2)} = 0 \end{cases}$$

# Эффективность поиска по цветовым гистограммам

- Превосходство признаков HistSP вкпе с функцией  $D_{\text{HistSP}}$  над классическими гистограммами с манхеттенской метрикой

Палитра	Число диапазонов
$nHSI^*(6, 2, 3)$	36
$nHSI^*(6, 2, 3)_{th}$	41
$nHSI^*(12, 2, 3)$	72
$nHSI^*(6, 4, 4)$	96
$nHSI^*(18, 2, 3)$	108
$nHSI^*(16, 6, 6)$	216
$nHSI^*(12, 4, 6)$	288
$nHSI^*(12, 6, 8)$	576
$nHSI^*(12, 6, 8)_{th}$	586
$nHSI^*(12, 12, 12)$	1728
$nHSI^*(17, 12, 15)$	3060
$nHSI^*(17, 12, 15)_{th}$	3077
$nHSI^*(18, 18, 18)$	5832



ImageDB-1000:  
1000 изображений,  
100 запросов

Добавление граничных условий позволило повысить показатели точности и полноты для схем с небольшим числом цветовых диапазонов (выигрыш в точности до 10%).




При равномерном квантовании оптимально использование порядка 500 цветовых диапазонов.

При равномерном квантовании с граничными условиями близкие по точности и полноте результаты при 41 диапазоне.

# Гистограммы или моменты? (1)



Stricker M., Orengo M. Similarity of Color Images. ... (3000 изображений)

index	sim. measure	rank of the image			max. rank
					
9 moments	$W_1$	4	5	8	8
	$W_2$	2	8	6	8
	$W_3$	4	6	9	9
cum. hist.	8/2/2 $L_\infty$	34	98	79	98
	16/4/4 $L_\infty$	3	57	42	57
	8/2/2 $L_1$	53	162	30	162
	16/4/4 $L_1$	33	354	8	354
	8/2/2 $L_2$	65	158	34	158
	16/4/4 $L_2$	15	306	11	306
histogram	8/2/2 $L_1$	138	394	48	394
	16/4/4 $L_1$	4	132	6	132
	8/2/2 $L_2$	71	541	102	541
	16/4/4 $L_2$	10	1358	75	1358



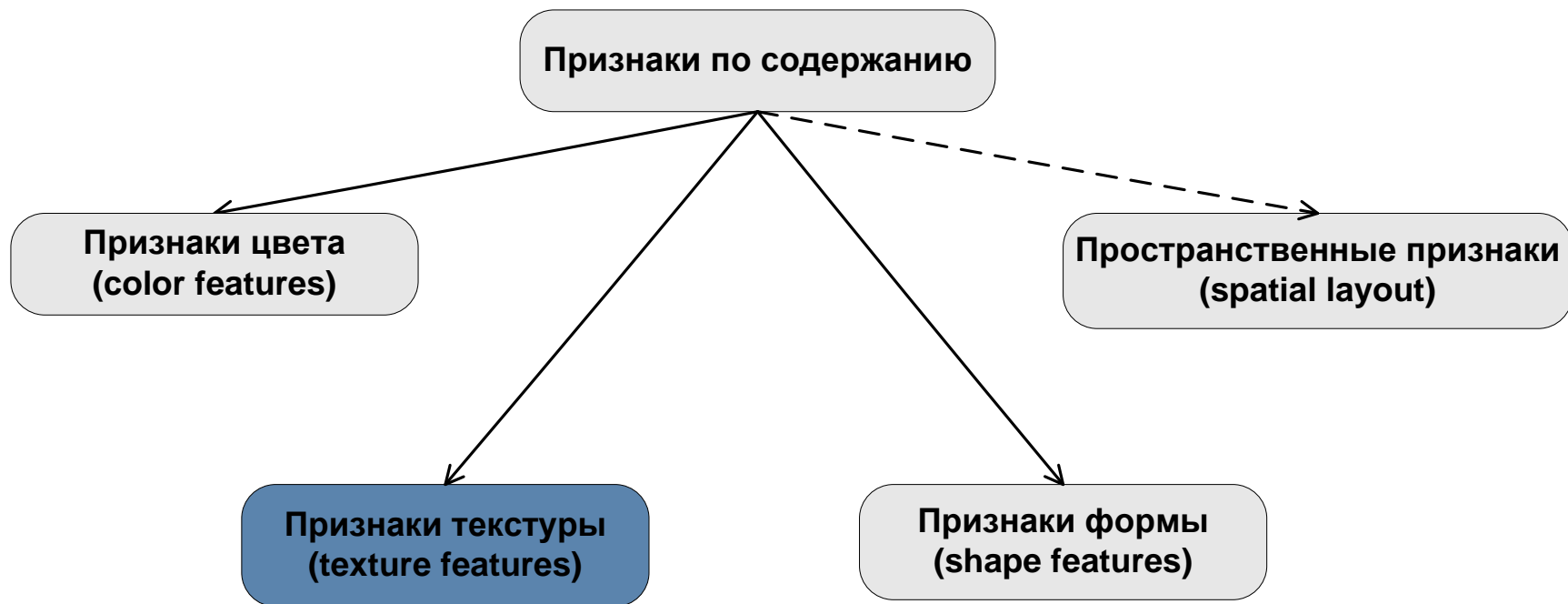
# Гистограммы или моменты? (2)

База Corel Photo Set (285 изображений)

	Полнота	Точность
ColorHist	56,77 %	23,02 %
ColorMoment	55,98 %	25,06 %



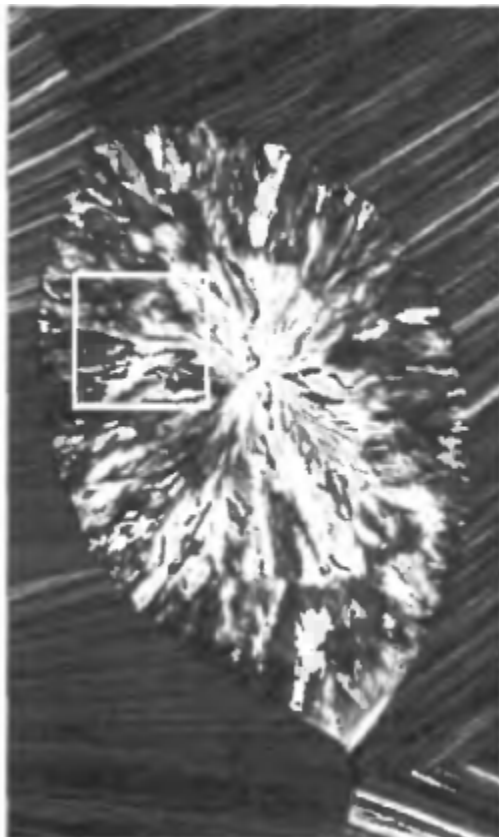




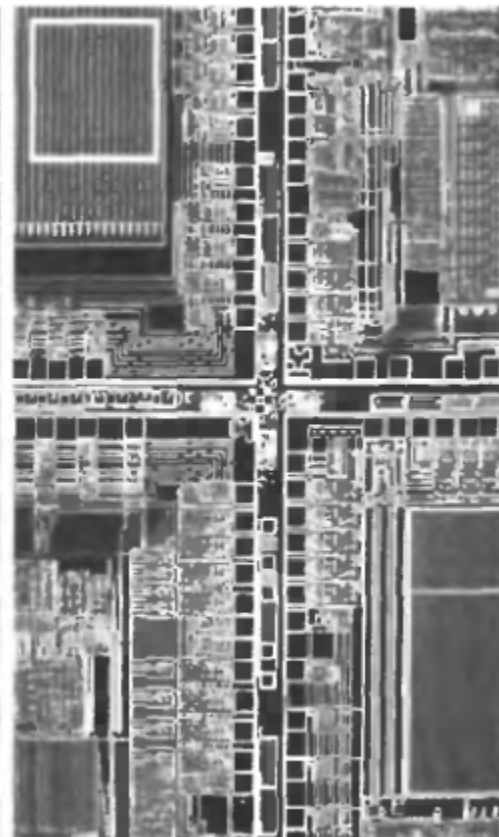
# Текстура



Гладкая

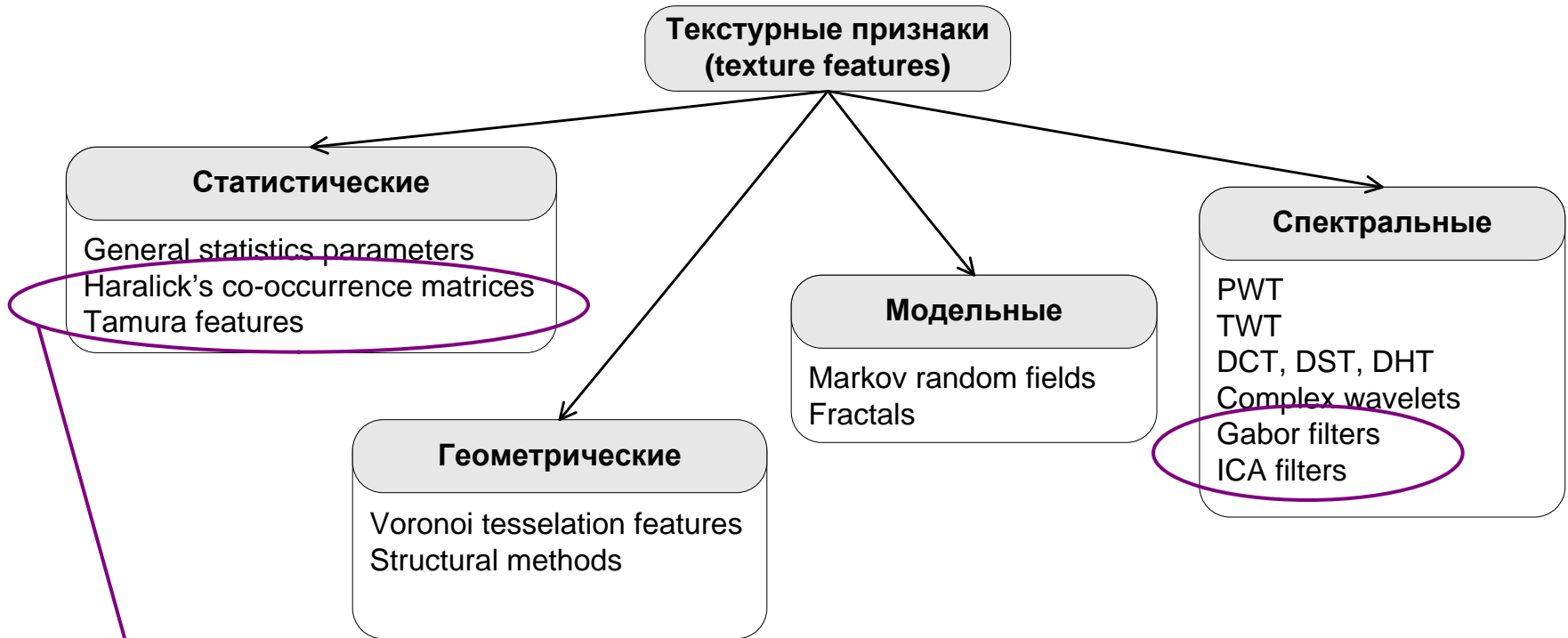


Грубая



Периодическая

# Текстура



- Матрицы смежности – Haralick's co-occurrence matrices
- Признаки Tamura – Tamura features (Tamura image)

# Матрицы смежности

## Grey Level Co-occurrence Matrices (GLCM):

Матрица частот пар пикселей определенной яркости, расположенных на изображении определенным образом относительно друг друга.

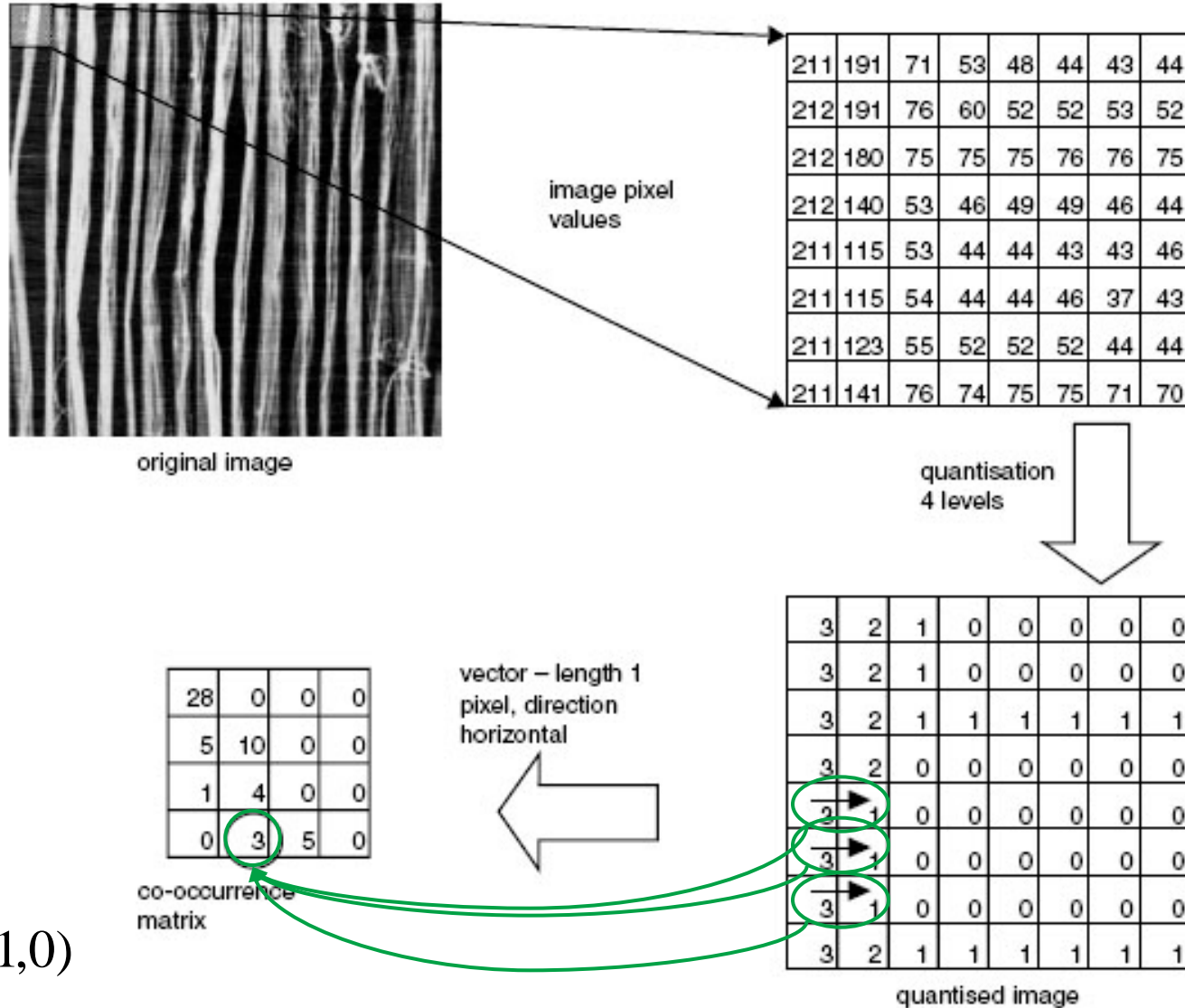
$$C(i, j) = \sum_{p=1}^N \sum_{q=1}^M \begin{cases} 1, & \text{если } I(p, q) = i, I(p + \Delta x, q + \Delta y) = j \\ 0, & \text{иначе} \end{cases}$$

$(\Delta x, \Delta y)$  – параметр сдвига, задающий взаимное расположение пикселей;

$I(p, q)$  – уровень яркости пикселя изображения, расположенного в точке  $(p, q)$ .



# Матрицы смежности: пример



$$(\Delta x, \Delta y) = (1, 0)$$



# Матрицы смежности: характеристики

Статистические параметры, вычисленные по матрицам:

$$Energy = \sum_i \sum_j C^2(i, j) \quad - \text{минимален, когда все элементы равны}$$

$$Entropy = -\sum_i \sum_j C(i, j) \log_2 C(i, j) \quad - \text{мера хаотичности, максимален, когда все элементы равны}$$

$$Contrast = \sum_i \sum_j (i - j)^2 C(i, j) \quad - \text{мал, когда большие элементы вблизи главной диагонали}$$

$$Inverse \text{ Difference Moment} = \sum_i \sum_j \frac{C(i, j)}{1 + (i - j)^2} \quad - \text{мал, когда большие элементы далеки от главной диагонали}$$



# Признаки Tamura

Характеристики, существенные для зрительного восприятия:

- Зернистость (coarseness)
- Контрастность (contrast)
- Направленность (directionality)
- Линейность (line-likeness)
- Регулярность (regularity)
- Грубость (roughness)

Tamura image:

Coarseness-coNtrast-Directionality

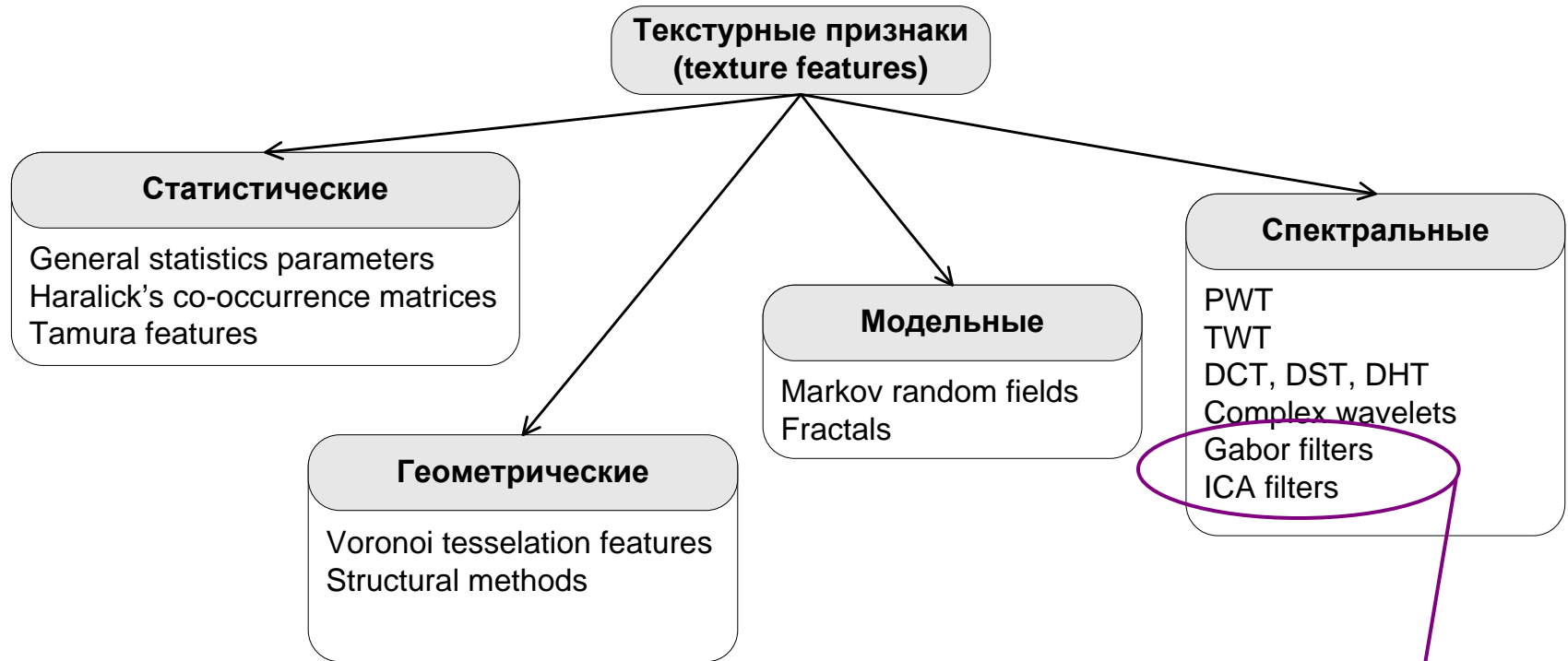
– точки в трехмерном пространстве CND

Признаки:

- Евклидово расстояние в 3D (QBIC)
- 3D гистограмма (Mars)



# Текстура: спектральные



- Вейвлет-признаки, фильтры Габора
- Фильтры ICA



# Вейвлет-признаки

Вейвлет-анализ – разложение сигнала по специальному базису:

$$f(x) = \sum_{j,k} \alpha_k \psi_{j,k}(x)$$

$$\psi_{j,k} = 2^{j/2} \varphi(2^j x - k)$$

$$j, k \in \mathbb{Z}, \quad \varphi(x) \in L^2(\mathbb{R})$$

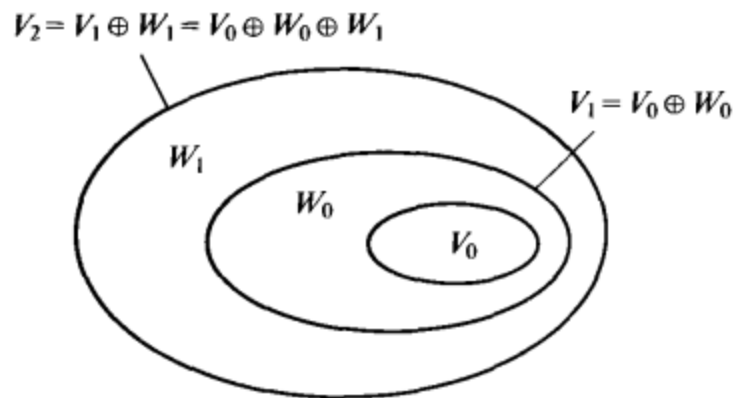
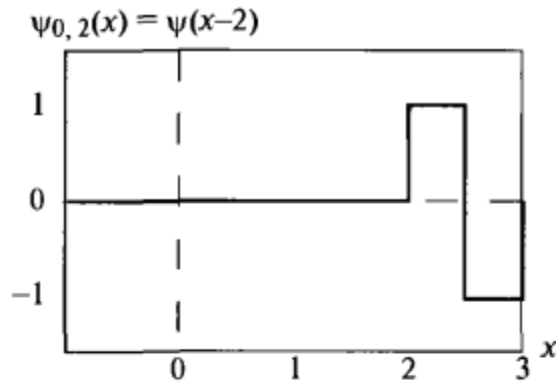
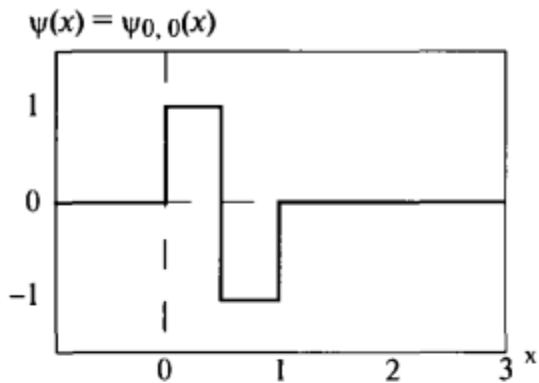
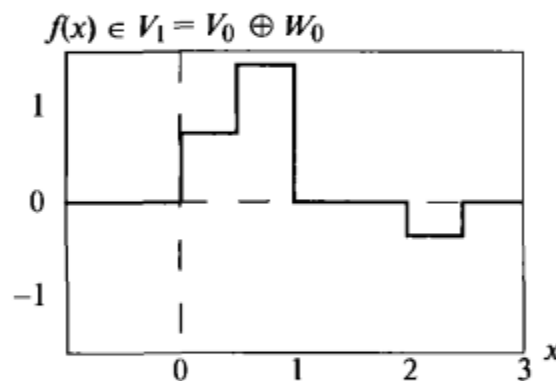
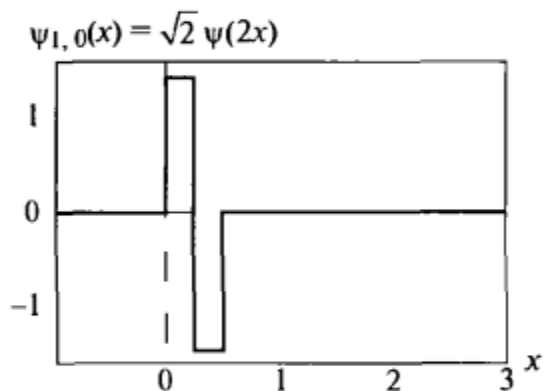


Рис. 7.11. Взаимосвязь функциональных пространств, порождаемых масштабирующей функцией и вейвлет-функцией.



Масштабирующая функция Хаара:

$$\varphi(x) = \begin{cases} 1 & 0 \leq x < 1; \\ 0 & \text{в остальных случаях} \end{cases}$$



Вейвлет-функция Хаара:

$$\psi(x) = \begin{cases} 1, & 0 \leq x < 0,5; \\ -1, & 0,5 \leq x < 1; \\ 0 & \text{в остальных случаях.} \end{cases}$$

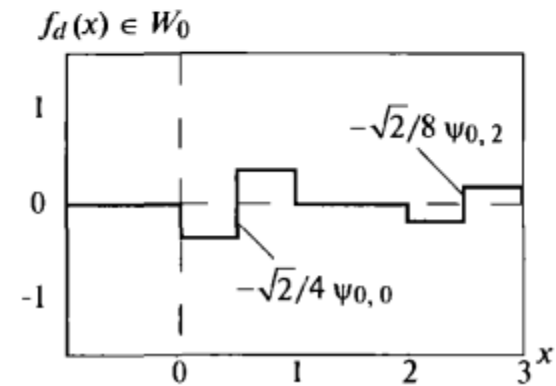
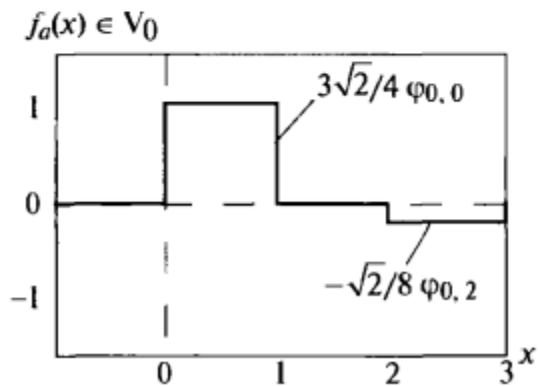
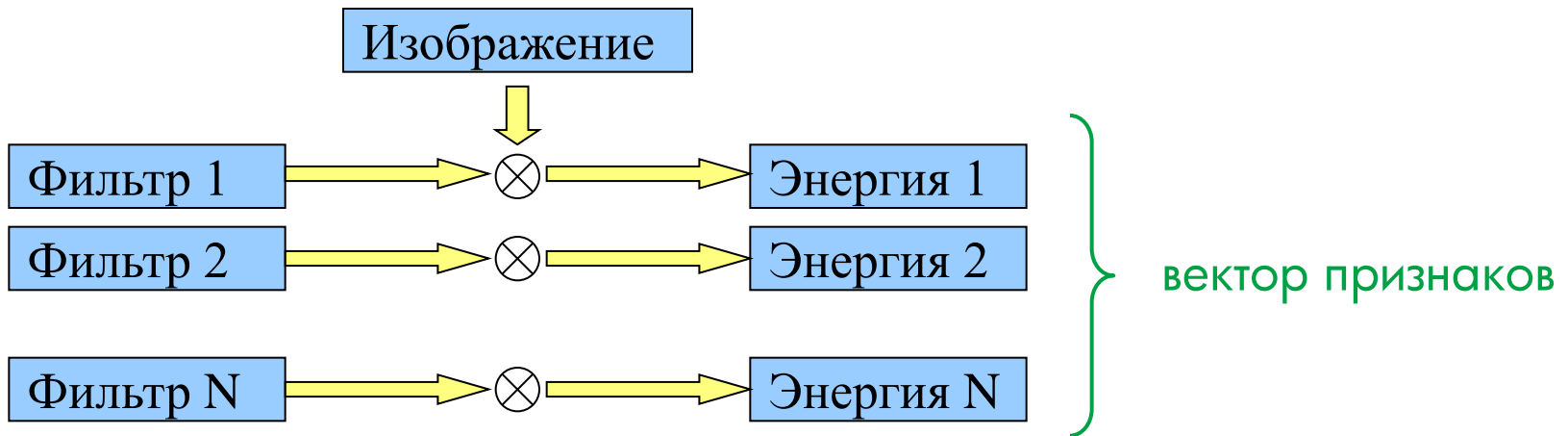


Рис. 7.12. Вейвлеты Хаара в подпространствах  $W_1$  и  $W_2$ .



# Вейвлет-признаки

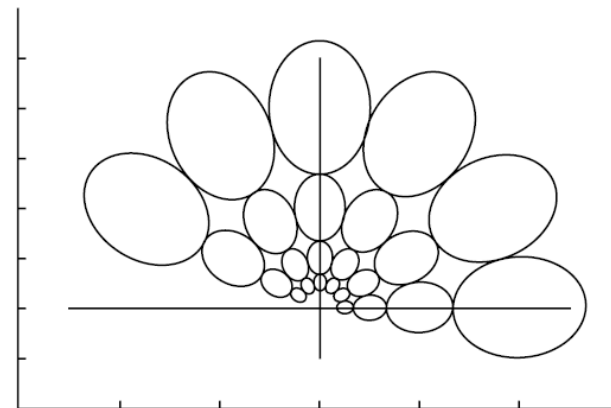
Набор базисных функций – банк фильтров



# Фильтры Габора

Масштабирующая функция: функция Габора

$$g(x, y) = \left( \frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi j W x \right]$$



Набор фильтров:

$$g_{mn}(x, y) = a^{-m} g(x', y'), \quad a > 1, \quad m, n = \text{integer}, \quad m = 0, 1, \dots, S-1,$$

$$x' = a^{-m} (x \cos \Theta + y \sin \Theta),$$

$$y' = a^{-m} (-x \sin \Theta + y \cos \Theta),$$

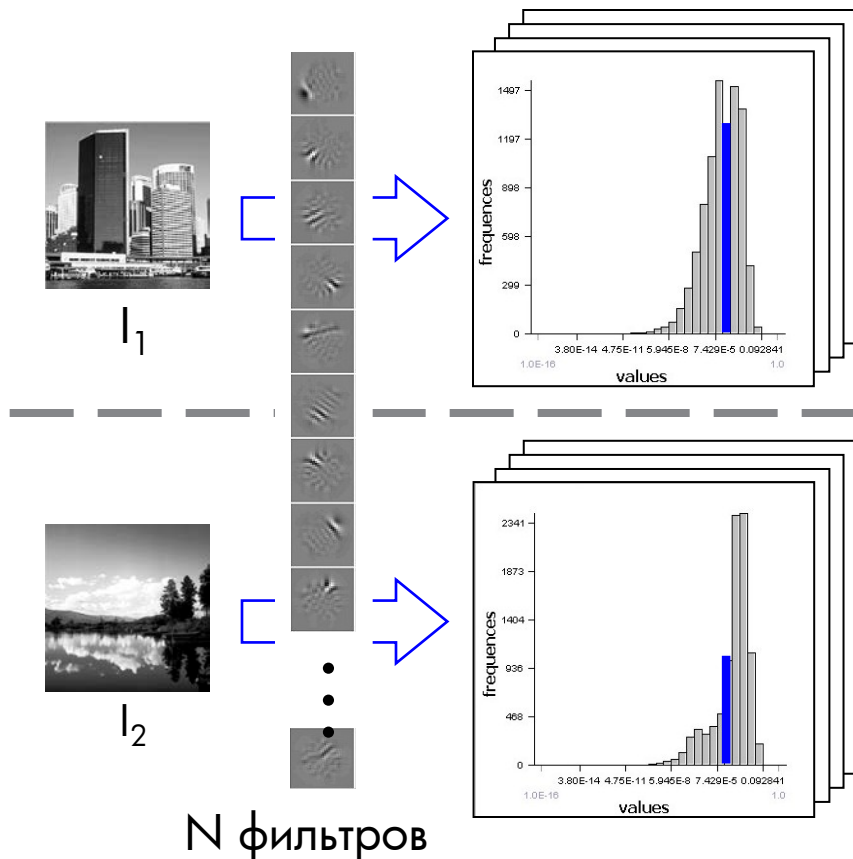
$$\Theta = n\pi / K$$

$$a = (U_h / U_l)^{-1/(S-1)}$$

$K$  – общее число направлений,  
 $S$  – число масштабов,  
 $U_h, U_l$  – максимум и минимум рассматриваемых частот.

# Фильтры ИСА

Фильтры получены при помощи анализа независимых компонент



$$KL_H(H_1, H_2) = \sum_{b=1}^B (H_1(b) - H_2(b)) \log \frac{H_1(b)}{H_2(b)}$$

$$\text{dist}(I_1, I_2) = \sum_{i=1}^N KL_H(H_{1i}, H_{2i})$$

H. Borgne, A. Guerin-Dugue, A. Antoniadis. Representation of images for classification with independent features. Pattern Recognition Letters, vol. 25, p. 141-154, 2004

# Сравнение текстурных признаков

**Table 6: TRECVID evaluation – mean average precision retrieval**

Feature	Single	Combined with HSV
Gabor-2-4	3.93%	4.31%
Co-occurrence homogeneity	2.85%	3.03%
Tamura standard all	2.57%	3.43%
Tamura CND	1.65%	2.72%
Tamura coarseness-2	0.97%	2.49%

В контексте задачи поиска

**Table 7: ImageCLEF retrieval results**

Feature	Mean average precision
Gabor-2-4	35.3%
Co-occurrence homogeneity	19.8%
Tamura standard all	20.7%
Tamura CND	18.4%
Tamura coarseness-2	14.5%

P. Howarth, S. Rüger. Robust texture features for still image retrieval.  
In Proc. IEE Vis. Image Signal Processing, vol. 152, No. 6, December 2006



# Сравнение текстурных признаков (2)

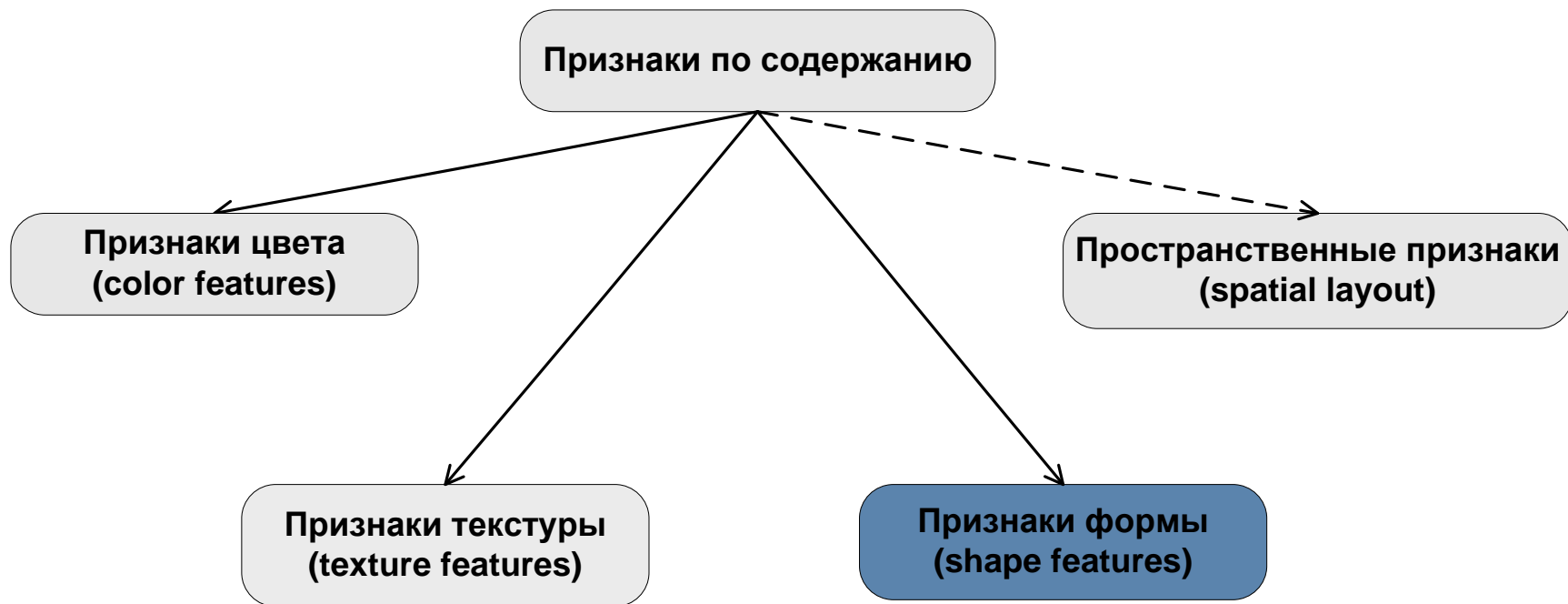
## Фильтры Габора v. s. фильтры ICA

Эксперименты по классификации изображений:

- Коллекция ангиографических снимков
  - Фильтры ICA лучше на 13%
- Коллекция текстур Brodatz
  - Фильтры ICA лучше на 4%

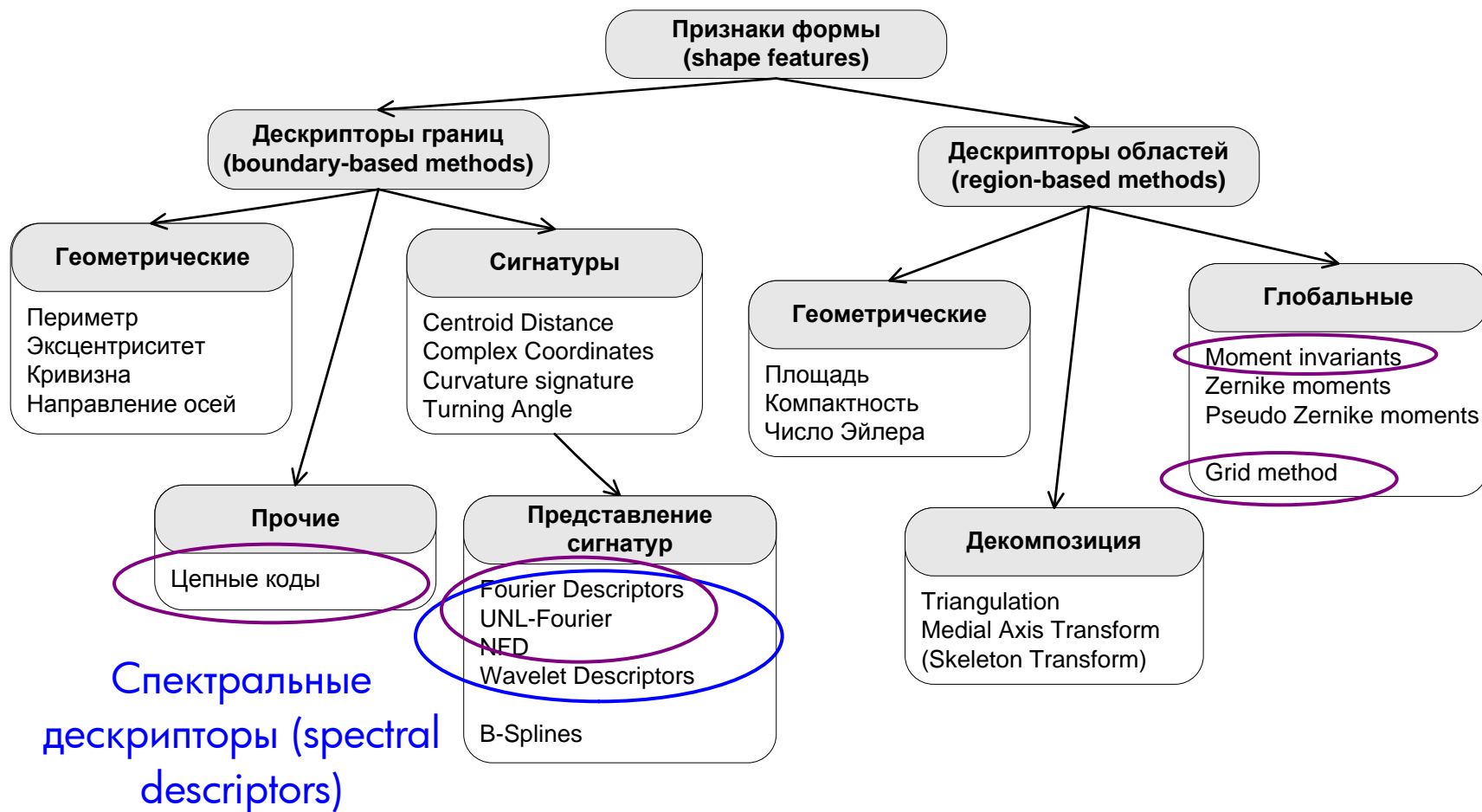
Snitkowska, E. Kasprzak, W. Independent Component Analysis of Textures in Angiography Images. Computational Imaging and Vision, vol. 32, pages 367-372, 2006.







# Форма объектов

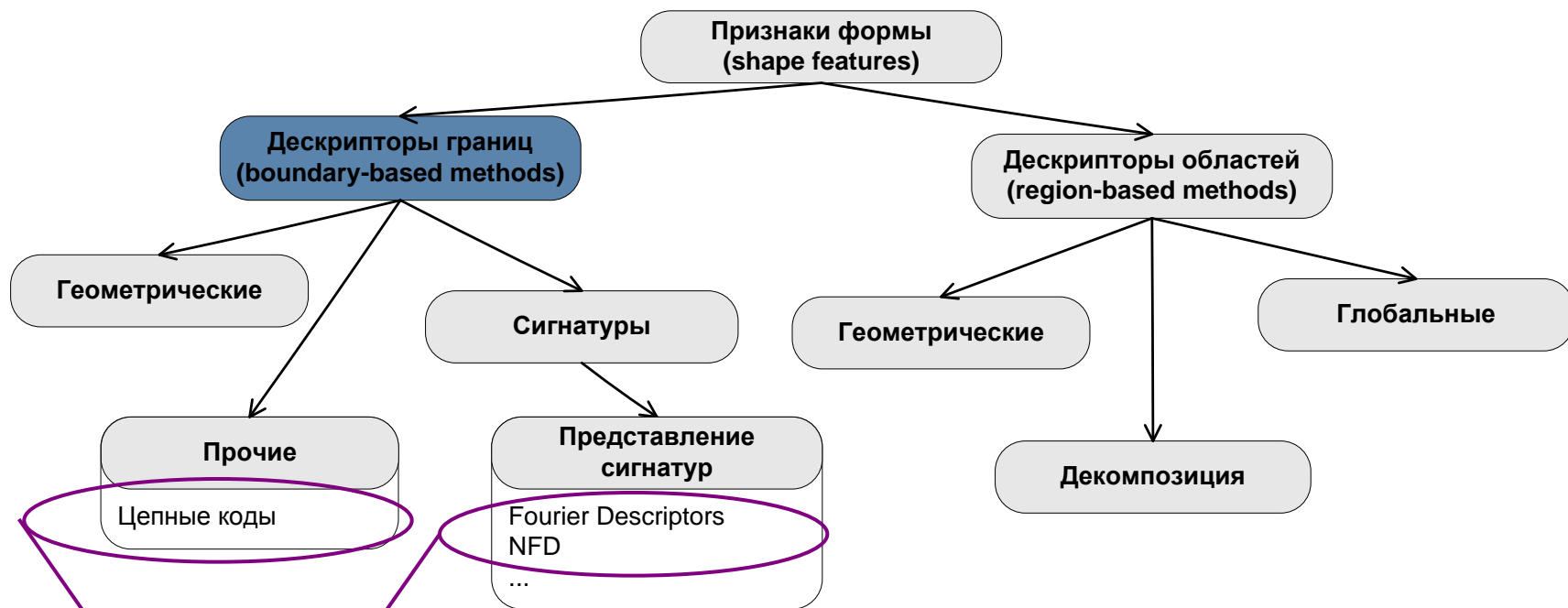


# Требования к признакам формы

- Инвариантность к параллельному переносу
- Инвариантность к изменению масштаба
- Инвариантность к повороту
- Устойчивость к незначительным изменениям формы
- Простота вычисления
- Простота сравнения



# Форма объектов: границы



- Цепные коды (Chain Codes)
- Дескрипторы Фурье (Fourier Descriptors)



# Дескрипторы Фурье

## 1. Вычисление сигнатуры (2D -> 1D):

- Расстояние от центра от центра до границы
- Комплексные координаты:  $z(t) = x(t) + iy(t)$
- ...

## 2. Вычисление коэффициентов Фурье ( $s(t)$ – сигнатура):

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t) e^{-j2\pi nt / N}$$

## 3. Нормализация (NFD – Normalized Fourier Descriptors):

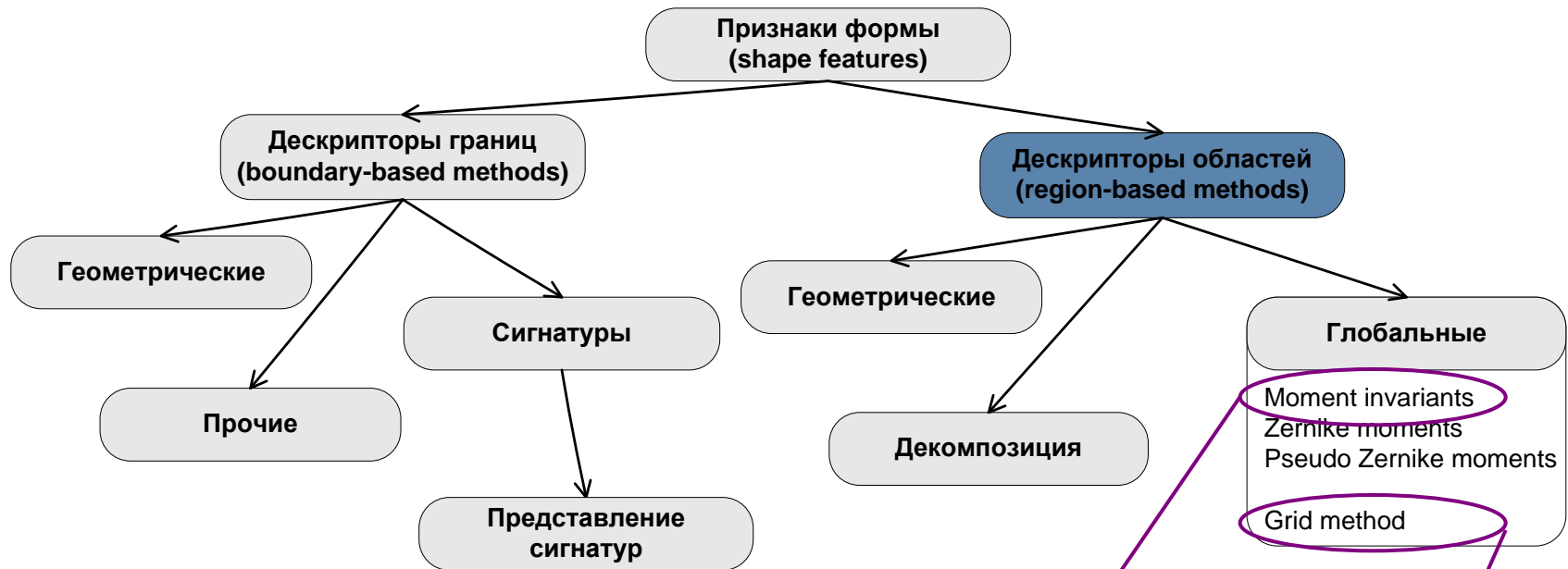
$$\frac{|u_1|}{|u_0|}, \frac{|u_2|}{|u_0|}, \dots, \frac{|u_{N-1}|}{|u_0|}$$

## 4. Сравнение:

$$d = \left( \sum_{n=0}^{N_c} |f_I^n - f_J^n|^2 \right)^{\frac{1}{2}}$$

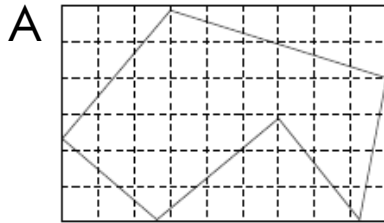


# Форма объектов: области

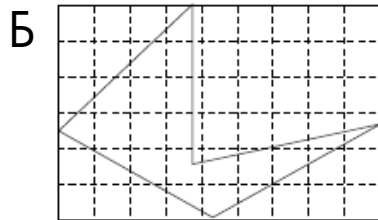


- Грид-метод (Grid-method)
- Инвариантные моменты (Moment invariants)

# Грид-метод



A: 001111000 011111111 111111111 111111111 111110111 0111000011

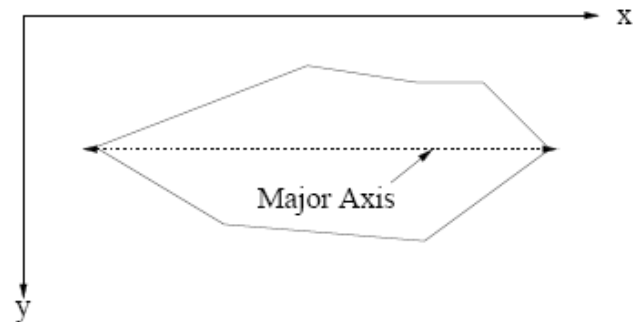


Б: 001100000 011100000 111100000 111101111 111111110 001111000

## Инвариантность:

Нормализация по главной оси:

- направление;
- размер;
- позиционирование на гриде.



# Инвариантные моменты

Момент порядка (р+q) двумерной непрерывной функций:

$$m_{pq} = \iint x^p y^q f(x, y) dx dy$$

Центральные моменты для  $f(x, y)$  – дискретного изображения:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y), \quad \bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

Вектор признаков:

С использованием нормированных центральных моментов был выведен набор из 7 инвариантных к параллельному переносу, повороту и изменению масштаба моментов.





# Сравнение признаков формы

Table 1. Average retrieval efficiency values for different methods.  $T$  is the short list size of retrieved images

Methods	$T=5$	$T=10$	$T=15$	$T=20$
Reduced chain code	55.1%	47.6%	50.0%	60.6%
Fourier descriptors (FD)	72.2%	76.9%	75.9%	74.9%
UNL features	81.3%	79.9%	83.7%	89.3%
Moment invariants (MI)	84.7%	86.3%	86.8%	87.7%
Zernike moments	66.9%	66.5%	70.4%	78.2%
Pseudo-Zernike moments	66.9%	66.5%	70.4%	78.2%
MI and FD	93.8%	87.3%	87.1%	89.6%
MI and UNL	93.3%	89.2%	89.3%	91.1%

Mehetre B. M., Kankanhalli M. S., Lee W. F. Shape measures for content based image retrieval: a comparison. *Inf. Processing and Management*, vol. 33, No. 3, pages 319-337, 1997.



# Заключение

- Большой выбор различных способов представления изображений
  - Цвет: гистограммы или статистическая модель?
  - Текстура: фильтры Габора, фильтры ICA
  - Форма: дескрипторы Фурье, инвариантные моменты
- При сравнении изображений часто необходимо комбинировать различные признаки

