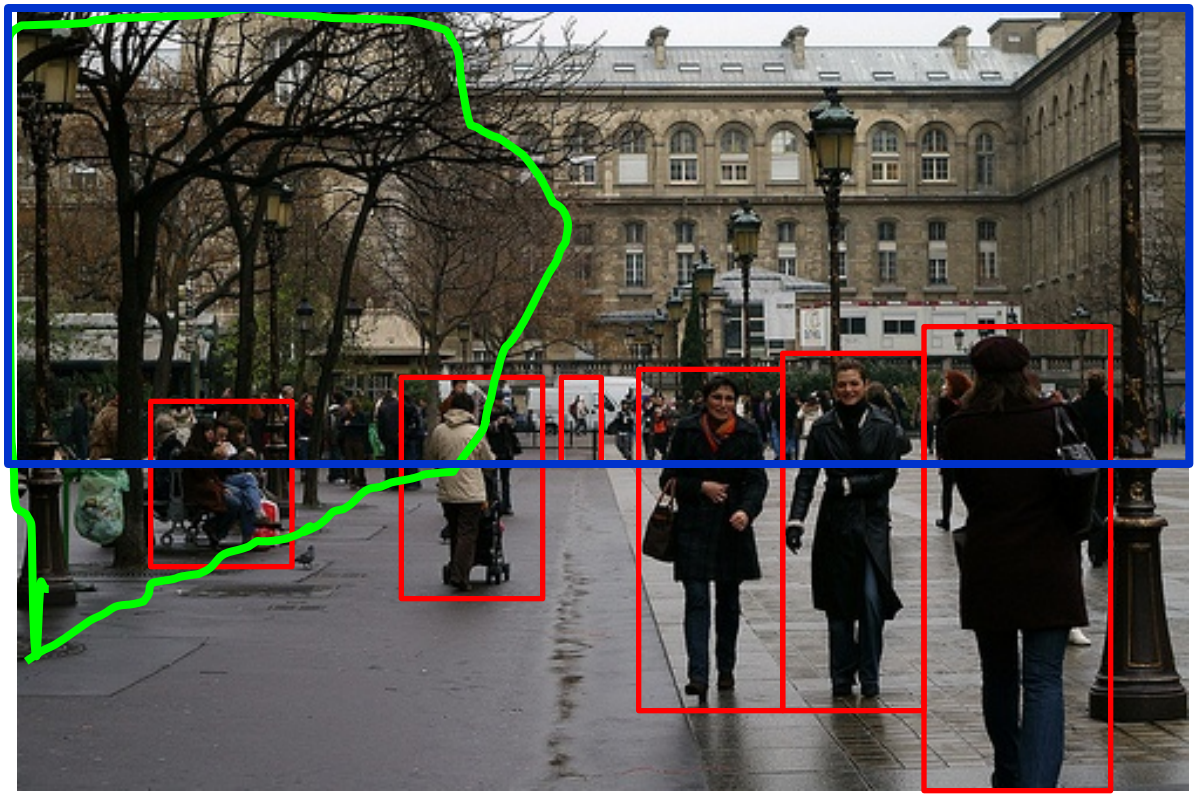
Оптимизация энергии в задачах компьютерного зрения и алгоритмы на графах

Мини-курс, Computer Science Club, Санкт-Петербург, 2010

Виктор Лемпицкий

Microsoft Research

Что такое компьютерное зрение?



Изображение сцены

Компьютерная графика

Компьютерное зрение

Описание сцены

Из истории вопроса...

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Subgoal for July

Analysis of s ... the following set:

balls

brick ... textures

cylinders.

Each face will be of uniform and distinct color and/or texture.

Background will be homogeneous.

План на август:

1. ~~Сегментация на объекты~~
2. Распознавание отдельных объектов

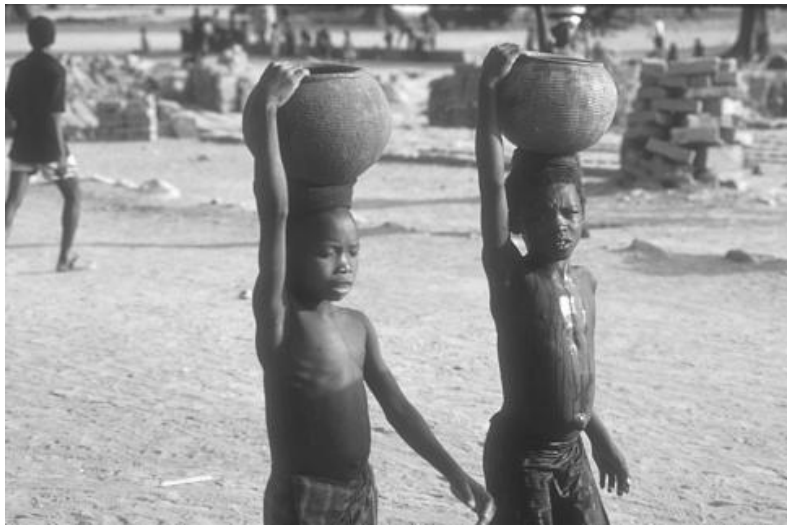
Extensions for August

The first priority will be to handle objects of the same sort but with complex surfaces and backgrounds, e.g. cigarette pack with writing and bands of different color, or a cylindrical battery.

Then extend class of objects to objects like tools, cups, etc.

Сегментация

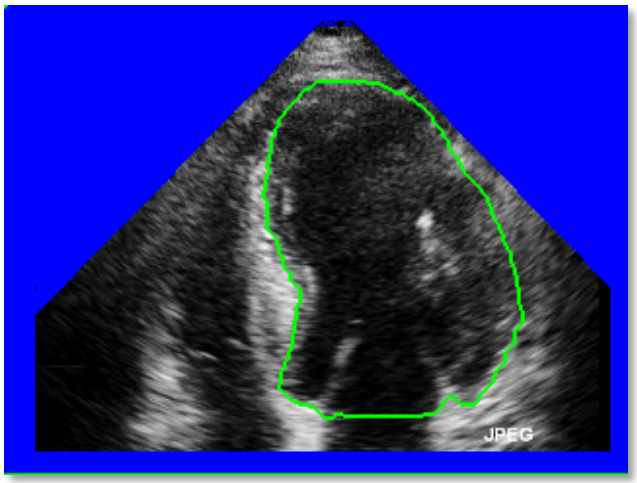
Пример из Berkeley Segmentation Dataset:

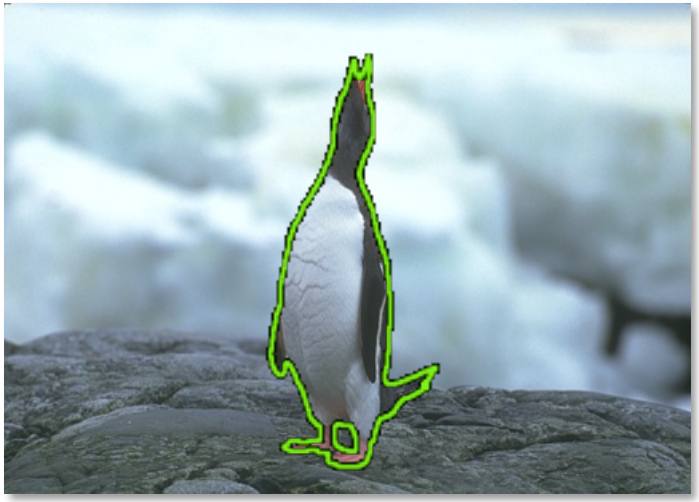


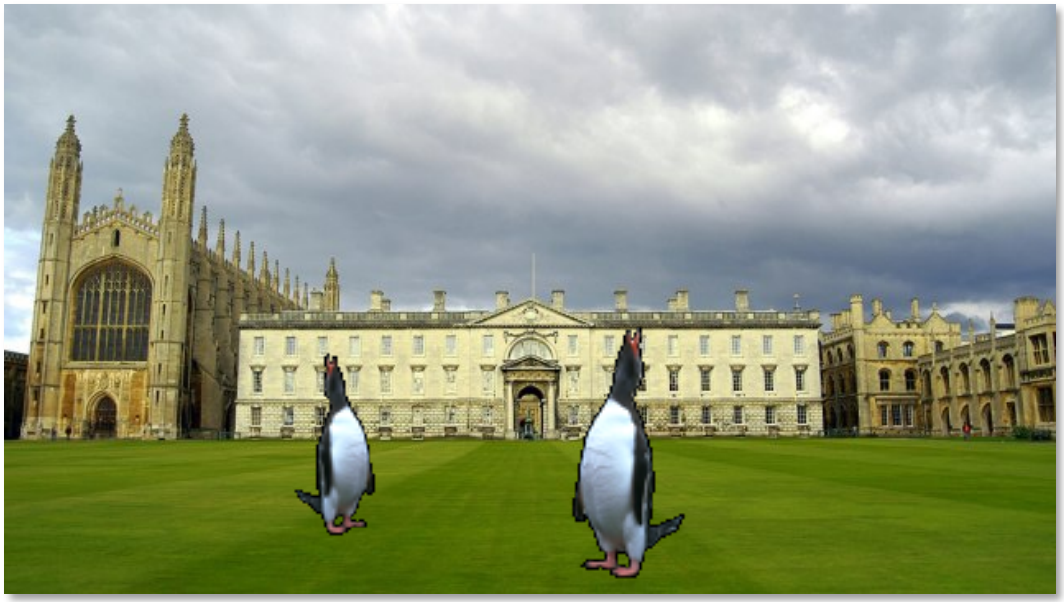
Оба ответа верные! Задача плохо определена... ☹

Бинарная сегментация

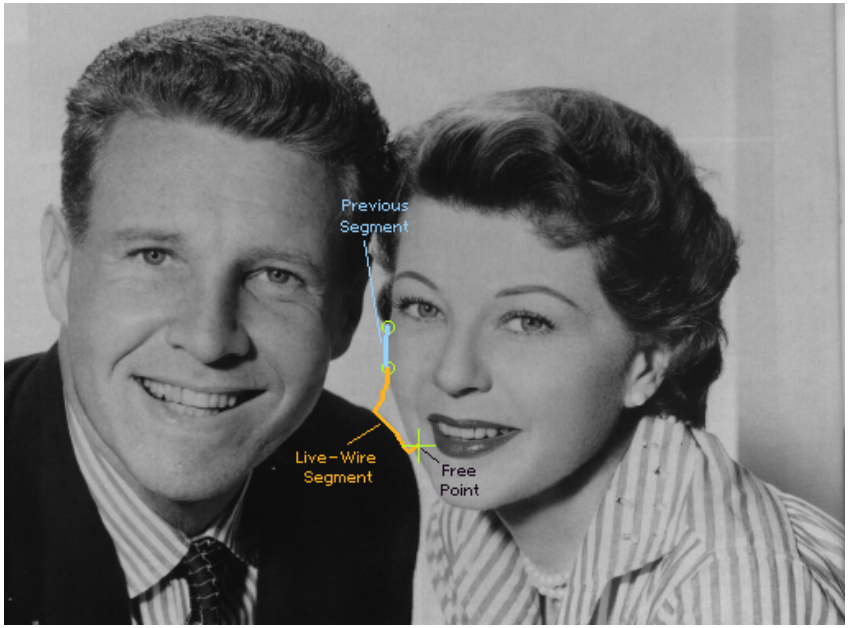
- ...или разбиение на фон/объект

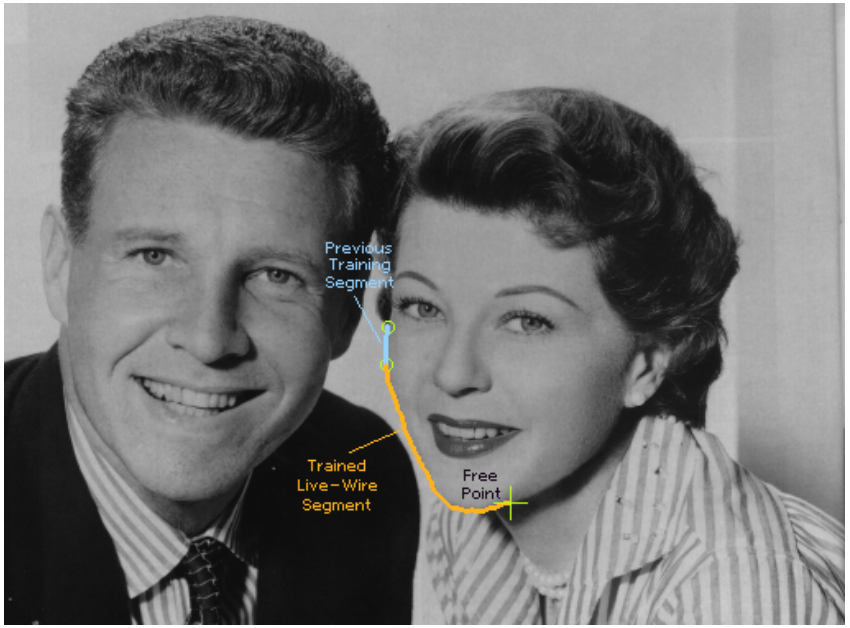






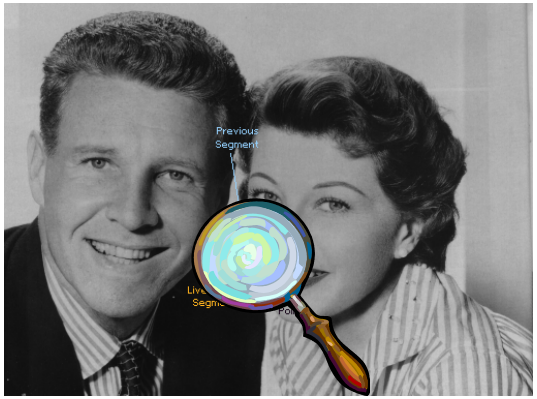
“Умные ножницы”
[Mortensen & Barett, 1995]

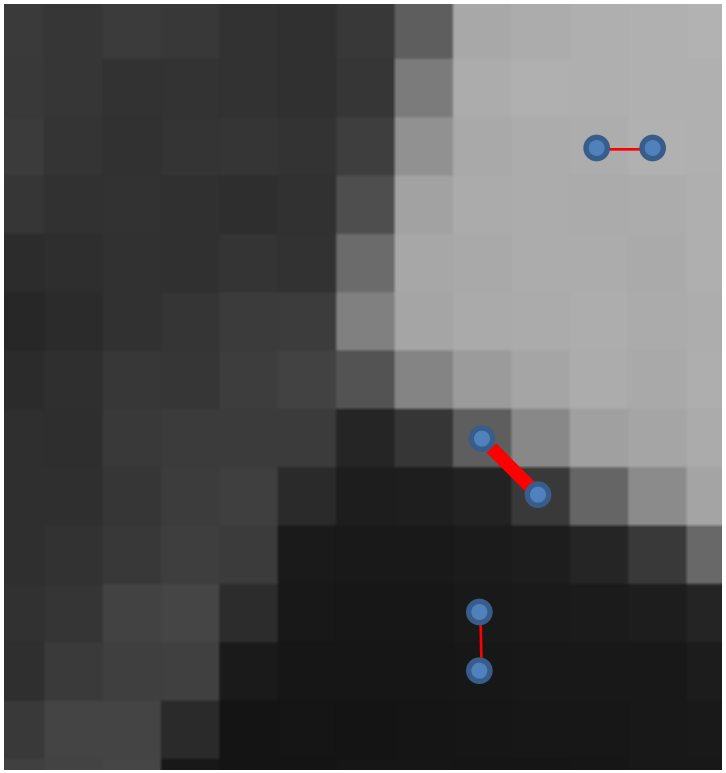


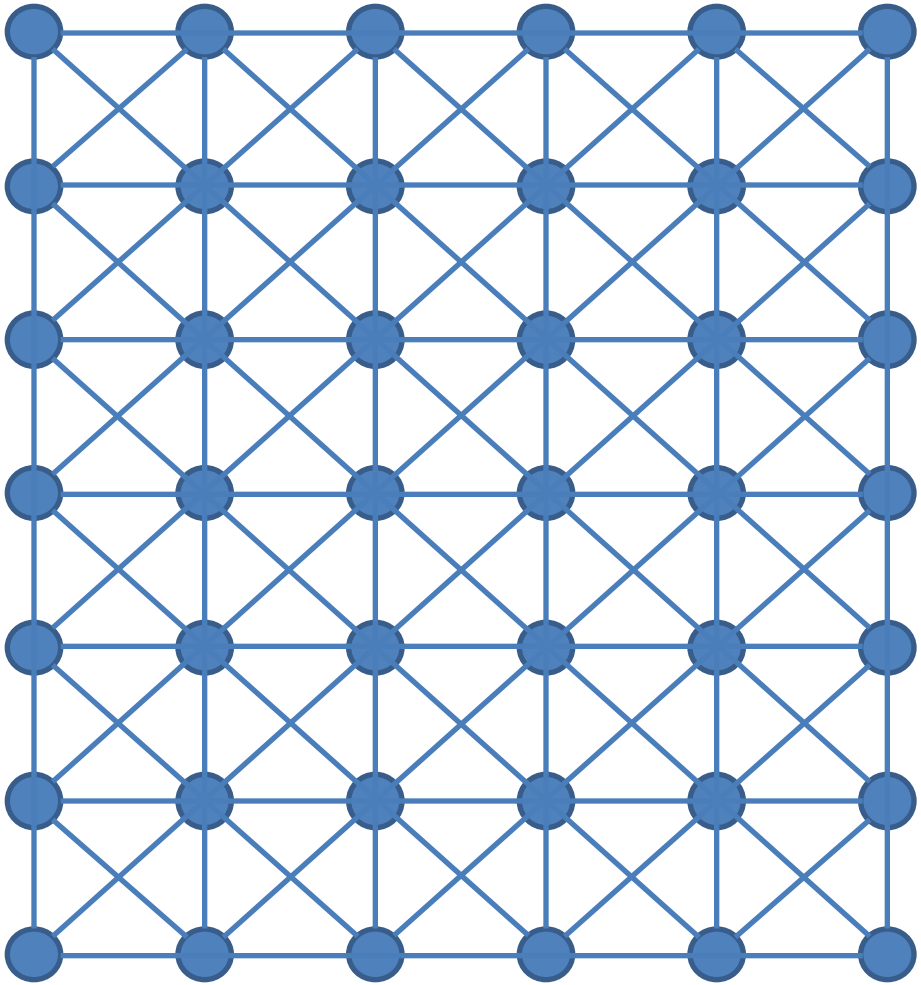


aka “magnetic lasso” (Photoshop), “livewire”

От картинки к графу
[Mortensen & Barett, 1995]



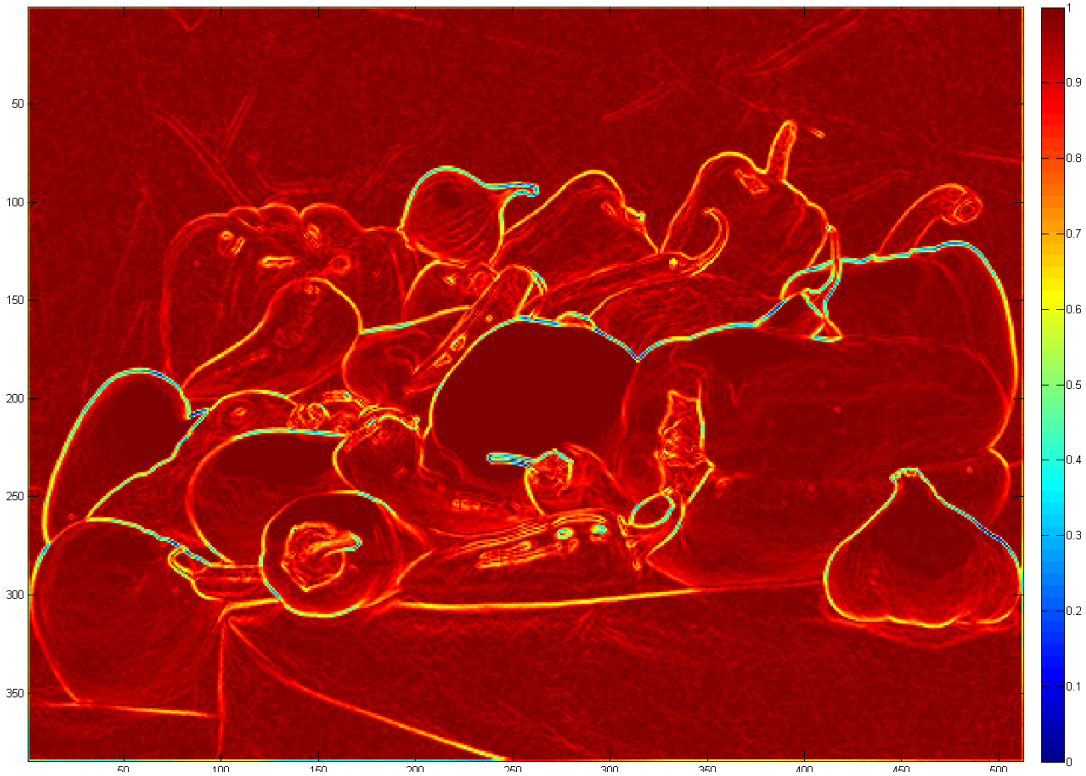




$$l(\boldsymbol{p}, \boldsymbol{q}) = \omega_Z \cdot f_Z(\boldsymbol{q}) + \omega_G \cdot f_G(\boldsymbol{q}) + \omega_D \cdot f_D(\boldsymbol{p}, \boldsymbol{q})$$

Выбор весов для графа [Mortensen & Barett, 1995]

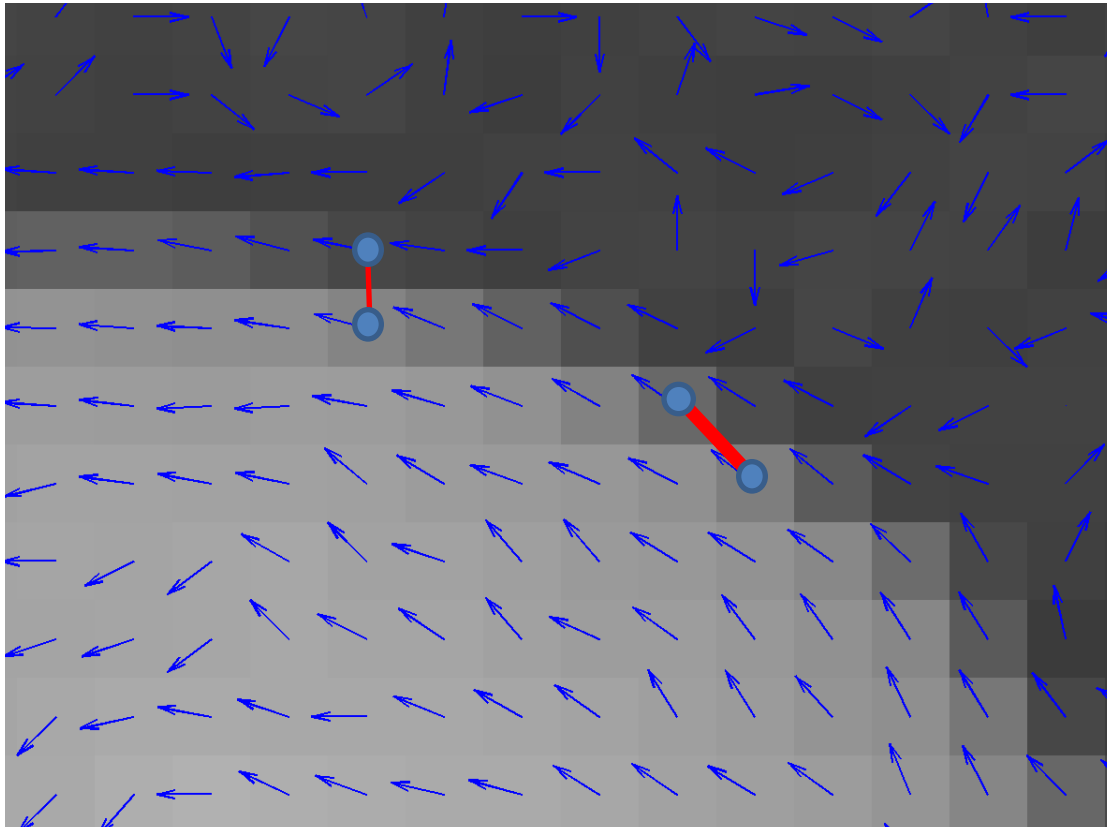
$$l(\boldsymbol{p}, \boldsymbol{q}) = \omega_Z \cdot f_Z(\boldsymbol{q}) + \omega_G \cdot f_G(\boldsymbol{q}) + \omega_D \cdot f_D(\boldsymbol{p}, \boldsymbol{q})$$



$$f_G = \frac{\max(G) - G}{\max(G)} = 1 - \frac{G}{\max(G)}$$

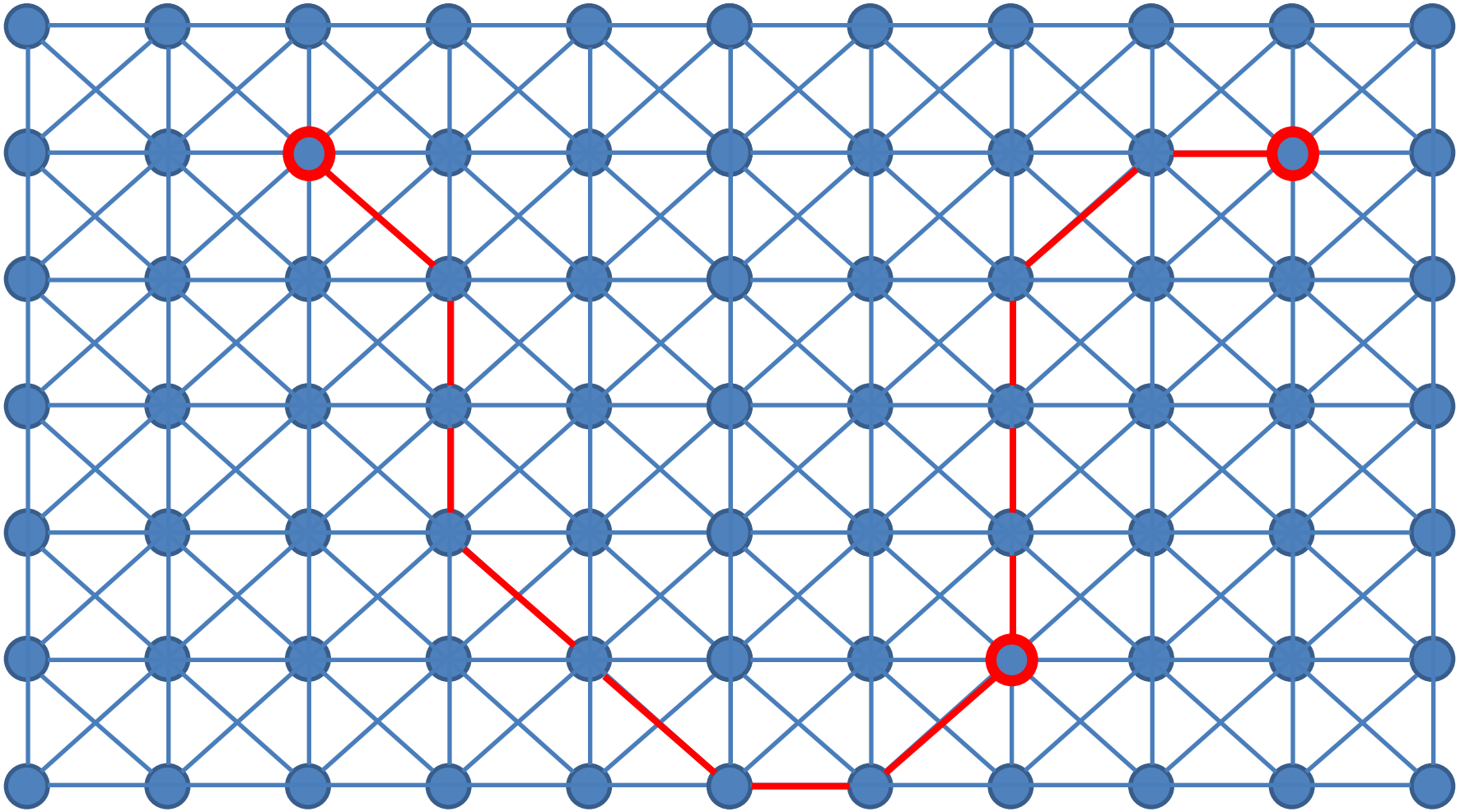
Выбор весов для графа [Mortensen & Barett, 1995]

$$l(\boldsymbol{p}, \boldsymbol{q}) = \omega_Z \cdot f_Z(\boldsymbol{q}) + \omega_G \cdot f_G(\boldsymbol{q}) + \omega_D \cdot f_D(\boldsymbol{p}, \boldsymbol{q})$$



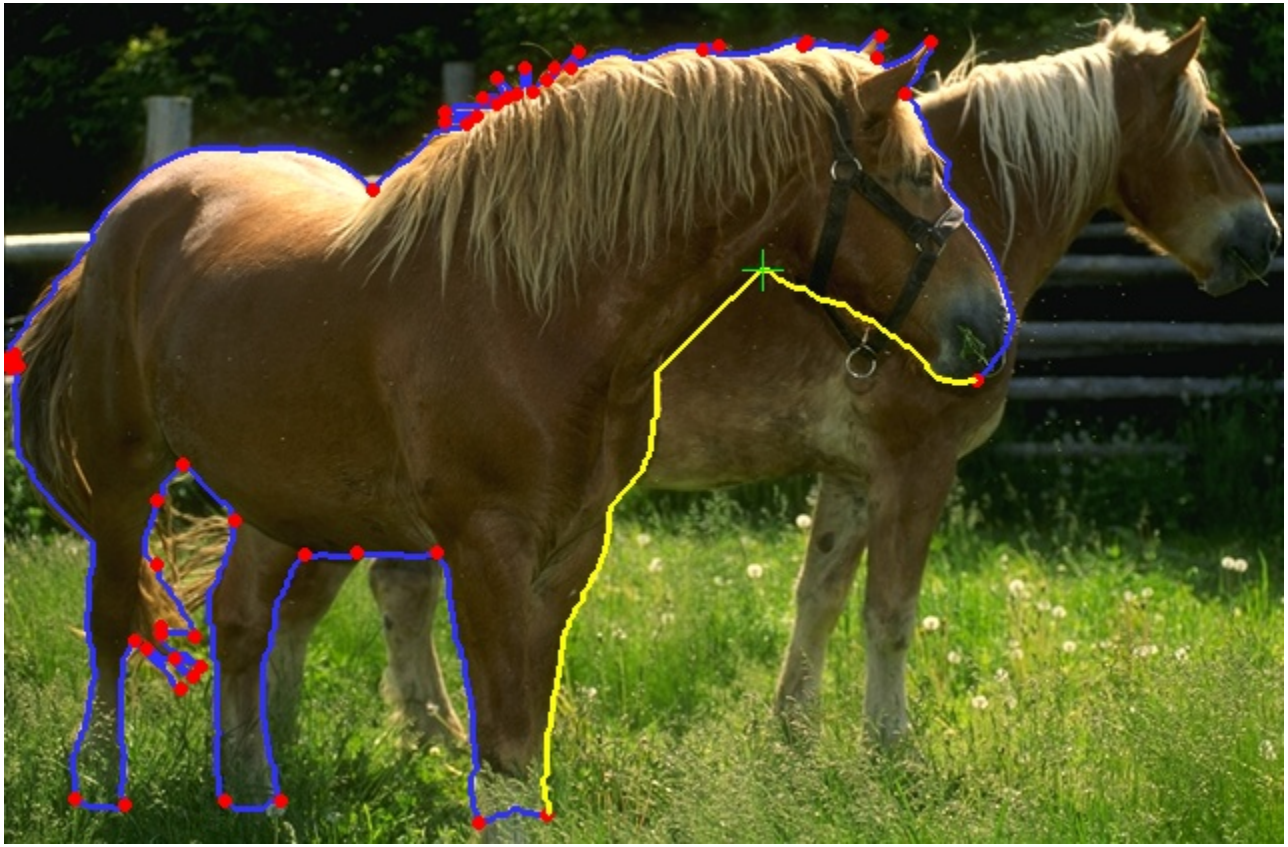
$$f_D(\boldsymbol{p}, \boldsymbol{q}) = \frac{2}{3\pi}\{\operatorname{acos}[d_{\boldsymbol{p}}(p, q)] + \operatorname{acos}[d_{\boldsymbol{q}}(p, q)]\}$$

Сегментация кратчайшими путями [Mortensen & Barett, 1995]



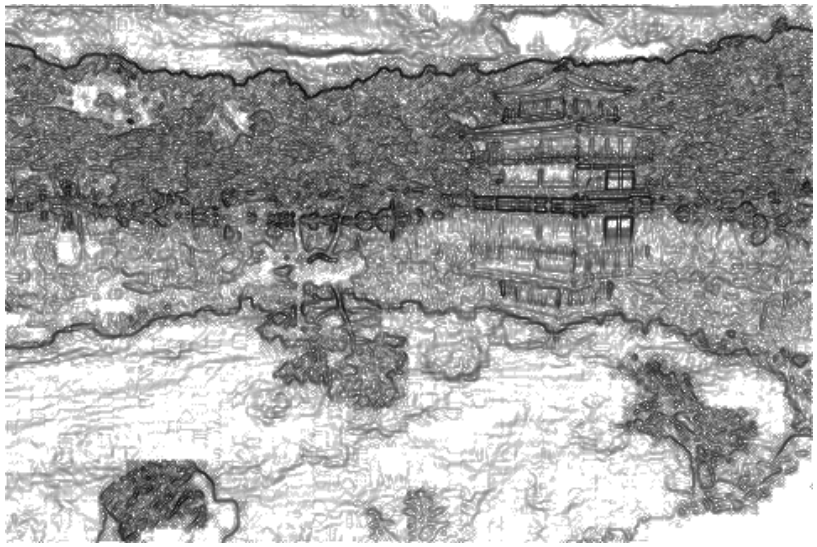
Граница = цепь кратчайших путей

“Умные ножницы”
[Mortensen & Barett, 1995]

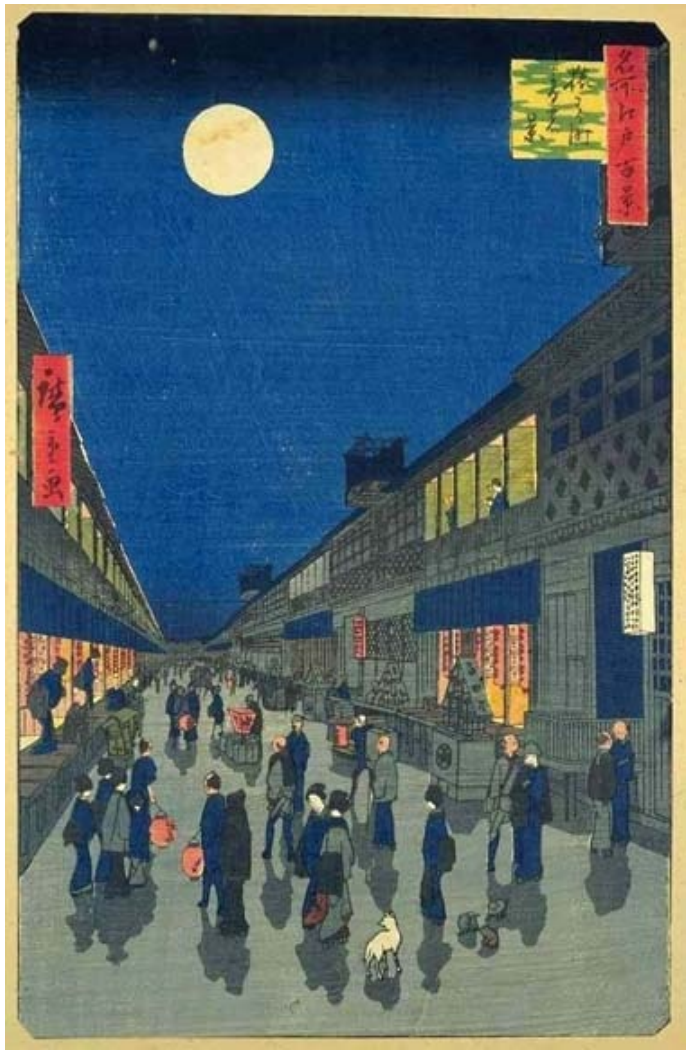


images from [Mortensen&Barett,1999]

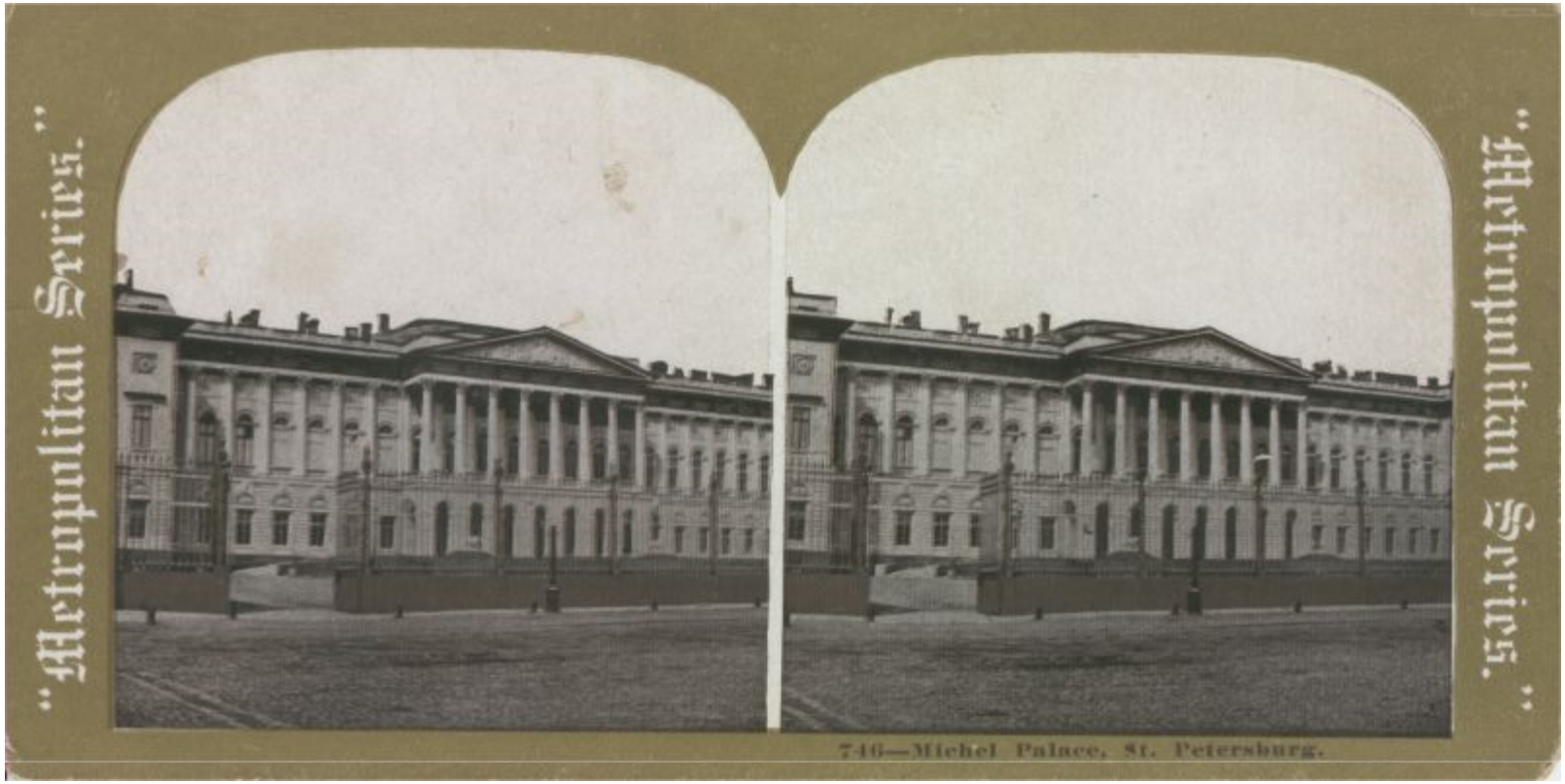
Seam Carving: изменение размера [Avidan & Shamir, 2007]



Seam Carving
[Avidan & Shamir, 2007]



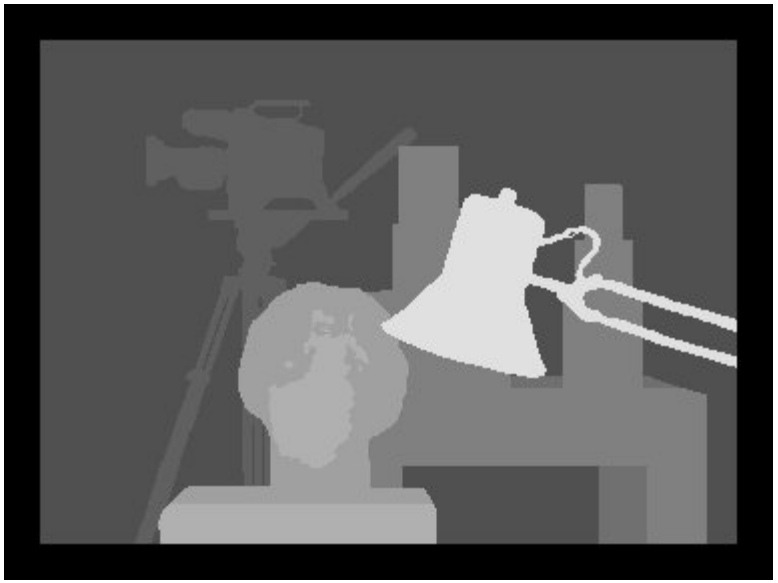
Стереопары



from Agnes Svoboda Morris collection

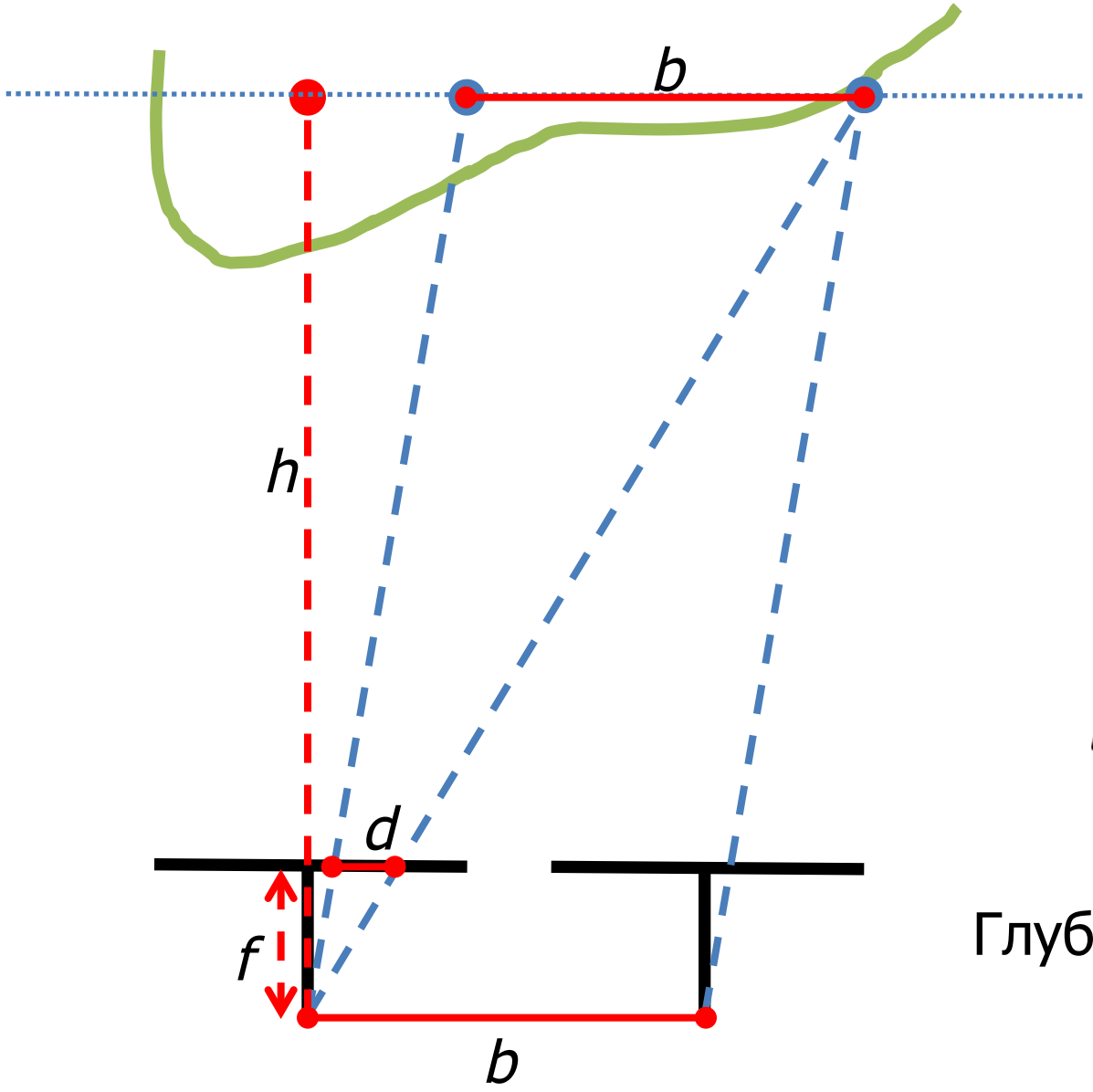
Стереосопоставление

Тестовая стереопара
(University of Tsukuba)



Карта смещений

Немного геометрии

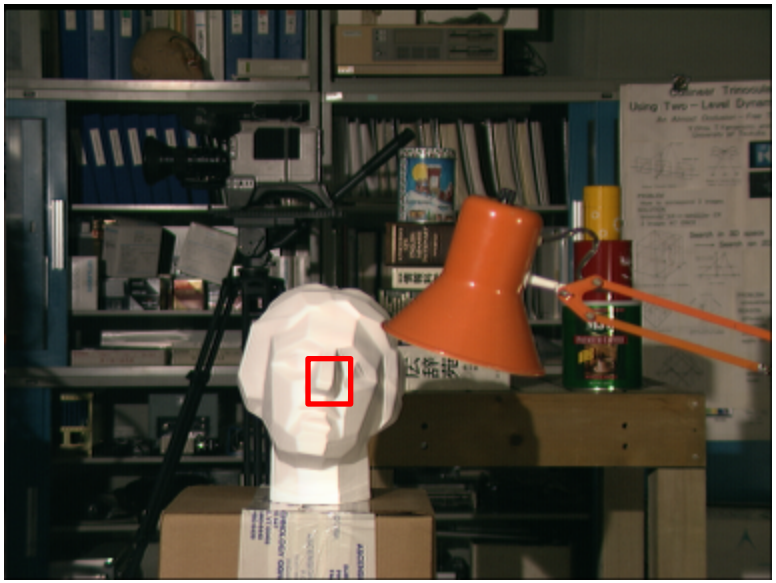


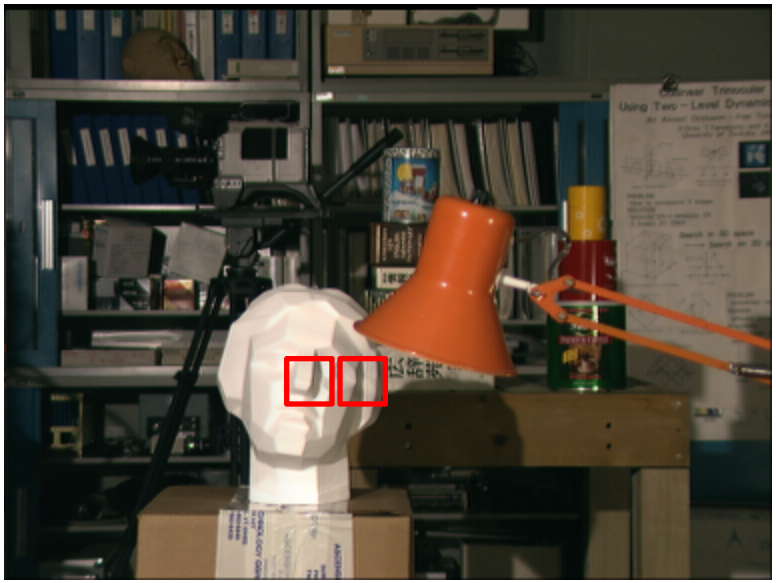
$$\frac{d}{b} = \frac{f}{h}$$

$$h = bf\,d^{-1}$$

Глубина $\propto$ 1/Смещение

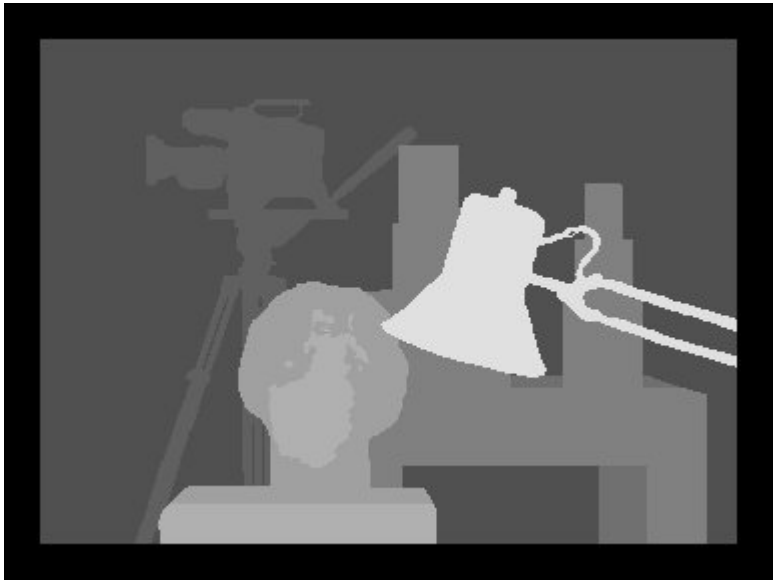
Стерео: локальный подход

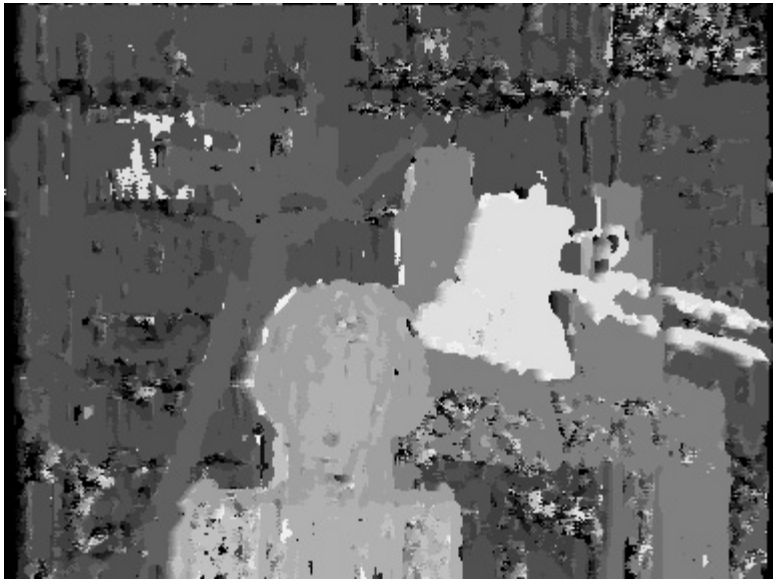


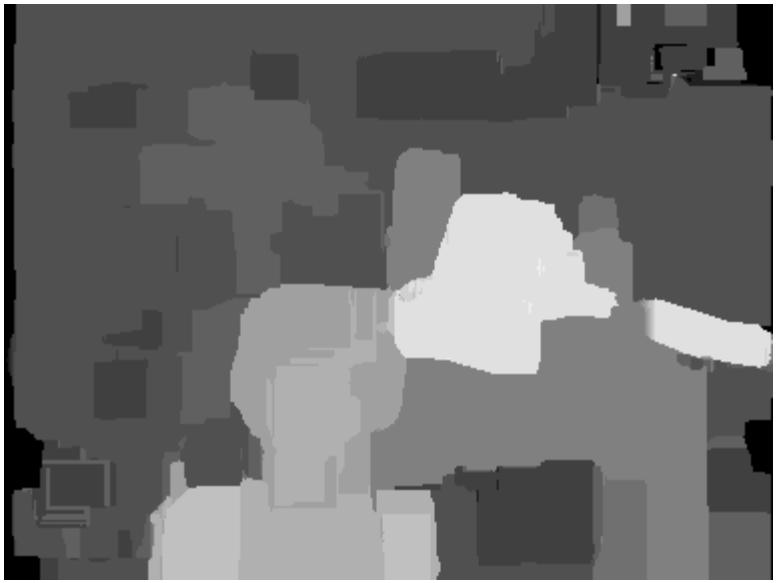


$$d(x,y) = \underset{d}{argmin} \sum_{i=-K}^{K} \sum_{j=-K}^{K} ||L(x+i,y+j) - R(x-d+i,y+j)||_2^2$$

Стерео: локальный подход



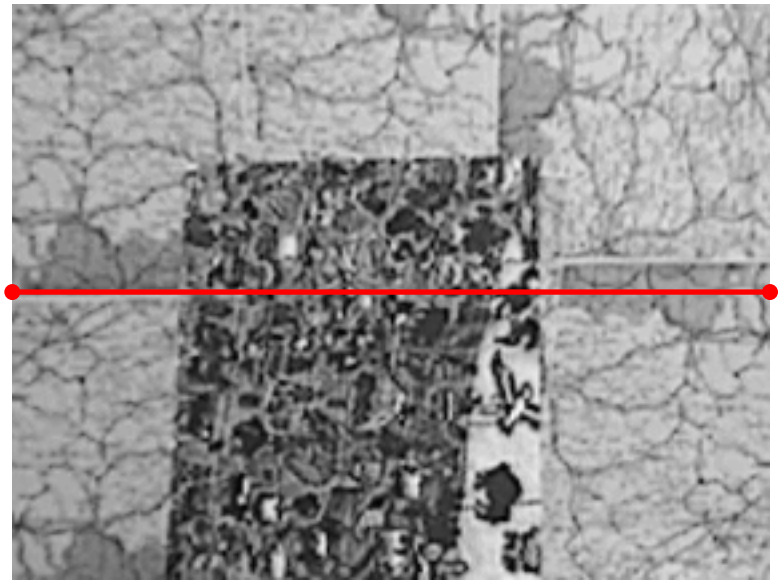


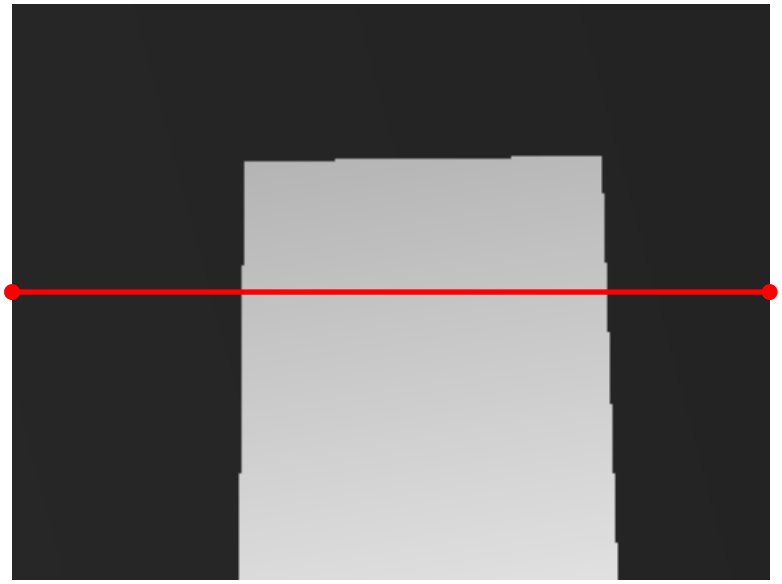


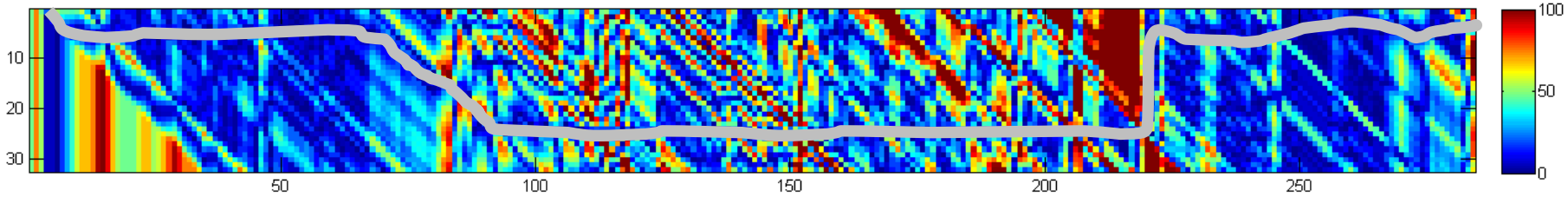
Результаты

Сканлайн

Stereopair from Middlebury stereo webpage

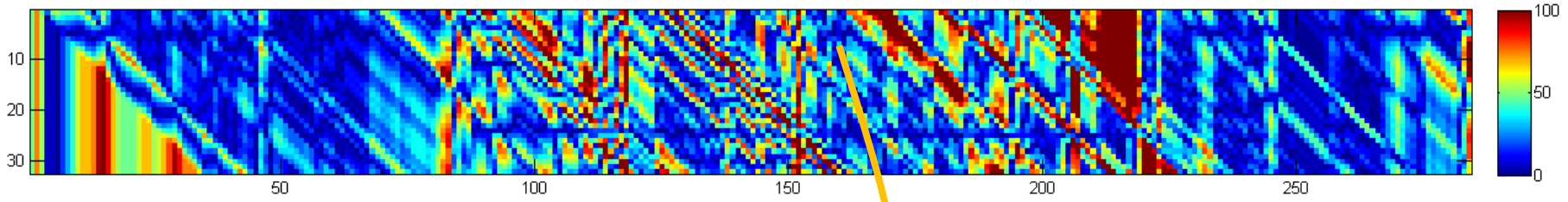


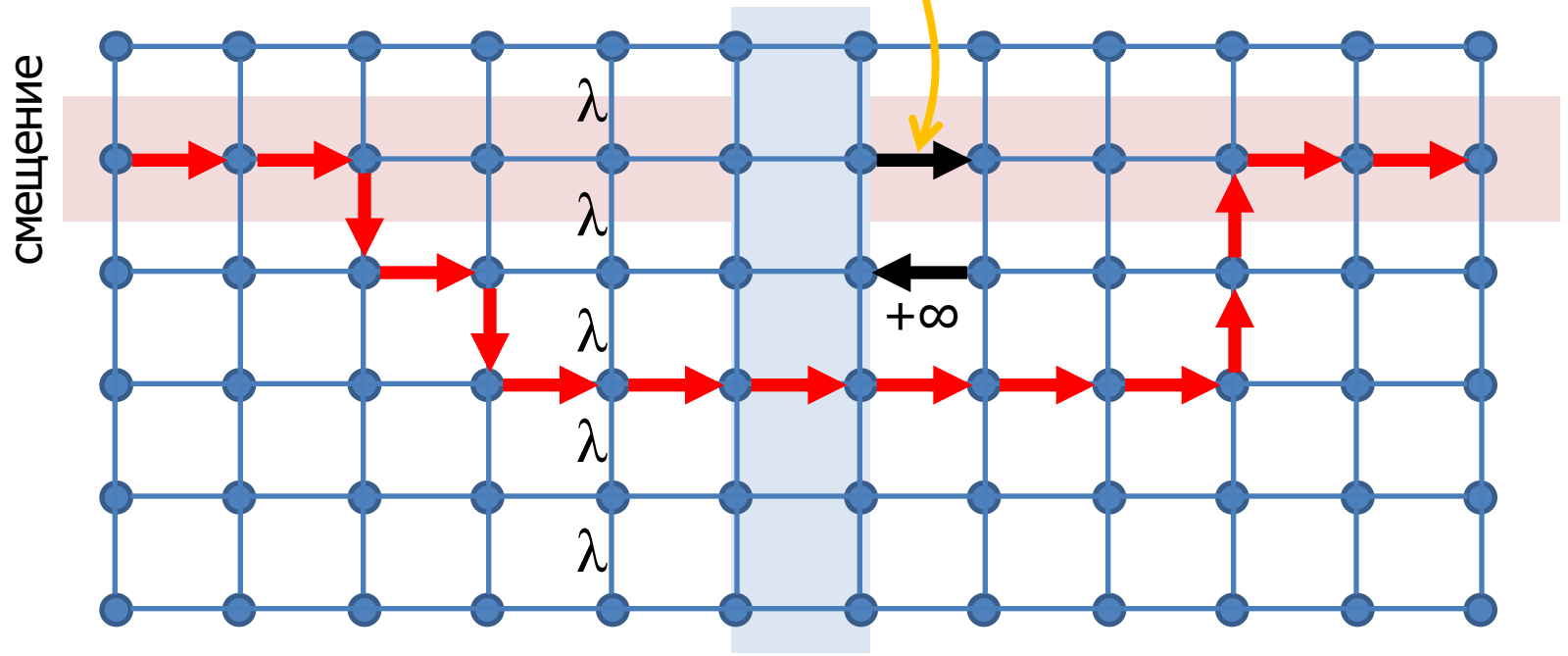




$$f(x,d) = |L(x,y_0) - R(x-d,y_0)|$$

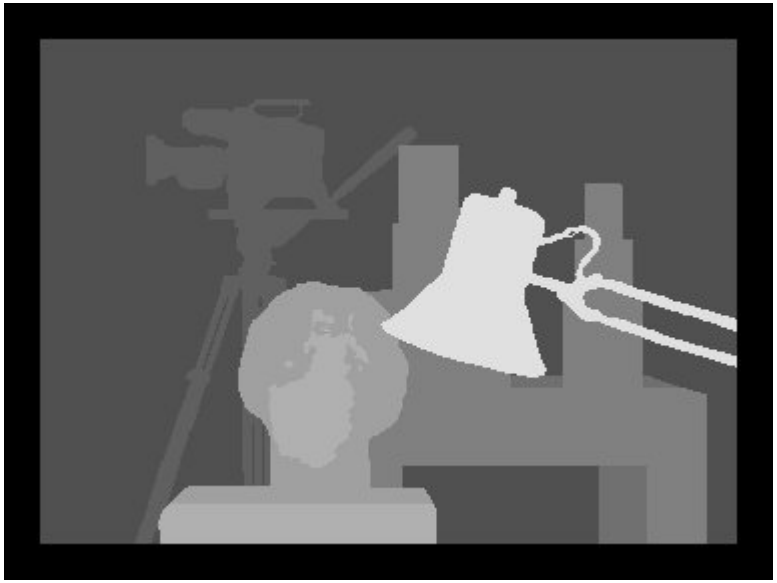
От стерео к кратчайшим путям





$$\{d_1, \ldots, d_W\} = \underset{d_1,\ldots,d_W}{argmin} \sum_{i=1}^{W} |L(x_i, y_0) - R(x_i - d_i, y_0)| + \lambda \sum_{i=1}^{W-1} |d_i - d_{i+1}|$$

Стерео и энергии



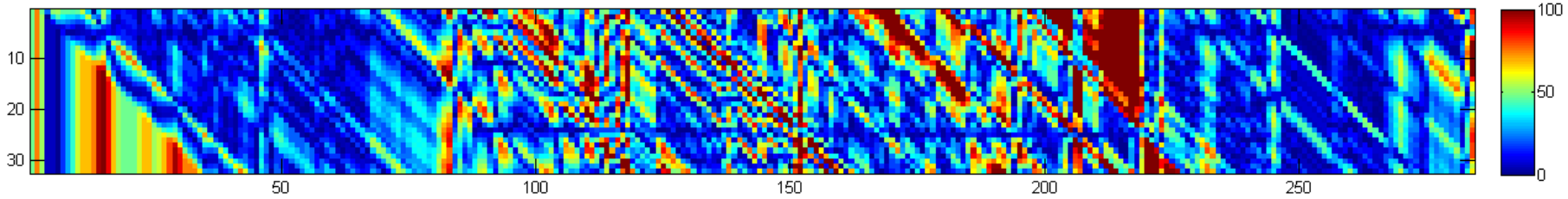
Построили алгоритм для:

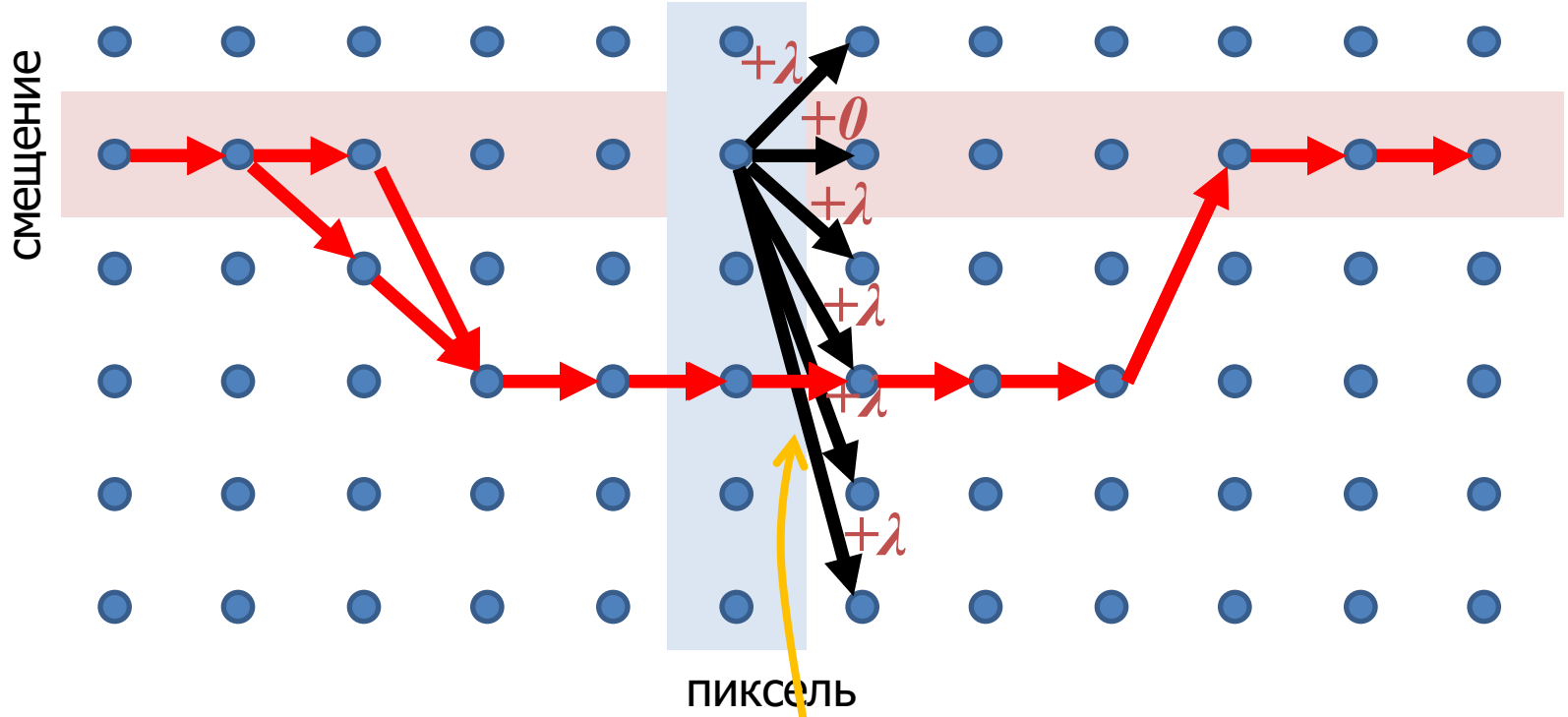
$$\{d_1, \ldots, d_W\} = \underset{d_1,\ldots,d_W}{argmin} \sum_{i=1}^{W} |L(x_i, y_0) - R(x_i - d_i, y_0)| + \lambda \sum_{i=1}^{W-1} |d_i - d_{i+1}|$$

Хотим:

$$\{d_1, \ldots, d_W\} = \underset{d_1,\ldots,d_W}{argmin} \sum_{i=1}^{W} |L(x_i, y_0) - R(x_i - d_i, y_0)| + \lambda \sum_{i=1}^{W-1} [d_i \neq d_{i+1}]$$

Новая конструкция графа

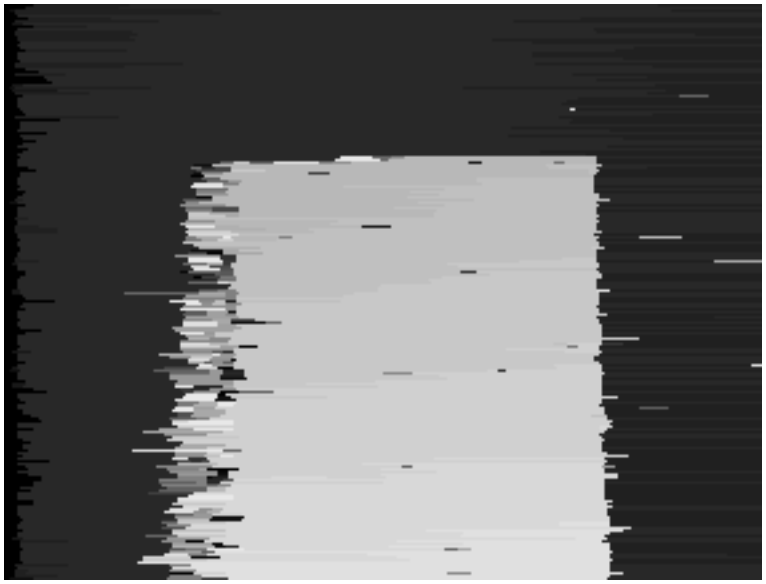




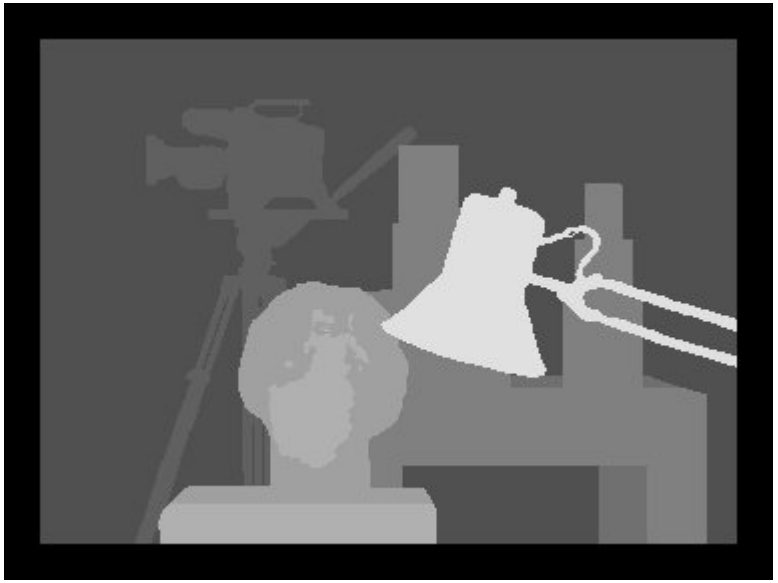
$$\{d_1, \ldots, d_W\} = \underset{d_1,\ldots,d_W}{argmin} \sum_{i=1}^{W} |L(x_i, y_0) - R(x_i - d_i, y_0)| + \lambda \sum_{i=1}^{W-1} [d_i \neq d_{i+1}]$$

Результаты

Оптимизация вдоль сканлайна



Ground truth:



Динамическое программирование

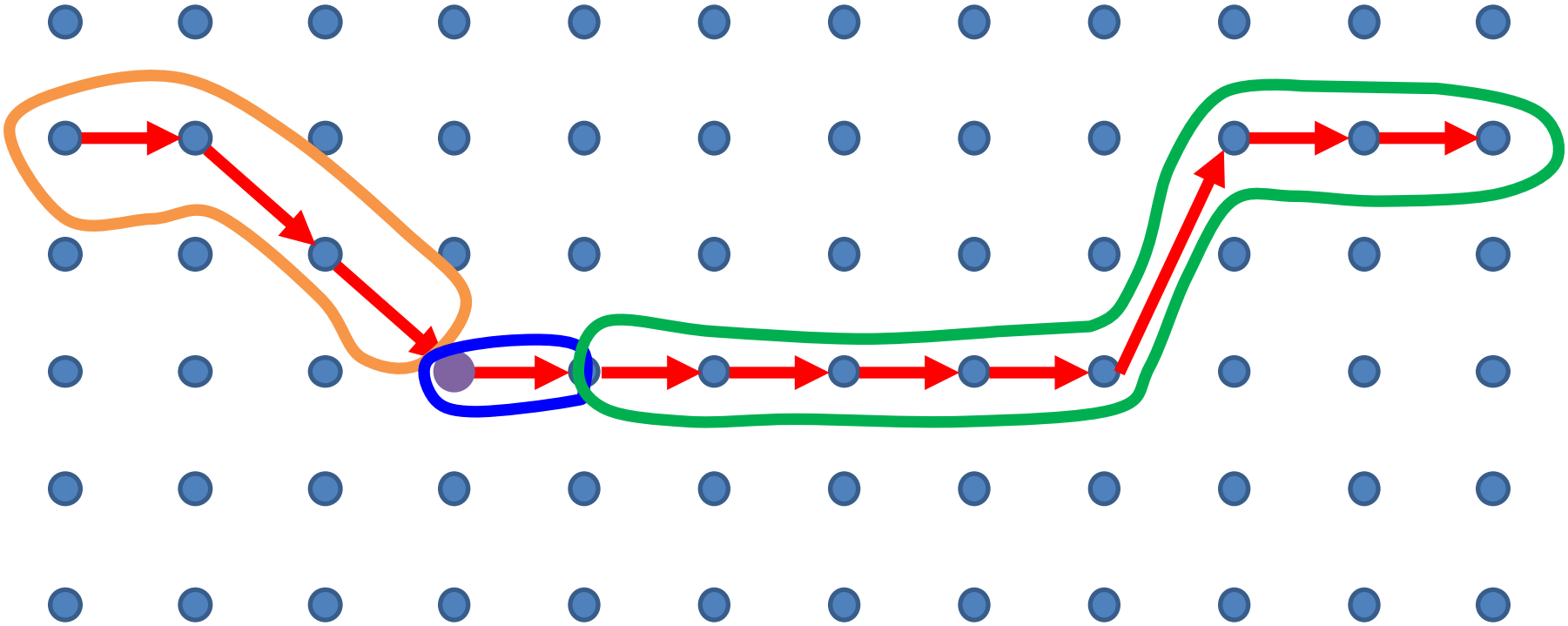
$$\{d_1, \dots, d_W\} = \underset{d_1,\dots,d_W}{argmin} \sum_{i=1}^{W} |L(x_i, y_0) - R(x_i - d_i, y_0)| + \lambda \sum_{i=1}^{W-1} [d_i \neq d_{i+1}]$$

$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1})$$

$$U_i(d_i) = |L(x_i, y_0) - R(x_i - d_i, y_0)|$$

$$V(d_i, d_{i+1}) = \lambda [d_i \neq d_{i+1}]$$

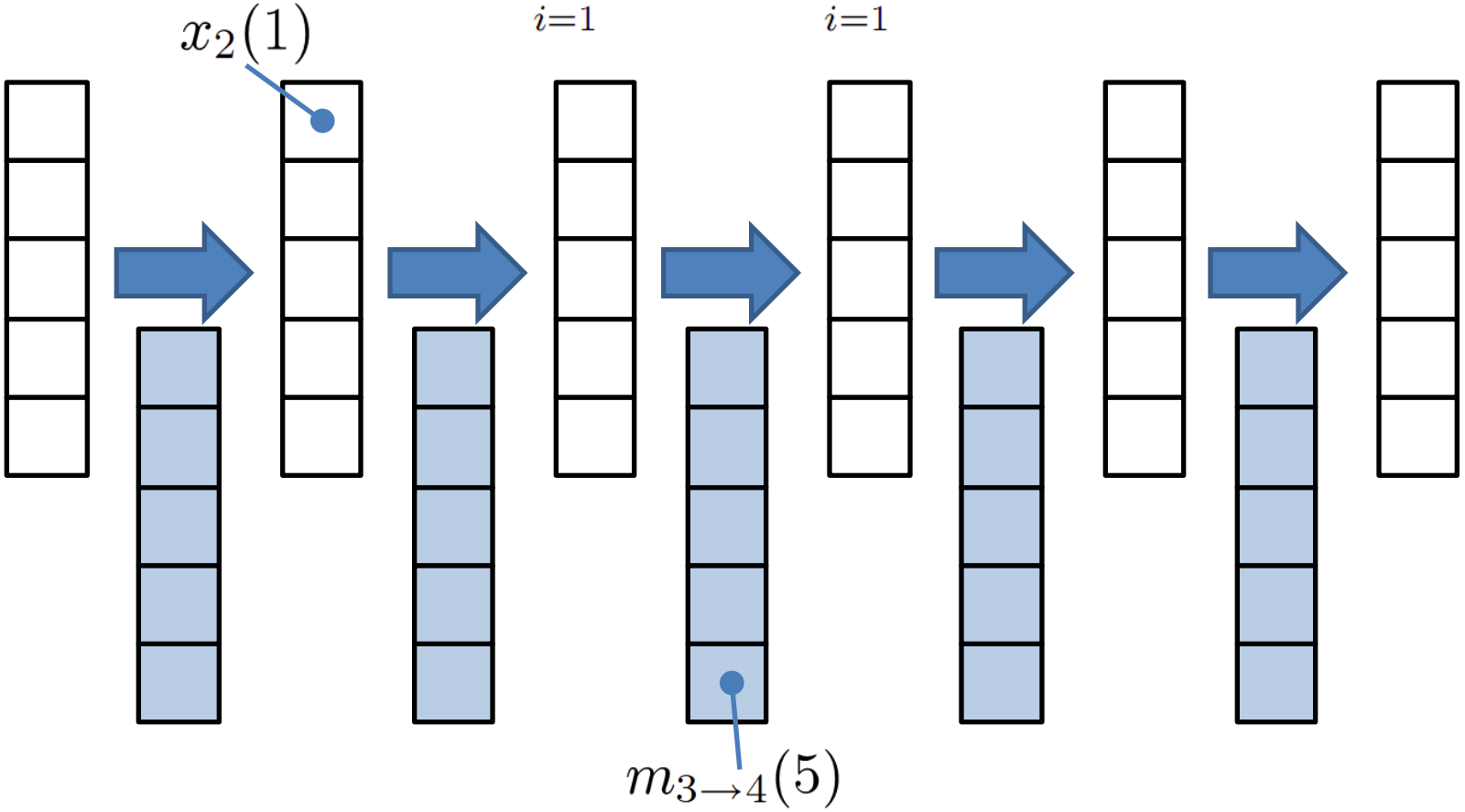
Динамическое программирование



$$E(\mathbf{d}) = E_1^{t-1}(d_1, \ldots, d_{t-1}) + V(d_{t-1}, d_t) + U_t(d_t) + V(d_t, d_{t+1}) + E_{t+1}^{W}(d_{t+1}, \ldots, d_W)$$

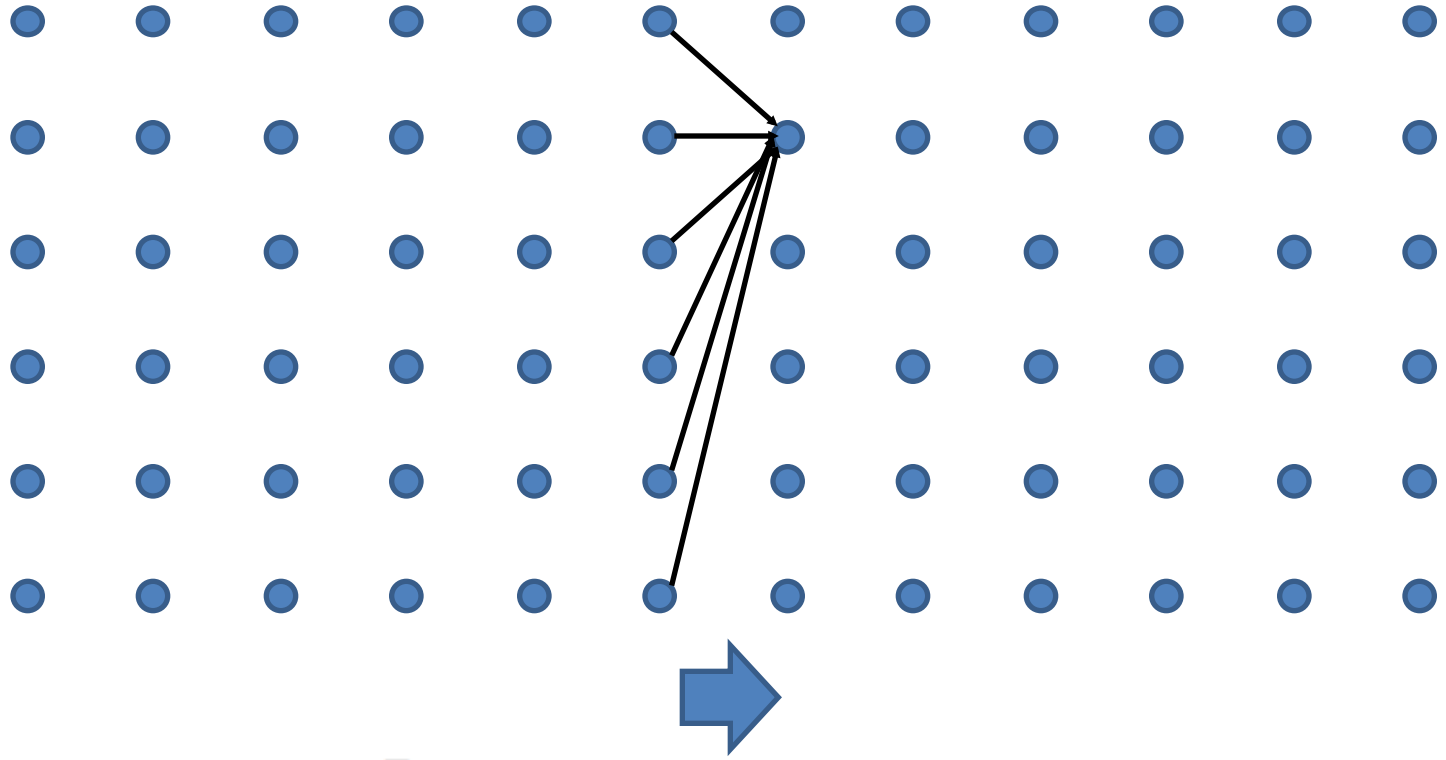
Передача сообщений

$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1})$$



$$m_{i\to i+1}(j) = \min_{k=1}^{D} \left(m_{i-1\to i}(k) + U_i(k) + V(k,j) \right)$$

Передача сообщений



$$m_{i \to i+1}(j) = \min_{k=1}^{D} \left(m_{i-1 \to i}(k) + U_i(k) + V(k, j) \right)$$

Передача сообщений

$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1})$$

$$E_a^b(d_a, \ldots, d_b) = \sum_{i=a}^{b} U_i(d_i) + \sum_{i=a}^{b-1} V(d_i, d_{i+1})$$

$$m_{i \to i+1}(j) = \min_{k=1}^{D} \left(m_{i-1 \to i}(k) + U_i(k) + V(k, j) \right)$$

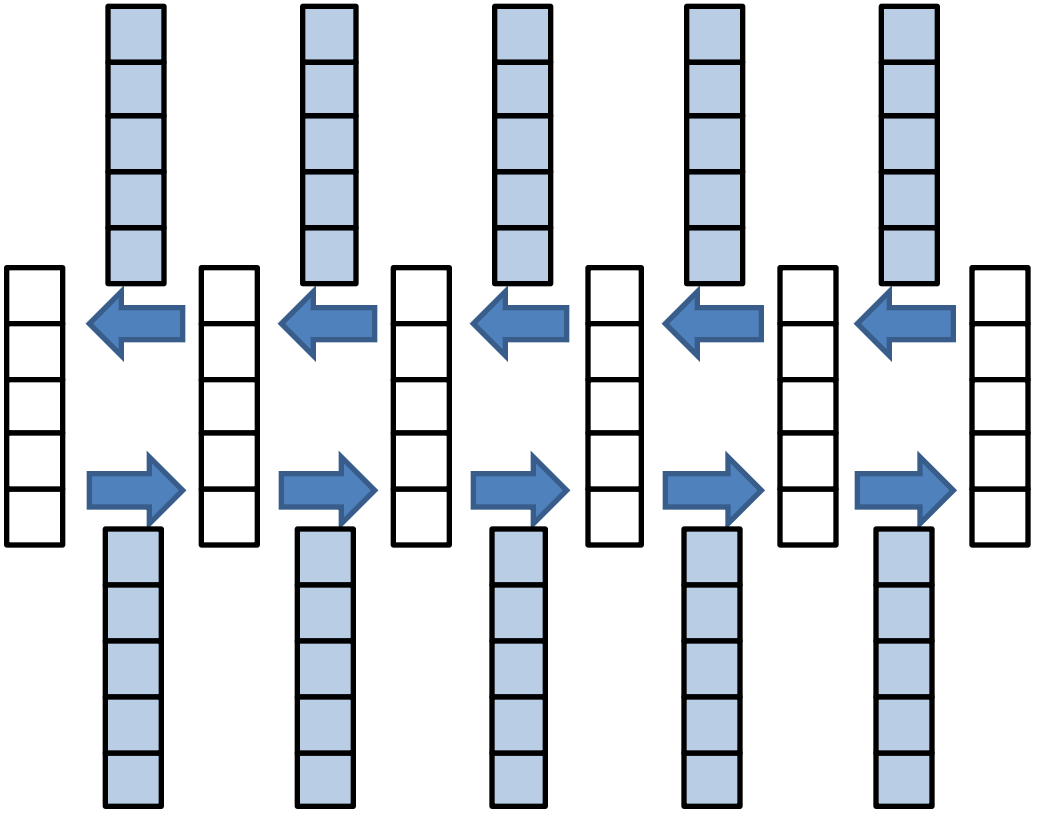
$$m_{0 \to 1}(\cdot) = 0$$

$$m_{t-1 \to t}(d_t) = \underset{d_1, \ldots d_{t-1}}{argmin} \left(E_1^{t-1}(d_1, \ldots, d_{t-1}) + V(d_{t-1}, d_t) \right)$$

Док-во: индукция

Передача сообщений

$$m_{i \to i-1}(j) = \min_{k=1}^{D} \left(m_{i+1 \to i}(k) + U_i(k) + V(k,j) \right)$$



$$m_{i \to i+1}(j) = \min_{k=1}^{D} \left(m_{i-1 \to i}(k) + U_i(k) + V(k,j) \right)$$

Распространение сообщений

$$E(\mathbf{d}) = E_1^{t-1}(d_1, \ldots, d_{t-1}) + V(d_{t-1}, d_t) + U_t(d_t) + V(d_{t+1}, d_t) + E_{t+1}^{W}(d_{t+1}, \ldots, d_W)$$

$$m_{t-1 \to t}(d_t) = \underset{d_1, \ldots d_{t-1}}{argmin} \left(E_1^{t-1}(d_1, \ldots, d_{t-1}) + V(d_{t-1}, d_t) \right)$$

$$m_{t+1 \to t}(d_t) = \underset{d_{t+1}, \ldots d_W}{argmin} \left(E_{t+1}^{W}(d_{t+1}, \ldots, d_W) + V(d_t, d_{t+1}) \right)$$

$$m_{t-1 \to t}(d_t) + U_t(d_t) + m_{t+1 \to t}(d_t) = \underset{\mathbf{d}, \mathbf{d}(t) = d_t}{argmin} E(\mathbf{d})$$

Вывод: распространение сообщений находит минимум энергии – в отсутствии «ничьих» просто берем оптимальное d_t в каждой вершине.

Стоимость передачи сообщения

Задача: сколько операций надо для подсчета сообщения в нашем случае

$$m_{i \to i+1}(j) = \min_{k=1}^{D} \left(m_{i-1 \to i}(k) + U_i(k) + V(k, j) \right)$$

$$V(i, j) = \lambda \cdot [i \neq j]$$

$$V(i, j) = \lambda \cdot \min(1, \frac{|i - j|}{K})$$

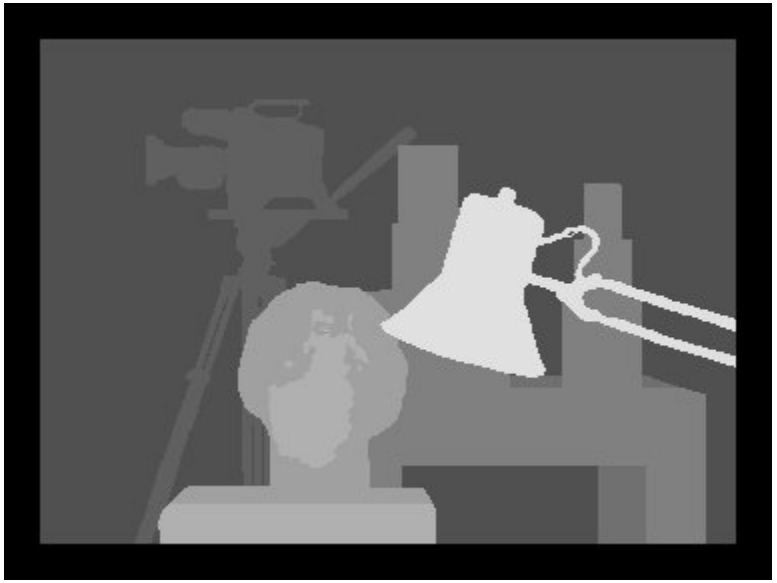
А в этих случаях?:

$$V(i, j) = \lambda \cdot \min(1, \frac{(i - j)^2}{K^2})$$

Что получилось

Оптимизация вдоль сканлайна

Ground truth:



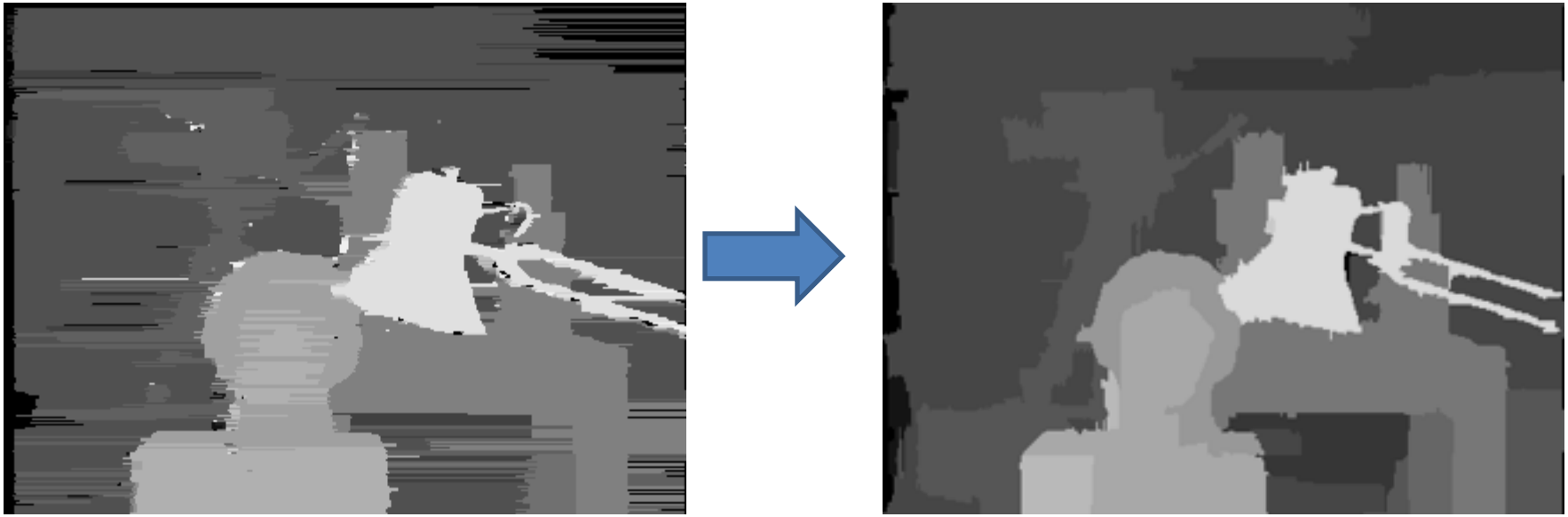
$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1})$$

Проблема: сканлайны друг с другом не связаны

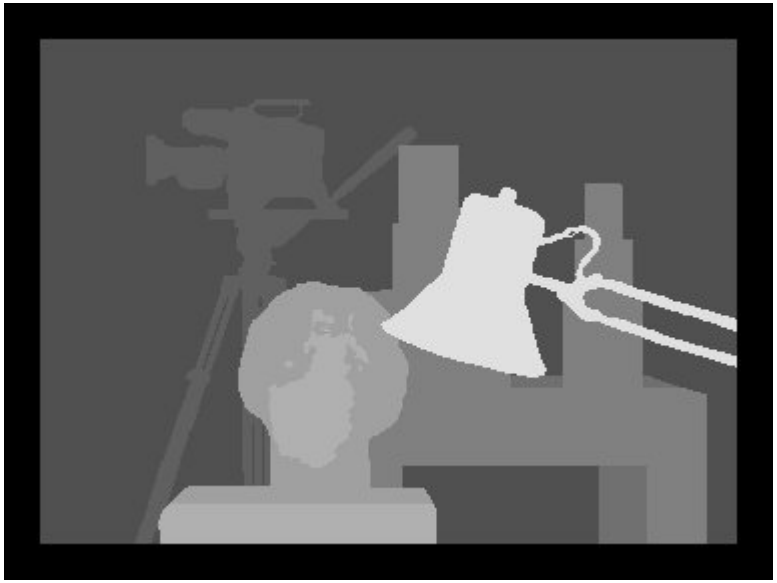
От строк к дереву
[Veksler, 2005]

$$\left\{ E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1}) \right\}$$

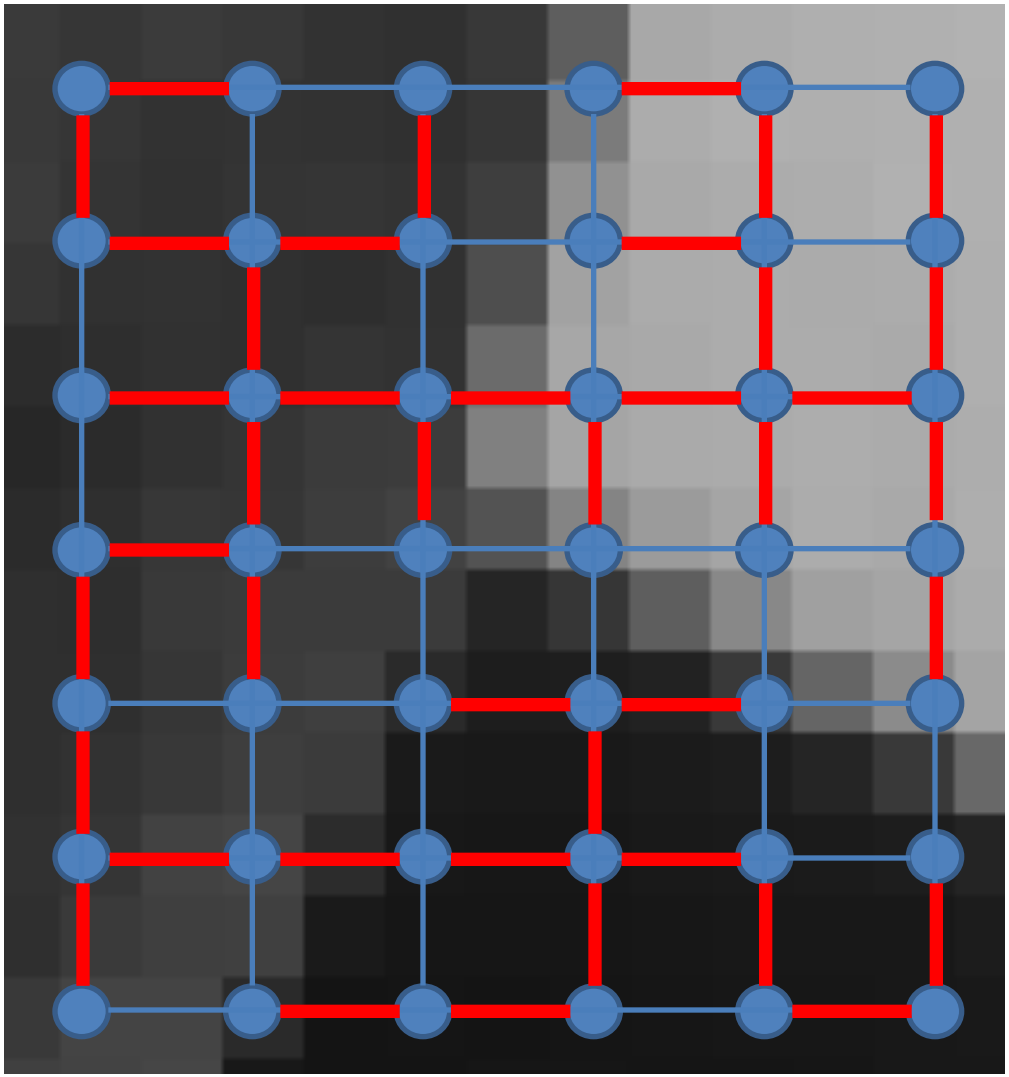
$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{T}} V(d_p, d_q)$$



Выбор дерева



Перепады цвета коррелируют с перепадами глубины

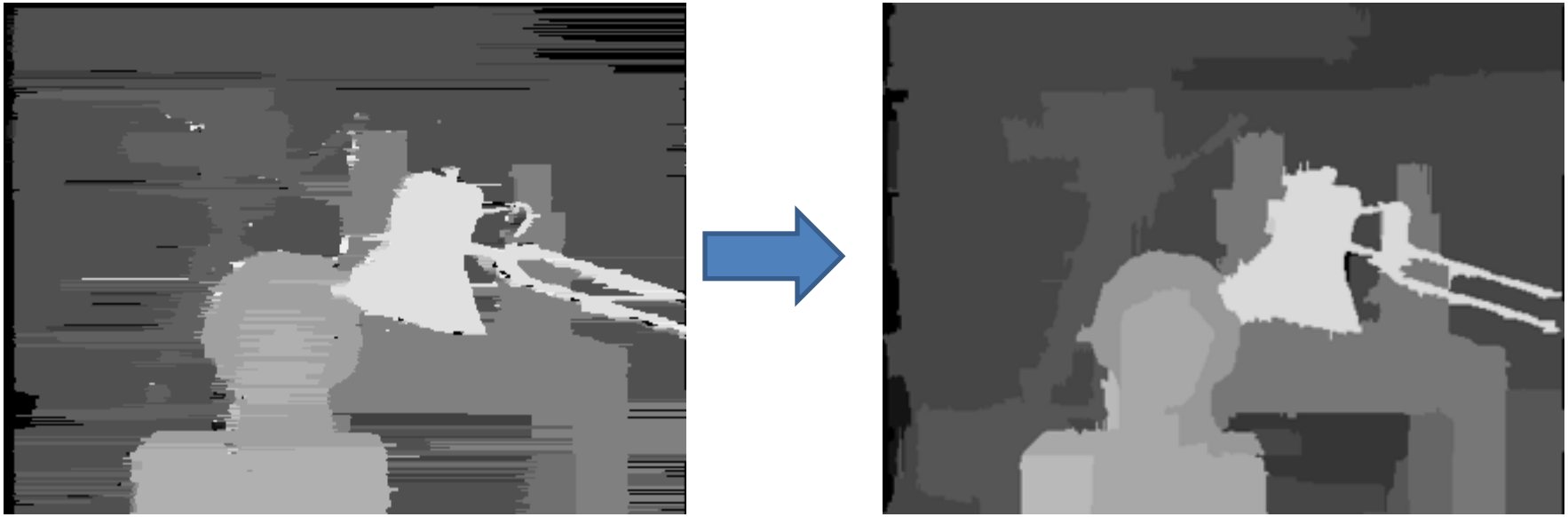


Минимальное остовное дерево!

От строк к дереву
[Veksler, 2005]

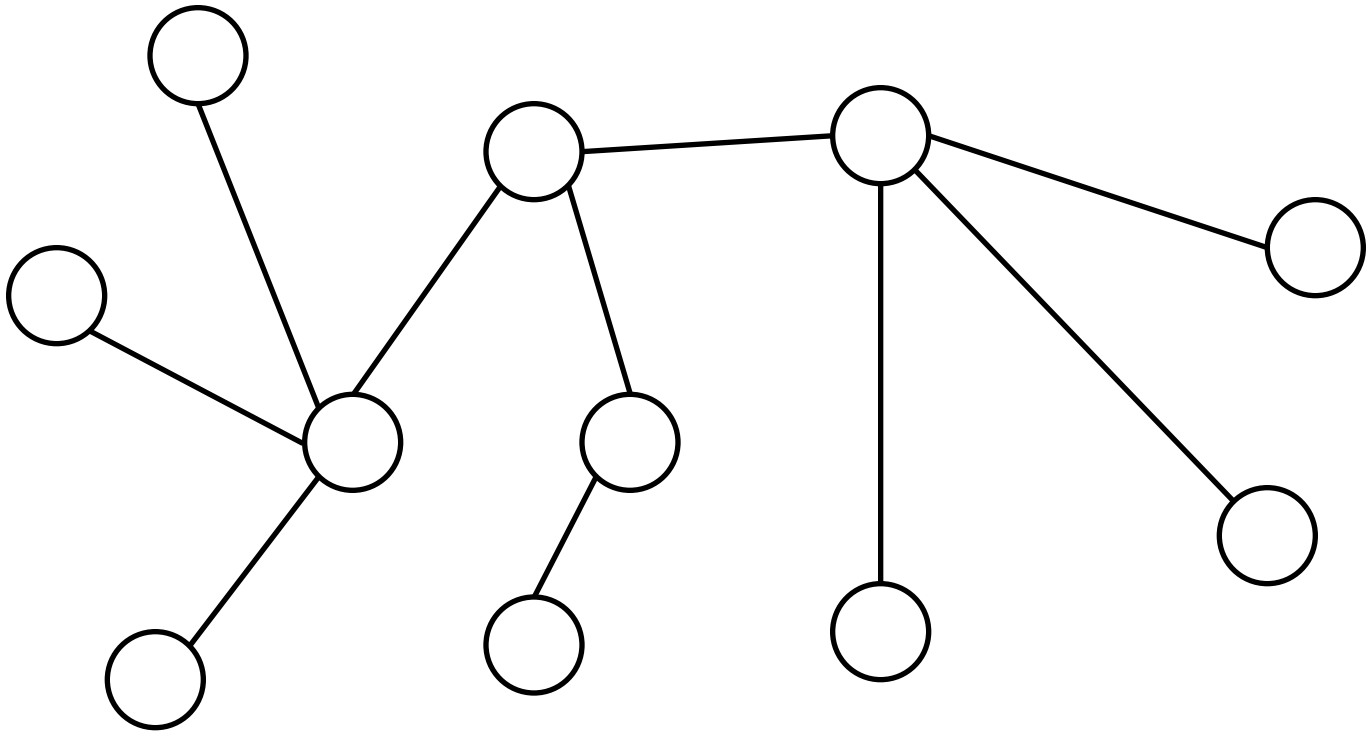
$$\left\{ E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1}) \right\}$$

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{T}} V(d_p, d_q)$$



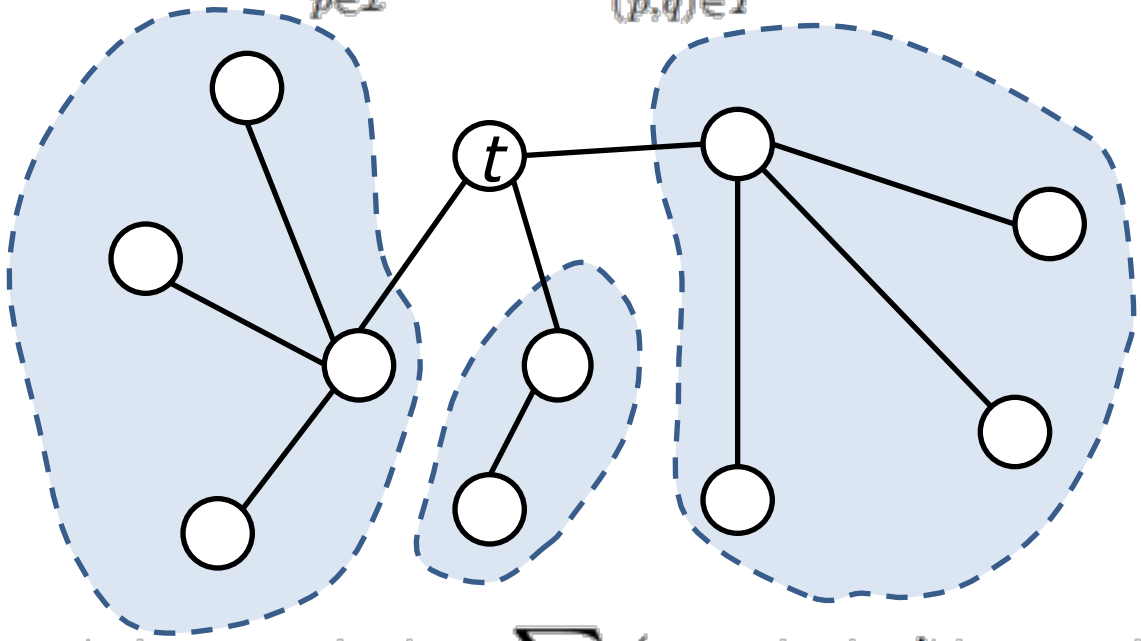
От строк к дереву

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{T}} V(d_p, d_q)$$



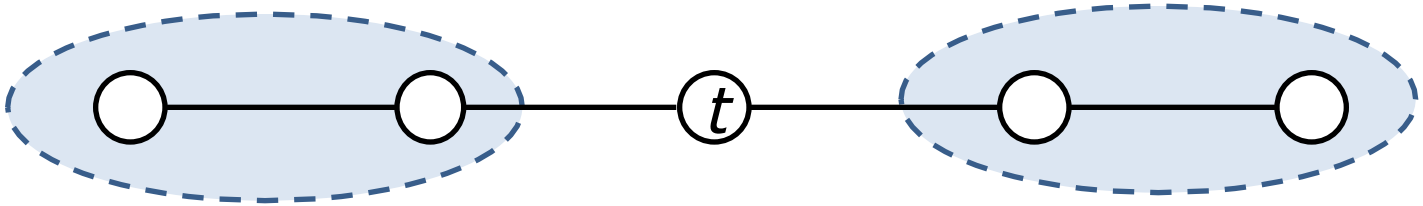
Динамическое программирование

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{T}} V(d_p, d_q)$$

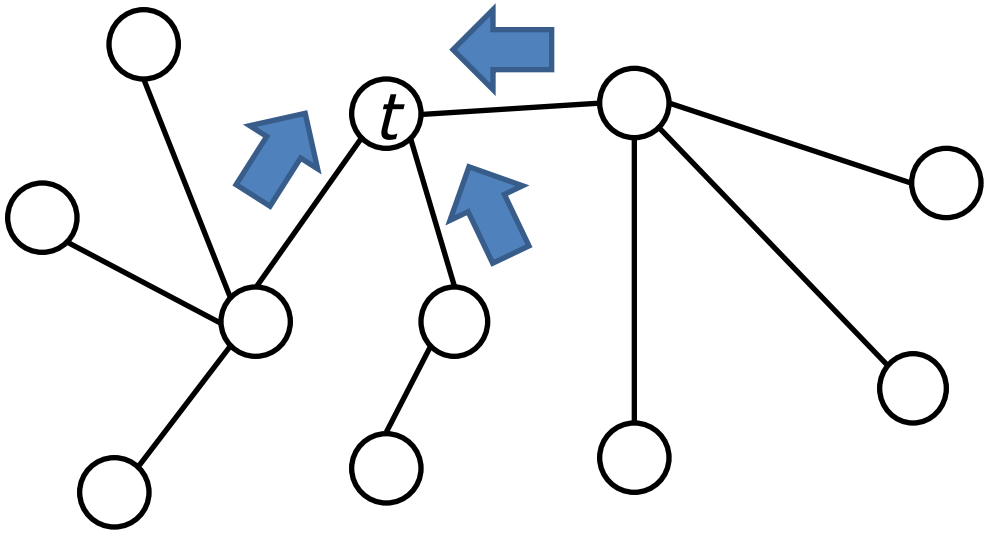


$$E(\mathbf{d}) = U_t(d_t) + \sum_{t'} \left(E_{\mathcal{T}'}(\mathbf{d}(\mathcal{T}')) + V(d_{t'}, d_t) \right)$$

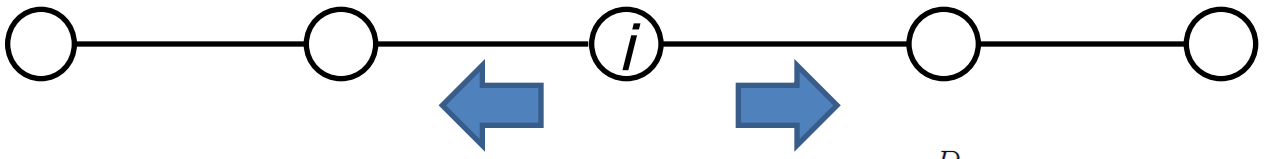
$$E(\mathbf{d}) = E_1^{t-1}(d_1, \ldots, d_{t-1}) + V(d_{t-1}, d_t) + U_t(d_t) + V(d_{t+1}, d_t) + E_{t+1}^{W}(d_{t+1}, \ldots, d_W)$$



Передача сообщений



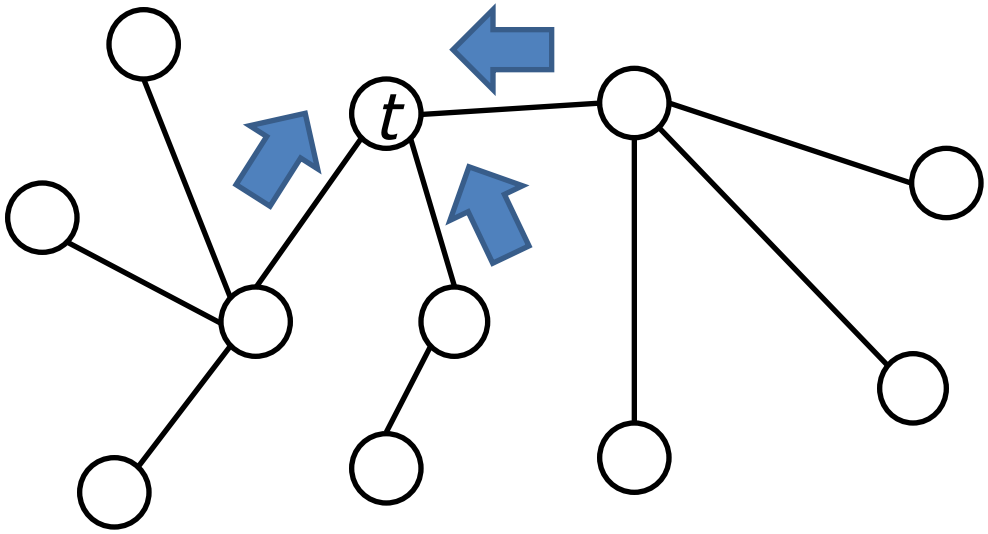
$$m_{t' \to t}(j) = \min_{k=1}^{D} \left(\sum_{t'' \neq t} m_{t'' \to t'}(k) + U_{t'}(k) + V(k,j) \right)$$



$$m_{i \to i-1}(j) = \min_{k=1}^{D} \left(m_{i+1 \to i}(k) + U_i(k) + V(k,j) \right)$$

$$m_{i \to i+1}(j) = \min_{k=1}^{D} \left(m_{i-1 \to i}(k) + U_i(k) + V(k,j) \right)$$

Передача сообщений



$$m_{t' \to t}(j) = \min_{k=1}^{D} \left(\sum_{t'' \neq t} m_{t'' \to t'}(k) + U_{t'}(k) + V(k,j) \right)$$

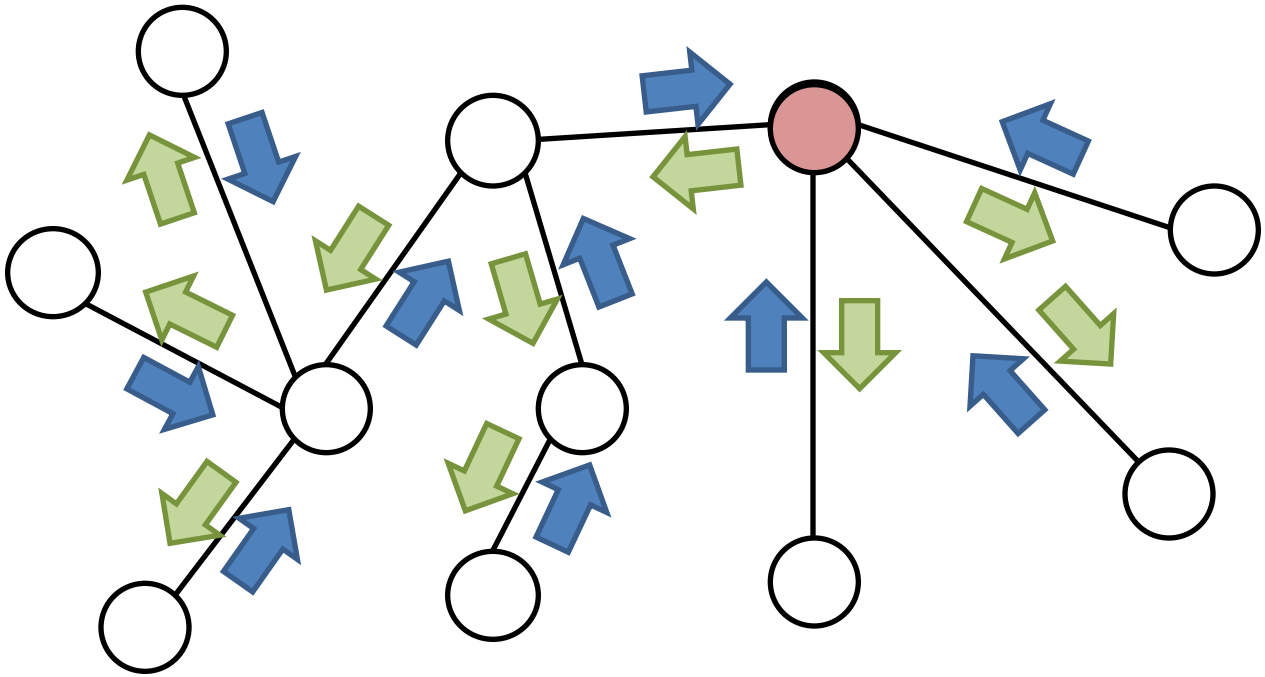
$$\sum_{t'} m_{t' \to t}(d_t) + U_t(d_t) = \underset{\mathbf{d}, \mathbf{d}(t)=d_t}{argmin}\, E(\mathbf{d})$$

Доказательство:

1. Назначить t корнем
2. Провести индукцию от листьев

Пересылка сообщений: расписание

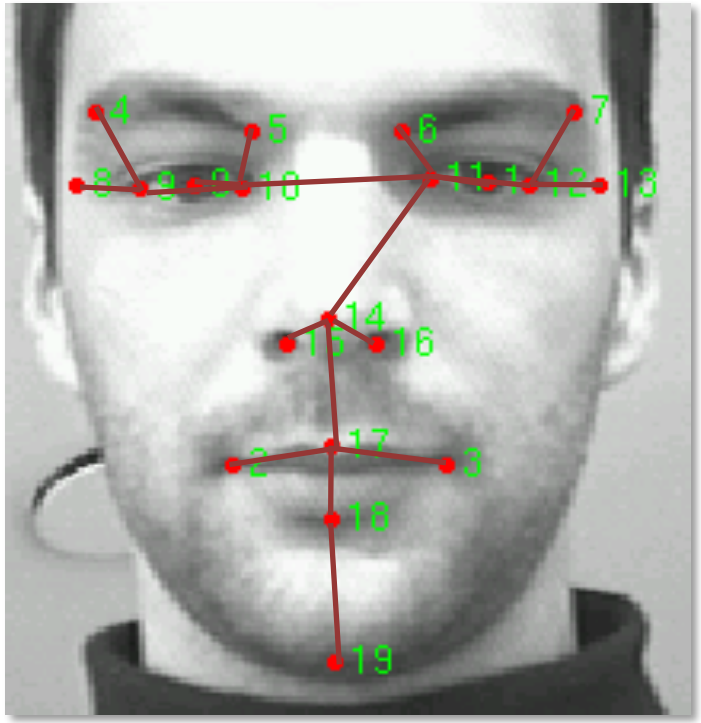
$$m_{t' \to t}(j) = \min_{k=1}^{D} \left(\sum_{t'' \neq t} m_{t'' \to t'}(k) + U_{t'}(k) + V(k,j) \right)$$



Сложность: всего O(ED) !

Pictorial structures [Felzenszwalb Huttenlocker 05]

Image from BioID/FGNet dataset



Обучается на тренировочных данных

$$E(\mathbf{x}) = \sum_{p \in \mathcal{V}} U^p(x_p) + \sum_{p,q \in \mathcal{E}} ||x_p - x_q - l_{pq}||^2$$

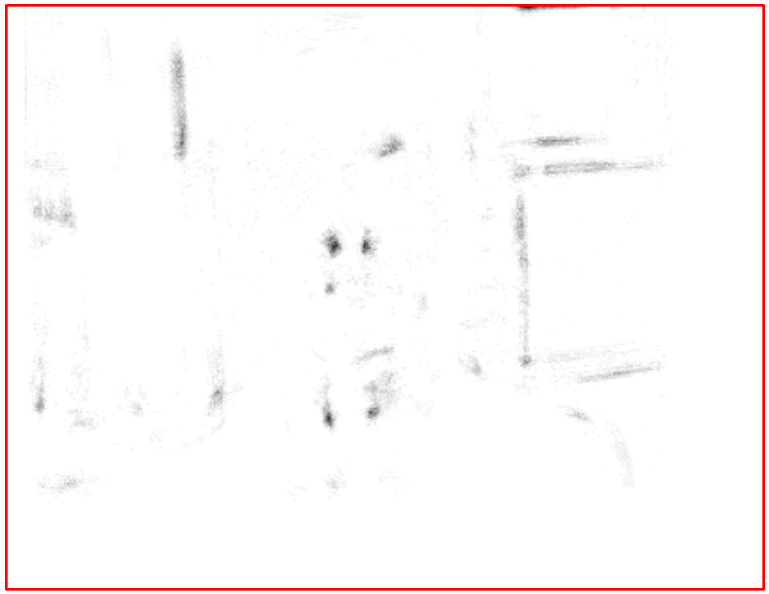
$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{T}} V(d_p, d_q)$$

Pictorial structures

$$E(\mathbf{x}) = \sum_{p \in \mathcal{V}} U^p(x_p) + \sum_{p,q \in \mathcal{E}} ||x_p - x_q - l_{pq}||^2$$

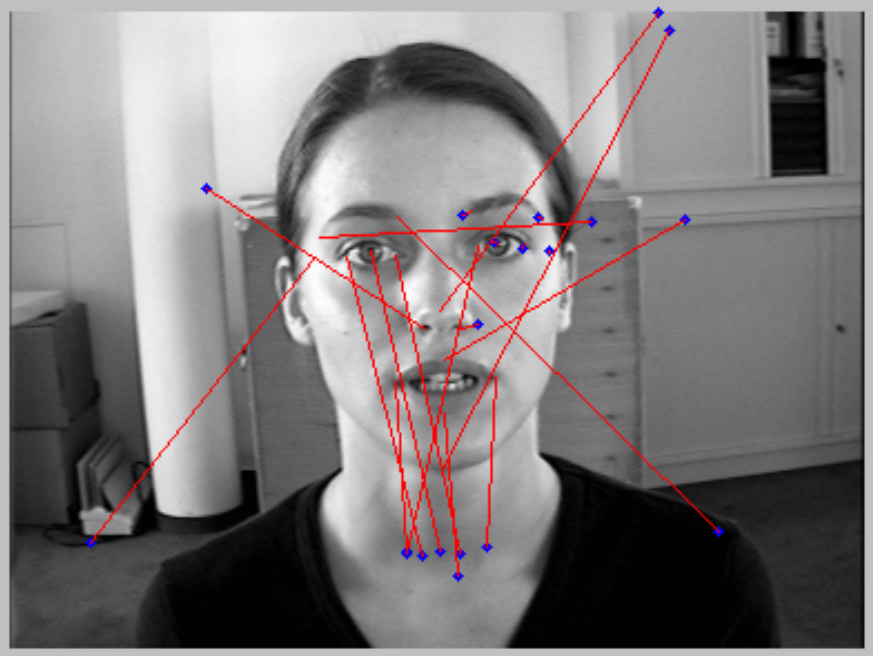
Угол правой брови

Левая ноздря



Pictorial structures

Независимый поиск точек



Pictorial structure result

$$m_{t' \to t}(j) = \min_{k=1}^{D} \left(\sum_{t'' \neq t} m_{t'' \to t'}(k) + U_{t'}(k) + V(k, j) \right)$$

Квадратичная функция

Идея [Felzenszwalb&Huttenlocher 05]: как быстро передать сообщение

Обобщаем дальше?

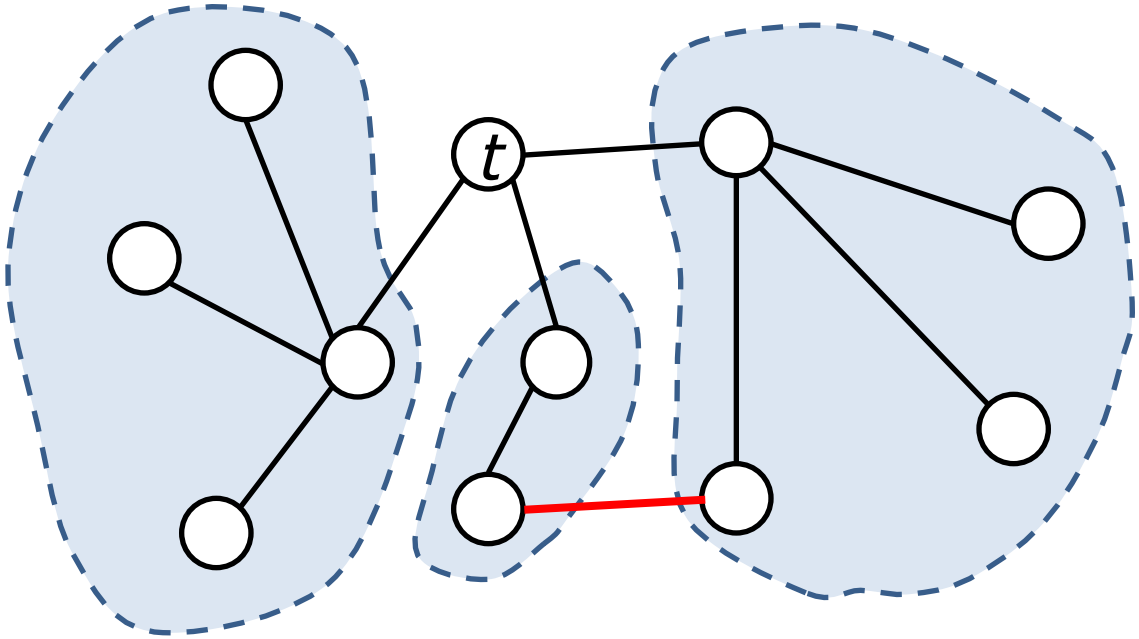
$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1})$$

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{T}} V(d_p, d_q)$$

Хотим: $$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{L}} V(d_p, d_q)$$

Динамическое программирование?

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{L}} V(d_p, d_q)$$



$$m_{t' \to t}(j) = \min_{k=1}^{D} \left(\sum_{t'' \neq t} m_{t'' \to t'}(k) + U_{t'}(k) + V(k, j) \right)$$

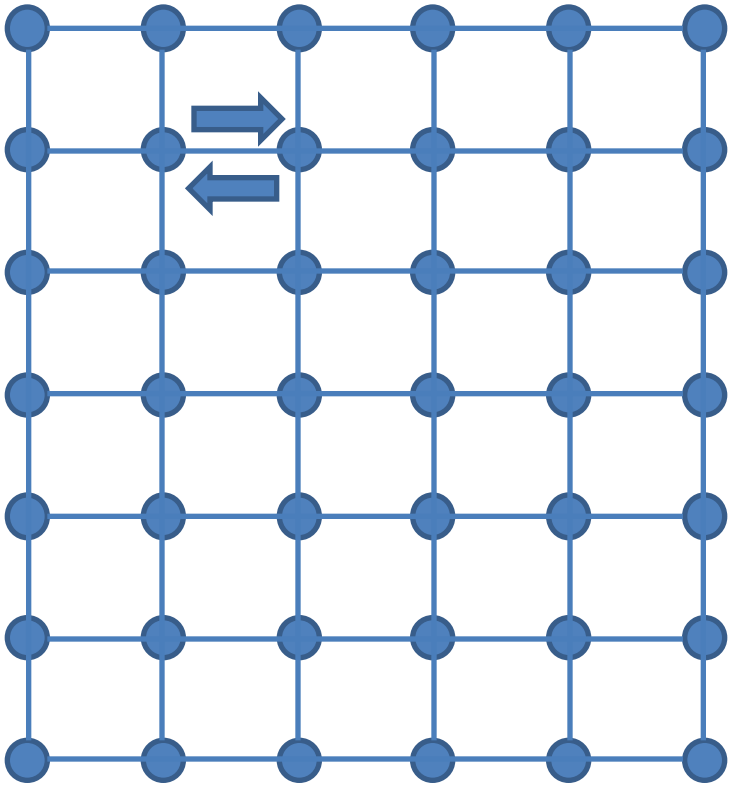
Пересылка сообщений

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{L}} V(d_p, d_q)$$

$$m_{t' \to t}(j) = \min_{k=1}^{D} \left(\sum_{t'' \neq t} m_{t'' \to t'}(k) + U_{t'}(k) + V(k,j) \right)$$

[Pearl 1988]: “Loopy Belief Propagation”

- Сообщения передаются много раундов
- Эвристика: нет гарантий на сходимость/близость к минимуму
- Результат зависит от расписания/инициализации
- Эмпирически, часто дает очень хороший результат

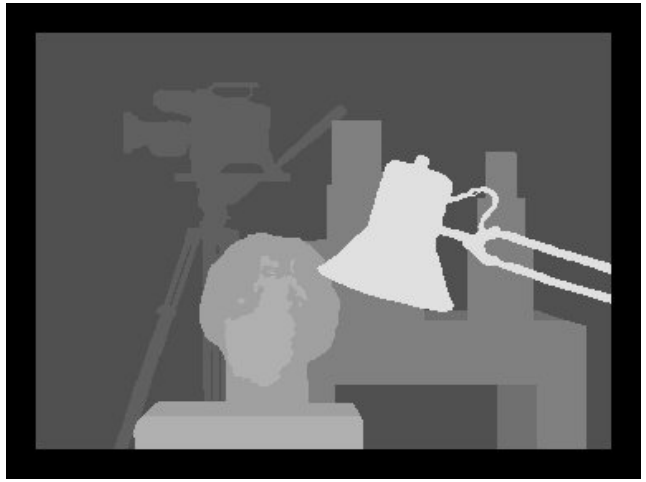


Полная модель: результат

$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1})$$

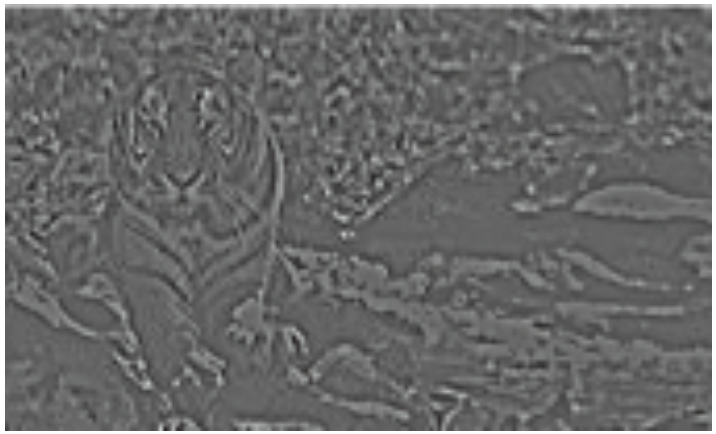
$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_i(d_i) + \sum_{(p,q) \in \mathcal{T}} V(d_p, d_q)$$

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_i(d_i) + \sum_{(p,q) \in \mathcal{L}} V(d_p, d_q)$$



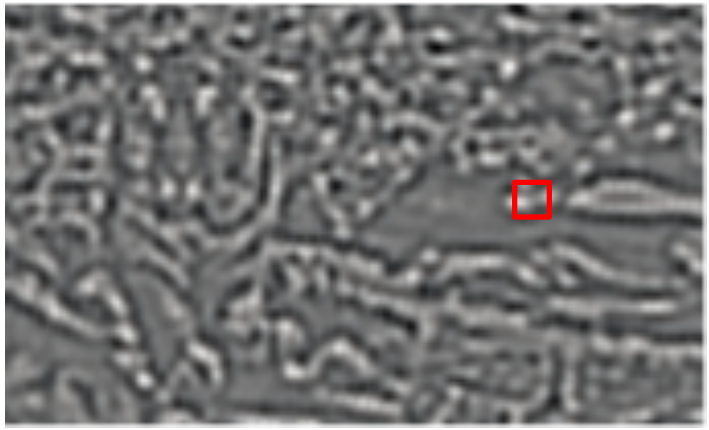
Увеличение разрешения [Freeman,Paztor,Carmichael,Jones 2000]

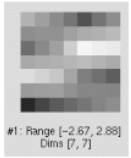
Средние частоты

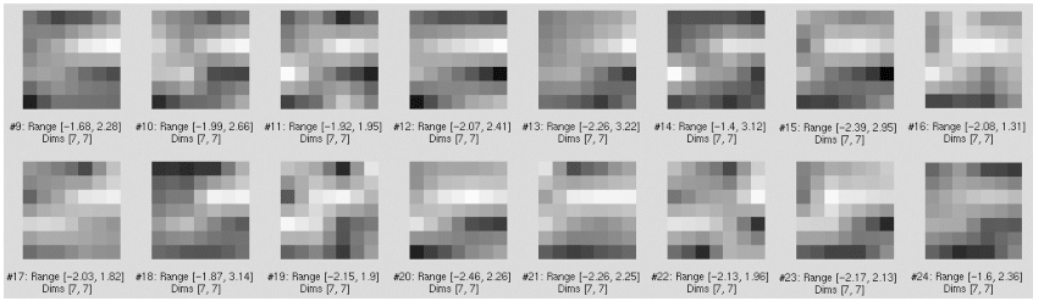


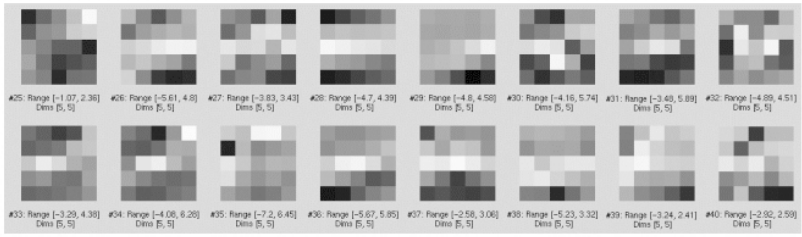
Высокие частоты = ?

Увеличение разрешения [Freeman,Paztor,Carmichael,Jones 2000]

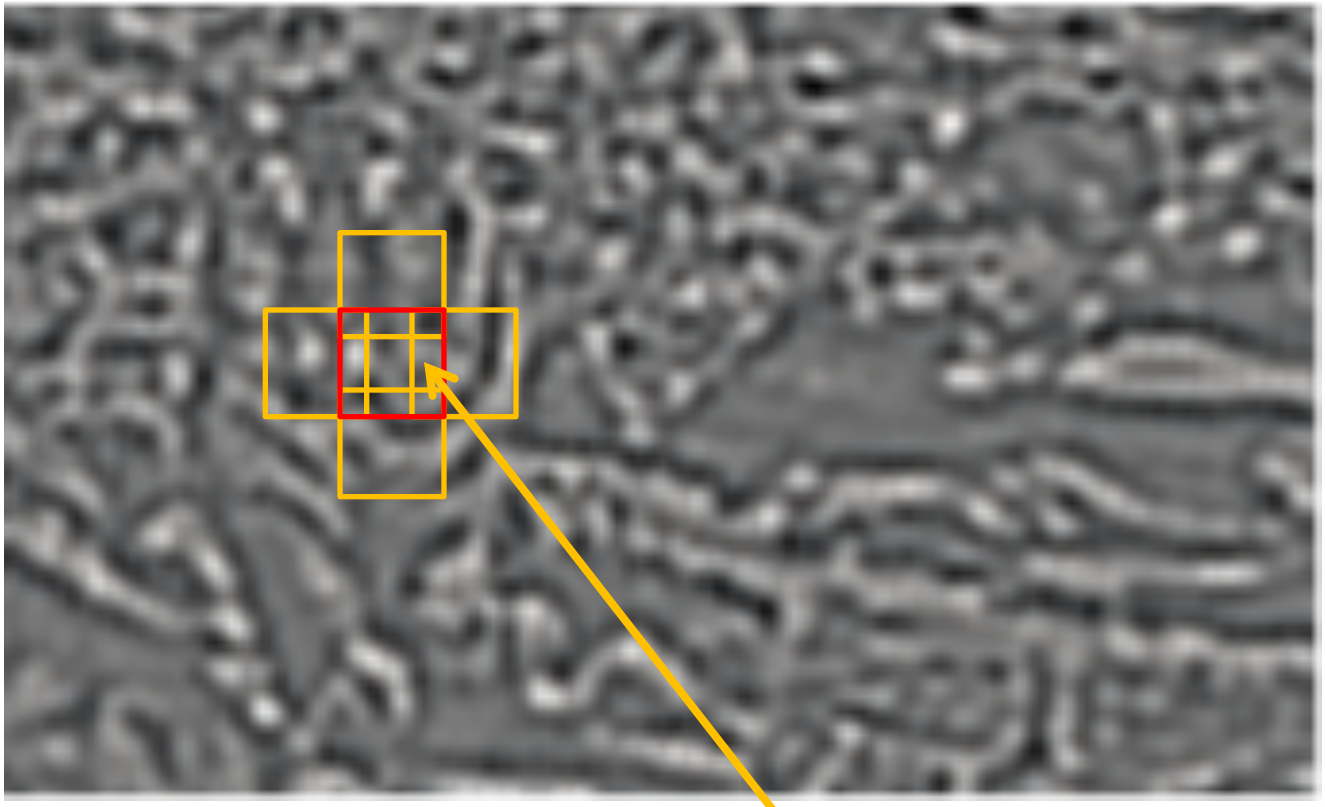






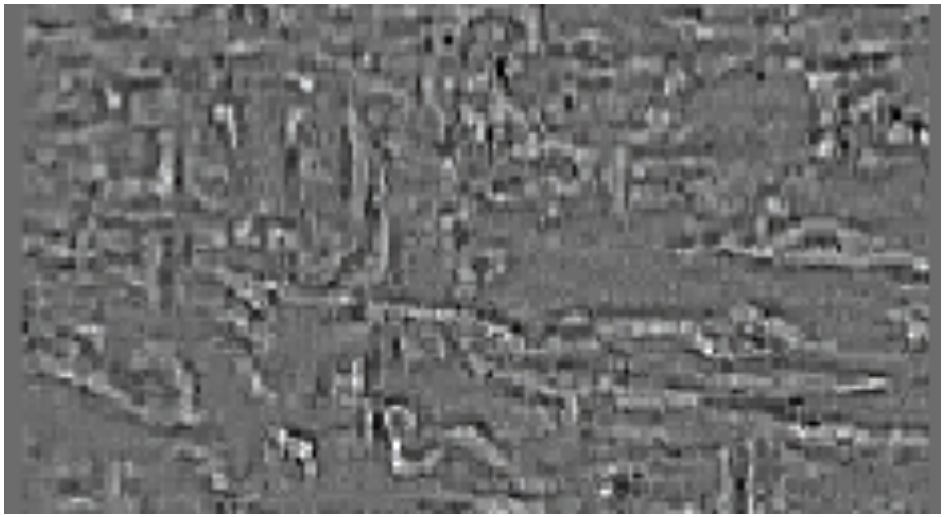


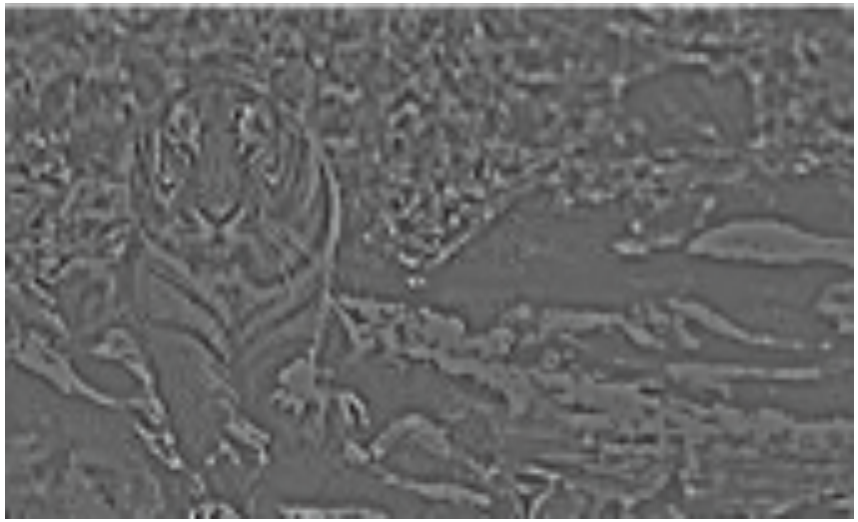
Увеличение разрешения [Freeman,Paztor,Carmichael,Jones 2000]



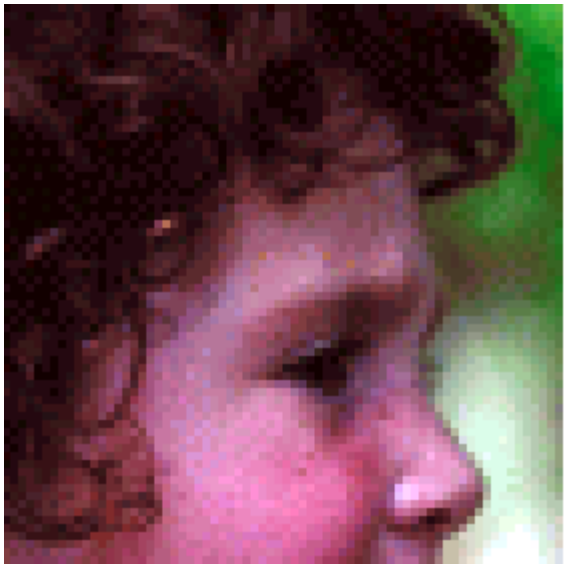
$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{L}} V(d_p, d_q)$$

Увеличение разрешения [Freeman,Paztor,Carmichael 2000]

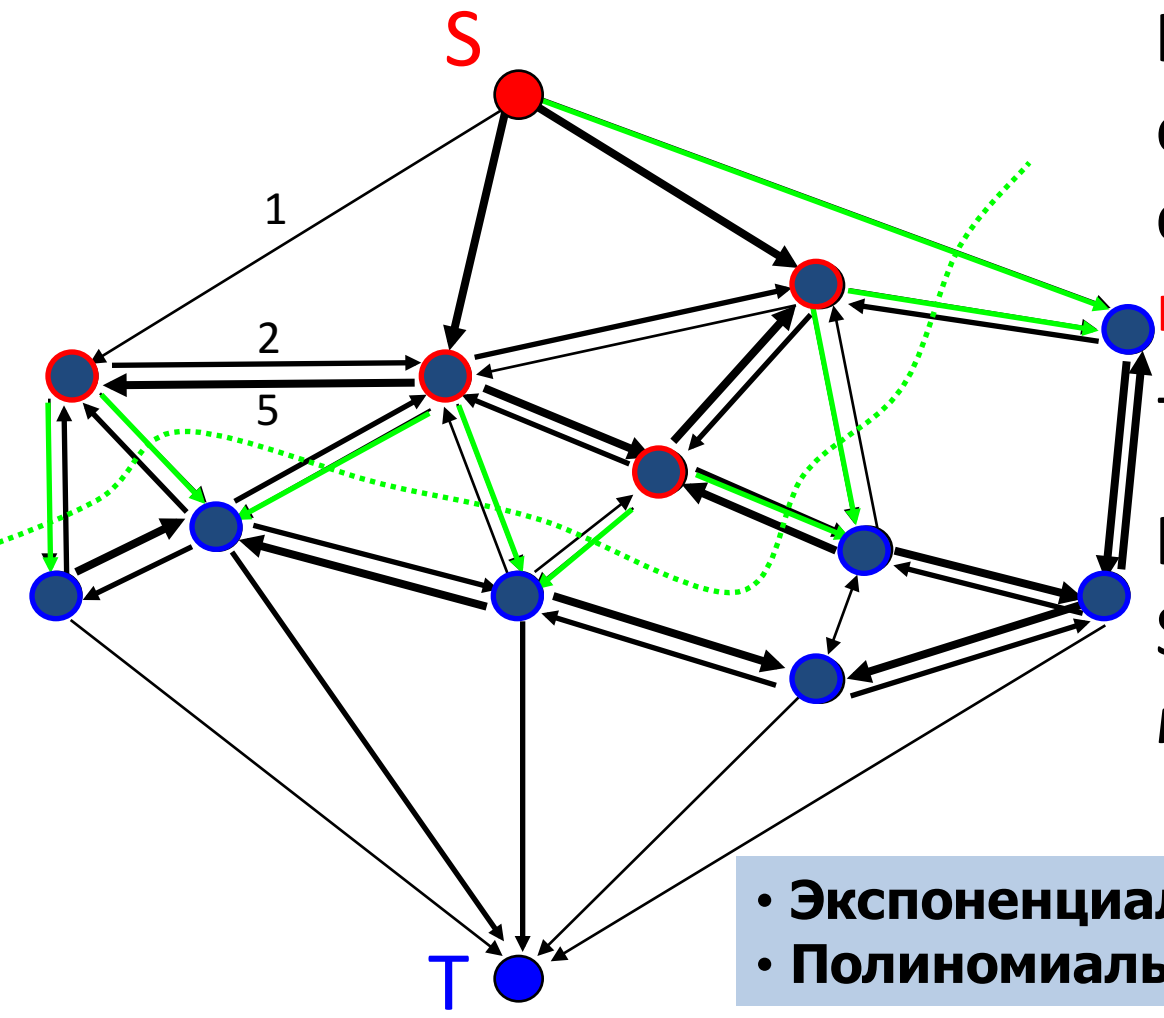




Увеличение разрешения [Freeman,Paztor,Carmichael,Jones 2000]



Минимальный st-разрез на графе [Ford and Fulkerson 60]

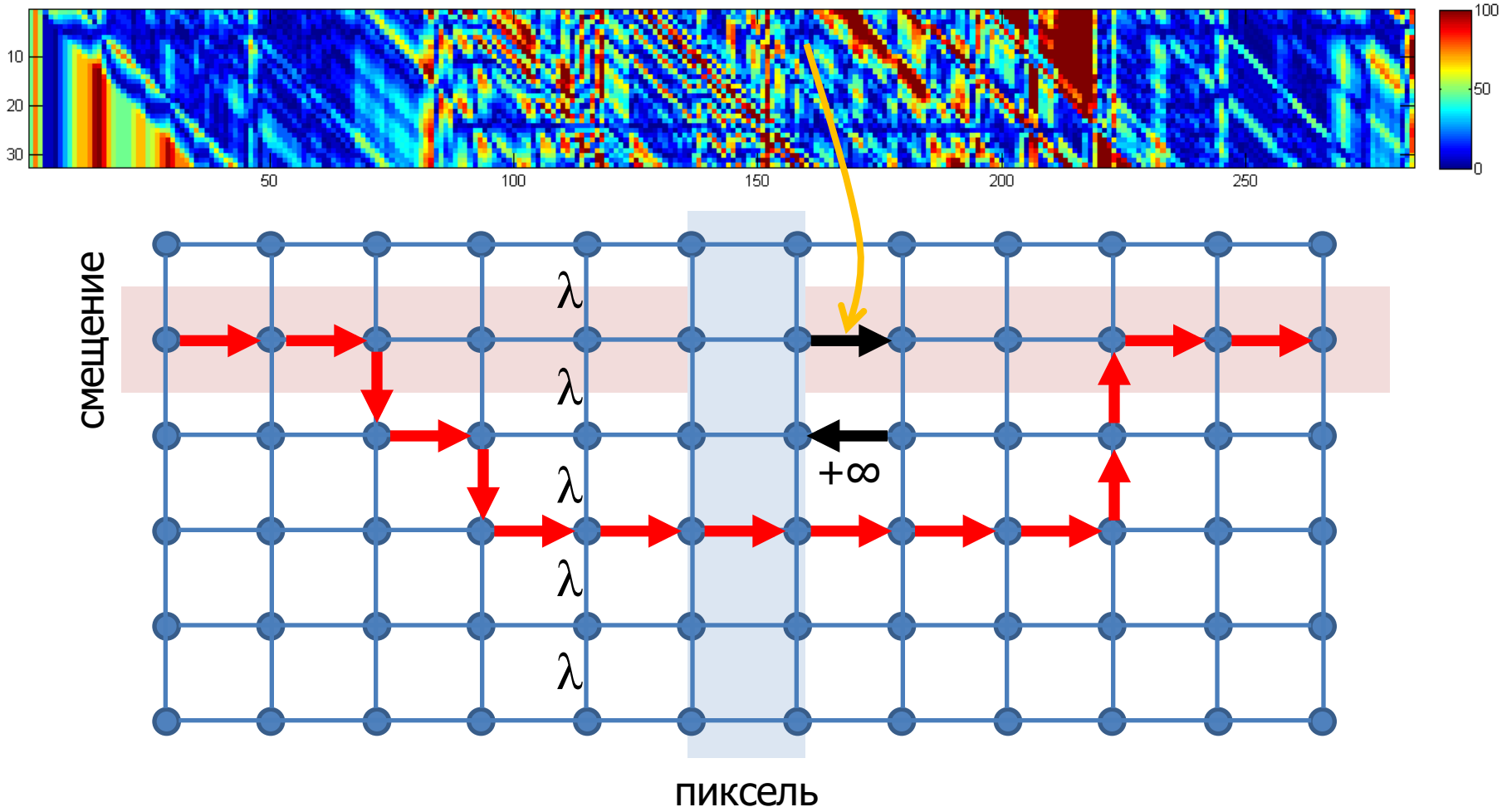


На входе:
ориентированный граф с выделенными истоком и стоком (сеть) + веса дуг ≥ 0.

На выходе:
ST-разрез на графе с минимальным весом.

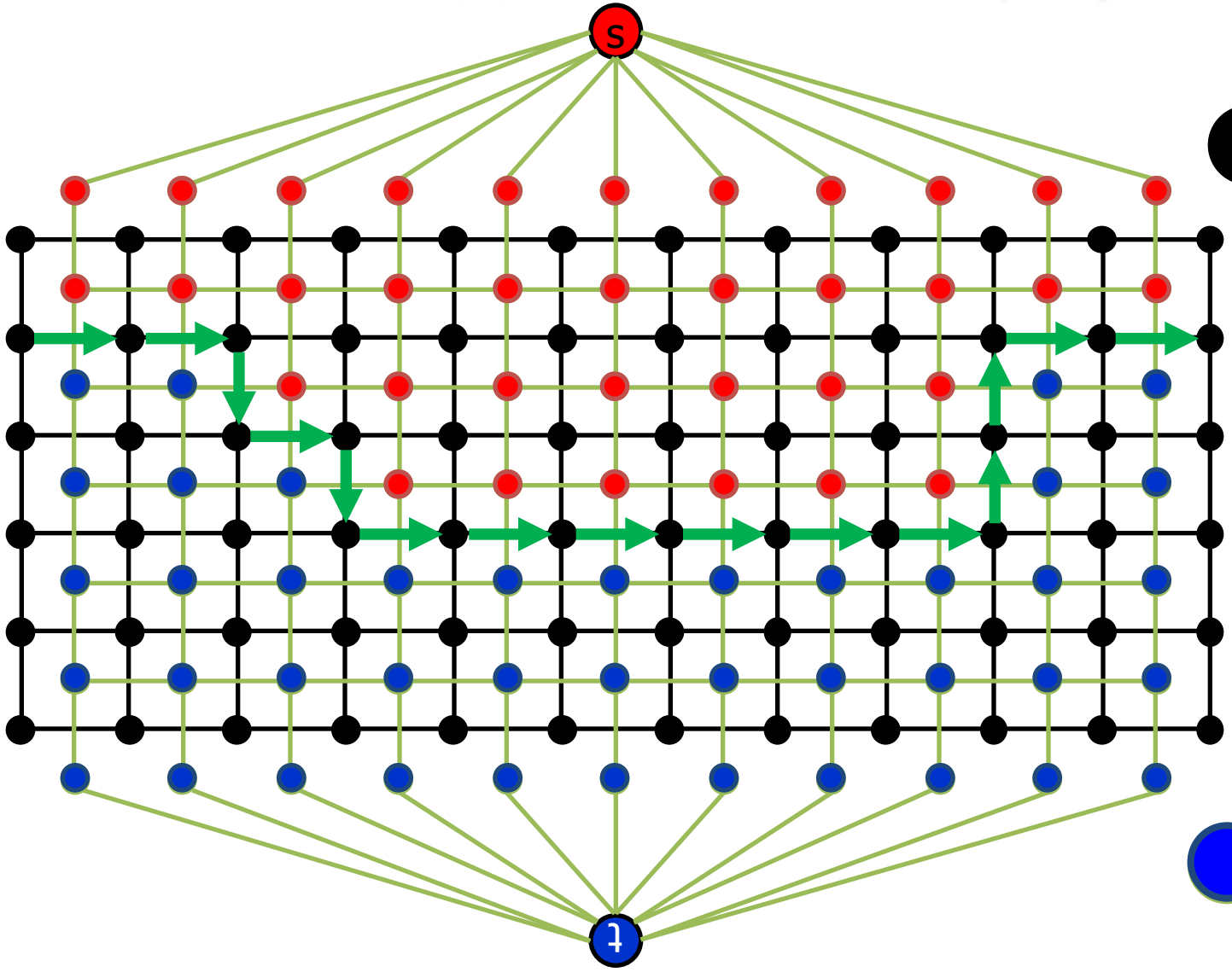
- **Экспоненциальное** количество разрезов.
- **Полиномиальное** время поиска.

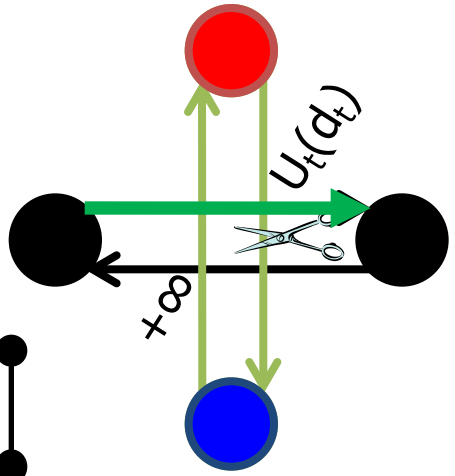
Сканлайн: напоминание

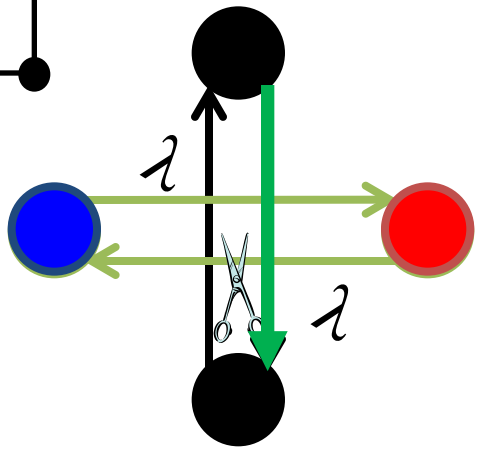


$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} \lambda |d_i - d_{i+1}|$$

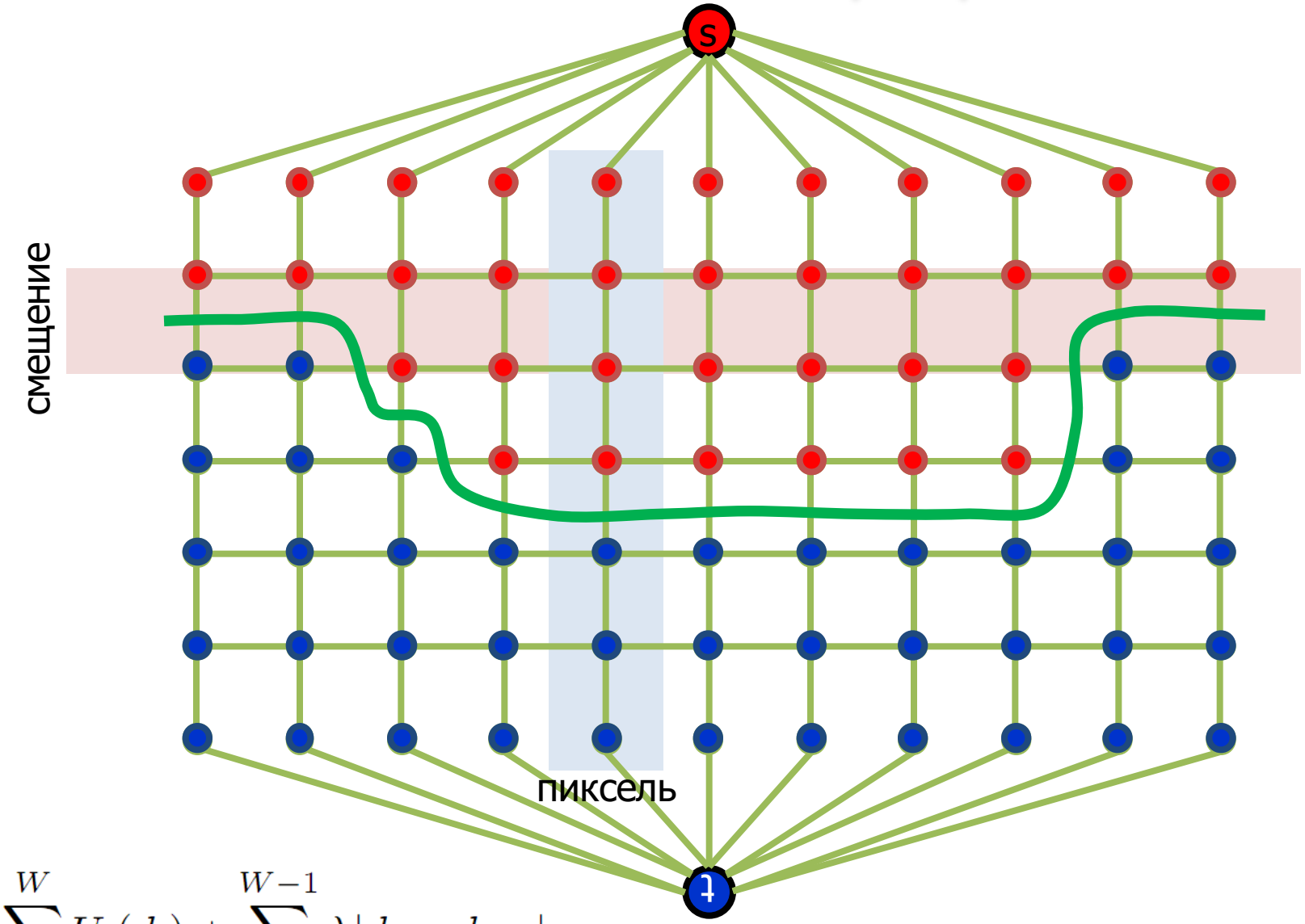
Двойственный граф





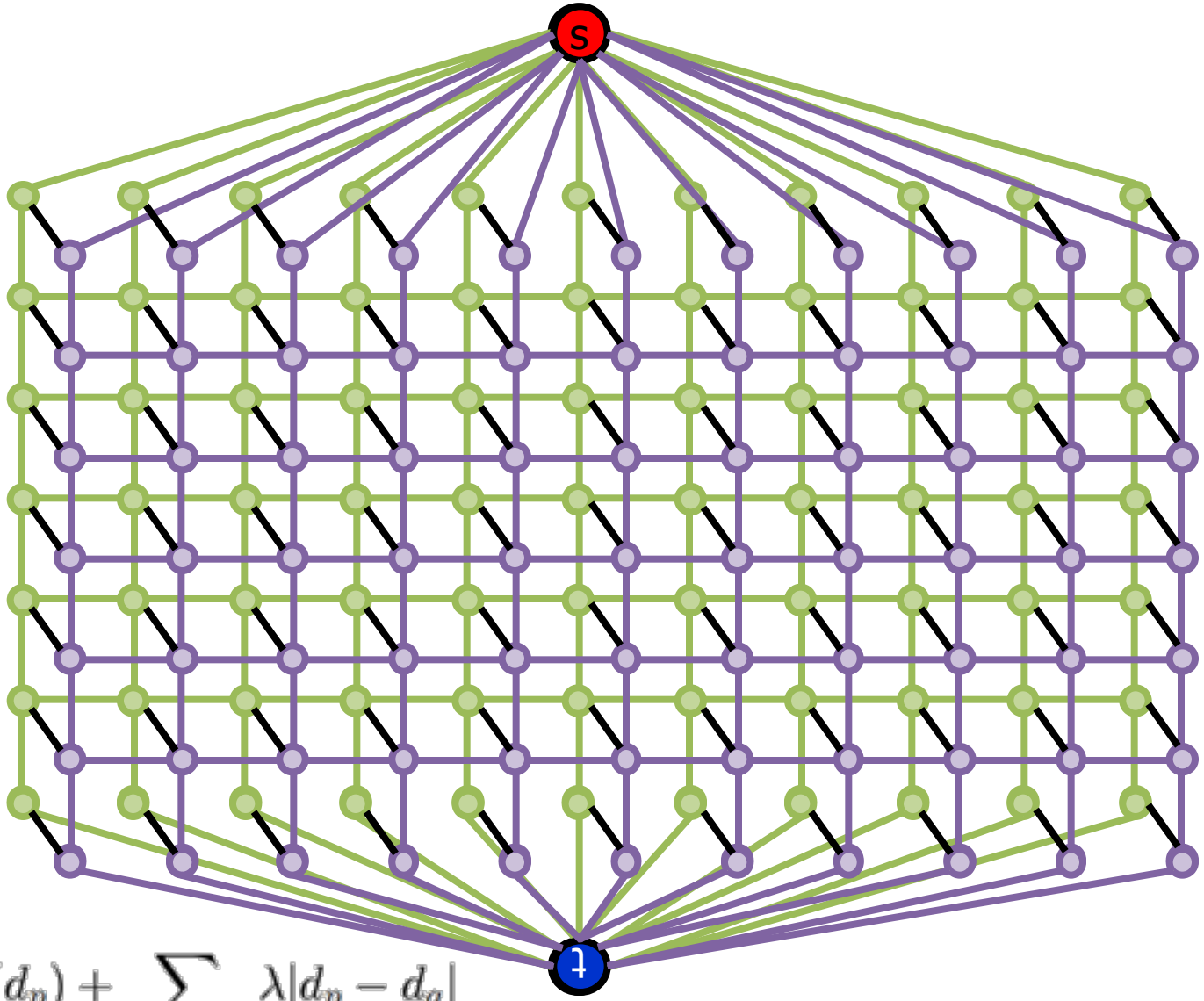


Оптимизация с помощью разреза



$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} \lambda |d_i - d_{i+1}|$$

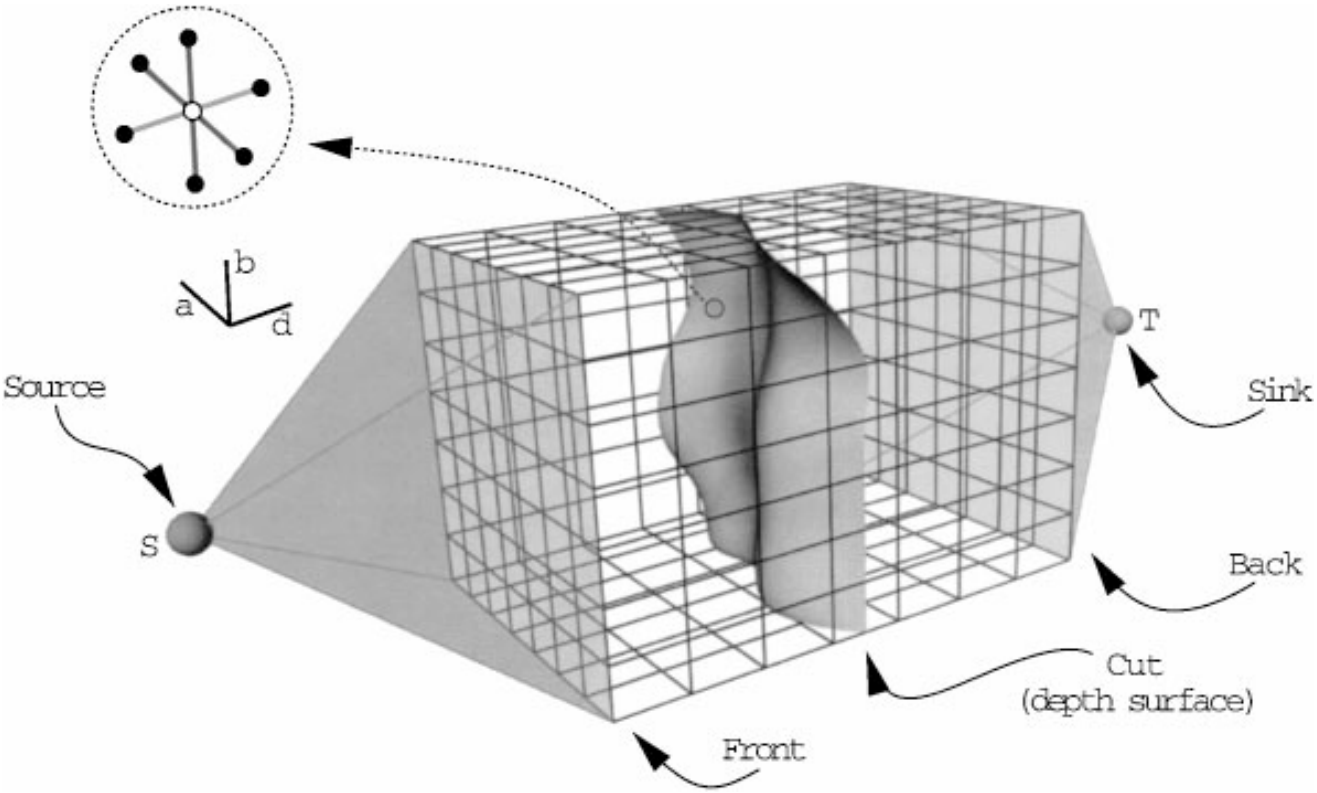
Оптимизация с помощью разреза



$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{L}} \lambda |d_p - d_q|$$

Стерео с помощью разреза [Roy&Cox 1998, Ishikawa&Geiger 1998]

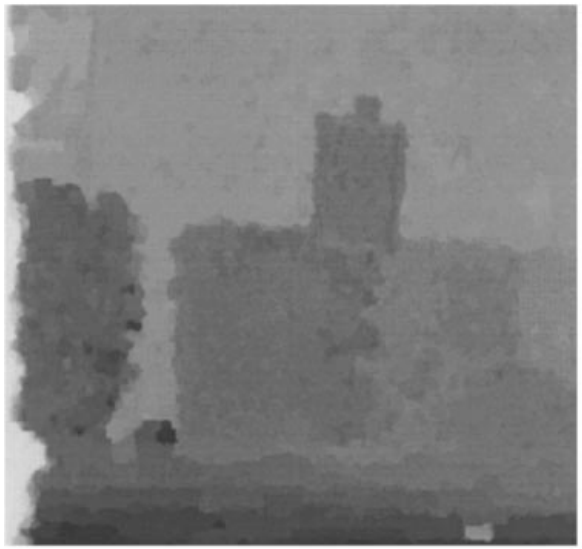
$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_p(d_p) + \sum_{(p,q) \in \mathcal{L}} \lambda |d_p - d_q|$$

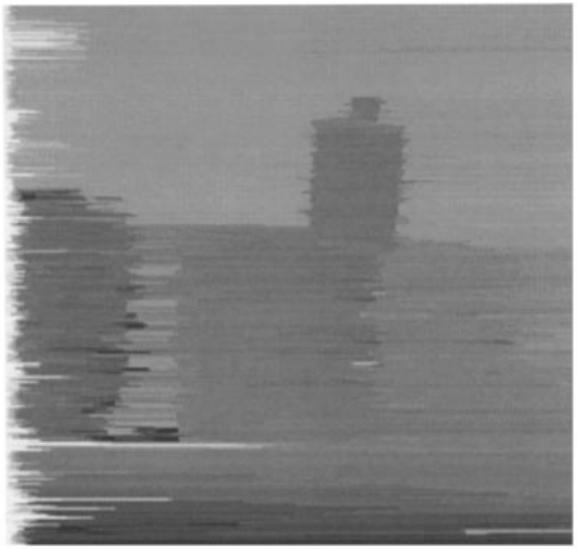


Стерео с помощью разреза [Roy&Cox 1998]

shrub-15

shrub-3





Стерео с помощью разреза [Roy&Cox 1998]

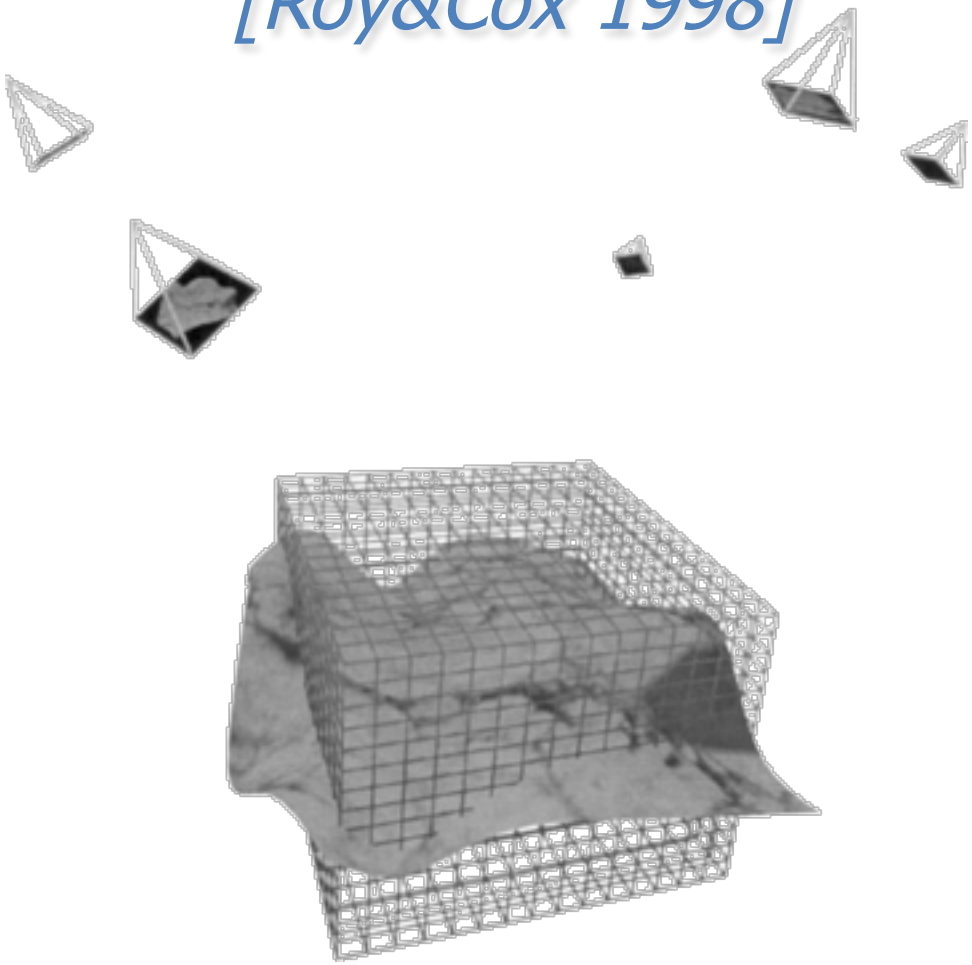


Image from[Roy, 2002]

- Обобщение на несколько изображений
- … но снятых с «одной и той же стороны»

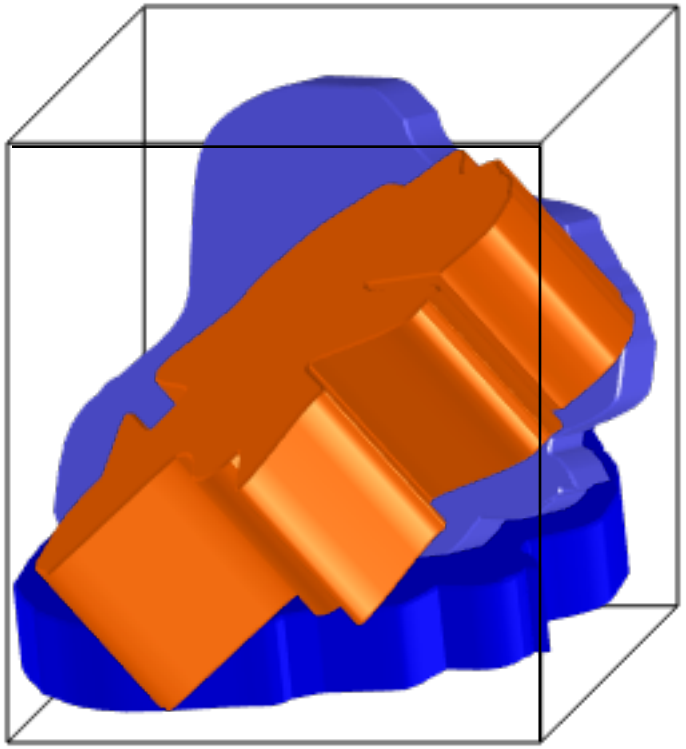
«Всесторонняя» реконструкция

- Трехмерные данные (напр. лазерные сканы)
- Наборы фотографий

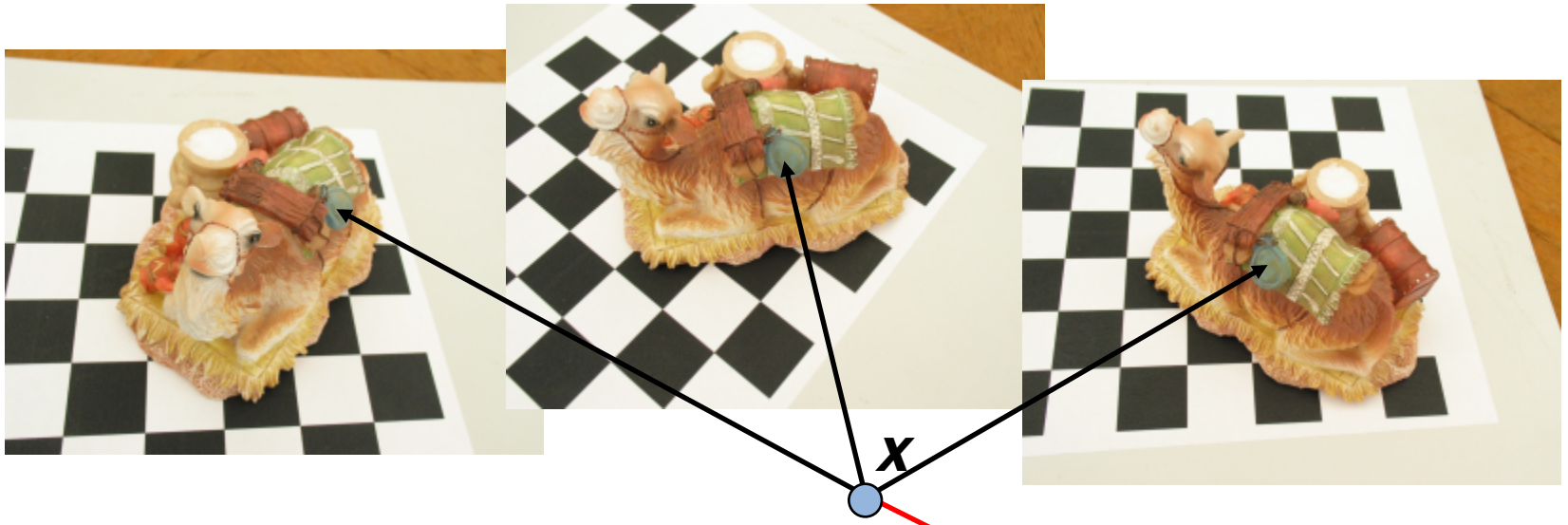
Предполагается наличие регистрации и ограничивающего объема

Энергетический подход к геометрической реконструкции

- Конструируется пространство «всех» поверхностей
- Вводится **функционал-энергия** на пространстве поверхностей
- Энергия = соответствие данным + регуляризация
- Результат реконструкции – минимум функционала



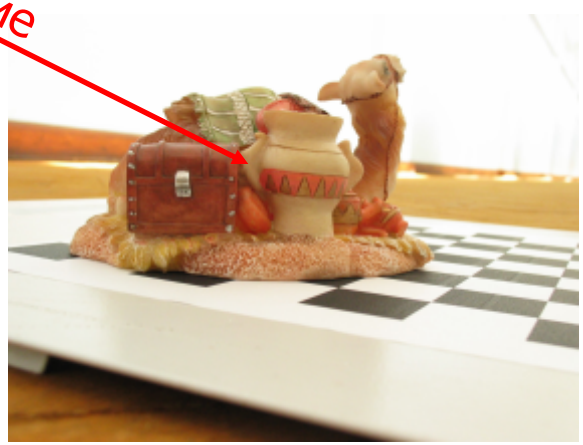
Фотосостоятельность



- **Фотосостоятельность:**

Точка, лежащая на поверхности, имеет схожие цвета проекции

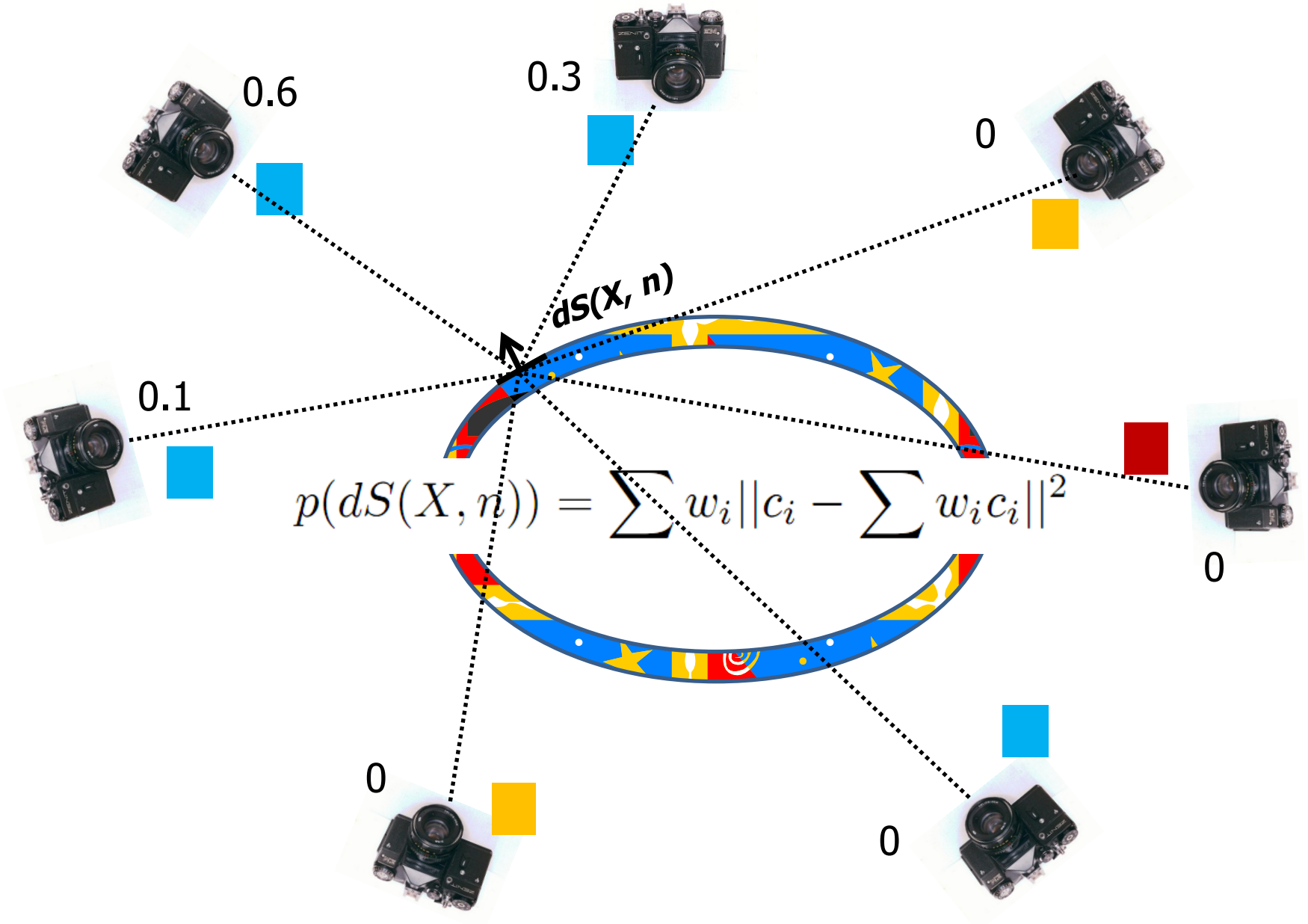
- Закрытия: невидимые точки не подчиняются фотосостоятельности



Точная *видимость* неизвестна => нужны приближенные оценки

Ориентированная фотосостоятельность

[Lempitsky, Boykov, Ivanov 2006]



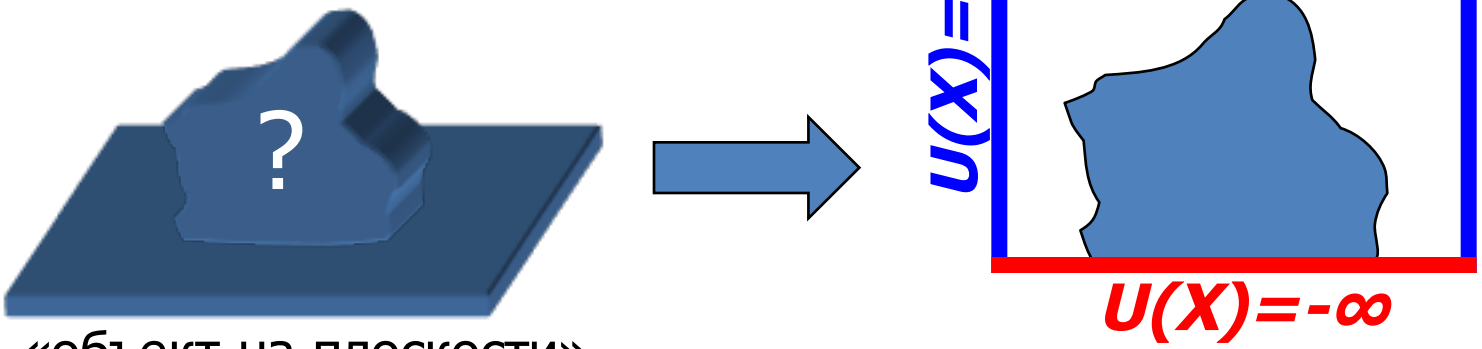
$$p(dS(X,n)) = \sum w_i ||c_i - \sum w_i c_i||^2$$

Энергия

$$E(M) = \underbrace{\int_M p(X, n_{dS})\, dS}_{\text{общая фотосостоятельность}}$$

U(X) может быть:

- равномерным
- учитывающим силуэты
- учитывающим жесткие условия:



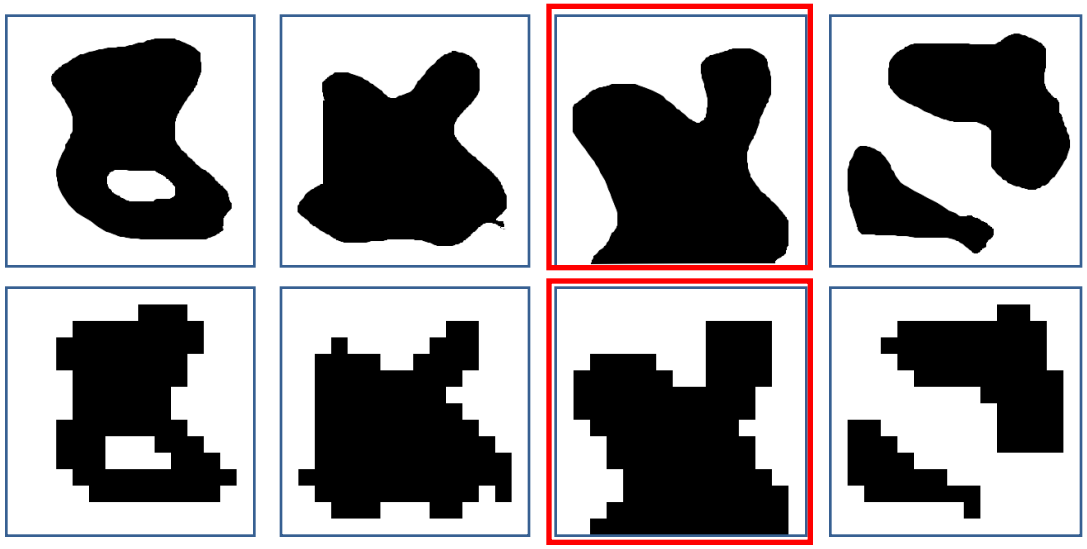
«объект на плоскости»

Дискретная глобальная оптимизация

$$E(M) = \int_M p(X, n_{dS})\, dS + \int_{int\, M} U(X)\, dV$$

Пространство «всех» поверхностей

Конечное подмножество



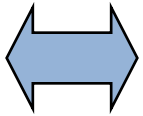
Выбор дискретного множества

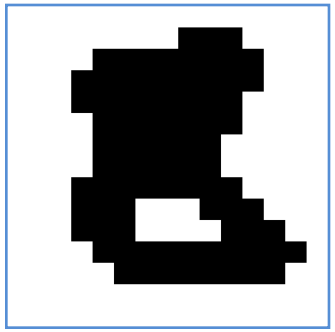
$$E(M) = \int_M p(X, n_{dS})\, dS + \int_{int\, M} U(X)\, dV$$

Цель: выбрать достаточно плотное множество

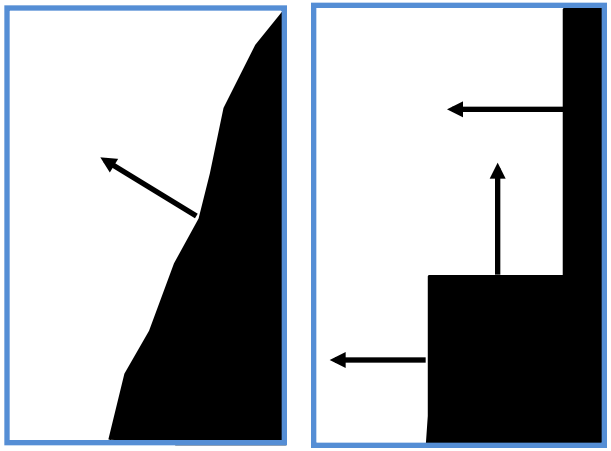
Очевидный выбор: воксельный подход.

Истинный глобальный минимум





Ближайшая поверхность



Увеличение

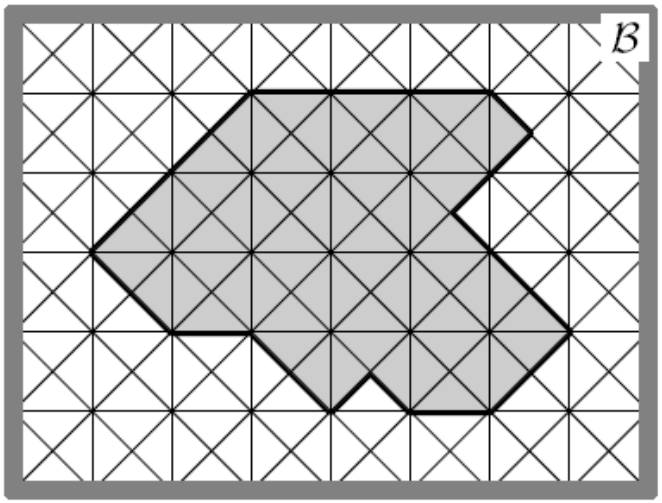
Плохая аппроксимация n_{dS}!

Вывод: недостаточно плотный набор ориентаций

Дискретизация функционала

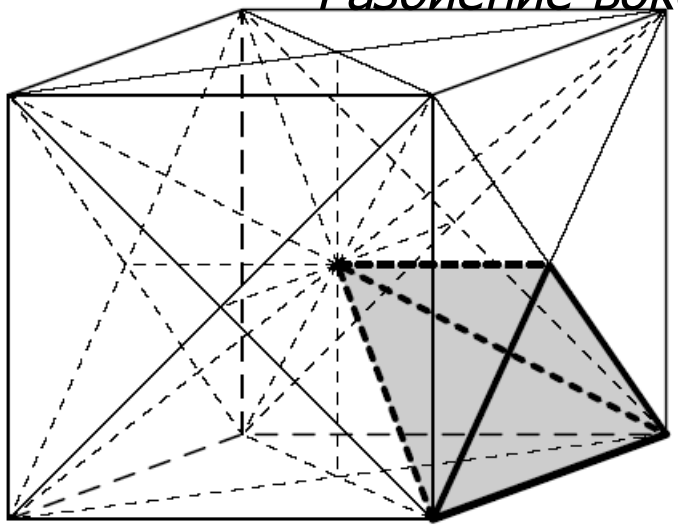
$$E(M) = \int_M p(X, n_{dS})\, dS + \int_{int\, M} U(X)\, dV$$

Лучшая дискретизация в 2D случае:



Аналогичная дискретизация в 3D случае:

Разбиение вокселя

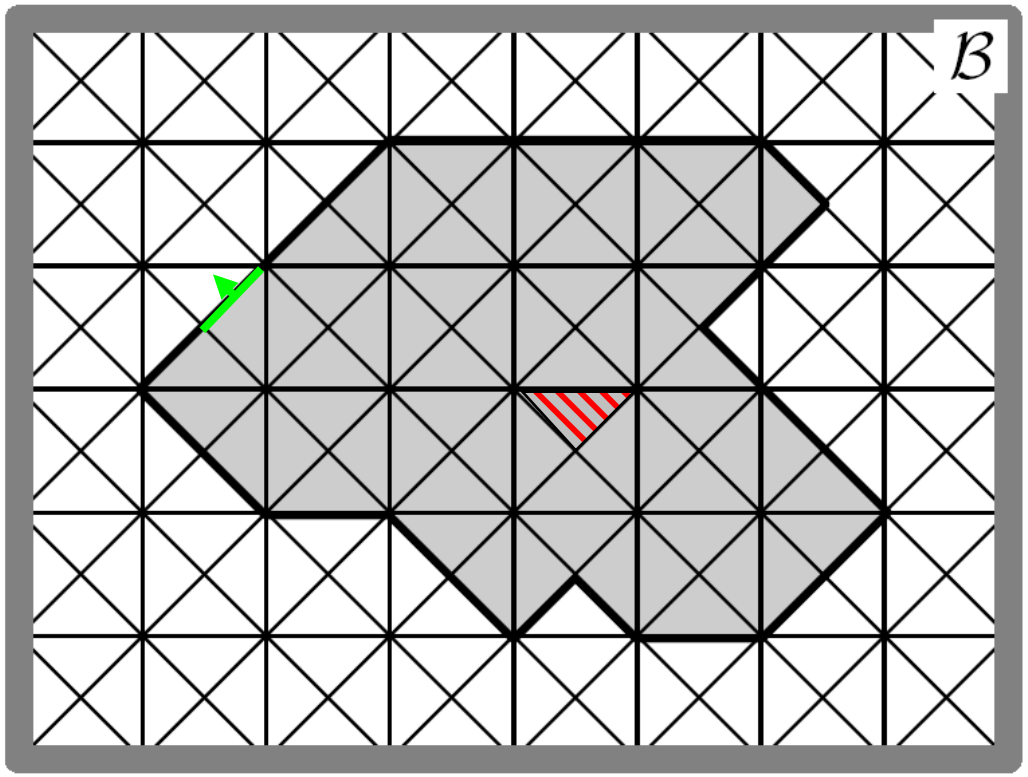


- Каждый воксель разбивается на 24 тетраэдра
- Дискретное множество = поверхности из тетраэдров
- 18 ориентаций поверхности (вместо 6)

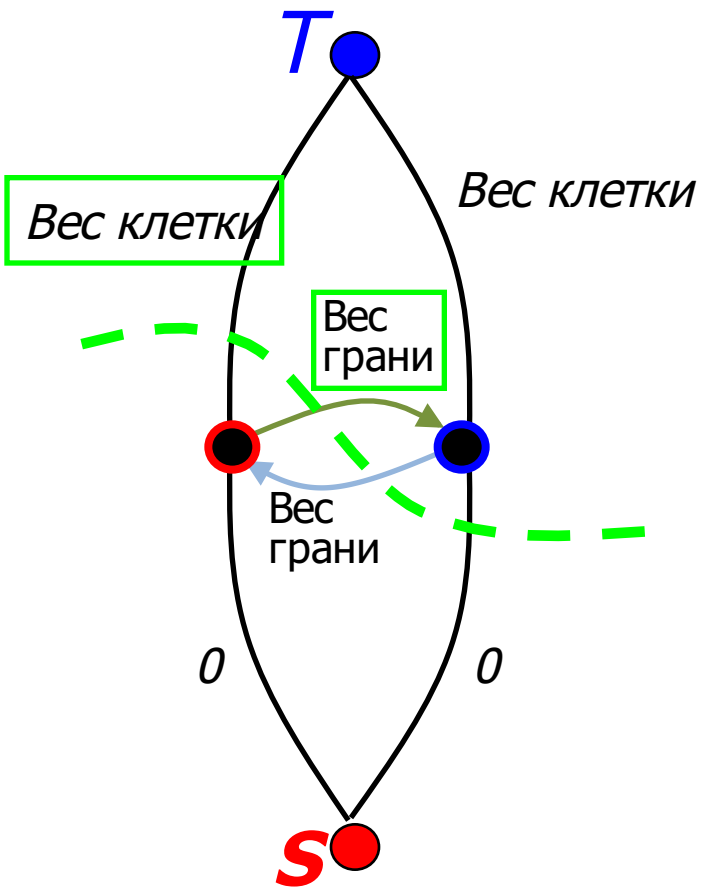
Функционал на дискретном множестве

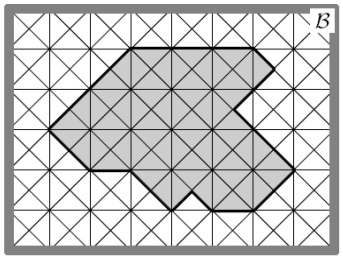
$$E(M) = \int_M p(X, n_{dS})\, dS + \int_{int\, M} U(X)\, dV$$

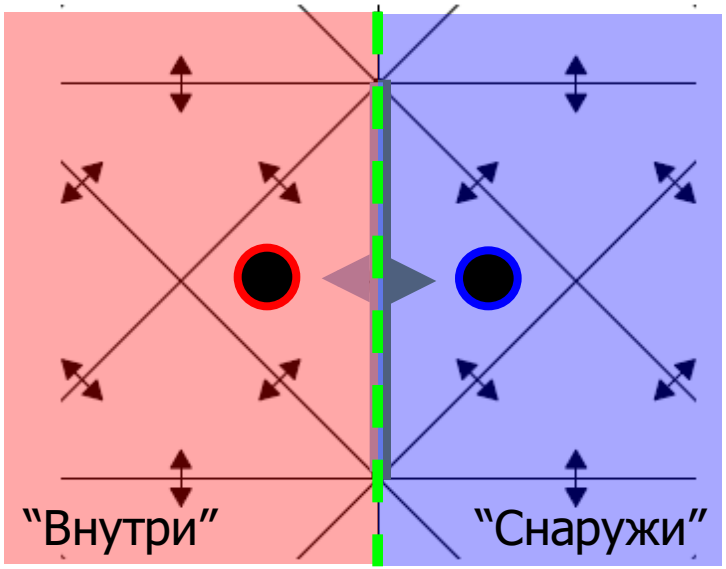
Функционал находится как сумма *весов граней* и *весов клеток*



Сведение к разрезу на графе







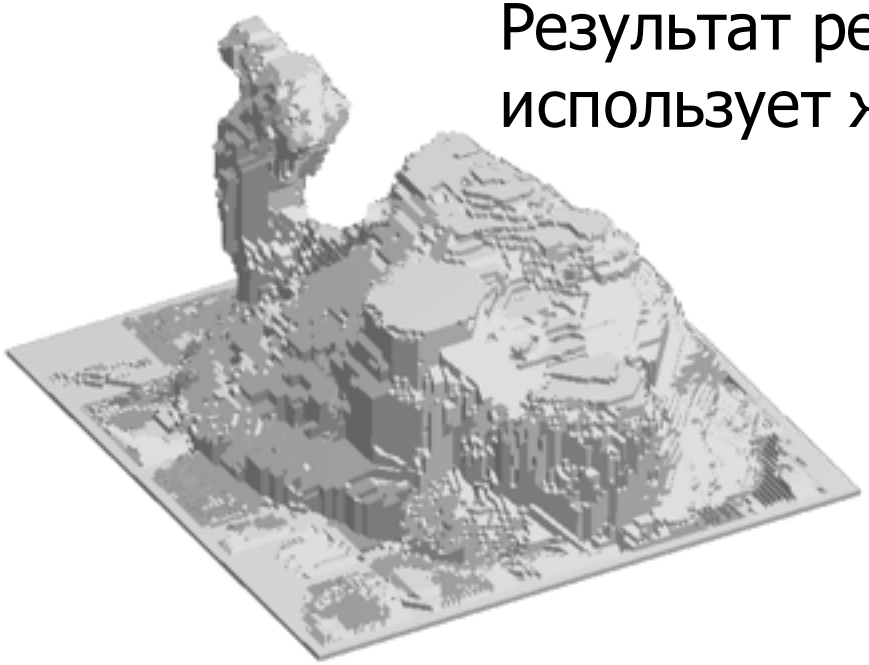
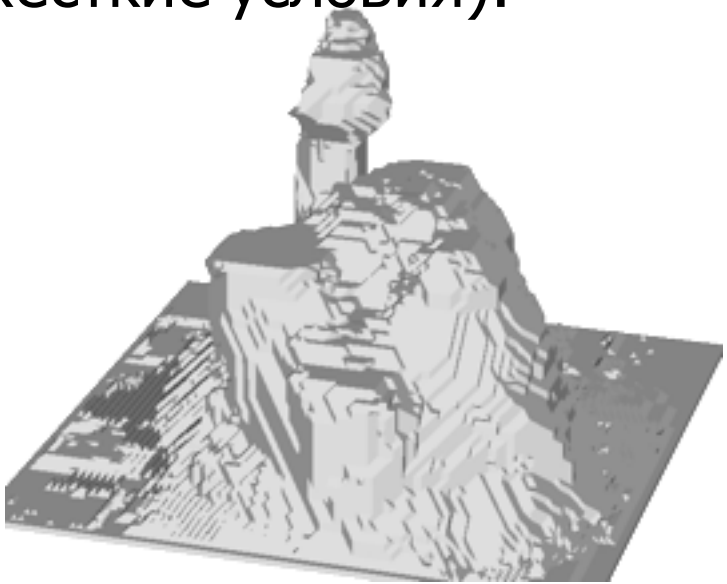
Вес разреза = значение энергии

Минимальный разрез ⇔ Глобальный минимум энергии

Реконструкция: пример
[Lempitsky,Boykov,Ivanov 2006]

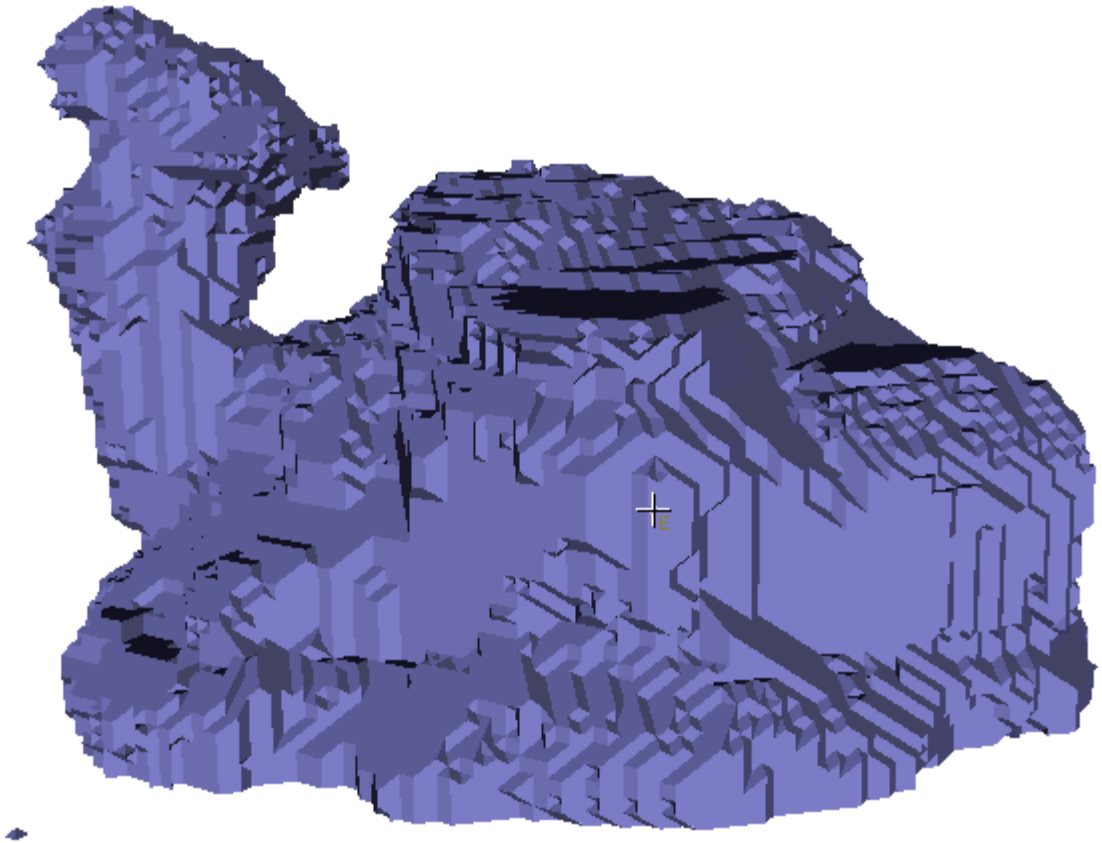
3 из 16 изображений

Результат реконструкции ($U(x)$ использует жесткие условия):

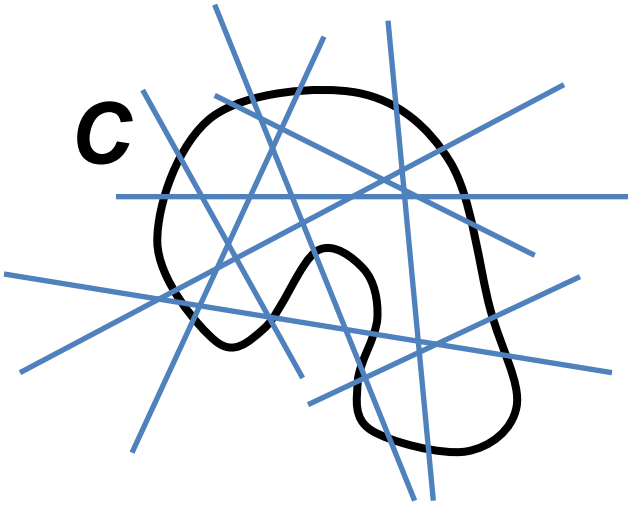
Улучшенный U(x) [Boykov,Lempitsky 2006]

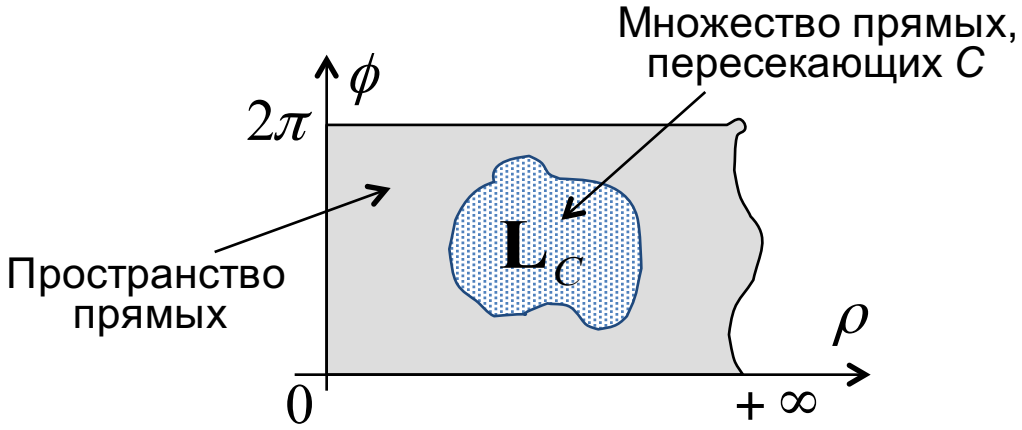
11/25 16:02 MSRC-1291309 t-victor AutoScreenRecorder



Bitmanagement
Software GmbH

Формула Коши-Крофтона (slide from Yuri Boykov)



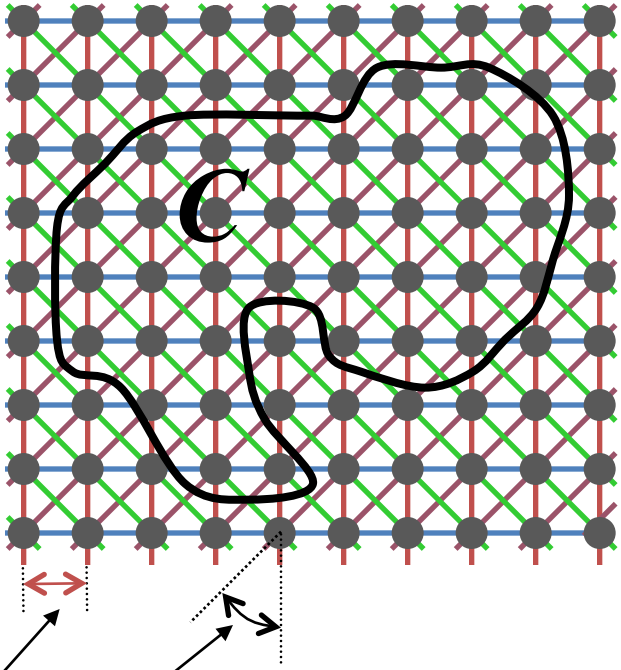


Евклидова длина **C** :

$$\| C \|_{\varepsilon} = \frac{1}{2} \int n_L \cdot d\rho \cdot d\phi$$

количество пересечений

Стоимость разреза аппроксимирует длину (slide from Yuri Boykov)



$$\| C \|_{\varepsilon} \approx \frac{1}{2} \sum_{k} n_k \cdot \Delta\rho_k \cdot \Delta\phi_k = \| C \|_{gc}$$

Евклидова длина

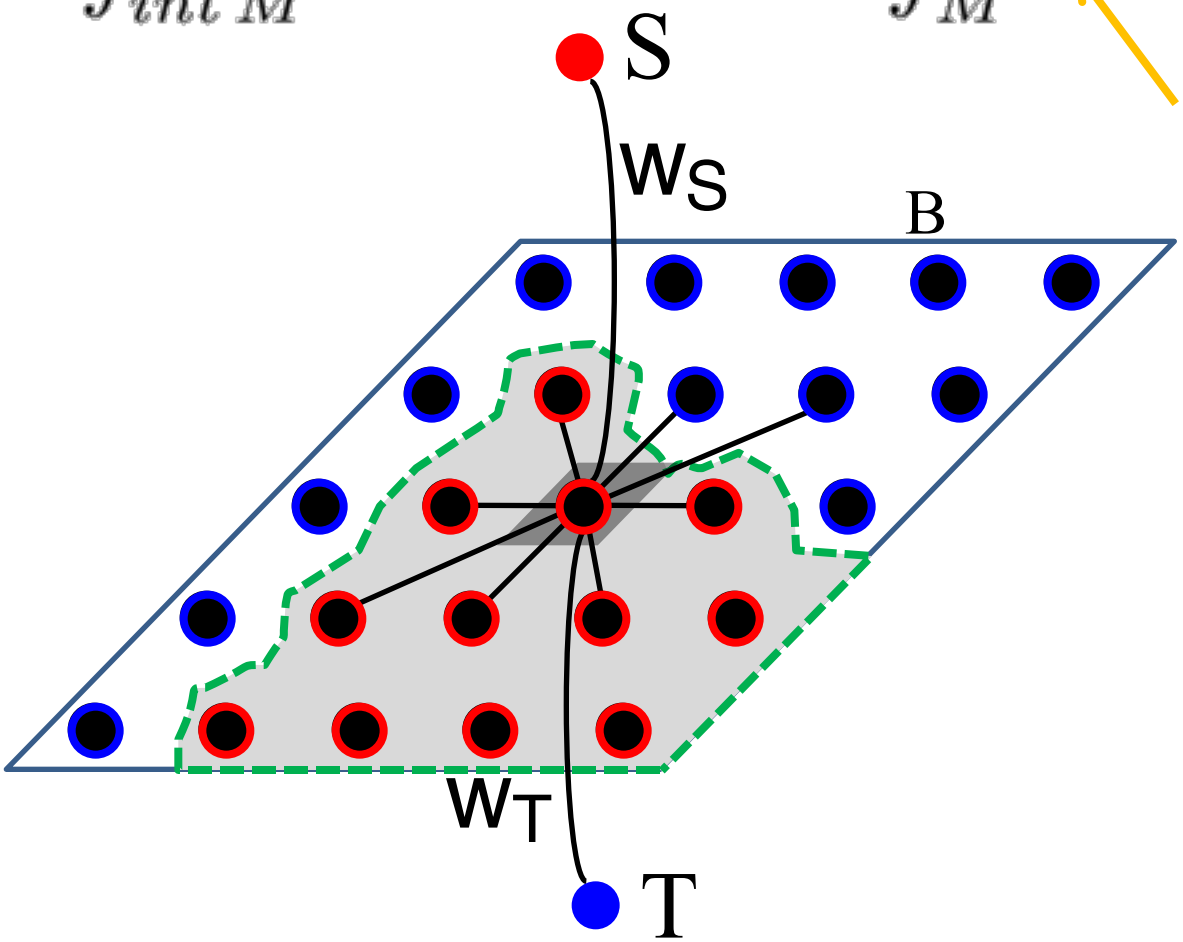
Количество пересеченных граней C

Вес ребра в графе:

$$w_k = \frac{\Delta\rho_k \cdot \Delta\phi_k}{2}$$

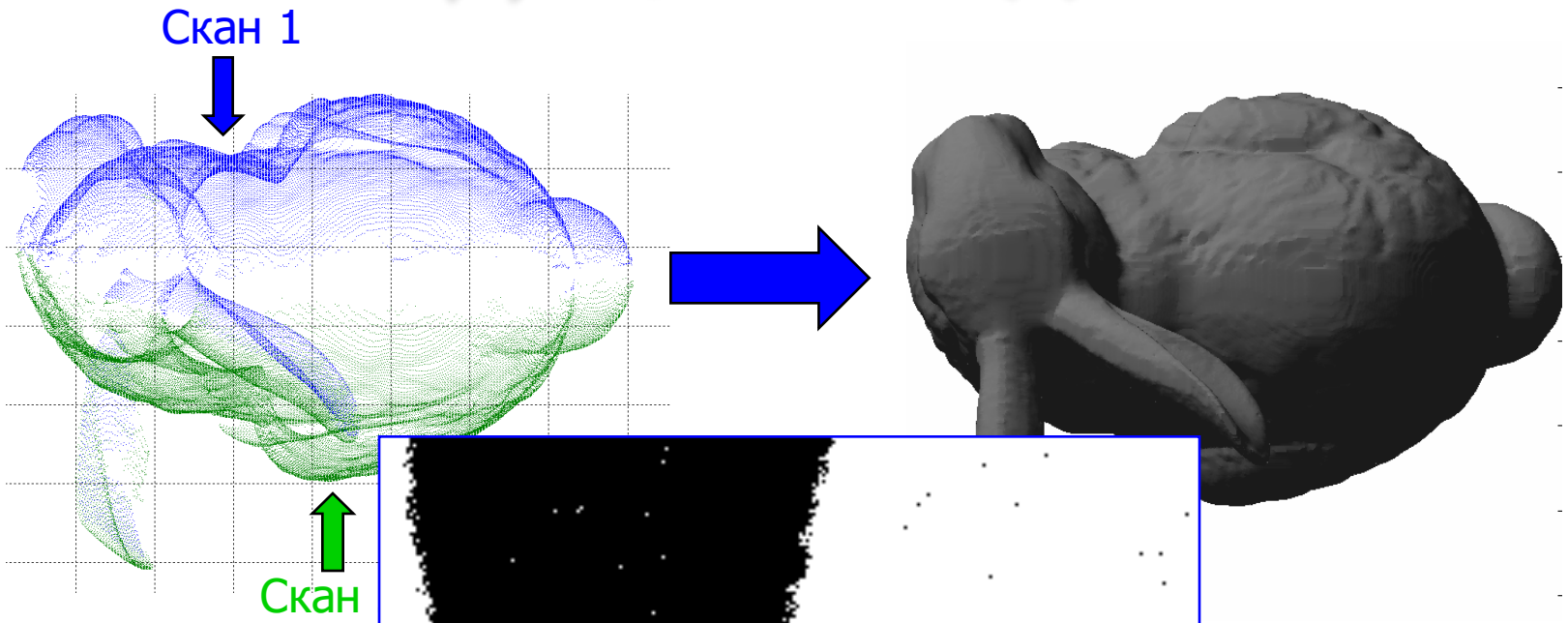
«Гео-разрез»
[Boykov&Kolmogorov 2003]

$$\int_{int\,M} U(X)\,dV + \lambda \int_{M} dS \rightarrow \min$$



Можно заменить на риманову метрику

Реконструкция по 3Д сканам



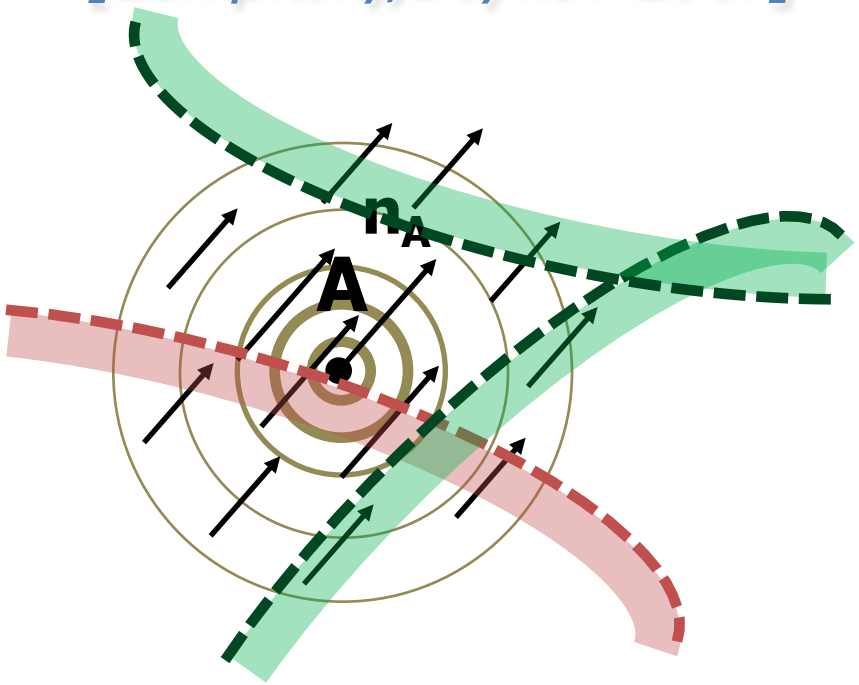
На входе: точки нками нормалей

На выходе: пов)

Сложности: шум бки регистрации

От точек к функционалу

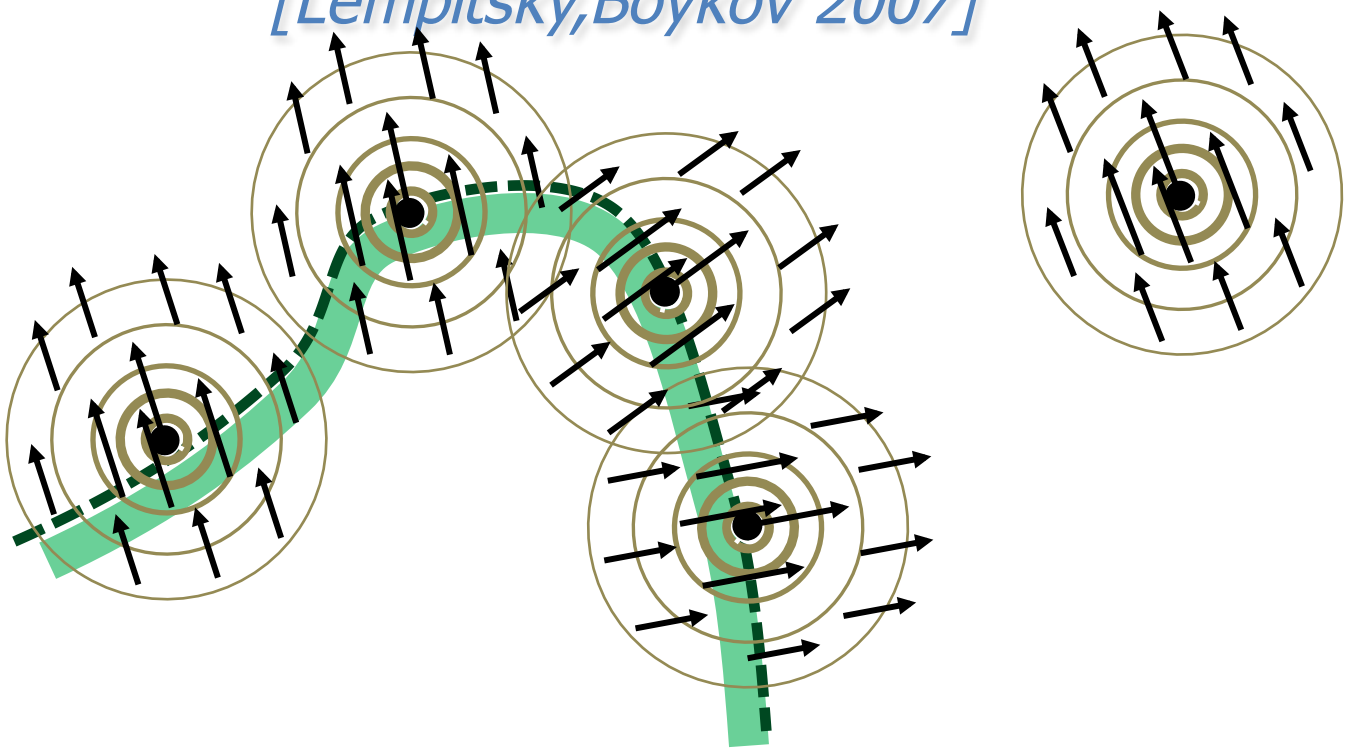
[Lempitsky,Boykov 2007]



- Неопределенность моделируется векторным полем
- Минимизация потока «притягивает» поверхность к точке: $\int_M (n_M(S), v_A(X))\, dS \to \max$

Формулировка энергии

[Lempitsky,Boykov 2007]

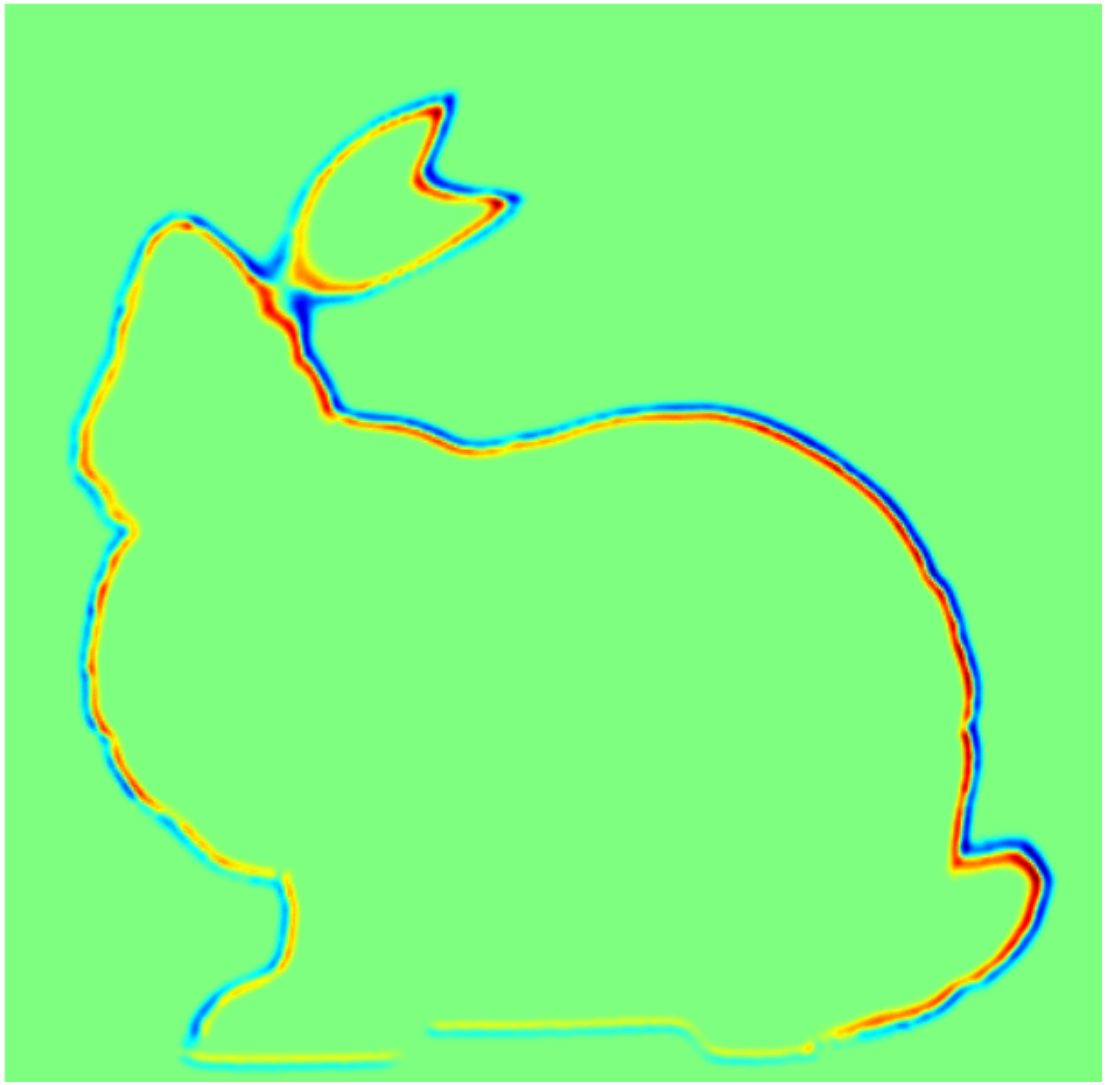


Энергия = сумма потоков векторных полей + регуляризация

$$-\int_M (n_M(S), \sum_A v_A(X))\, dS \;+\; \lambda \int_M dS \;\to \min$$

или $$-\int_{int\, M} div \sum_A v_A(X)\, dV \;+\; \lambda \int_M dS \;\to \min$$

Энергия (пример)



Срез ограничивающего объема, цвет соответствует дивергенции

Поверхность георазреза

Размер сетки:
551x544x428

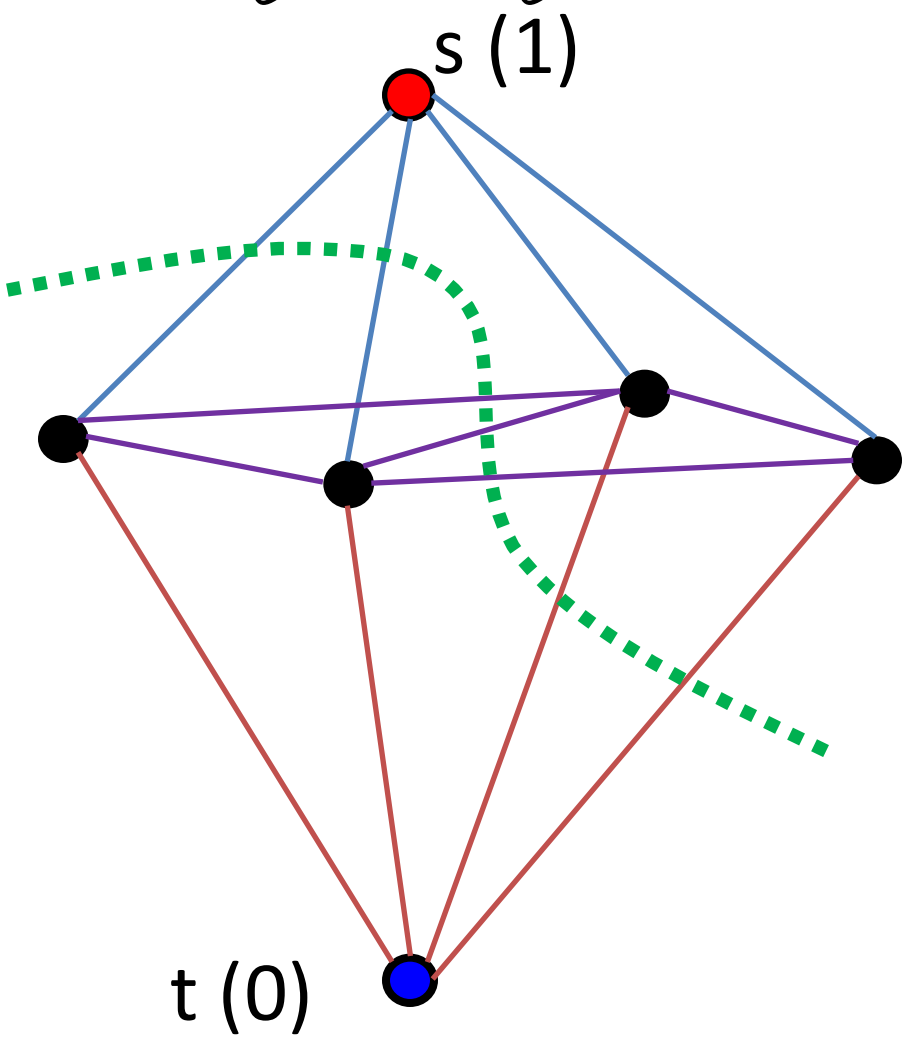
[Lempitsky, Boykov 07]:
Как использовать особенности графа, чтобы посчитать поток

Пример 2

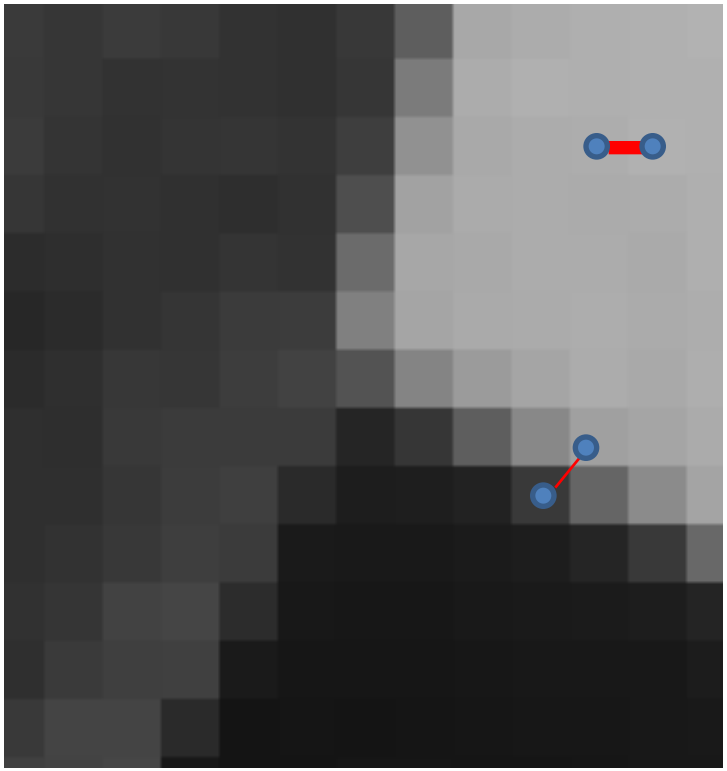
Реализация энергии на графе

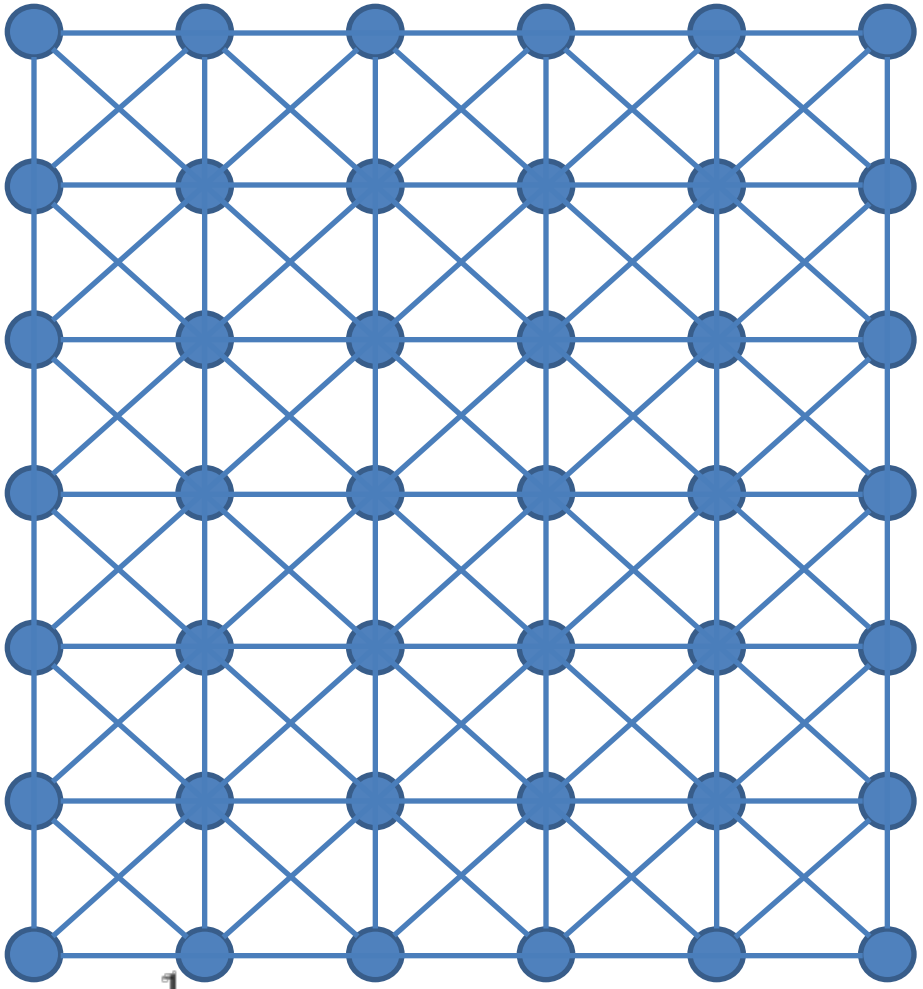
$$\mathbf{x} \in \{0,1\}^{\mathcal{V}} \quad E(\mathbf{x}) = \sum_{p \in \mathcal{V}} F^p \cdot x_p + \sum_{p \in \mathcal{V}} B^p \cdot (1 - x_p) + \sum_{p,q \in \mathcal{E}} P^{pq} |x_p - x_q|$$

$F^p \geq 0$, $B^p \geq 0$, $P^{pq} \geq 0$



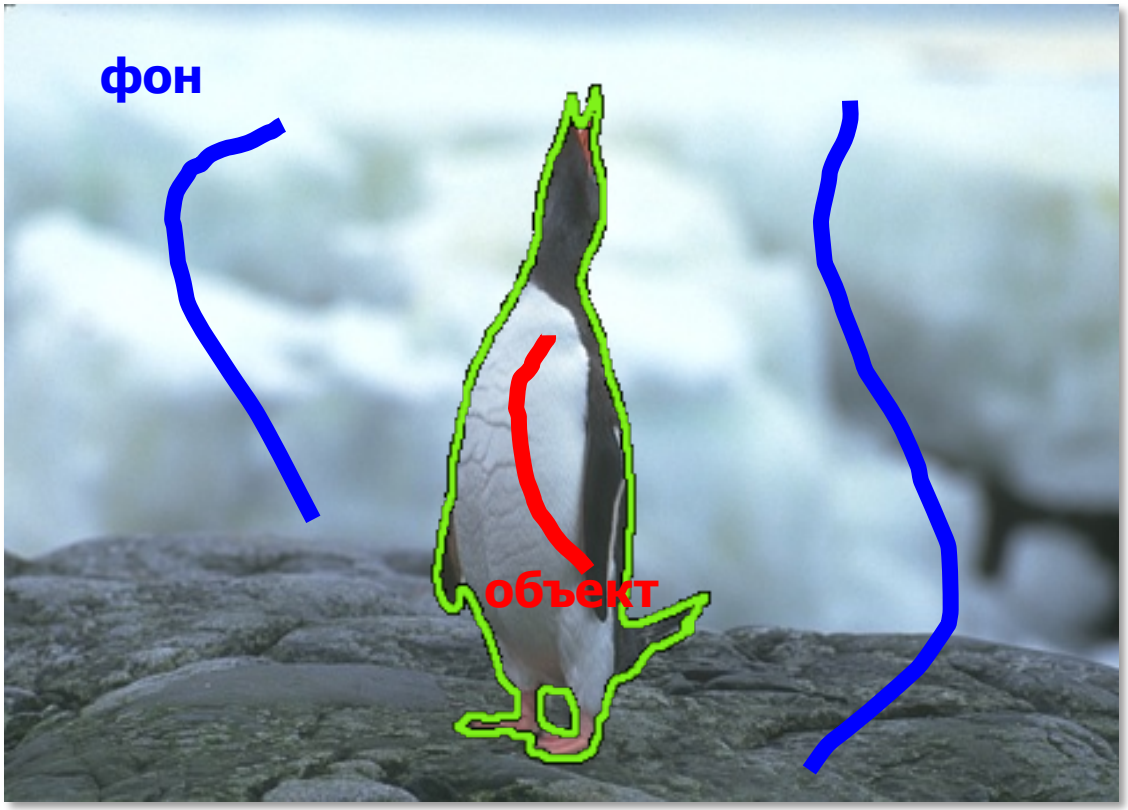
Сегментация с помощью разреза [Boykov & Jolly '01]



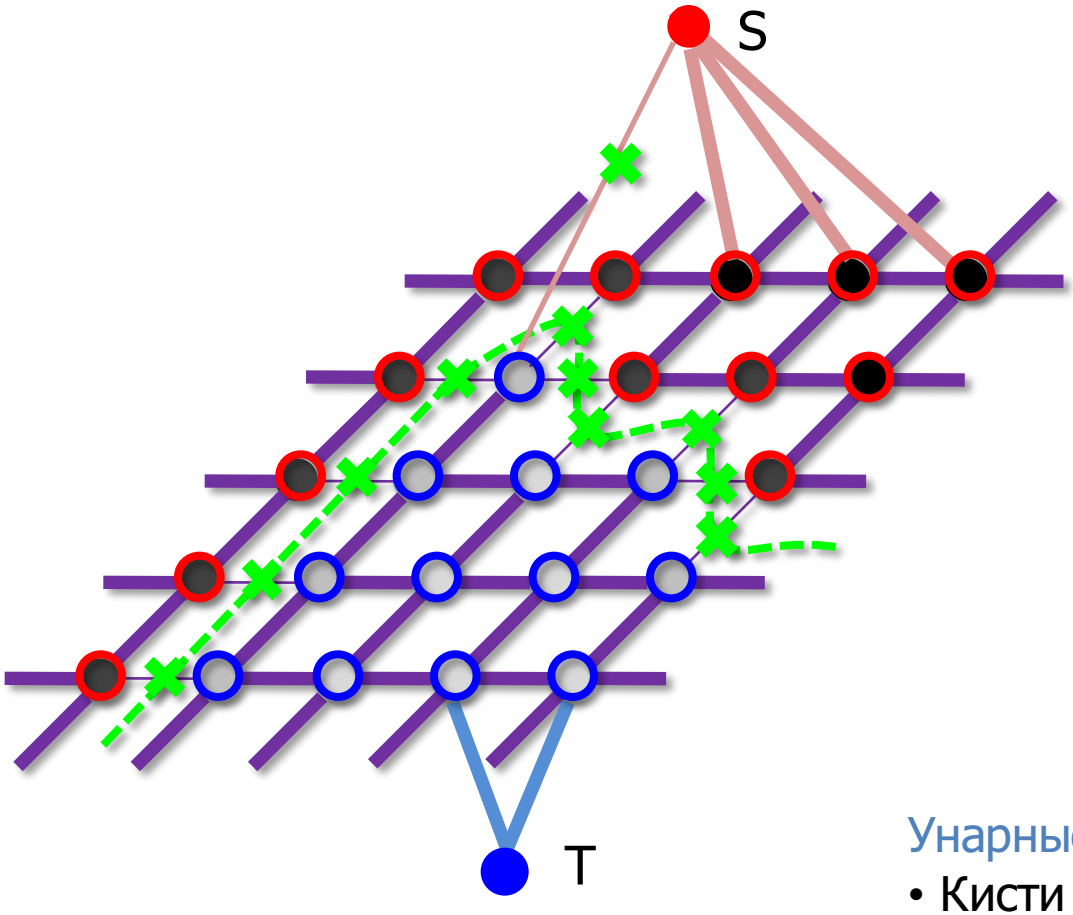


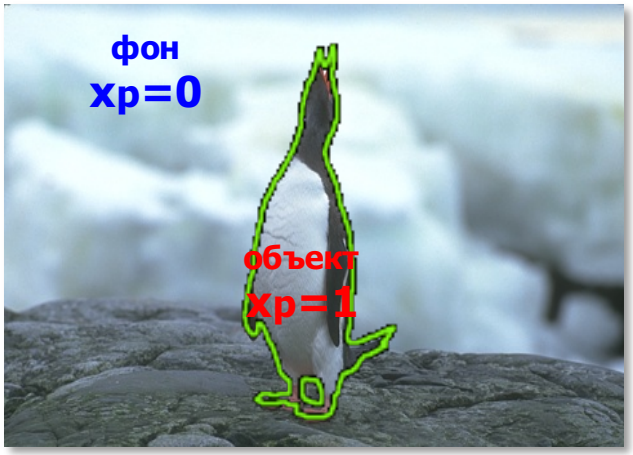
$$B_{\{p,q\}} \propto exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{dist(p,q)}$$

«Кистевой» интерфейс [Boykov & Jolly '01]



Сегментация с помощью разреза [Boykov & Jolly '01]





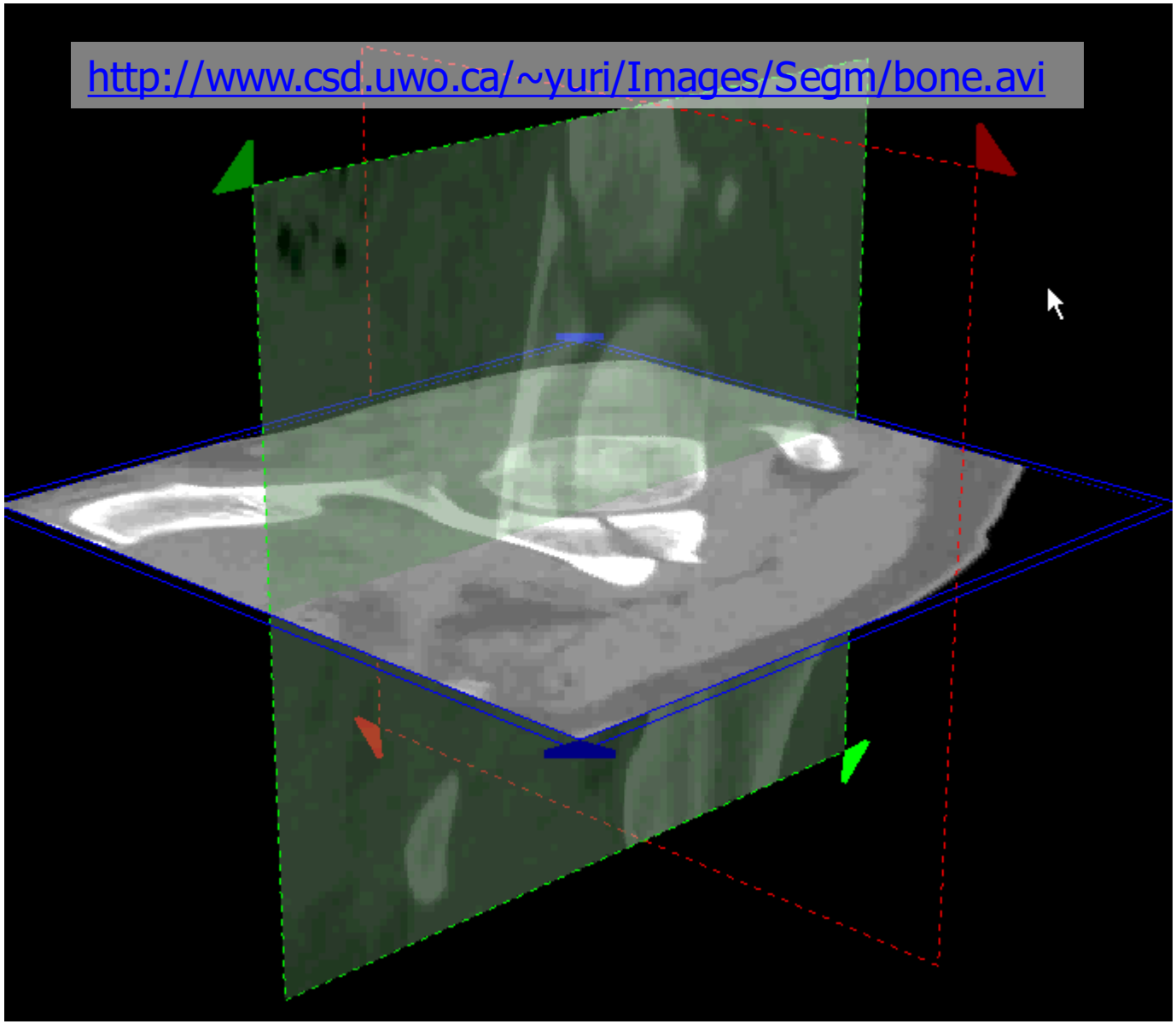
Унарные члены:
- Кисти
- Вероятностная модель

Парные члены:
- Перепады цвета

$$E(\mathbf{x}) = \sum_{p \in \mathcal{V}} F^p \cdot x_p + \sum_{p \in \mathcal{V}} B^p \cdot (1 - x_p) + \sum_{p,q \in \mathcal{E}} P^{pq} |x_p - x_q|$$

Пример сегментации

Видео из [Boykov, Kolmogorov, 2003]



Реализация более общих энергий [Kolmogorov & Zabih 2004]

$\mathbf{x} \in \{0,1\}^{\mathcal{V}}$ $\quad E(\mathbf{x}) = \sum_{p \in \mathcal{V}} F^p \cdot x_p + \sum_{p \in \mathcal{V}} B^p \cdot (1 - x_p) + \sum_{p,q \in \mathcal{E}} P^{pq} \cdot |x_p - x_q|$

≥ 0 $\quad \geq 0$ $\quad \geq 0$

$$E(\mathbf{x}) = \sum_{p \in \mathcal{V}} U_p(x_p) + \sum_{p,q \in \mathcal{N}} V_{pq}(x_p, x_q) \quad = ?$$

	$x_q = 0$	$x_q = 1$
$x_p = 0$	**A =$V_{pq}(0,0)$**	**B = $V_{pq}(0,1)$**
$x_p = 1$	**C = $V_{pq}(1,0)$**	**D = $V_{pq}(1,1)$**

=

A	**A**	+	**0**	**0**	+	**0**	**D-C**	+	**0**	**B+C-A-D**
A	**A**		**C-A**	**C-A**		**0**	**D-C**		**0**	**0**

Достаточное условие: $V_{pq}(0,0) + V_{pq}(1,1) \leq V_{pq}(0,1) + V_{pq}(1,0)$

α-Расширение [Boykov, Veksler, Zabih, 1998] (слайд Юрия Бойкова)

$$E(\mathbf{x}) = \sum_{p \in \mathcal{V}} U_p(x_p) + \sum_{p,q \in \mathcal{N}} V_{pq}(x_p, x_q), \quad \mathbf{x} \in \mathbf{L}^{\mathcal{V}}$$

начальное решение

● -расширение

● -расширение

● -расширение

● -расширение

● -расширение

● -расширение

● -расширение

Каждый шаг – бинарная оптимизация

α-Расширение [Boykov, Veksler, Zabih, 1998]

$$E(\mathbf{x}) = \sum_{p \in \mathcal{V}} U_p(x_p) + \sum_{p,q \in \mathcal{N}} V_{pq}(x_p, x_q), \quad \mathbf{x} \in \mathbf{L}^{\mathcal{V}}$$

$$\mathbf{y} \in \{0,1\}^{\mathcal{V}}$$

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

$$E^f(\mathbf{y}) = E(\mathbf{x}^c(\mathbf{y})) = \sum_{p \in \mathcal{V}} U_p^f(y_p) + \sum_{p,q \in \mathcal{N}} V_{pq}^f(y_p, y_q)$$

$$V_{pq}^f(0,0) + V_{pq}^f(1,1) \leq V_{pq}^f(0,1) + V_{pq}^f(1,0)$$

$$V_{pq}(x_p^0, x_q^0) + V_{pq}(\alpha, \alpha) \leq V_{pq}(x_p^0, \alpha) + V_{pq}(\alpha, x_q^0)$$

$$V(i,j) \leq V(i,k) + V(k,j)$$

α-Расширение [Boykov, Veksler, Zabih, 1998]

$$V(i,j) \le V(i,k) + V(k,j)$$

$$V(i,j) = \lambda \cdot [i \neq j]$$

$$V(i,j) = \lambda \cdot \min(1, \frac{|i-j|}{K})$$

~~$$V(i,j) = \lambda \cdot \min(1, \frac{(i-j)^2}{K^2})$$~~

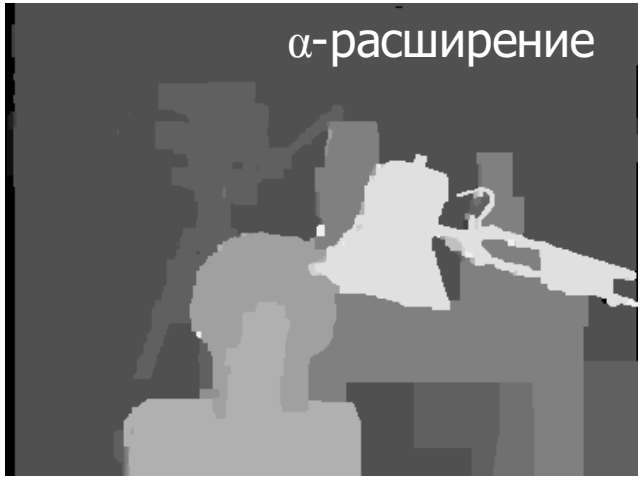
- Гарантии на сходимость
- Гарантии на близость к глобальному минимуму
- На практике быстрее и аккуратнее передачи сообщений

Сравнение моделей/методов

Results from Middlebury benchmark pages

$$E(\mathbf{d}) = \sum_{i=1}^{W} U_i(d_i) + \sum_{i=1}^{W-1} V(d_i, d_{i+1})$$

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_i(d_i) + \sum_{(p,q) \in \mathcal{T}} V(d_p, d_q)$$



110.20%

$$E(\mathbf{d}) = \sum_{p \in \mathcal{I}} U_i(d_i) + \sum_{(p,q) \in \mathcal{L}} V(d_p, d_q)$$

100.15%

Сшивка изображений [Agarwala et al. 2004]

Сшивка изображений с помощью α-расширений [Agarwala et al. 2004]

$$E(\mathbf{x}) = \sum_{p \in \mathcal{V}} U_p(x_p) + \sum_{p,q \in \mathcal{N}} V_{pq}(x_p, x_q), \quad \mathbf{x} \in \mathbf{L}^{\mathcal{V}}$$

$$V_{pq}(x_p, x_q) = ||I_{x_p}(p) - I_{x_q}(p)|| + ||I_{x_p}(q) - I_{x_q}(q)||$$

Related work and Resources (including those used to prepare the slides)

Berkeley Segmentation Dataset, http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/

Eric N. Mortensen, William A. Barrett: Intelligent scissors for image composition. SIGGRAPH 1995: 191-198

Eric N. Mortensen, William A. Barrett: Interactive Segmentation with Intelligent Scissors. Graphical Models and Image Processing 60(5): 349-384 (1998)

Shai Avidan, Ariel Shamir: Seam carving for content-aware image resizing. ACM Trans. Graph. 26(3): 10 (2007)

Middlebury stereo page, http://vision.middlebury.edu/stereo/

Daniel Scharstein, Richard Szeliski: A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. International Journal of Computer Vision 47(1-3): 7-42 (2002)

Olga Veksler: Stereo Correspondence by Dynamic Programming on a Tree. CVPR (2) 2005: 384-390

Related work and Resources (including those used to prepare the slides)

BioID dataset, http://www.bioid.com/support/downloads/software/bioid-face-database.html

F. Felzenszwalb, Daniel P. Huttenlocher: Pictorial Structures for Object Recognition. International Journal of Computer Vision 61(1): 55-79 (2005)

Pearl, J. Probabilistic Reasoning in Intelligent Systems, San Mateo: Morgan Kaufmann, 1988

Middlebury MRF page, http://vision.middlebury.edu/MRF/

Richard Szeliski, Ramin Zabih, Daniel Scharstein, Olga Veksler, Vladimir Kolmogorov, Aseem Agarwala, Marshall F. Tappen, Carsten Rother: A Comparative Study of Energy Minimization Methods for Markov Random Fields. ECCV (2) 2006: 16-29

William T. Freeman, Egon C. Pasztor, Owen T. Carmichael: Learning Low-Level Vision. International Journal of Computer Vision 40(1): 25-47 (2000)

William T. Freeman, Thouis R. Jones, and Egon C. Pasztor, *Example-based super-resolution*, IEEE Computer Graphics and Applications, March/April, 2002

Related work and Resources (including those used to prepare the slides)

Ford, L. R.; Fulkerson, D. R. (1956). "Maximal flow through a network". Canadian Journal of Mathematics 8: 399–404.

Yuri Boykov, Vladimir Kolmogorov: An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision. IEEE Trans. Pattern Anal. Mach. Intell. 26(9): 1124-1137 (2004)
Graph Cut code: http://www.cs.ucl.ac.uk/staff/V.Kolmogorov/software.html

Sébastien Roy, Ingemar J. Cox: A Maximum-Flow Formulation of the N-Camera Stereo Correspondence Problem. ICCV 1998: 492-502

Hiroshi Ishikawa, Davi Geiger: Occlusions, Discontinuities, and Epipolar Lines in Stereo. ECCV (1) 1998: 232-248

Victor S. Lempitsky, Yuri Boykov, Denis V. Ivanov: Oriented Visibility for Multiview Reconstruction. ECCV (3) 2006: 226-238

Yuri Boykov and Victor Lempitsky: Form Photohulls to Photoflux Optimization. BMVC, Edinburgh, 2006

Related work and Resources (including those used to prepare the slides)

Victor S. Lempitsky, Yuri Boykov: Global Optimization for Shape Fitting. CVPR 2007

The Stanford 3D scanning repository, http://graphics.stanford.edu/data/3Dscanrep/

Yuri Boykov, Marie-Pierre Jolly: Demonstration of Segmentation with Interactive Graph Cuts. ICCV 2001: 741

Vladimir Kolmogorov, Ramin Zabih: What Energy Functions Can Be Minimized via Graph Cuts? IEEE Trans. Pattern Anal. Mach. Intell. 26(2): 147-159 (2004)

Yuri Boykov, Olga Veksler, Ramin Zabih: Fast Approximate Energy Minimization via Graph Cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23(11): 1222-1239 (2001)

ECCV 2006 Tutorial on Graph Cuts versus Level Sets: http://www.csd.uwo.ca/~yuri/Abstracts/eccv06-tutorial.html

Aseem Agarwala, Mira Dontcheva, Maneesh Agrawala, Steven M. Drucker, Alex Colburn, Brian Curless, David Salesin, Michael F. Cohen: Interactive digital photomontage. ACM Trans. Graph. 23(3): 294-302 (2004)