Отчёт по второму заданию курса графических моделей «Алгоритмы минимизации парно-сепарабельной энергии для задачи склеивания панорам»

Кафедра математических методов прогнозирования

Бобров Евгений, 417 группа

19 апреля 2017

Минимизация энергии от бинарных переменных

Задача склеивания изображений состоит в построении одного составного изображения на основе набора исходных изображений. В рамках данного задания предполагается, что все исходные изображения выровнены друг относительно друга. В этих условиях задачу можно решать при помощи минимизации энергии, переменные которой соответствуют номеру изображения, из которого взят конкретный пиксель.

Рассмотрим неориентированный граф $G = (V, E)$. Каждой вершине $i \in V$ поставим в соответствие бинарную переменную $x_i \in \{0, 1\}$. Рассмотрим следующую *энергию* Q (функцию $Q : \{0, 1\}^{|V|} \to \mathbb{R}$):

$$Q(x) = \sum_{i \in V} \theta_i(x_i) + \sum_{\{i,j\} \in E} \theta_{ij}(x_i, x_j)$$

Здесь $\theta_i(x_i)$ – унарные потенциалы, $\theta_{ij}(x_i, x_j)$ – парные потенциалы. Энергия, состоящая из унарных, парных, и константного потенциалов называется парно-сепарабельной. Задача минимизации энергии Q состоит в поиске значения вектора

переменных $x = \{x_i\}, i \in V$, на котором энергия $Q(x)$ принимает минимальное значение.

Необходимые условия на потенциалы энергии Q, позволяющие перейти от задачи минимизации энергии к задаче поиска минимального разреза графа G:

1. Унарные потенциалы неотрицательны $\forall i \in V : \theta_i(0) \geq 0, \theta_i(1) \geq 1$.

2. Парные потенциалы неотрицательны и равны 0 при равенстве связанных переменных: $\forall \{i, j\} \in E, \theta_{ij}(0,0) = \theta_{ij}(1,1) = 0, \theta_{ij}(0,1) \geq 0, \theta_{ij}(1,0) \geq 0$

В этом случае энергию $Q(x)$ можно записать в следующем виде:

$$Q(x) = \sum_{i \in V} (x_i \theta_i(1) + (1 - x_i)\theta_i(0)) + \sum_{\{i,j\} \in E} (x_i(1 - x_j)\theta_{ij}(1,0) + x_j(1 - x_i)\theta_{ij}(0,1))$$

По энергии $Q(x)$ построим ориентированный граф $G = (V, E)$, разрез которого будет соответствовать присвоению значений переменным x.

Минимизация энергии от K-значных переменных

Рассмотрим неориентированный граф $G = (V, E)$. Каждой вершине $i \in V$ поставим в соответствие дискретную переменную $y_i \in \{1, ..., K\}$. Рассмотрим следующую энергию W (функцию $W : \{1, ..., K\}^{|V|} \to \mathbb{R}$):

$$W(y) = \sum_{i \in V} \psi_i(y_i) + \sum_{\{i,j\} \in E} \psi_{ij}(y_i, y_j)$$

Здесь $\psi_i(y_i)$ – унарные потенциалы, $\psi_{ij}(y_i, y_j)$ – парные потенциалы. Задача минимизации энергии $W(y)$ состоит в поиске значения вектора переменных $y = \{y_i\}, i \in V$, на котором энергия $W(y)$ принимает минимальное значение.

Для уменьшения числа параметров, парные потенциалы можно параметризовать следующим образом: $\psi_{ij}(y_i, y_j) = c_{ij} d(y_i, y_j)$. Где c_{ij} – коэффициенты, соответствующие рёбрам графа, а $d(y_i, y_j)$ – расстояние между метками соседних переменных.

Алгоритм альфа-расширения и репараметризация

Приведем формальное описание каждого шага алгоритма α-расширение. Пусть есть текущая разметка y и выбрана «расширяемая» метка $\alpha \in \{1, ..., K\}$. Будем строить энергию бинарных переменных $E(x)$ по энергии $W(y)$.

- Граф энергий $E(x)$ и $W(y)$ одинаков $G = (V, E)$.
- Значение бинарной переменной $x_i = 0$ соответствует $y_i = y_i$ (значение переменной y_i не меняется); $x_i = 1$ соответствует $y_i = \alpha$.
- Унарные потенциалы:

$$\theta_i(0) = \psi_i(y_i), \theta_i(1) = \psi_i(\alpha)$$

- Парные потенциалы:

$$\theta_{ij}(0,0) = \psi_{ij}(y_i, y_j) = c_{ij} d(y_i, y_j),\ \theta_{ij}(0,1) = \psi_{ij}(y, \alpha) = c_{ij} d(y, \alpha)$$

$$\theta_{ij}(1,0) = \psi_{ij}(\alpha, y) = c_{ij} d(\alpha, y),\ \theta_{ij}(1,1) = \psi_{ij}(\alpha, \alpha) = 0$$

Применим следующую последовательность эквивалентных преобразований, не изменяющих энергию $W(y)$:

$$\psi_i(\alpha) := \psi_i(\alpha) + c_{ij} d(y_i, y_j)$$

$$\psi_j(y_j) := \psi_j(y_j) + c_{ij} d(y_i, \alpha) - c_{ij} d(y_i, y_j)$$

$$\psi_i(y_i) := \psi_i(y_i) + c_{ij}d(y_i, y_j) - c_{ij}d(y_i, \alpha)$$

$$d(\alpha, y_j) := d(\alpha, y_j) + d(y_i, \alpha) - d(y_i, y_j)$$

$$\{c_{ij}d(\alpha, y_j) := c_{ij}d(\alpha, y_j) + c_{ij}d(y_i, \alpha) - c_{ij}d(y_i, y_j)\}$$

$$d(y_i, y_j) = d(y_i, \alpha) = d(\alpha, \alpha) = 0$$

Условие неотрицательности парных потенциалов выполнено, когда $d(\alpha, y_j) + d(y_i, \alpha) - d(y_i, y_j) \geq 0$, то есть $d(y_i, y_j) \leq d(\alpha, y_j) + d(y_i, \alpha)$. Последнее условие является неравенством треугольника, и выполнено, когда d – метрика в дискретном пространстве $\{1...K\}$, и $c_{ij} > 0 \ \forall i, j$

Таким образом задача минимизации энергии от K-значных переменных сведена к задаче минимизации энергии от бинарных переменных. Последняя же решается алгоритмами поиска минимального разреза построенного по данной энергии графа.

Эксперименты

В качестве метрики для парных потенциалов была выбрана индикаторная функция, тогда применяются так называемые *потенциалы Поттса*:

$$\psi_{ij}(y_i, y_j) = c_{ij}d(y_i, y_j) = c_{ij}[y_i \neq y_j]$$

Коэффициенты c_{ij}, соответствующие рёбрам графа, будем рассчитывать как среднеквадратичное отклонение разностей соседних вершин по всем изображениям $\{1...K\}$, усреднённое по трём цветовым каналам.

Съёмка проводилась со статичной камерой на штативе и со съёмкой на телефон от руки. Также были смонтированы скриншоты из видеолекций ШАДа.

Рис. 1: Статичная камера, два изображения

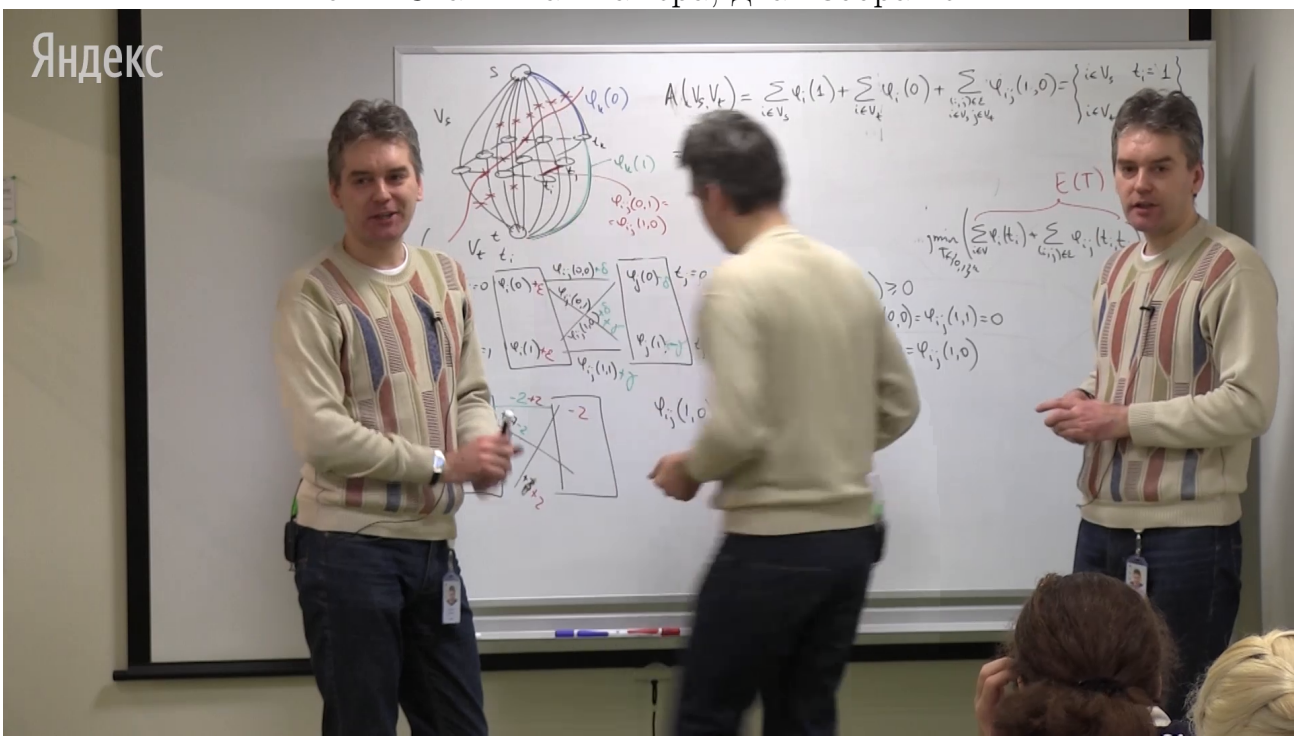

Рис. 2: Скриншоты лекции, три изображения

Рис. 3: Съёмка от руки, три изображения

Рис. 4: Статичная камера, восемь изображений