

# Методы поиска изображений по содержанию

---

Наталья Васильева  
HP Labs, Russia;  
СПбГУ  
[nvassilieva@hp.com](mailto:nvassilieva@hp.com)

29 ноября 2007

# План

---

## I. Обзор методов поиска изображений

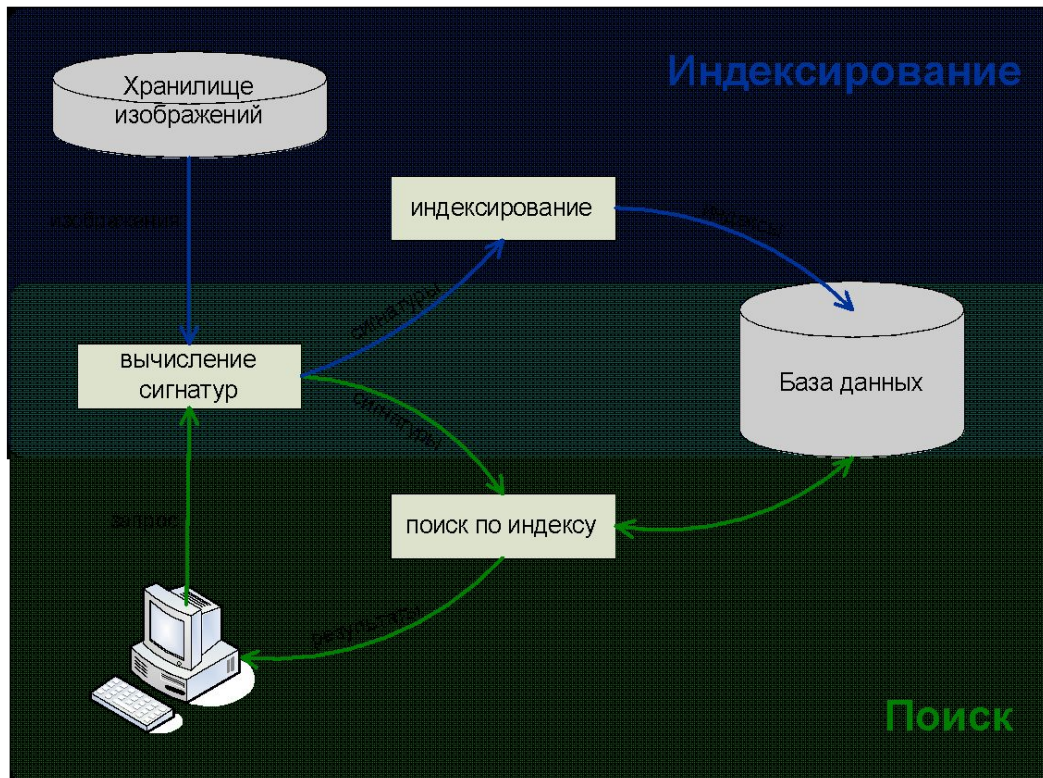
1. Основные направления исследований
2. Уровни содержания изображения
3. Цвет
4. Текстура
5. Форма объектов

## II. Синтез данных в контексте CBIR

1. Существующие решения и их недостатки
2. Альтернативные подходы
3. Поиск в частично-аннотированной базе
4. WTGF: **W**eighted **T**otal with **G**ravitation **F**unction
5. Адаптивный поиск

# CBIR: направления исследований

## Поиск по содержанию – Content Based Image Retrieval (CBIR)



Традиционная архитектура систем CBIR

- Выделение признаков изображений
- Многомерное индексирование
- Проектирование систем поиска

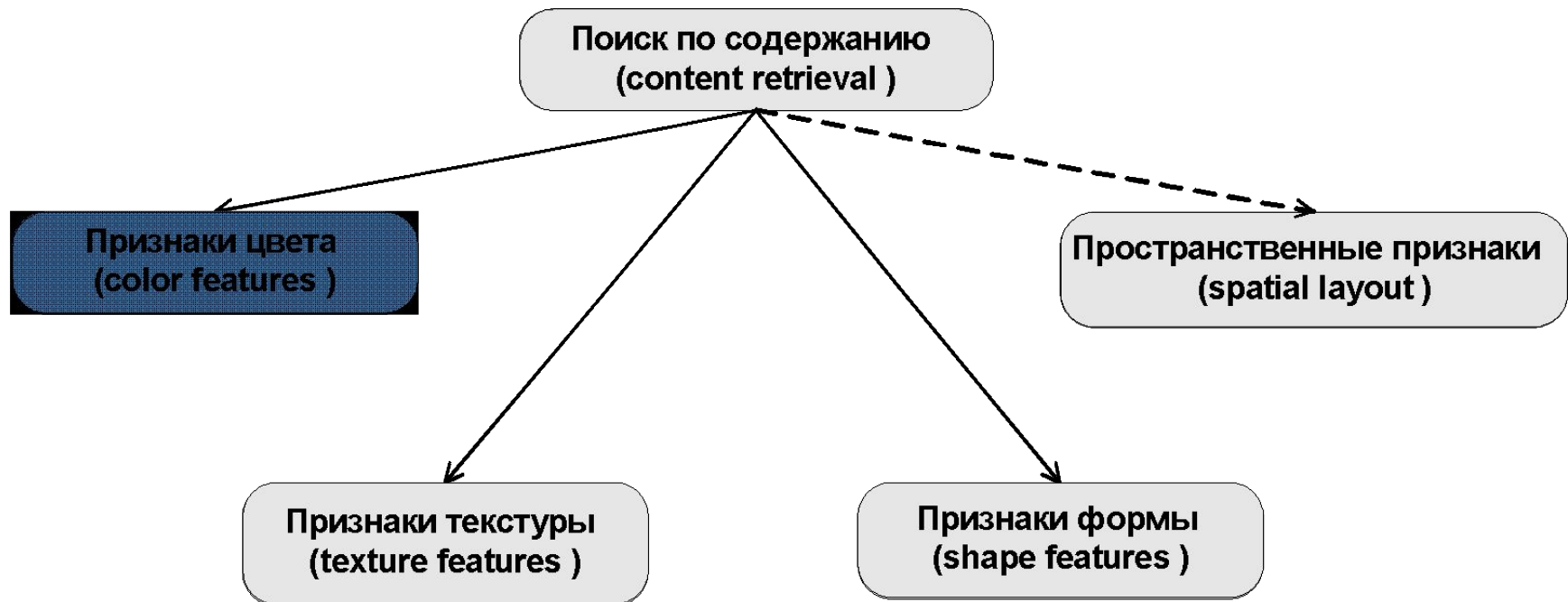
# Уровни содержания изображения

---

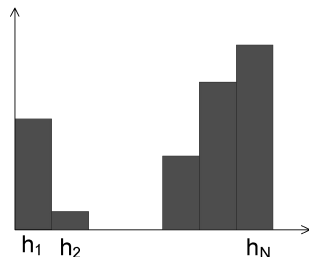


# Поиск по содержанию

---



# Цвет



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

Метрики:  $L_1, L_2, L_\infty$

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

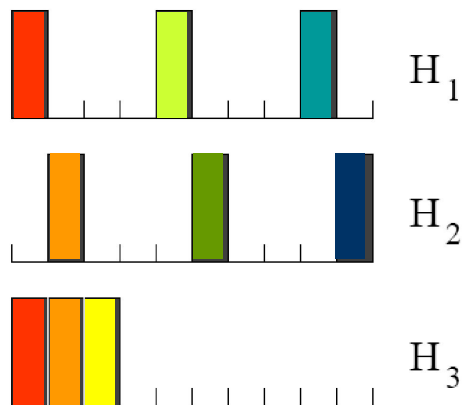
$$F(I) = (E_1^I, E_2^I, E_3^I, \\ \sigma_1^I, \sigma_2^I, \sigma_3^I, \\ s_1^I, s_2^I, s_3^I)$$

Метрики:  $\sim L_1$

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

# Цветовые гистограммы – недостатки

## 1. Не учитывается схожесть цветов:



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

- Кумулятивные гистограммы

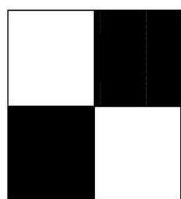
- $$d(H_1, H_2) = \sqrt{(H_1 - H_2) \cdot A \cdot (H_1 - H_2)^T}$$

A – матрица с коэффициентами  
«схожести» 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

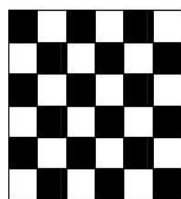
2. Не учитывается пространственное расположение цветов:



A



B

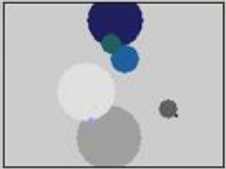



C

$$H_A = H_B = H_C$$



img239



```
{160, 160, 160, 0.2425520833333333,77,99};  
{224, 224, 224, 0.2184895833333333,60,65};  
{32, 32, 96, 0.2080208333333333,81,13};  
{32, 96, 160, 0.1088541666666666,78,41};  
{32, 96, 96, 0.0799479166666666,78,30};  
{96, 96, 96, 0.0706770833333333,120,78};  
{160, 160, 224, 0.0302604166666666,64,87};  
{32, 32, 32, 0.0139583333333333,126,83};  
{160, 224, 224, 0.0107291666666666,67,77};
```

$$f_i^A = (a_i, b_i, c_i, weight_i^A, x_i^A, y_i^A)$$

$i = 1..N$  – число цветов;

$(a_i, b_i, c_i)$  – параметры цвета  $i$ ;

$weight_i$  – количество цвета  $i$  на изображении  $A$ ;

$(x_i, y_i)$  – координаты центра цветового пятна.

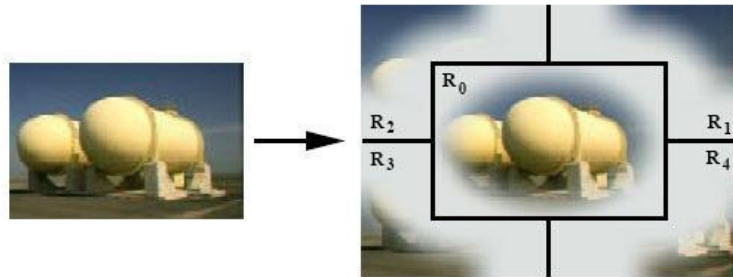
Васильева Н., Новиков Б. Построение соответствий между низкоуровневыми характеристиками и семантикой статических изображений. Труды RCDL'2005.



# Пространственное расположение цветов

---

- Разбиение изображения на фиксированные блоки
- «Нечеткие области»

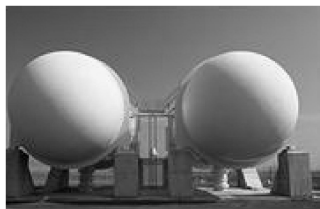


Stricker M., Dimai A. Spectral Covariance and Fuzzy Regions for Image Indexing. Machine Vision and Applications, vol. 10., p. 66-73, 1997




- Сегментация



# 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)

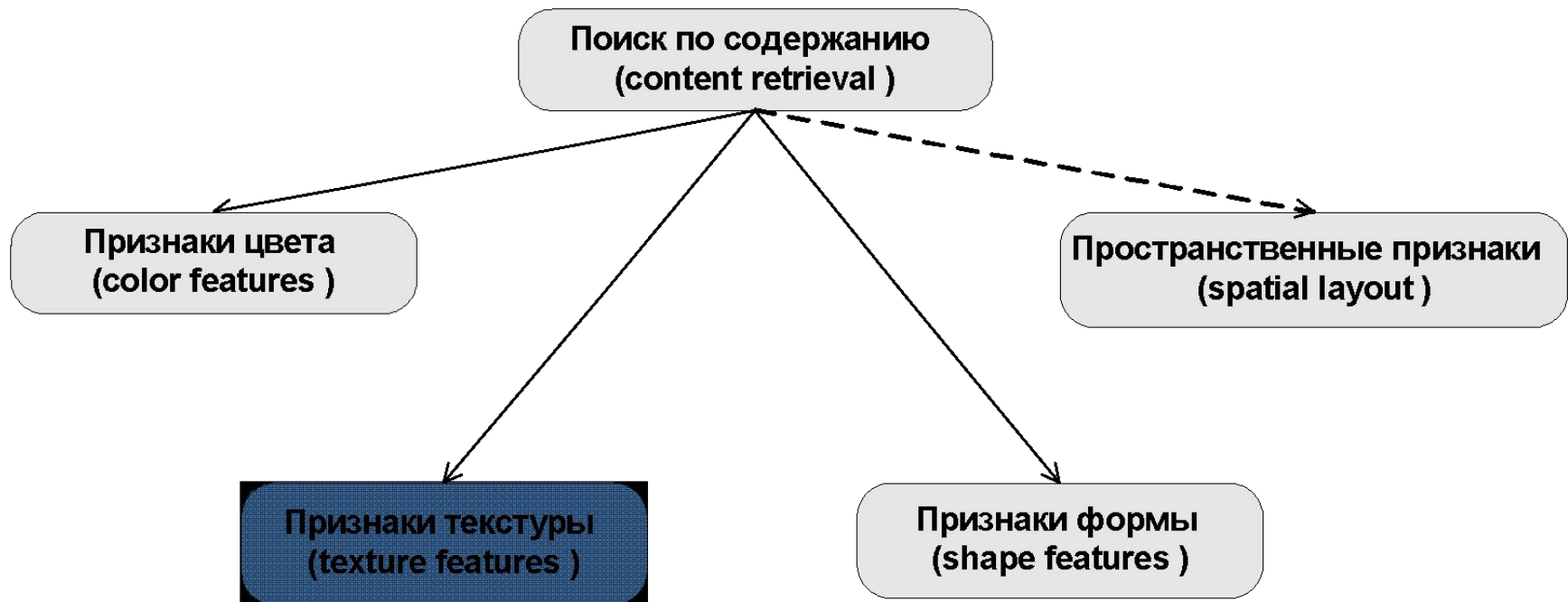
---

База Corel Photo Set (285 изображений)  
эксперимент в рамках дипломной работы М. Теплых

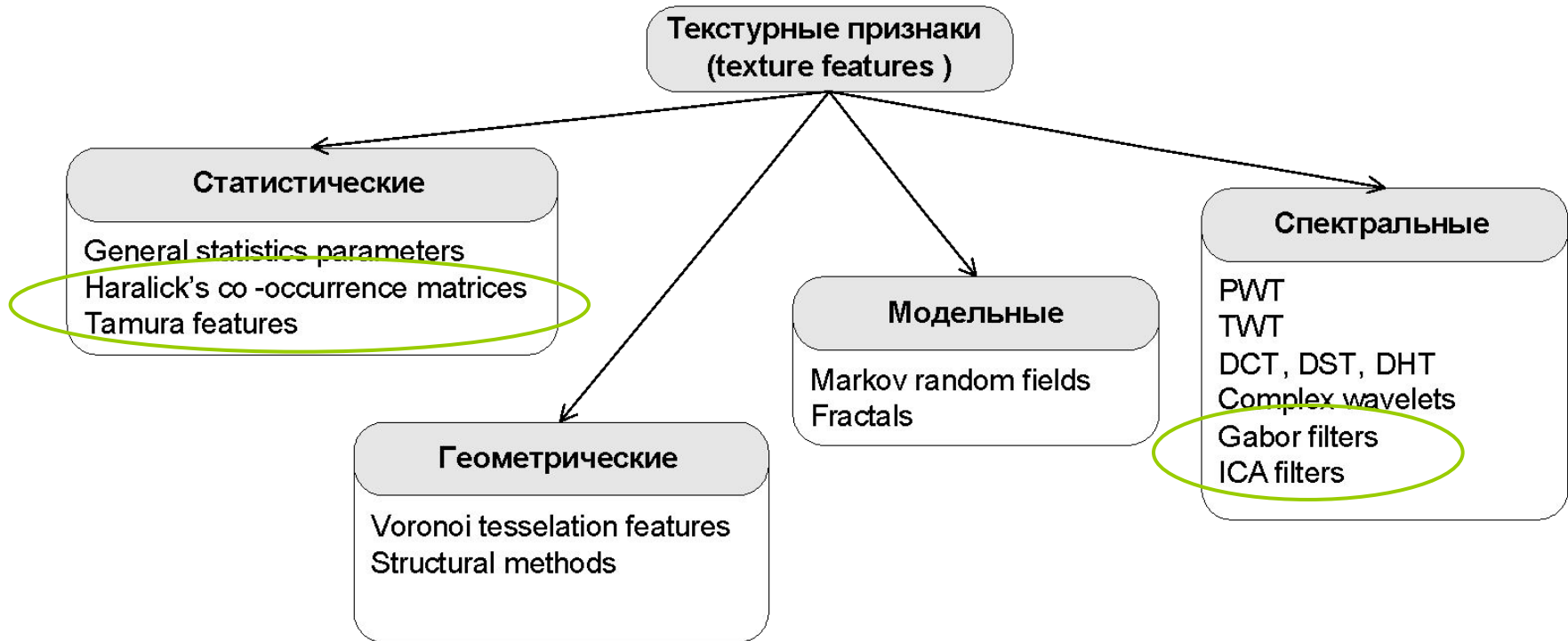
	Полнота	Точность
ColorHist	56,77 %	23,02 %
ColorMoment	55,98 %	25,06 %

# Поиск по содержанию: текстура

---



# Текстура: статистические



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

# Матрицы смежности

---

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

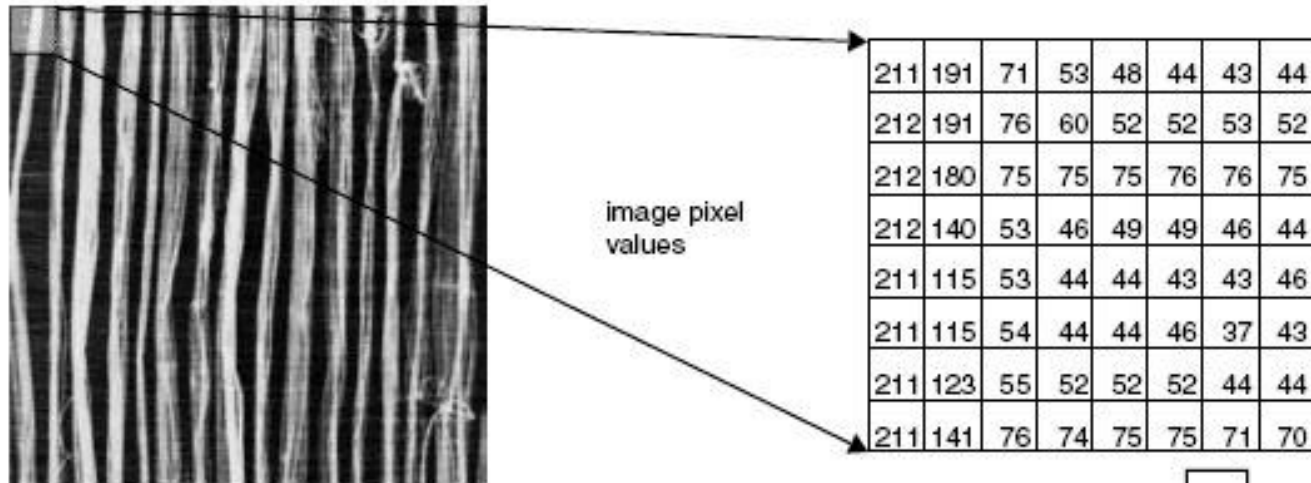
Матрица частот пар пикселей определенной яркости, расположенных на изображении определенным образом относительно друг друга.

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

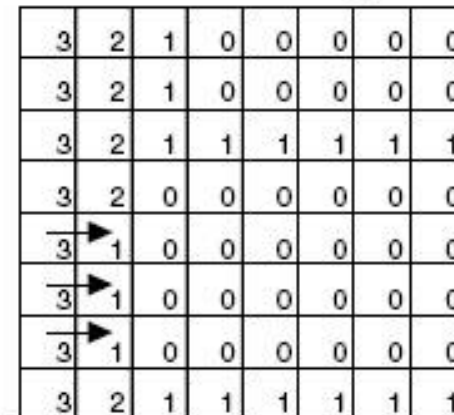
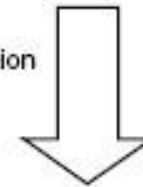
$(\Delta x, \Delta y)$  – параметр сдвига, задающий взаимное расположение пикселей;

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

# Матрицы смежности: пример



quantisation  
4 levels

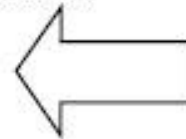


quantised image

28	0	0	0
5	10	0	0
1	4	0	0
0	3	5	0

co-occurrence  
matrix

vector – length 1  
pixel, direction  
horizontal



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

# 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)

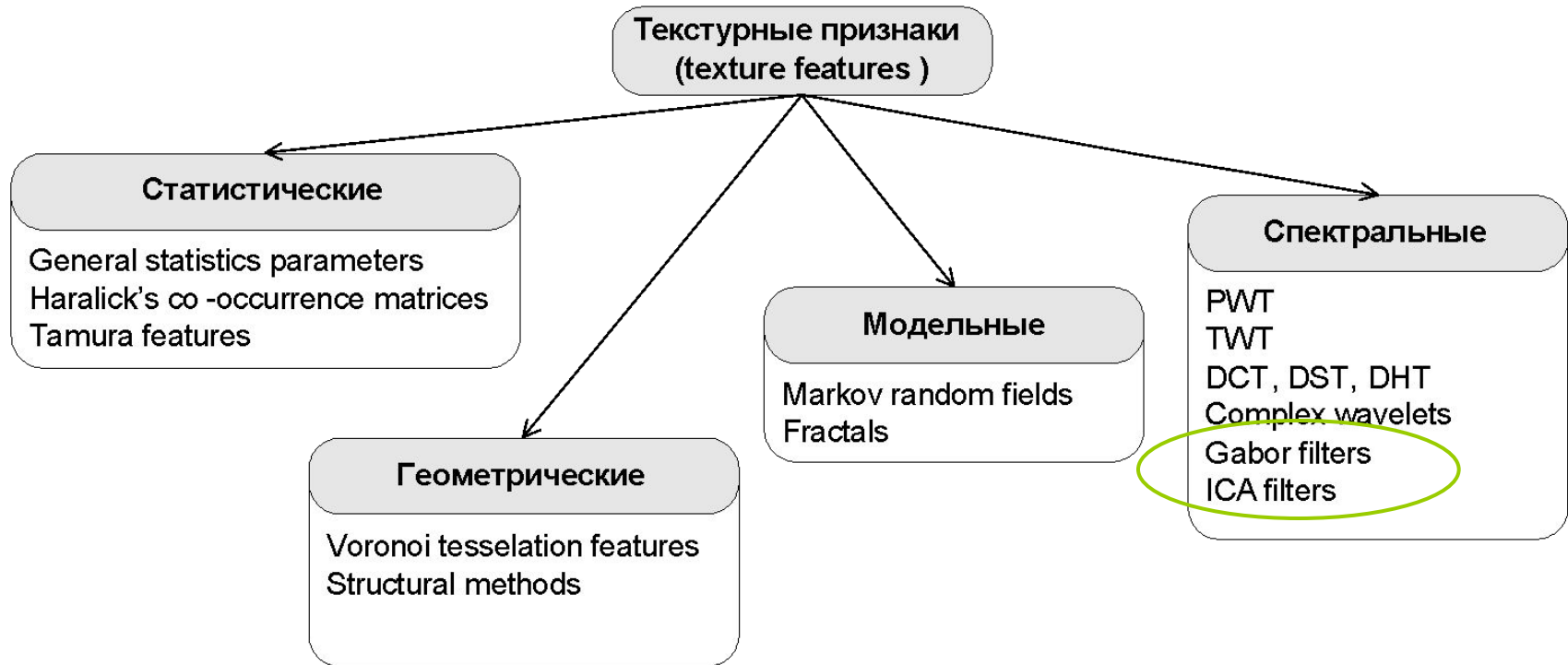
Tamura image:

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:

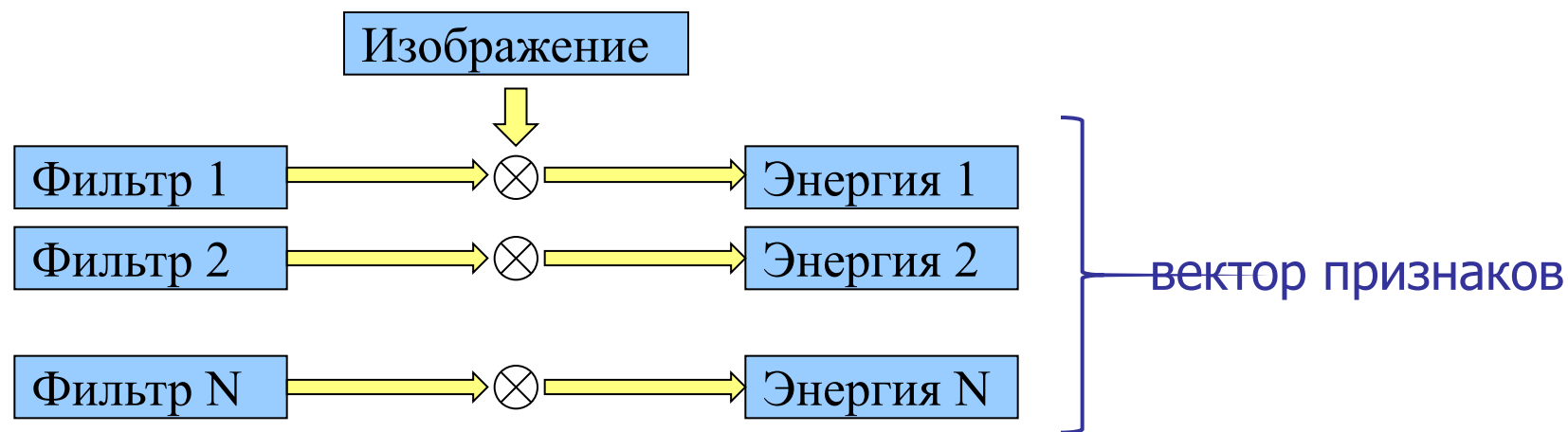
$$f(x) = \sum_{j,k} \alpha_k \psi_{j,k}(x)$$

Базисные функции:

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

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

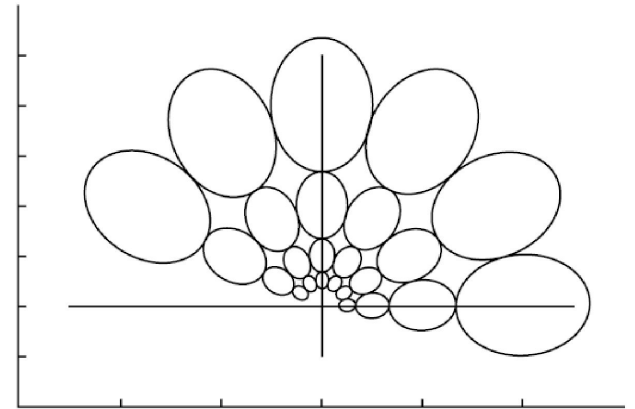
Набор базисных функций – банк фильтров



# Фильтры Габора

Порождающий вейвлет: 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:

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

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

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

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

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

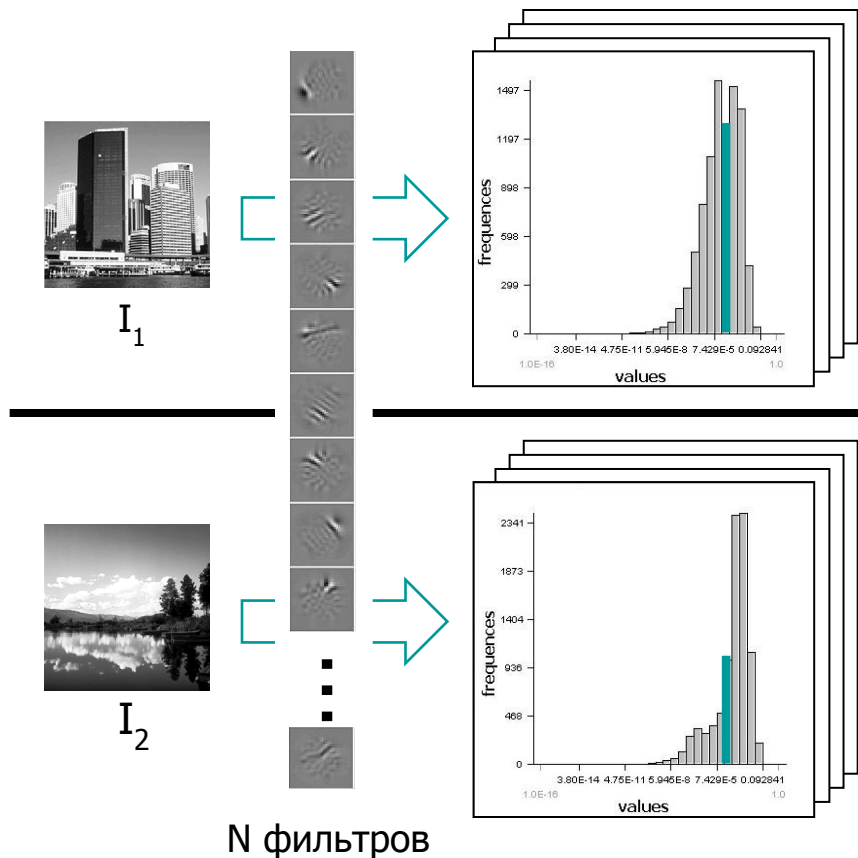
$K$  – общее число направлений,

$S$  – число масштабов,

$U_h, U_l$  – максимум и минимум рассматриваемых частот.

# Фильтры ИСА

Фильтры получены при помощи анализа независимых компонент



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

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

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

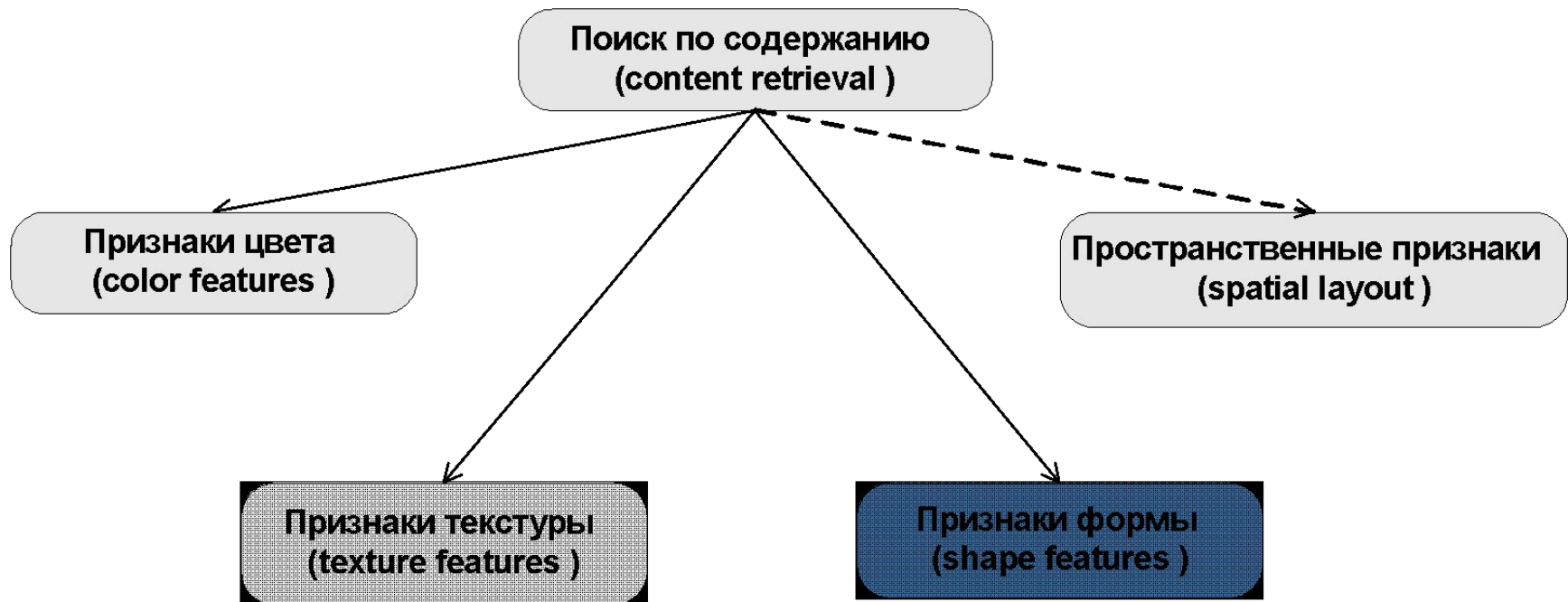
Эксперименты по классификации изображений:

- 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.

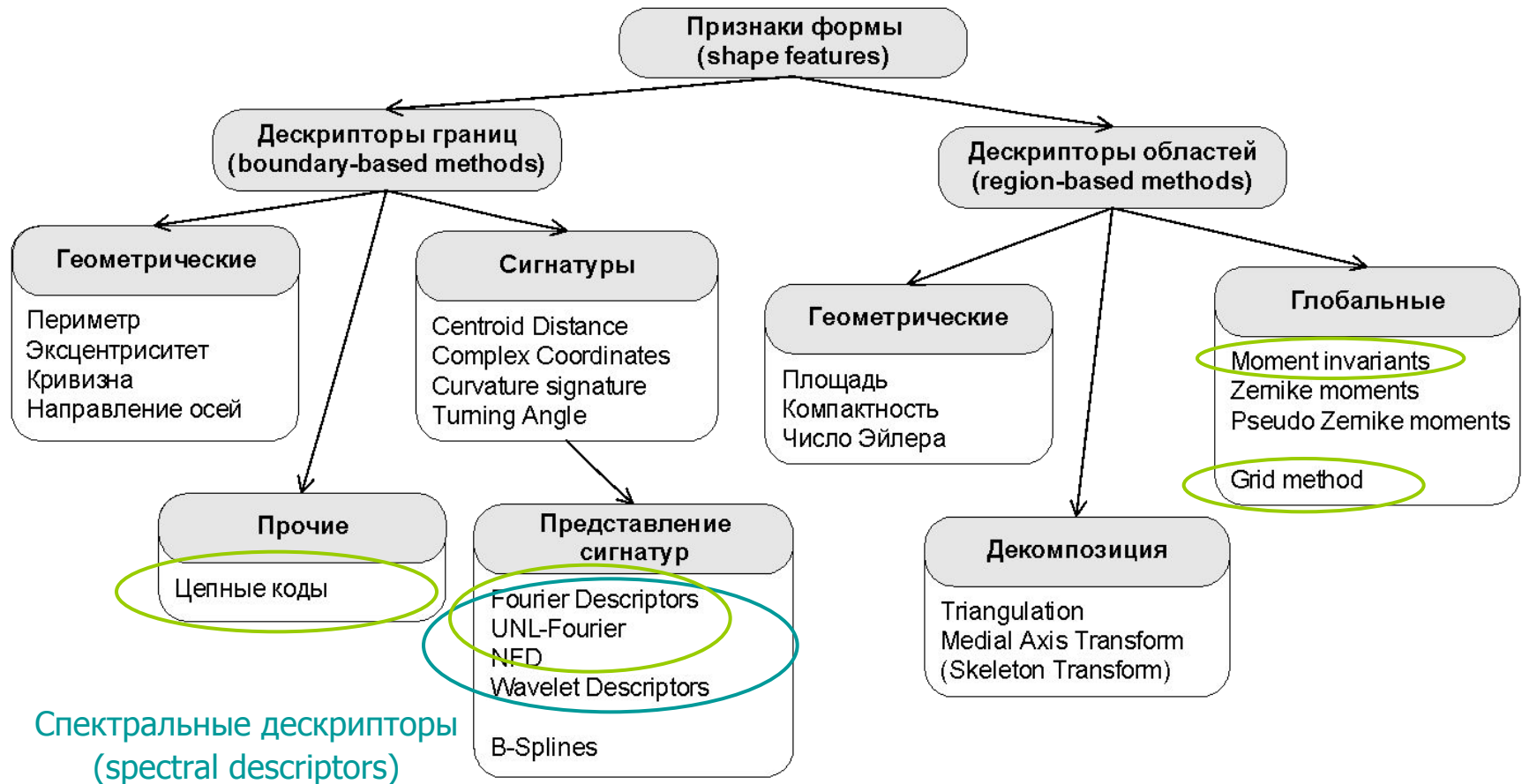
# Поиск по содержанию: форма

---





# Форма объектов

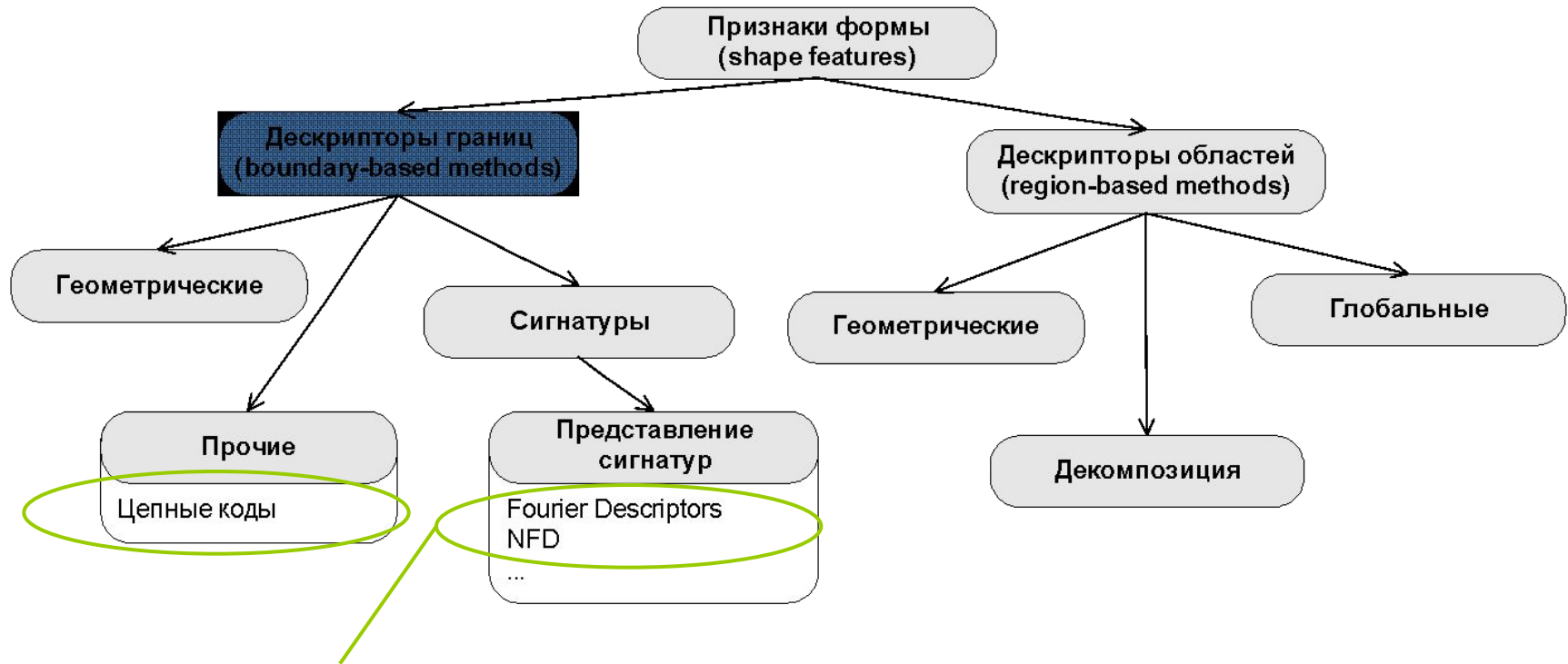


# Требования к признакам формы

---

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

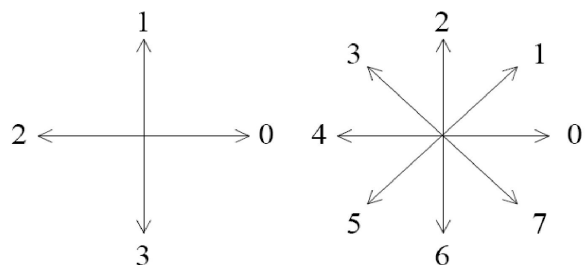
# Форма объектов: границы



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

# Цепные коды

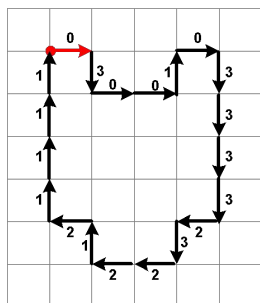
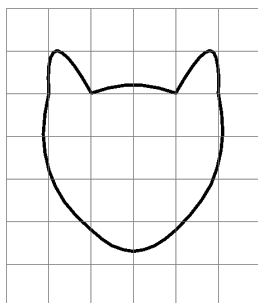
Нумерация направлений для 4-связного и 8-связного цепных кодов:



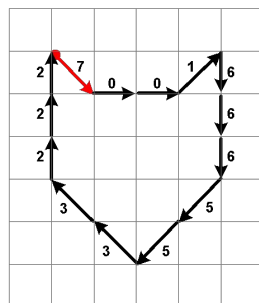
А: 03001033332322121111

Б: 70016665533222

Пример:



А



Б



Инвариантность к выбору начальной точки: минимальный код



Инвариантность к повороту: разности цифр кода

# Дескрипторы Фурье

---

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

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

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

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

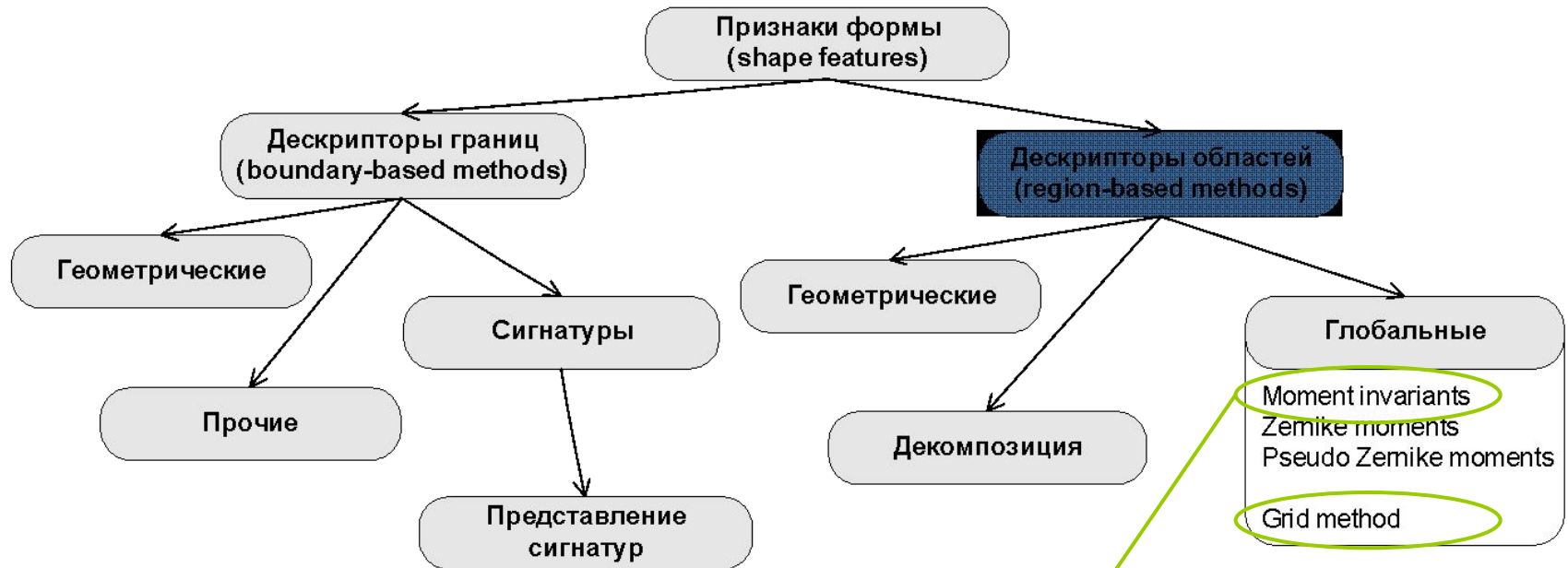
## 3. Нормализация (NFD – Normalized Fourier Descriptors):

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

## 4. Сравнение:

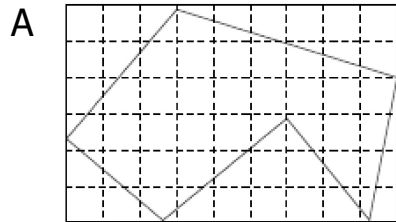
$$d = \left( \sum_{n=0}^{N_c} |f_I^n - f_J^n|^2 \right)^{\frac{1}{2}}$$

# Форма объектов: области

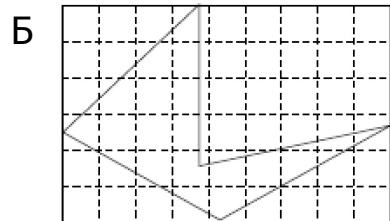


- Грид-метод (Grid-method)
- Инвариантные моменты (Moment invariants)

# Грид-метод



A: 001111000 011111111 111111111 111111111 111110111 0111000011

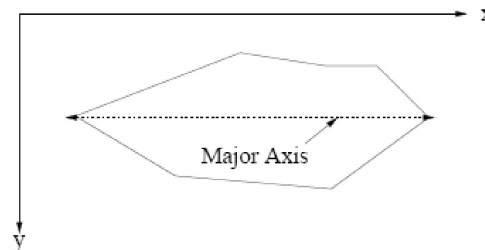


Б: 001100000 011100000 111100000 111101111 111111110 001111000

## Инвариантность:

Нормализация по главной оси:

- направление;
- размер;
- позиционирование на гриде.



# Инвариантные моменты

---

Момент порядка (p+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.

# Признаки в системах поиска

	Цвет	Текстура	Форма
QBIC	Гистограммы (HSV) $dist^2 = H_1 A H_2^T$	Tamura Image, Euclid dist	Геометрические для границ + моменты
VisualSEEk	Гистограммы (HSV), Color Sets, Location info		
Netra	Гистограммы (HSV), Color codebook, кластеризация	Фильтры Габора	Fourier-based (Фурье)
Mars	Гистограммы, HSV $dist = 1 - \sum_{i=1}^N \min(H_1(i), H_2(i))$	Tamura Image, 3D Histo	MFD (Фурье)

# План

---

## I. Обзор методов поиска изображений

1. Основные направления исследований
2. Уровни содержания изображения
3. Цвет
4. Текстура
5. Форма объектов

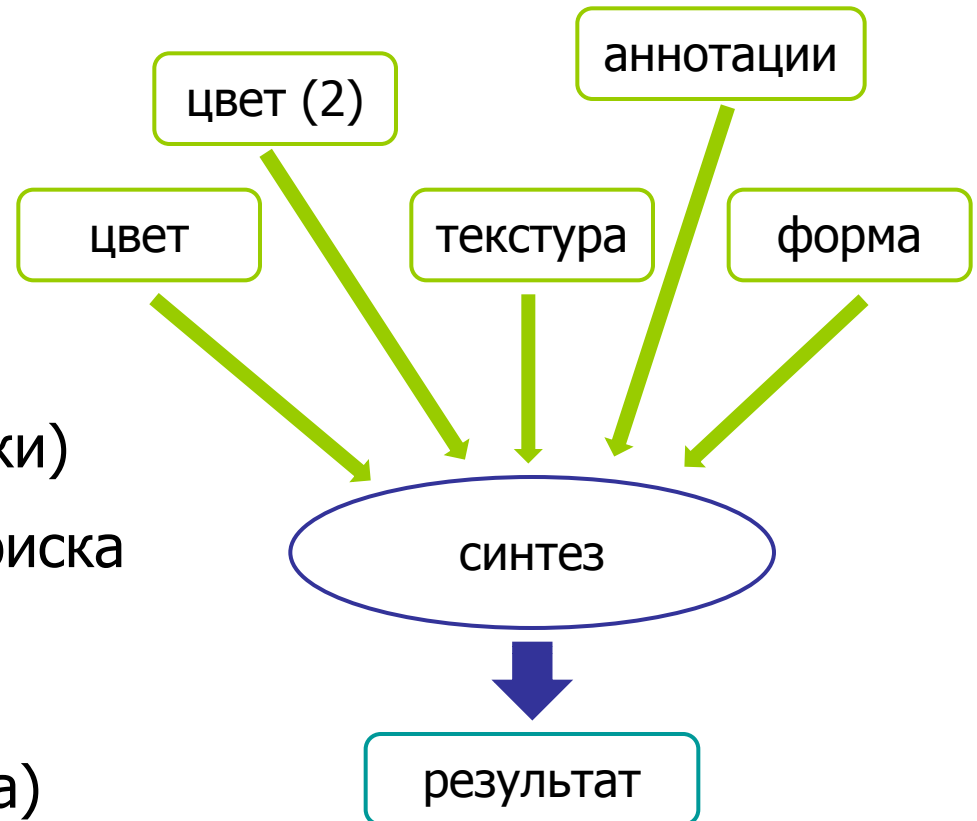
## II. Синтез данных в контексте CBIR

1. Существующие решения и их недостатки
2. Альтернативные подходы
3. Поиск в частично-аннотированной базе
4. WTGF: **W**eighted **T**otal with **G**ravitation **F**unction
5. Адаптивный поиск

# Синтез данных в контексте CBIR

---

- Комбинированный поиск (различные характеристики)
- Уточнение результатов поиска (разные алгоритмы)
- Дополнение результатов поиска (разные множества)



# Существующие недостатки

---

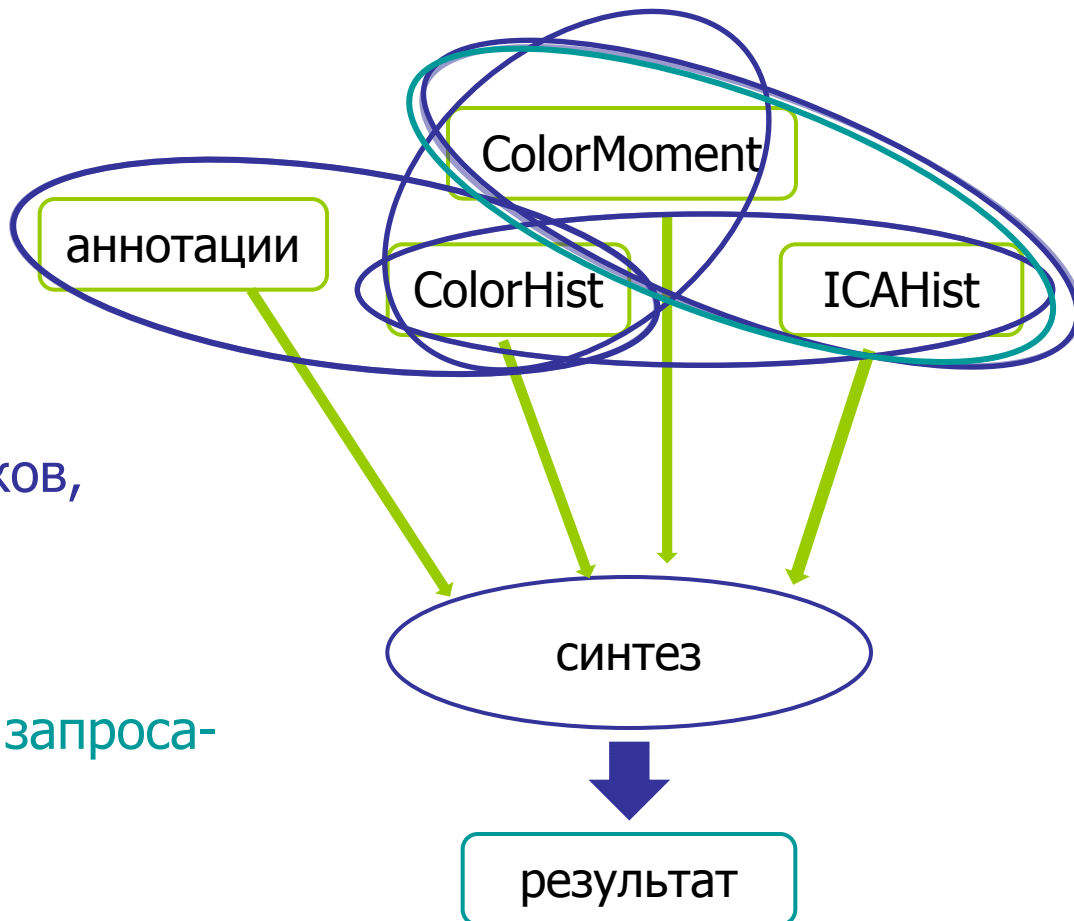
- CombMax, CombMin, CombSum
- CombAVG
- $\text{CombMNZ} = \text{CombSUM} * \text{number of nonzero similarities}$
- ProbFuse
- HSC3D
- Линейная комбинация (CombSum с весами)

## Недостатки:

- Не учитываются веса источников
- Если учитываются:
  - линейная зависимость итогового ранга элемента от его рангов в различных источниках и весов источников
- Не учитываются особенности запроса-образца

# Альтернативные подходы

---



- Учитывать веса источников, нелинейная зависимость результата от весов
- Учитывать особенности запроса-образца

# Синтез ранжированных списков с весами

$\omega_i$  – вес  $i$ -го списка;  $r_k^i$  – ранг  $k$ -го элемента в списке  $i$

$\omega_1$   $(x_{1,1}^1, r_{1,1}^1), (x_{1,2}^1, r_{1,2}^1), \dots, (x_{1,n}^1, r_{1,n}^1)$

$\omega_2$   $(x_{2,1}^2, r_{2,1}^2), (x_{2,2}^2, r_{2,2}^2), \dots, (x_{2,n}^2, r_{2,n}^2)$

...

$\omega_m$   $(x_{m,1}^m, r_{m,1}^m), (x_{m,2}^m, r_{m,2}^m), \dots, (x_{m,n}^m, r_{m,n}^m)$

$r_k^0 = f(\Omega, R_k)$ , где  
 $\Omega$  – множество весов всех списков,  
 $R_k$  – множество рангов элемента  $k$

## Существующие решения:

- CombMax, CombMin, CombSum
- CombAVG
- CombMNZ = CombSUM \* number of nonzero similarities
- ProbFuse
- HSC3D





# Свойства функции синтеза

---

- 1) Симметричность
- 2) Монотонность по каждому из аргументов
- 3) Функции для определения ранга объекта

Функция ранга  $([0..1], [0..1])^N \rightarrow [0..1]$

Функция веса  $[0..1]^N \rightarrow [0..1]$

- 4) MinMax условие /CombMin, CombMax, CombAVG/:

$$\min\{r_x^{(\alpha_1)}, r_x^{(\alpha_2)}, \dots, r_x^{(\alpha_N)}\} \leq r_x^{(0)} \leq \max\{r_x^{(\alpha_1)}, r_x^{(\alpha_2)}, \dots, r_x^{(\alpha_N)}\}$$

- 5) Дополнительное свойство (аналог HSC3D):  
условие взвешенной стабилизации элементов с высоким рангом  
(правило конусов)

# Weighted Total with Gravitation Function

---

Модернизация CombAVG, в качестве веса - стабилизационная (гравитационная) функция:

$$r_x^{(0)} = \frac{\sum_i g(r_x^{(\alpha_i)}, w^{(\alpha_i)}) \cdot r_x^{(\alpha_i)}}{\sum_i g(r_x^{(\alpha_i)}, w^{(\alpha_i)})}$$

где

$$g(r_x^{(\alpha_i)}, w^{(\alpha_i)}) = (w^{(\alpha_i)})^2 \cdot \left( r_x^{(\alpha_i)} + \frac{1}{12} \right)^4$$

# Эксперименты: метод оценки

---

Параметры  $R_{overlap}$   
 $N_{overlap}$ :

$$R_{overlap}(x) = \frac{M \cdot R^{(0)}(x)}{\sum_i^M R^{(\alpha_i)}(x)}$$

$$N_{overlap}(x) = \frac{M \cdot N^{(0)}(x)}{\sum_i^M N^{(\alpha_i)}(x)}$$

Lee J. H. Analyses of multiple evidence combination. SIGIR '97: Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval. New York, NY, USA: ACM Press, p. 267-276, 1997.

# Описание эксперимента I

## Данные:

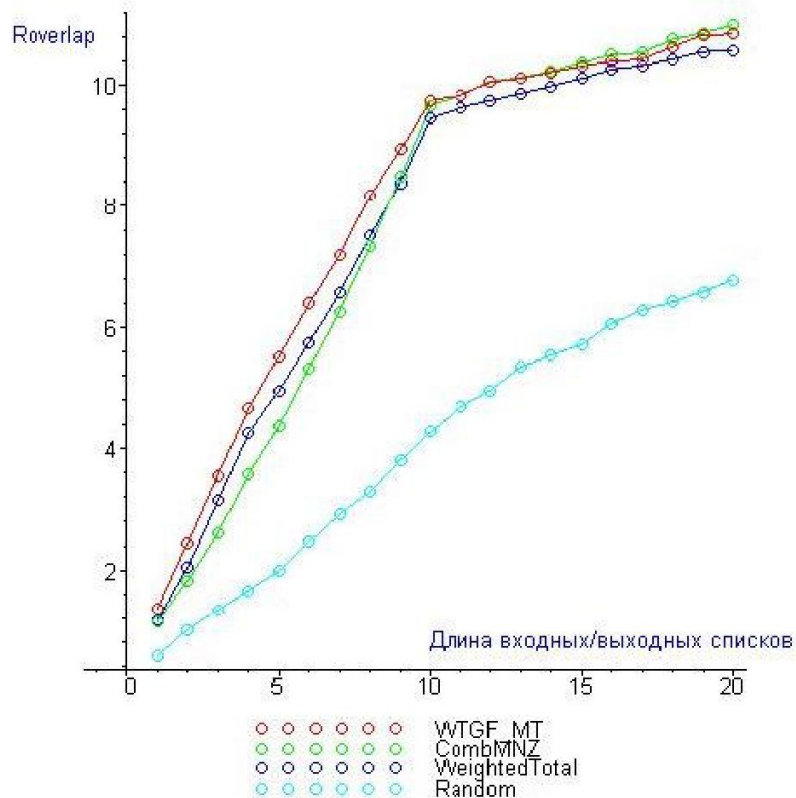
Коллекция Flickr (~15000)

## Методы:

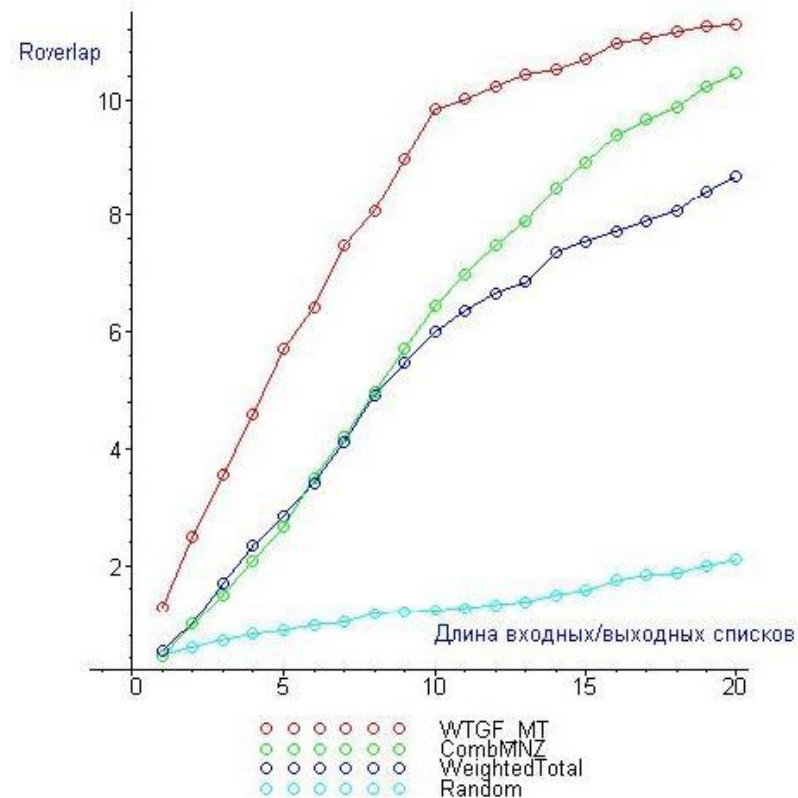
- Random с условиями MinMax
- CombMNZ
- WTGF\_MT
- WeightedTotal



# Результаты эксперимента I: $R_{\text{overlap}}$



а) Зависимость Roverlap от размера списка при  $\delta=0.03$  для 10 входных списков



б) Зависимость Roverlap от размера списка при  $\delta=0.07$  для 10 входных списков

# Описание эксперимента II

---

## Данные:

Коллекция Corel Photo Set (285)

## Участники синтеза (попарное смешивание):

- цветовые гистограммы с пространственной информацией (ColorHist )
- статистические признаки цвета (ColorMoment )
- текстурные признаки на основе фильтров ICA (ICAHist)

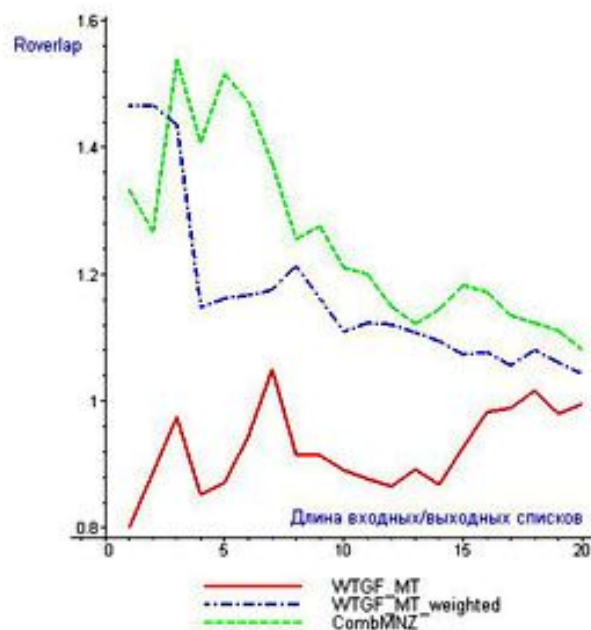
## Методы:

- CombMNZ
- WTGF\_MT
- WTGF\_MT\_weighted

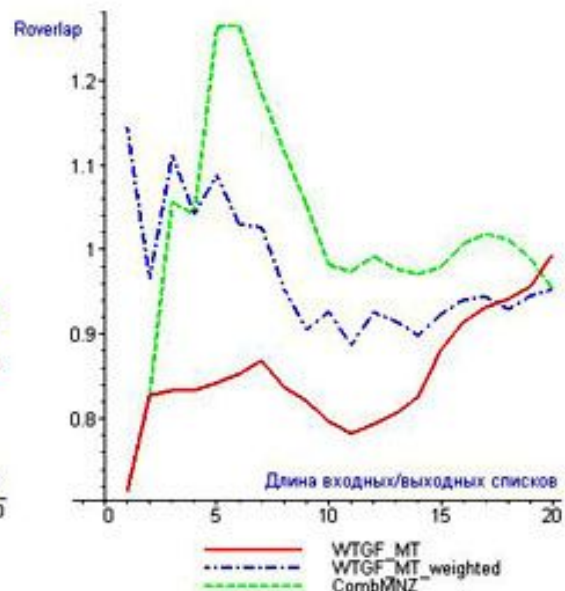
# Результаты эксперимента II

Графики зависимости значений Roverlap от размера списков для различных функций синтеза применительно к различным методам поиска по содержанию:

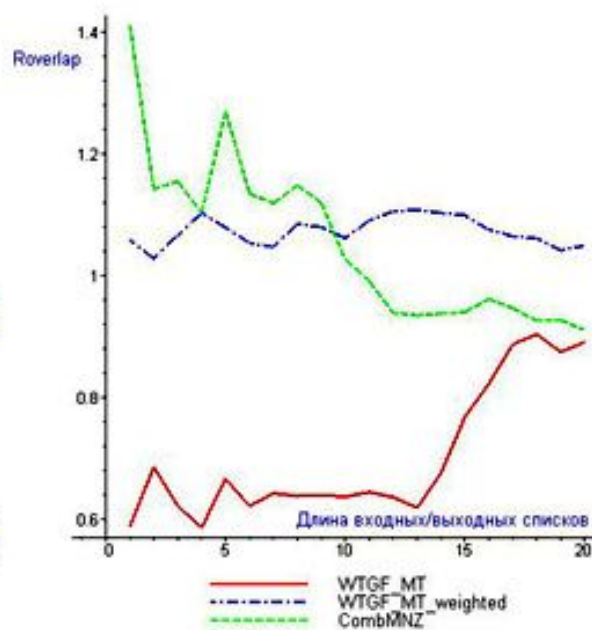
a) ColorHist и ColorMoment; b) ColorHist и ICAHist; c) ColorMoment и ICAHist.



a) ColorHist + ColorMoment



b) ColorHist + ICAHist



c) ColorMoment + ICAHist

# Адаптивный поиск

---

$$a \cdot C + (1 - a) \cdot T$$



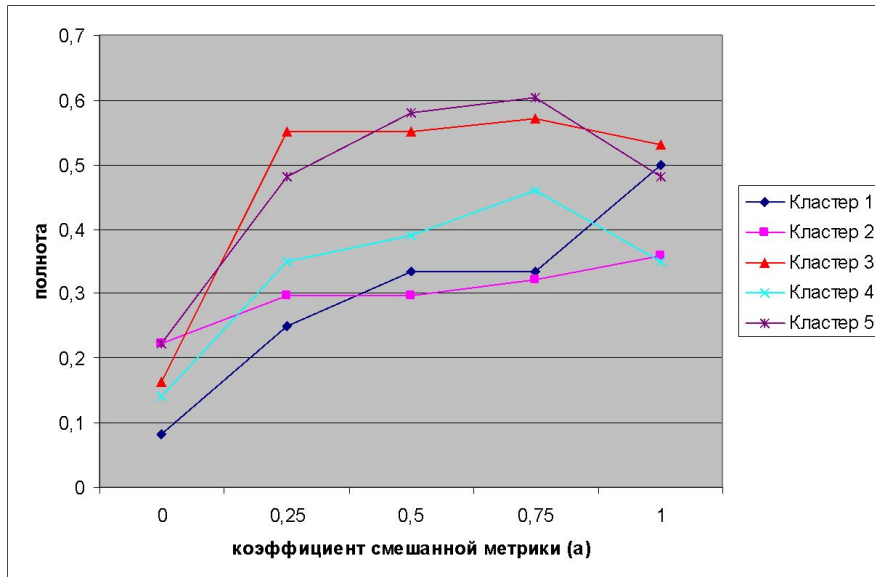


# Описание эксперимента III

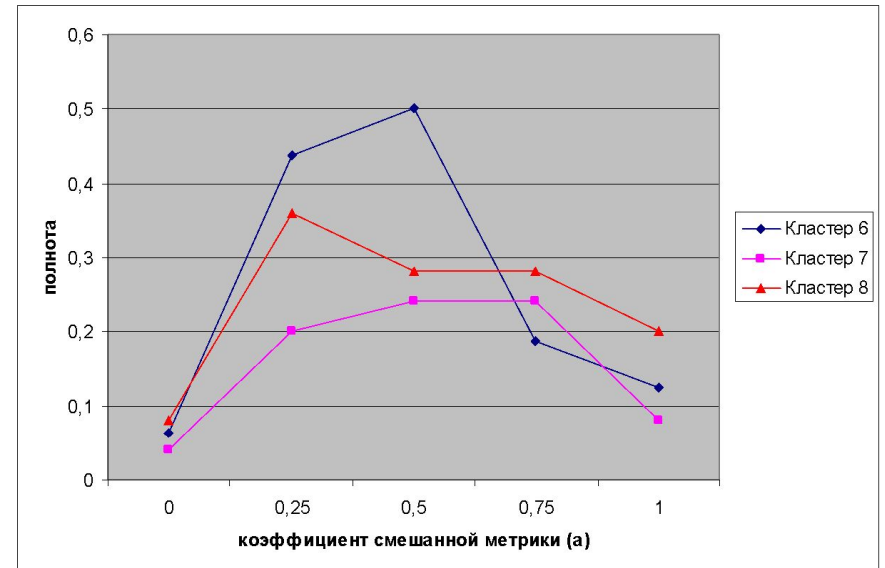
---

- Характеристики
  - Цвет – статистическое представление
  - Текстура – свертки с фильтрами ICA
- По оценкам ассессоров изображения разбиты на классы
- Выбор метрики для класса:
  - Каждое изображение – запрос для поиска с использованием смешанной метрики
  - Коэффициенты: 0, 0.25, 0.5, 0.75, 1
  - Чем больше суммарная (по всем изображениям класса) полнота, тем лучше метрика

# Результаты: зависимость полноты



Кластеры с преобладанием характеристики цвета.



Кластеры со смещением соотношения характеристик в сторону текстуры.

# Результаты: кластеры

---

Описание	Размер	Метрика
Небо	7	$1 \cdot C + 0 \cdot T$
Животные	10	$1 \cdot C + 0 \cdot T$
Облака	8	$0.75 \cdot C + 0.25 \cdot T$
Озера	9	$0.75 \cdot C + 0.25 \cdot T$
Поля, луга	10	$0.75 \cdot C + 0.25 \cdot T$
Листва	5	$0.5 \cdot C + 0.5 \cdot T$
Небоскребы	6	$0.5 \cdot C + 0.5 \cdot T$
Группы людей	5	$0.25 \cdot C + 0.75 \cdot T$

# Результаты: примеры

---

$$1 \cdot C + 0 \cdot T$$



$$0.5 \cdot C + 0.5 \cdot T$$



$$0.25 \cdot C + 0.75 \cdot T$$



# Методы синтеза: выводы (1)

---

- Методы синтеза применимы к задаче поиска изображений и позволяют существенно улучшить результаты поиска.
- WTGF:
  - большое количество источников;
  - невысокая степень перекрытия источников;
  - источники с различными весами.
- CombMNZ:
  - равнозначные источники;
  - высокая степень перекрытия источников.
- Предложенная схема поиска по частично аннотированной базе оправдала себя.

# Методы синтеза: выводы (2)

---

- Возможно выделить классы изображений, для которых большее значение имеет та или иная характеристика.
- Можно ли выделить общие признаки для изображений одного класса?
- Позволит ли адаптивный подход улучшить результат поиска?

# Заключение

---

- Большой выбор различных алгоритмов поиска по каждой из характеристик в отдельности
  - Цвет: гистограммы или статистическая модель?
  - Текстура: фильтры Габора, фильтры ICA
  - Форма: дескрипторы Фурье, инвариантные моменты
- Необходимо комбинировать методы поиска по различным характеристикам
  - Выбор метода синтеза зависит от конкретной задачи (что с чем смешиваем)
  - Важно учитывать веса источников
  - Адаптивный подход?