

Курс «Анализ изображений и видео, ч.2.»

# Лекция №1 «Введение в графические модели»

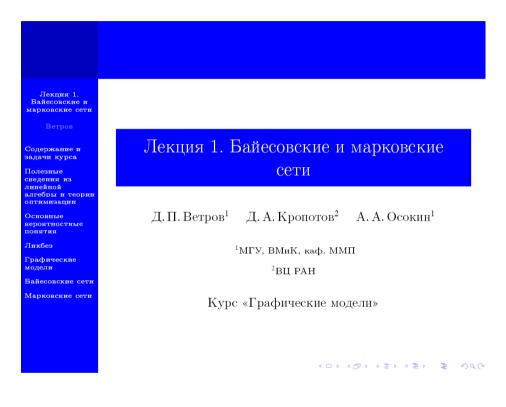
### Антон Конушин

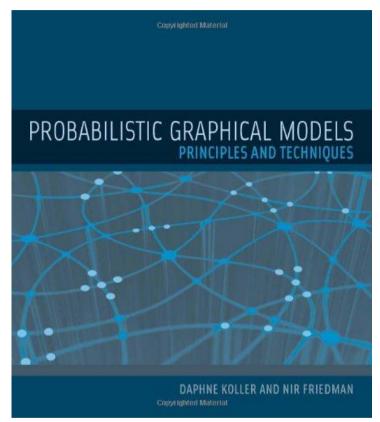
Заведующий лабораторией компьютерной графики и мультимедиа ВМК МГУ

10 февраля 2017 года

### Рекомендации







Курс Дмитрия Ветрова в ШАД

http://www.amazon.com/Probabilistic-Graphical-Models-Principles-Computation/dp/0262013193

Мы рассмотрим некоторые базовые понятия, которые нужны для построения моделей распознавания изображений

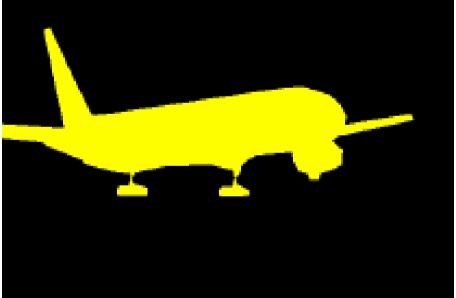
### Связи между пикселами



Рассмотрим задачу сегментации изображений

- Какие пиксели изображения принадлежат самолёту?
- $y_i$  метки всех пикселей,  $i \in [1, N]$





Есть ли связи (зависимости) между соседними пикселами y<sub>i</sub> , y<sub>i</sub> при решении этой задачи?

## Случайные поля



- Есть N переменных  $Y = [y_1, ..., y_N]$  , описывающих нашу задачу
- Y называют *случайным полем* (Random field)
- Варианты:
  - Непрерывное случайное поле  $y_i \in \mathbb{R}$
  - Дискретное случайное поле  $y_i \in \{1,...K\} \ K > 2$
  - Бинарное  $y_i \in \{0,1\}$
- Мы будем рассматривать в основном дискретные поля. Многие задачи хорошо на них ложатся и для них предложены эффективные алгоритмы.

## Случайные поля



- Совместное распределение P(Y) описывает вероятность конфигурации переменных с учётом их зависимостей
- Многие задачи можно сформулировать как

$$Y = \underset{Y}{\operatorname{arg}} \max P(Y)$$

• Чаще используется такая формулировка:

$$Y = \underset{Y}{\operatorname{arg}} \max P(Y \mid X)$$

где X – наблюдаемые переменные

• Необходимо учитывать зависимости между переменными, но оставить задачу вычислимой на практике

### Совместные вероятностные распределения



• Универсальный способ (для дискретных переменных): таблица

$$P(y_1, y_2)$$

	1	2	3	4	5
1					
2					
3				•	
4					
5					

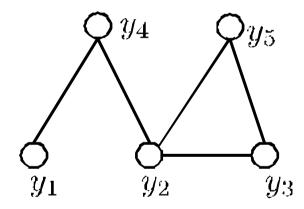
$$P(y_1 = 4, y_2 = 3)$$

- Проблемы:
  - Экспоненциальный рост размера таблицы (  $10^6$  переменных  $4^{10^6}$  элементов в таблице)
  - Экспоненциальный рост данных, необходимых для эмпирической оценки
- Вывод: необходимы априорные предположения о структуре таблицы (т.е. вероятностного распределения)

### Графическая модель



- Построим граф G=(V,E), V вершины, E рёбра, такой что:
  - Все переменные из Y проиндексированы вершинами V
  - Ребро E=(V<sub>i</sub>V<sub>j</sub>) присутствует, если между у<sub>i</sub> и у<sub>j</sub> есть непосредственная (а не опосредственная) вероятностная зависимость



Отсутствие ребра между у<sub>і</sub> и у<sub>ј</sub> означает, что они условно независимы, т.е.:

$$P(y_i, y_j \mid Y_{\{i,j\}}) = P(y_i \mid Y_{\{i,j\}})P(y_i \mid Y_{\{i,j\}})$$

Записав совместное распределение Р(Y) получим:

$$P(Y) = P(y_i, y_j \mid Y_{\{i,j\}}) P(Y_{\{i,j\}}) = P(y_i \mid Y_{\{i,j\}}) P(y_j \mid Y_{\{i,j\}}) P(Y_{\{i,j\}})$$

Переменные у<sub>і</sub> и у<sub>ј,</sub> не соединённые рёбрами, входят в разные множители совместного распределения

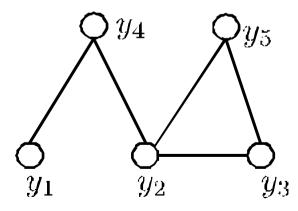
### Факторизация



В общем виде совместное распределение P(Y) записывается с помощью неотрицательных потенциальных функций, называемых «факторами», определённых на максимальных кликах D<sub>i</sub>

$$P(\mathbf{y}) = \frac{1}{Z} \prod_{i} \Phi_{i}(\mathsf{D}_{i}) \quad Z = \sum_{\{Y\}} \prod_{i} \Phi_{i}(D_{i}) \quad \Phi_{i}(D_{i}) > 0$$

нормализующая константа (partition function)



$$P(Y) = \frac{1}{Z}\Phi(y_1, y_4)\Phi(y_2, y_4)\Phi(y_2, y_3, y_5)$$

### Факторизация



$$P(\mathbf{y}) = \frac{1}{Z} \prod_{i} \Phi_{i}(\mathsf{D}_{i}) \quad Z = \sum_{\{Y\}} \prod_{i} \Phi_{i}(D_{i}) \quad \Phi_{i}(D_{i}) > 0$$

нормализующая константа (partition function)

- Факторы **Ф** не обязаны иметь вероятностную природу. Они показывают «сродство» (affinities) между переменными. Чем больше значение фактора, тем выше вероятностей данной конфигурации переменных
- Факторы обычно задаёт пользователь исходя из априорных сведений о задаче.
- Реже мы можем настроить факторы по данным.

## Конструктивный подход



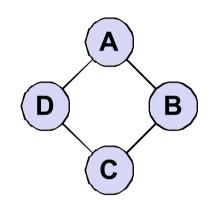
Получаем конструктивный подход к построению моделей, решающих нашу задачу:

- Проводим параметризацию модели (определяем случайное поле Y)
- Исходя из априорной информации определяем структуру зависимостей между переменными
- Задаём или настраиваем факторы
- Проводим вывод (inference), определяя наиболее вероятную конфигурацию переменных

### Простой пример



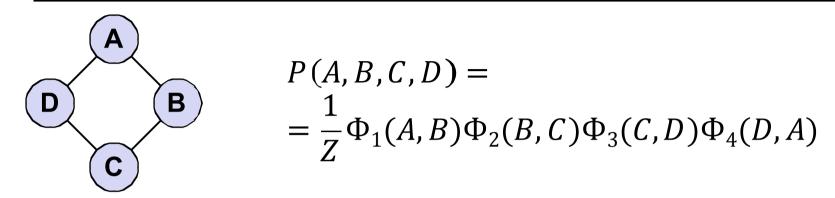
- Alice (A), Bob(B), Charles (C), Debbie (D) –
   четыре студента (наше поле Y)
- Лектор сделал ошибку на занятии. Каждый из студентов может догадаться об ошибке и понять тему правильно (=1). Или не догадаться и ошибаться (=0) (бинарное поле)



- Задача понять, кто из студентов узнает об ошибке, а кто нет?
- Мы знаем, что студенты работают только в определенных парах AB, BC, CD, DA
- Зададим графическую модель для данной задачи
  - 4 переменные A,B,C,D (случайное поле)
  - 4 ребра, отражающие зависимости

### Факторизация нашего примера



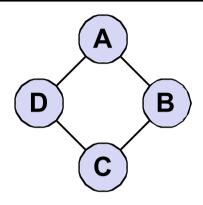


Настроим факторы по данным, насколько часто совпадают ответы студентов:

Видно, что (A,B), (B,C), (D,A) склонны соглашаться друг с другом (принимать одинаковые значения). А (C,D) обычно противоречат друг другу

### Вывод в нашем примере





$$\phi_1(A,B)$$
  $\phi_2(B,C)$ 

$$\phi_3(C,D)$$
  $\phi_4(D,A)$ 

$$P(A, B, C, D) =$$
  
=  $\frac{1}{Z} \Phi_1(A, B) \Phi_2(B, C) \Phi_3(C, D) \Phi_4(D, A)$ 

$ \  \   Large Assignment$			nt	Unnormalized	Normalized
$a^0$	$b^0$	$c^0$	$d^0$	300,000	0.04
$a^0$	$b^0$	$c^0$	$d^1$	300,000	0.04
$a^0$	$b^0$	$c^1$	$d^0$	300,000	0.04
$a^0$	$b^0$	$c^1$	$d^1$	30	$4.1 \cdot 10^{-6}$
$a^0$	$b^1$	$c^0$	$d^0$	500	$6.9 \cdot 10^{-5}$
$a^0$	$b^1$	$c^0$	$d^1$	500	$6.9 \cdot 10^{-5}$
$a^0$	$b^1$	$c^1$	$d^0$	5,000,000	0.69
$a^0$	$b^1$	$c^1$	$d^1$	500	$6.9 \cdot 10^{-5}$
$a^1$	$b^0$	$c^0$	$d^0$	100	$1.4 \cdot 10^{-5}$
$a^1$	$b^0$	$c^0$	$d^1$	1,000,000	0.14
$a^1$	$b^0$	$c^1$	$d^0$	100	$1.4\cdot 10^{-5}$
$a^1$	$b^0$	$c^1$	$d^1$	100	$1.4\cdot 10^{-5}$
$a^1$	$b^1$	$c^0$	$d^0$	10	$1.4\cdot 10^{-6}$
$a^1$	$b^1$	$c^0$	$d^1$	100,000	0.014
$a^1$	$b^1$	$egin{array}{c} c^1 \ c^1 \end{array}$	$d^0$	100,000	0.014
$\mid a^1 \mid$	$b^1$	$c^1$	$d^1$	100,000	0.014

Не смотря на то, что А и В склонны соглашаться друг с другом, другие их связи вынуждают принять разные решения!

### Польза факторизации



Факторизация – ключевая идея графических моделей. Посчитаем экономию на примере.

Пример: 
$$P(y_1, y_2, y_3, y_4, y_5, y_6) =$$

$$\frac{1}{Z}\Phi_1(y_1,y_2)\Phi_2(y_2,y_3)\Phi_3(y_3,y_4,y_5)\Phi_4(y_4,y_5,y_6)$$

Неотрицательные факторы

*«разделительная функция»* **NB**: для данных факторов она константа!

Пусть 
$$y_i \in \{1, 2, 3, 4\}$$

Было 
$$4^6 = 4096$$

Стало 
$$16+16+64+64 = 160$$

### «Разделительная функция»



$$P(y_1, y_2, y_3, y_4, y_5, y_6) =$$

$$\frac{1}{Z}\Phi_1(y_1,y_2)\Phi_2(y_2,y_3)\Phi_3(y_3,y_4,y_5)\Phi_4(y_4,y_5,y_6)$$

$$y_i \in \{1, 2, 3, 4\}$$

$$Z = \sum_{i=1}^{4} \sum_{j=1}^{4} \sum_{j=1}^{4} \sum_{j=1}^{4} \Phi_1(y_1, y_2) \Phi_2(y_2, y_3) \Phi_3(y_3, y_4, y_5) \Phi_4(y_4, y_5, y_6)$$

В общем случае, для данной модели и данного у:

- Не можем посчитать вероятность Р(у)
- Можем сравнить с вероятностью у1

### Несколько понятий



- Поле Гиббса
- Фактор-граф
- Марковское случайное поле
- Функция энергии
- Условное случайное поле

### Поле Гиббса



• Распределение P(Y) является распределением Гиббса, параметризованном набором положительно определённых факторов  $\Phi_i(D_i)$ , если его можно представить в виде:

$$P(Y) = \frac{1}{Z}\widetilde{P}_{\Phi}(Y) \qquad \widetilde{P}_{\Phi}(Y) = \prod_{i} \Phi_{i}(D_{i}) \qquad Z = \sum_{Y}\widetilde{P}_{\Phi}(Y)$$

- Будем говорить, что распределение Гиббса P(Y) с факторами {Ф(D<sub>i</sub>)} факторизуется по графу G, если каждому набору переменных D<sub>i</sub> соответствует полный подграф (клика) в G
- (Y,G) поле Гиббса, если P(Y) факторизуется по графу G

#### Пример:

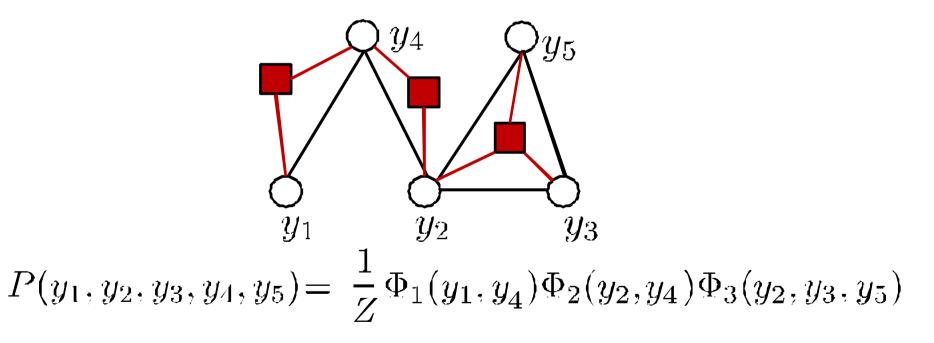
$$y_1$$
  $y_2$   $y_3$ 

$$P(y_1, y_2, y_3, y_4, y_5) = \frac{1}{Z} \Phi_1(y_1, y_4) \Phi_2(y_2, y_4) \Phi_3(y_2, y_3, y_5)$$

### Фактор-граф



- Фактор-граф неориентированный двудольный граф, в котором рёбрами соединены факторы и случайные переменные, входящие в фактор
- Фактор-граф даёт другое представление (визуализацию) зависимостей между переменными случайного поля



# Марковское случайное поле (МСП)

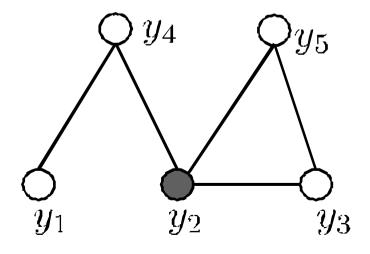


Графическая модель (Y,G) образует Марковское случайное поле, если выполняется свойство Марковости:

• Два любых множества вершин А,В условно независимы, если задано разделяющее множество вершин Т

$$P(Y_{V-T} | Y_T) = P(Y_A | Y_T) P(Y_B | Y_T)$$

#### Пример:



$$P(y_1, y_3, y_4, y_5 | y_2) =$$

$$P(y_1, y_4 | y_2)P(y_3, y_5 | y_2)$$

### Факторизация и условная независимость

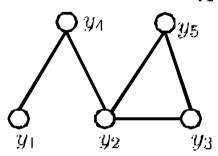


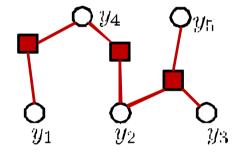
Можно показать, что X и Y условно независимы при условии Z тогда и только тогда, когда мы можем записать в следующем виде:

$$P(X,Y,Z) = \Phi 1(X,Z)\Phi_2(Y,Z)$$

Пример:

Pp: 
$$P(y_1, y_2, y_3, y_4, y_5) = \frac{1}{Z} \Phi_1(y_1, y_4) \Phi_2(y_2, y_4) \Phi_3(y_2, y_3, y_5)$$





Утверждение:

$$P(y_1, y_3, y_4, y_5, | y_2) = P(y_1, y_4 | y_2) P(y_3, y_5 | y_2)$$

т.е. 
$$(y_1,y_4) \perp \!\!\! \perp (y_3,y_5) \! \mid \! y_2$$

Общий случай:

не существует пути из 
$$Y_{\scriptscriptstyle 1}$$
 в  $Y_{\scriptscriptstyle 2}$  в обход  ${\it T} \ \Rightarrow \ Y_1 \perp \!\!\! \perp Y_2 \mid T$ 

**Таким образом:** задав фактор-граф, задаем структуру условной независимости переменных (*Марковское случайное поле*)

# Теорема Хаммерсли – Клиффорда



**Теорема:** Случайное поле Y является Марковским случайным полем на графе G, тогда и только тогда, когда оно является полем Гиббса на G.

#### Таким образом:

- Задав фактор-граф, задаём структуру условной независимости, т.е. Марковское случайное поле
- Задав Марковское случайное поле, получаем факторизацию совместной вероятности через произведение положительно определённых факторов, т.е. распределение Гиббса

Теорема была несколько раз доказана (1971,1974 и 1980)

### Функция энергии



- Запишем фактор  $\Phi(D)$  как  $\Phi(D) = \exp(-\varepsilon(D))$
- $\varepsilon(D) = -\ln \Phi(D)$  называется функцией энергии (energy function) или потенциалом (potential)
- Название пришло из статистической физики, в которой обычно вероятность состояния системы обратно пропорционально энергии
- Можем записать P(Y) в логарифмическом виде:

$$P(Y) \propto \exp\left(-\sum_{i=1}^{m} \varepsilon_i(D_i)\right)$$

• Максимизация P(Y) эквивалентна минимизации суммарной энергии системы!

$$\arg\max P(Y) = \arg\max \exp\left(-\sum_{i=1}^{m} \varepsilon_{i}(D_{i})\right) = \arg\min \sum_{i=1}^{m} \varepsilon_{i}(D_{i})$$

### Лог-линейная модель



- Определим признак (feature) f(D) как функцию из D в R
- Теперь мы можем задать лог-линейную модель Марковского поля (Y,G):
  - Определим набор признаков f<sub>i</sub>(D<sub>i</sub>), где каждый набор переменных D<sub>i</sub> соответствует клике в G
  - Выберем набор весов w<sub>i</sub>
  - и тогда

$$P(Y) = \frac{1}{Z} \exp\left(-\sum_{i=1}^{m} w_i f(D_i)\right)$$

## Conditional Random Field (CRF)



- Пусть Y набор целевых переменных, X набор наблюдаемых переменных
- Формализм графических моделей можно применить для записи P(Y|X)
- Условным случайным полем назовём (Y,X,G), в котором вершины G соответствуют Y, граф аннотирован факторами Ф<sub>і</sub>(D<sub>і</sub>), обусловленных наблюдаемыми переменными X

$$P(Y \mid X) == \frac{1}{Z} \prod_{i} \Phi_{i}(Y_{i} \mid X)$$

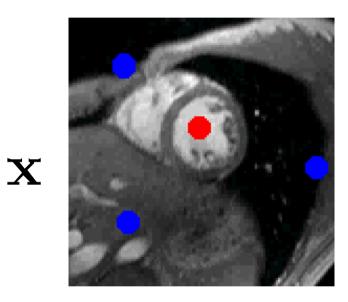
### Conditional Random Field

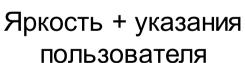


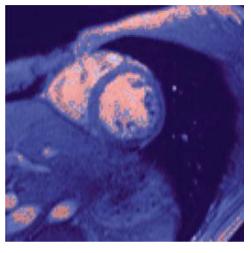
- Плюсом модели условного случайного поля является то, что нам нужно моделировать зависимости только между скрытыми переменными
- Мы можем использовать большой набор наблюдаемых переменных, зависимости между которыми сложны или вообще непонятны
- Поэтому в большинстве случаев используется или подразумевается именно CRF

## Пример: классификация пикселей

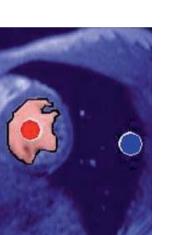








классификация



совместная классификация

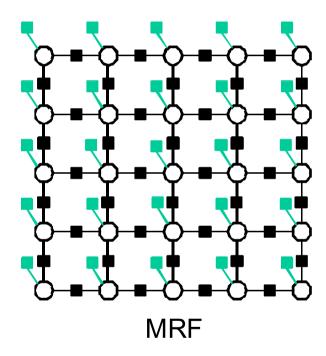
[Boykov and Jolly // ICCV2001]

Slide by Victor Lempitsky

### Сеточная модель изображения



Каждому пикселю изображения задаём метку у, и зависит метка пиксела только от своих соседей



$$\begin{split} &P(Y,X) = \\ &= \frac{1}{Z} \prod_{i} \Phi_{i}(y_{i},x_{i}) \prod_{(i,j) \in N} \Psi_{i,j}(y_{i},y_{j}) \end{split}$$

$$P(Y,X) = P(Y|X) =$$

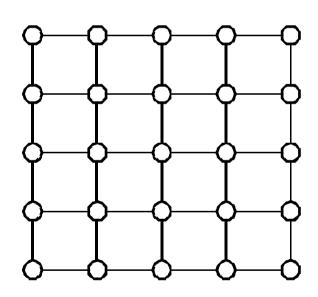
$$= \frac{1}{Z} \prod_{i} \Phi_{i}(y_{i}, x_{i}) \prod_{(i,j) \in N} \Psi_{i,j}(y_{i}, y_{j}) = \frac{1}{Z} \prod_{i} \Phi_{i}(y_{i}|X) \prod_{(i,j) \in N} \Psi_{i,j}(y_{i}, y_{j}|X)$$

### Факторы в модели изображения



$$P(Y | X) = \frac{1}{Z} \prod_{i} \Phi_{i}(y_{i} | X) \prod_{(i,j) \in N} \Psi_{i,j}(y_{i}, y_{j} | X)$$

$$E(Y | X) = -\sum_{i} \log(\Phi_{i}(y_{i} | X)) - \sum_{(i,j) \in N} \log(\Psi_{i,j}(y_{i}, y_{j} | X))$$

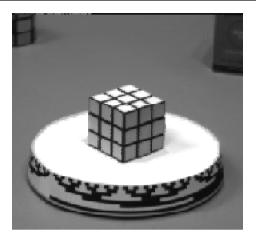


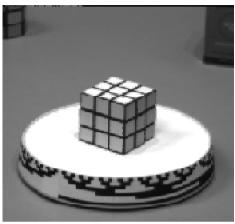
- Унарные (размер 1) зависимость метки от изображения (классификатор)
- Парные (размер 2) штраф за сходство или различие соседних переменных в зависимости от свойств изображения
- Более высоких порядков кривизну контура, площадь фигуры и т.д.

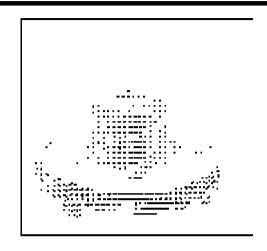
Обычно используют унарные и парные потенциалы, т.к. для потенциалов более высоких порядков оптимизация сложнее

## Примеры: попиксельная аннотация





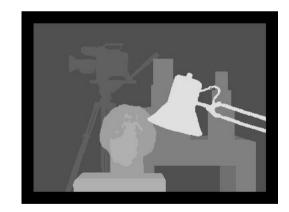




Вычисление оптического потока





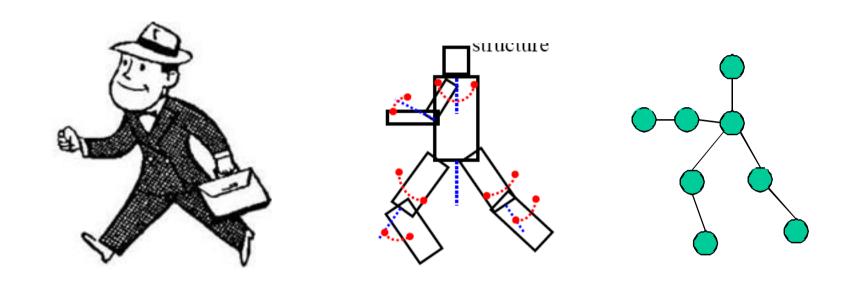


Стереосопоставление

Каждый пиксел изображения  $\mathbf{x}_i$  помечаем меткой  $\mathbf{y}_i$  – вектором оптического потока, диспаритетом, и т.д.

### Пример: разметка человека





- Какая будет «графическая модель» для задачи разметки частей тела человека или суставов?
- Положение головы человек непосредственно зависит от положения тела, т.к. голова «прикреплена» к телу
- Положение у<sub>голова</sub> можно описать дискретной переменной, перечислив возможные положения в изображении

# Пример: точки на лице



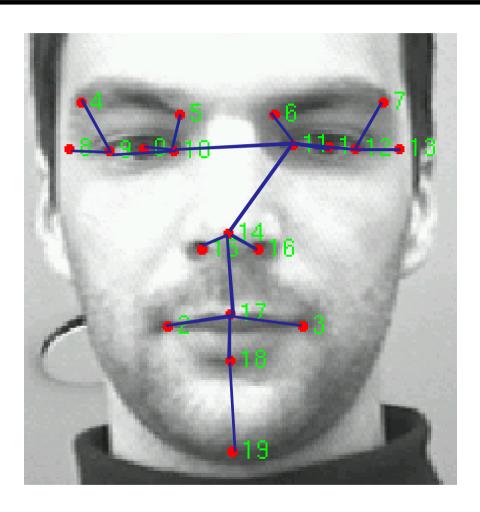


Image from BioID/FGNet dataset

### Методы вывода



Мы упомянем только несколько методов, в основном для дискретного случая:

- Покоординатный и градиентный спуск
- Динамическое программирование
- На основе разрезов графов для K=2
- Итеративный для К>2

### Простейшие алгоритмы



$$E(\mathbf{y}) = \sum_{C_i} E_i(Y_{C_i}) \to \min$$

1. Покоординатный спуск (aka Iterated Conditional Modes):

while change for i = 1 to N 
$$\hat{y}_i = \arg\min_{y_i} \sum_{C_i} E_i(Y_{C_i})$$
  $y_i \leftarrow \hat{y}_i$  endfor end while

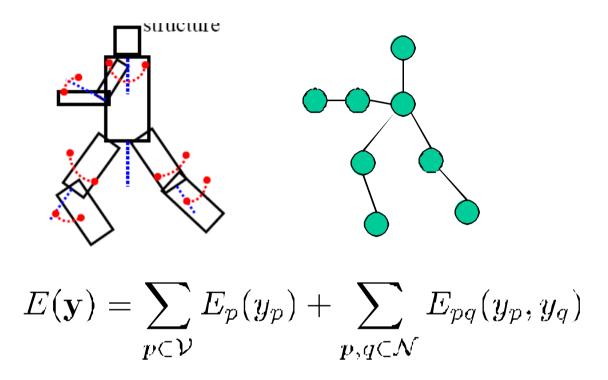
2. Градиентный спуск (для непрерывных у):

$$\mathbf{y} \leftarrow \mathbf{y} - \lambda_t \sum \frac{dE_i(Y_{C_i})}{d\mathbf{y}}$$

Для невыпуклых энергий (большинство интересных) локальные минимумы — существенная проблема

### Динамическое программирование

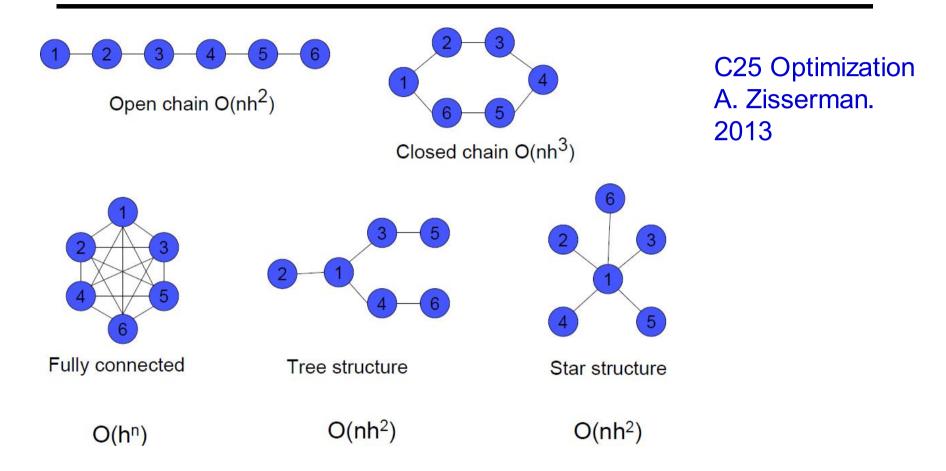




Для дискретных моделей в виде дерева динамическое программирование позволяет найти глобальное решение для любых энергий

### Динамическое программирование





На графах типа дерева энергию можно эффективно минимизировать с помощью динамического программирования (алгоритма Витерби)

### Методы на графах

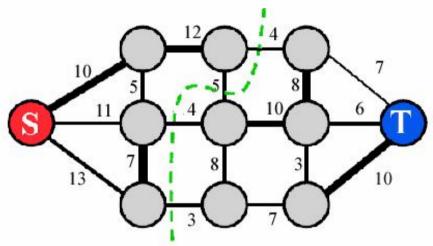


- Модели на основе CRF получили широкое распространение в зрении с появлением эффективных алгоритмов вывода для дискретных случаев
- Самым известным среди них является метод, в котором минимизация энергии сводится к задаче поиска минимального разреза / максимального потока в сети

## Минимальный st-разрез в графе



• Пусть G = (V, E) - направленный граф. Каждому ребру (u, v) присваивается некоторый неотрицательный вес – пропускная способность ребра c(u, v). В графе выделены 2 особые (терминальные) вершины – исток s и сток t. Такой граф называют сетью.



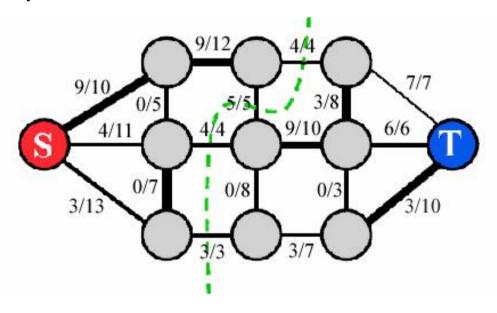
- st-разрез C = S,T это разбиение вершин из V на 2 непересекающихся множества S и T, такие, что  $s \in S, t \in T$
- Нужно получить st-разрез с минимальным весом

$$c(S,T) = \sum_{u \in S, v \in T, (u,v) \in E} c(u,v)$$

# Теорема Форда – Фалкерсона



Теорема Форда – Фалкерсона: величина максимального потока равна величине минимального st-разреза.

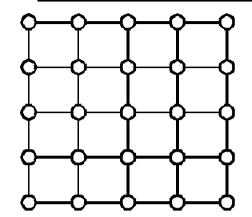


Задачи нахождения минимального разреза и максимального потока являются двойственными.

Есть эффективные (полиномиальной сложности) алгоритмы решения.

## Представление энергий на графе





Возьмём сеточную модель с бинарными переменными  $\mathbf{y} \in \{0,1\}^{|\mathcal{V}|}$ 

Преобразуем энергию так, чтобы её можно было отразить на граф

$$E(\mathbf{y}) = \sum_{p \in \mathcal{V}} E_p(y_p) + \sum_{p,q \in \mathcal{N}} E_{pq}(y_p, y_q)$$



$$E(\mathbf{y}) = \sum_{p \in \mathcal{V}} F^p y_p + \sum_{p \in \mathcal{V}} B^p (1 - y_p) + \sum_{p,q \in \mathcal{N}} P^{pq} |y_p - y_q|$$

Для парного потенциала нам нужно E(0,0) = E(1,1) = 0 и  $E(1,0) = E(0,1) = P^{pq} >= 0$ 

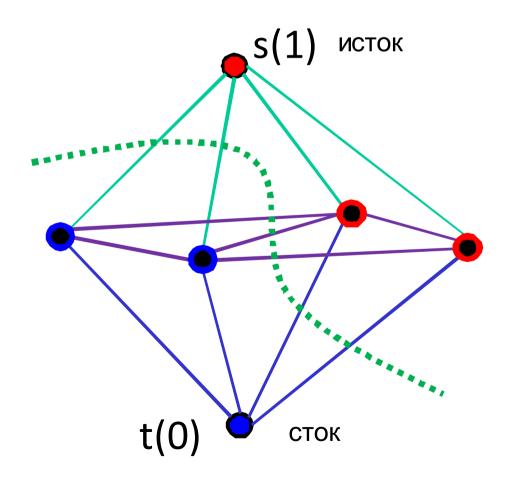
### Реализация энергии на графе



$$\mathbf{y} \in \{0, 1\}^{\mathcal{V}} \quad E(\mathbf{y}) = \sum_{p \in \mathcal{V}} F^{p} y_{p} + \sum_{p \in \mathcal{V}} B^{p} (1 - y_{p}) + \sum_{p, q \in \mathcal{N}} P^{pq} y_{p} - y_{q}$$

Такую энергию можно отобразить на сеть

Задав рёбрам веса E(0), E(1) и E(1,0)



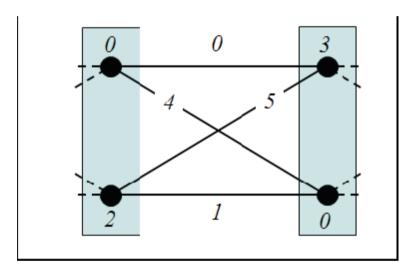
Если эти рёбра входят в разрез, тогда метка вершины 0

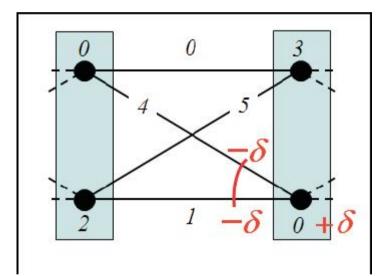
Если эти рёбра входят в разрез, тогда метка вершины 1

#### Репараметризация



- Рассмотрим парный потенциал, с 4 значениями (0,0),(0,1),
   (1,0), (1,1)
- Мы хотим сделать E(0,0) = E(1,1) = 0 и E(1,0) = E(0,1) > 0

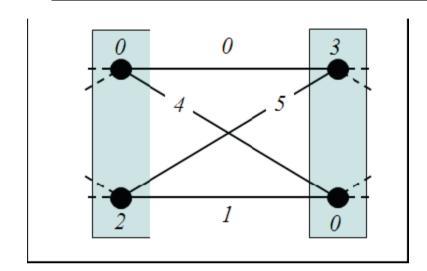


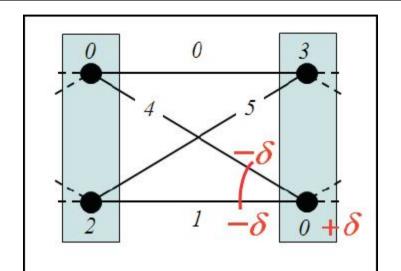


Вычитая одинаковые значения из двух ребер, сходящихся в одну вершину и прибавляя это же значение к весу самой вершины, мы получаем эквивалентный функционал энергии

## Субмодулярные энергии







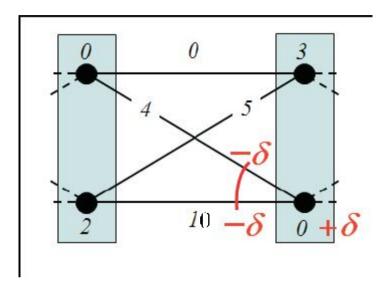
Если энергия удовлетворяет условию субмодулярности:

$$E_{pq}(0,0) + E_{pq}(1,1) \le E_{pq}(0,1) + E_{pq}(1,0)$$

Тогда мы можем провести такую репараметризацияю.

### Несобмодулярные энергии





Если условие нарушается, т.е.

$$E_{pq}(0,0) + E_{pq}(1,1) \ge E_{pq}(0,1) + E_{pq}(1,0)$$

тогда веса рёбер получаются отрицательными, и классические алгоритмы поиска минимального разреза становятся неприменимы

Есть варианты алгоритмов, когда энергия может быть сведена к субмодулярной путём замены части переменных на свои отрицания 1-у

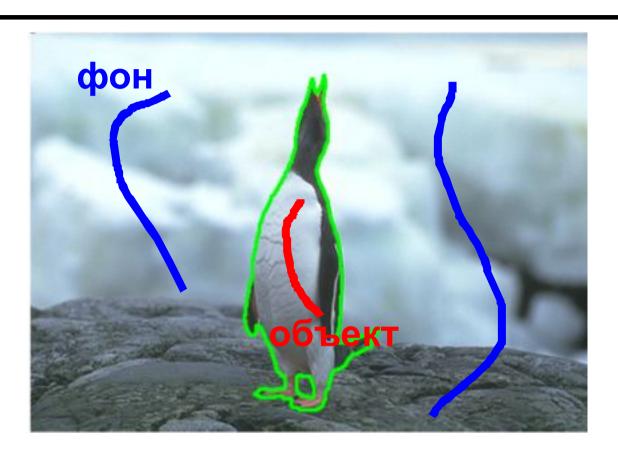
## Алгоритмы st-разреза



- Бойков и Колмогоров в 2001 провели эмпирическое сравнение различных версий алгоритмов, основанных на дополняющих путях и проталкивании предпотока на графах-решетках, типичных для задач компьютерного зрения
- Сделали собственную версию алгоритма, основанного на дополняющих путях
- Эмпирическая сложность близка к линейной относительно размера изображения.

#### Применение для бинарной сегментации





 $y_i \in \{0,1\}$ 

Energy 
$$_{MRF} = E(Y) = \sum_{p} \phi(I, y_{p}) + \sum_{\{p,q\} \in N} \psi(I, y_{p}, y_{q}) + const$$

Yuri Boykov and Marie-Pierre Jolly. Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D images. In International Conference on Computer Vision, (ICCV), vol. I, pp. 105-112, 2001.

#### Потенциалы и их визуализация



$$Energy_{MRF} = E(Y) = \sum_{p} (\phi(I \mid y_p) + \sum_{\{p,q\} \in N} (\phi(I \mid y_p, y_q) * \psi(y_p, y_q))) + const$$

p Унарное правдоподобие (Unary likelihood)

Член контраста (Contrast Term) Модель априорной вероятности Поттса (Potts Model Prior)

MAP solution 
$$Y^* = \arg\min_{Y} E(Y)$$



Данные (I)



Унарное правдоподобие (Unary likelihood)



Бинарные члены (Pair-wise Terms)



Решение

## Унарные потенциалы



• Яркости пикселей (вариант 1)

$$\begin{array}{ll} \phi(I\mid"obj")=-\ln\Pr(I_{_p}\mid O) & \Pr(I_{_p}\mid O) & -\text{ вероятность того,} \\ \phi(I\mid"bkg")=-\ln\Pr(I_{_p}\mid B) & \text{объекта яркость } I_{_p} \end{array}$$

Вероятность считается через гистограммы

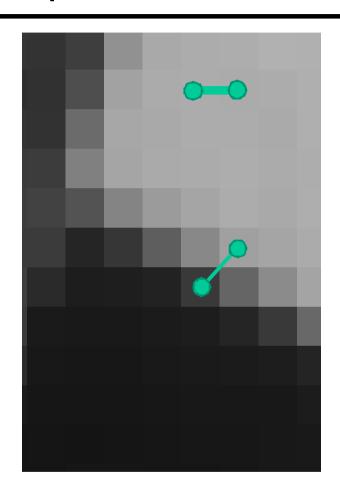
• Цвет пикселей (вариант 2)

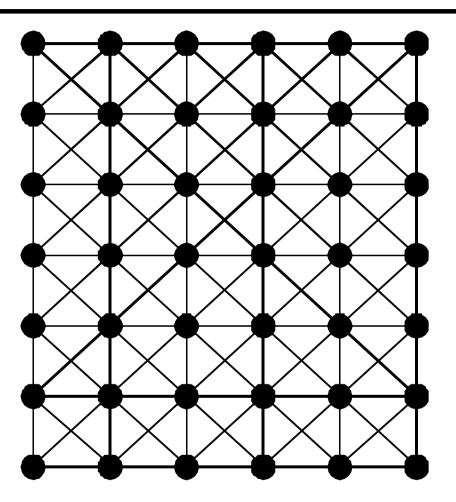
$$p(x_i | y_i = 0) = Ce^{-\frac{(x_i - color_{back})^2}{2\sigma^2}}$$
$$p(x_i | y_i = 1) = Ce^{-\frac{(x_i - color_{obj})^2}{2\sigma^2}}$$

Используем нормальное распределение для моделирования цвета пикселя

#### Парные потенциалы



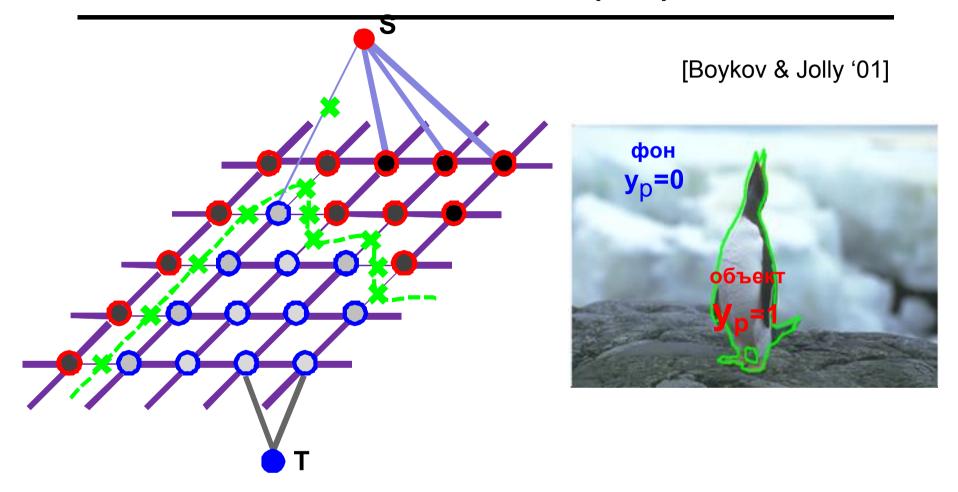




$$P^{pq} \propto \exp\left(-\frac{||I_p - I_q||^2}{2\sigma^2}\right) \cdot \frac{1}{||p - q||^2}$$

## Сегментация с помощью разреза





$$E(\mathbf{y}) = \sum_{p \in \mathcal{V}} F^p y_p + \sum_{p \in \mathcal{V}} B^p (1 - y_p) + \sum_{p,q \in \mathcal{N}} P^{pq} |y_p - y_q|$$

# Случай К>2 меток



$$\mathbf{y} = \arg\min_{\mathbf{y}} \sum_{p \in \mathcal{V}} E_p(y_p) + \lambda \sum_{p,q \in \mathcal{N}} E_{pq}(y_p, y_q)$$
$$y_p \in \{1, ..., K\} \quad K > 2$$

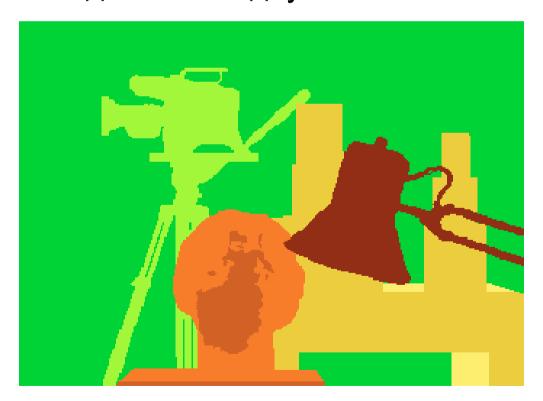
- Задача оптимизации для К > 2 является NPтрудной
- Но можно построить итерационный алгоритм, сходящийся к близкому к глобальному оптимуму ответу

### α-Расширение



$$\mathbf{y} = \arg\min_{\mathbf{y}} \sum_{p \in \mathcal{V}} E_p(y_p) + \lambda \sum_{p,q \in \mathcal{N}} E_{pq}(y_p, y_q)$$

Сведём к серии бинарных задач, выбирая на каждом шаге одну из меток



#### начальное решение

- •-расширение
- •-расширение
- •-расширение
- •-расширение
- -расширение
- •-расширение
- •-расширение

[Boykov, Veksler, Zabih, 1998]

# α-расширение – бинарная задача



$$E(\mathbf{y}) = \sum_{p \in \mathcal{V}} E_p(y_p) + \sum_{p,q \in \mathcal{N}} E_{pq}(y_p,y_q)$$
  $\mathbf{y} \in \{1\dots K\}^{|\mathcal{V}|}$   $\mathbf{y}^0$ – текущее решение  $\mathbf{z} \in \{0,1\}^{|\mathcal{V}|}$ 

$$\mathbf{y}^{\alpha}(\mathbf{z}) = \mathbf{y}^{0} \bullet (1 - \mathbf{z}) + \alpha \bullet \mathbf{z}$$

$$E^{\alpha}(\mathbf{z}) = E(\mathbf{y}^{\alpha}(\mathbf{z})) = \sum_{p \in \mathcal{V}} E_p^{\alpha}(z_p) + \sum_{p,q \in \mathcal{N}} E_{pq}^{\alpha}(z_p, z_q)$$

$$E_p(y_p^0)$$
 $E_p(\alpha)$ 

$$E_{pq}(y_p^0, y_q^0)$$
  $E_{pq}(y_p^0, \alpha)$   $E_{pq}(\alpha, y_q^0)$   $E_{pq}(\alpha, \alpha)$ 

#### Алгоритм α-расширения



```
\mathbf{y} = \arg\min_{\mathbf{y}} \sum_{p \in \mathcal{V}} E_p(y_p) + \lambda \sum_{p,q \in \mathcal{N}} E_{pq}(y_p, y_q)
   у = начальное решение;
   do
        change = false;
        for alpha = 1 to K
             y' = Pacuupehue(y, \alpha);
             if y' != y
                   change = true;
             y = y';
        endfor
   while change;
```

- Обычно «почти» сходится за два-три обхода по всем меткам
- Чуть лучший результат для случайного порядка обхода меток

# Условия на парный потенциал



$$E^{\alpha}(\mathbf{z}) = E(\mathbf{y}^{\alpha}(\mathbf{z})) = \sum_{p \in \mathcal{V}} E_p^{\alpha}(z_p) + \sum_{p,q \in \mathcal{N}} E_{pq}^{\alpha}(z_p, z_q)$$

Достаточное условие применимости разрезов графов:

$$E_{pq}(y_p^0, y_q^0)$$
  $E_{pq}(y_p^0, \alpha)$   $E_{pq}(\alpha, y_q^0)$   $E_{pq}(\alpha, \alpha)$ 

$$E_{pq}^{\alpha}(0,0) + E_{pq}^{\alpha}(1,1) \le E_{pq}^{\alpha}(0,1) + E_{pq}^{\alpha}(1,0)$$
$$E_{pq}(y_p^0, y_q^0) + E_{pq}(\alpha, \alpha) \le E_{pq}(y_p^0, \alpha) + E_{pq}(\alpha, y_q^0)$$

## Условия на парный потенциал



$$E_{pq}(y_p^0, y_q^0) + E_{pq}(\alpha, \alpha) \le E_{pq}(y_p^0, \alpha) + E_{pq}(\alpha, y_q^0)$$

Чтобы это выполнялось, парный потенциал должен быть метрикой:

- Симметричность E(a,b) = E(b,a)
- Аксиома тождества  $E(a,b) = 0 \Leftrightarrow a = b$
- Неравенство треугольника

$$E_{pq}(a,b) \le E_{pq}(a,c) + E_{pq}(c,b)$$

Для метрических потенциалов α- расширение сходится, но глобальность минимума не гарантируется

## Примеры метрических потенциалов



Достаточное условие:  $E_{pq}(a,a) = 0$ 

$$E_{pq}(a,b) \le E_{pq}(a,c) + E_{pq}(c,b)$$

Какие из указанных потенциалов удовлетворяют?:

$$E_{pq}(a,b) = \lambda[a \neq b]$$
 $E_{pq}(a,b) = \lambda|a-b|$ 
 $E_{pq}(a,b) = \lambda \min(|a-b|,T)$ 
 $E_{pq}(a,b) = \lambda(a-b)^2$ 

## Актуальные вопросы



- Как работать с факторами высоких порядков?
- Как правильно задавать структуру графа?
- Как настраивать все параметры модели?

### Резюме лекции



- Графические модели позволяют выразить общее решение для Y через локальные зависимости
- Дискретные модели позволяют описать очень многие задачи компьютерного зрения
- Для бинарных задач есть точные алгоритмы решения (разрез графов)
- Для дискретных задач с К>2 есть приближенные итерационные алгоритмы решения