Полносвязное CRF
с гауссовскими потенциалами

Влад Шахуро

3 марта 2017

CRF для многоклассовой сегментации

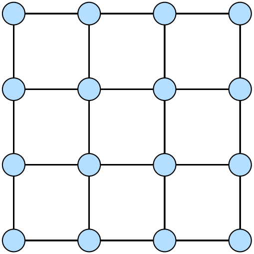
$$E(\boldsymbol{x}) = \sum_{i \in \mathcal{V}} \underbrace{\varphi_i(x_i)}_{\text{унарный}} + \sum_{(i,j) \in \mathcal{E}} \underbrace{\varphi_{ij}(x_i, x_j)}_{\text{парный}}$$

Унарный потенциал — выход классификатора: бустинг, нейросеть

Парный потенциал — штраф за разные метки соседних пикселей:

$$\varphi_{ij}(x_i, x_j) = \underbrace{[x_i \neq x_j]}_{\text{модель Поттса}} \left(w_1 + w_2 \exp\left(-\frac{\|I_i - I_j\|^2}{2\sigma^2} \right) \right)$$

4-связная система соседства

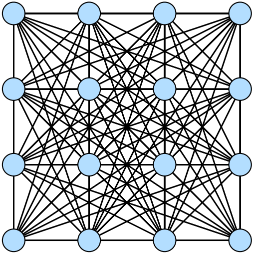


Учитываются только локальные связи пикселей

Границы объектов сглаживаются и уменьшаются

Оптимизация с помощью разреза графов: 1 сек для 50 000 переменных

Полносвязная система соседства



С разными весами учитываются дальние связи пикселей

Нет тендеции к уменьшению длины границы

Вычислительно сложно

- MCMC: 36 часов, не оптимум
- Разрез графов: > 72 часов, $O(|\mathcal{V}||\mathcal{E}|^2)$

Гауссовские парные потенциалы

$$\varphi_{ij}(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^{M} w_m \underbrace{K_m(\mathbf{f}_i, \mathbf{f}_j)}_{\text{гауссовское ядро}}$$

- μ — функция совместимости меток: модель Поттса или полуметрика, обучаемая на данных
- пиксель изображения описывается вектор-признаком $\mathbf{f}$

Гауссовские парные потенциалы

$$\varphi_{ij}(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^{M} w_m \underbrace{K_m(\mathbf{f}_i, \mathbf{f}_j)}_{\text{гауссовское ядро}}$$

$$= \mu(x_i, x_j)(w_1 \exp \underbrace{\left(-\frac{\|\mathbf{p}_i - \mathbf{p}_j\|}{2\sigma_1^2} - \frac{\|\mathbf{I}_i - \mathbf{I}_j\|}{2\sigma_2^2} \right)}_{\text{чувствительность к разнице цвета}}$$

$$+ w_2 \exp \underbrace{\left(-\frac{\|\mathbf{p}_i - \mathbf{p}_j\|}{2\sigma_3^2} \right)}_{\text{шумоподавление}})$$

- μ — функция совместимости меток: модель Поттса или полуметрика, обучаемая на данных
- пиксель изображения описывается вектор-признаком $\mathbf{f}$

Вариационный вывод

$$x^* = \arg\max_{x} P(x), \quad P(x) = \exp(-E(x))$$

Приближение самосогласованного поля (mean-field):

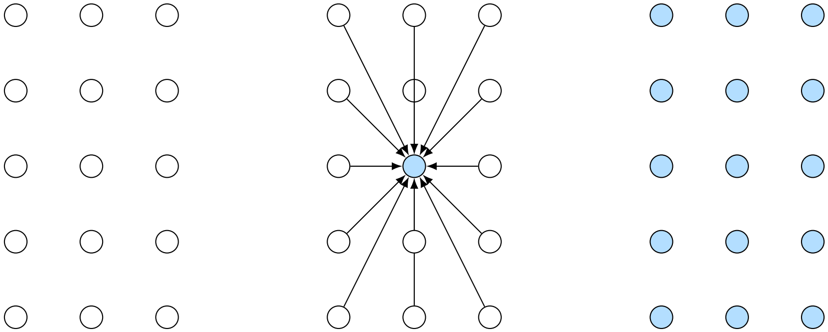
$$\begin{cases} KL(Q \,\|\, P) \to \min_Q \\ Q(x) = \prod_i Q_i(x_i) \end{cases}$$

$$x_i^* \approx \arg\max_{x_i} Q_i(x_i)$$

Существует итерационный процесс, который сходится к локальному минимуму задачи

Koller, Friedman. Probabilistic Graphical Models. Chapter 11.5

Передача сообщений



$Q(\boldsymbol{x})$ $Q'(\boldsymbol{x})$

Итерационный процесс

Инициализируем $Q_i(l) = \frac{1}{Z_i}\exp(-\varphi_i(l)) \quad \forall i,\ l$

Повторяем до сходимости:

1. Передача сообщений:
   $Q_i'^m(l) = \sum_{j \neq i} K_m(\mathbf{f}_i, \mathbf{f}_j) Q_j(l) \quad \forall i,\ l,\ m$

2. Суммирование сообщений по ядрам:
   $Q_i'(l) = \sum_{m=1}^{M} w_m Q_i'^m(l) \quad \forall i,\ l$

3. Суммирование сообщений с функцией совместимости:
   $\overline{Q}_i(l) = \sum_{l' \in \mathcal{L}} \mu(l, l') Q_i'(l') \quad \forall i,\ l$

4. Добавляем унарный потенциал:
   $Q_i(l) = \exp(-\varphi_i(x_i) - \overline{Q}_i(l)) \quad \forall i,\ l$

5. Нормализуем $Q_i(l), \quad \forall i,\ l$

Итерационный процесс

$O(Nd)$ Инициализируем $Q_i(l) = \frac{1}{Z_i}\exp(-\varphi_i(l)) \quad \forall i,\ l$

≈ 10 Повторяем до сходимости:

$O(N^2Md)$ Передача сообщений:
$Q_i'^m(l) = \sum_{j \neq i} K_m(\mathbf{f}_i, \mathbf{f}_j) Q_j(l) \quad \forall i,\ l,\ m$

$O(NMd)$ Суммирование сообщений по ядрам:
$Q_i'(l) = \sum_{m=1}^{M} w_m Q_i'^m(l) \quad \forall i,\ l$

$O(Nd)$ Суммирование сообщений с функцией совместимости:
$\overline{Q}_i(l) = \sum_{l' \in \mathcal{L}} \mu(l, l') Q_i'(l') \quad \forall i,\ l$

$O(Nd)$ Добавляем унарный потенциал:
$Q_i(l) = \exp(-\varphi_i(x_i) - \overline{Q}_i(l)) \quad \forall i,\ l$

$O(Nd)$ Нормализуем $Q_i(l), \quad \forall i,\ l$

Итерационный процесс

$O(Nd)$ Инициализируем $Q_i(l) = \frac{1}{Z_i}\exp(-\varphi_i(l)) \quad \forall i,\ l$

≈ 10 Повторяем до сходимости:

$O(N^2Md)$ Передача сообщений:
$Q_i'^m(l) = \sum_{j \neq i} K_m(\mathbf{f}_i, \mathbf{f}_j) Q_j(l) \quad \forall i,\ l,\ m$

$O(NMd)$ Суммирование сообщений по ядрам:
$Q_i'(l) = \sum_{m=1}^{M} w_m Q_i'^m(l) \quad \forall i,\ l$

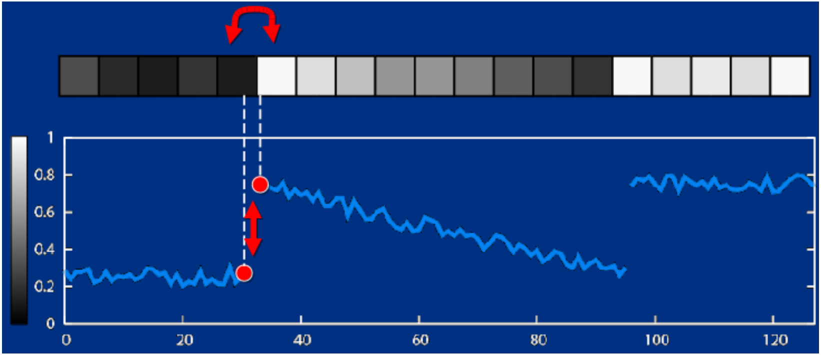
$O(Nd)$ Суммирование сообщений с функцией совместимости:
$\overline{Q}_i(l) = \sum_{l' \in \mathcal{L}} \mu(l, l') Q_i'(l') \quad \forall i,\ l$

$O(Nd)$ Добавляем унарный потенциал:
$Q_i(l) = \exp(-\varphi_i(x_i) - \overline{Q}_i(l)) \quad \forall i,\ l$

$O(Nd)$ Нормализуем $Q_i(l), \quad \forall i,\ l$

Билатеральный фильтр для 1d сигнала

$$BF[I]_i = \frac{1}{W_i} \sum_{j \in S} G_{\sigma_s}(\|i - j\|) G_{\sigma_r}(\|I_i - I_j\|) I_i$$



Билатеральная фильтрация эквивалентна гауссовской фильтрации в пространстве $S \times R$

Анализ итерационного процесса

$$Q_i^{\prime m}(l) = \sum_{j \neq i} K_m(\mathbf{f}_i, \mathbf{f}_j) Q_j(l) \quad \forall i,\ l,\ m$$

$K_m(\mathbf{f}_i, \mathbf{f}_j)$ — гауссовское ядро

Можем заменить сумму на гауссовскую свёртку:

$$Q^{\prime m}(l) = K_m \otimes Q(l), \quad \forall l,\ m$$

Существует структура данных, которая приближенно сворачивает сигнал высокой размерности с гауссианой за линейное время от количества сэмплов

Adams et al. Fast High-Dimensional Filtering Using the Permutohedral Lattice. Computer Graphics Forum, 2010

Итерационный процесс

$O(Nd)$ Инициализируем $Q_i(l) = \frac{1}{Z_i}\exp(-\varphi_i(l)) \quad \forall i,\ l$

≈ 10 Повторяем до сходимости:

$O(NMd)$ Передача сообщений:
$Q'^m(l) = K_m \otimes Q(l), \quad \forall l,\ m$

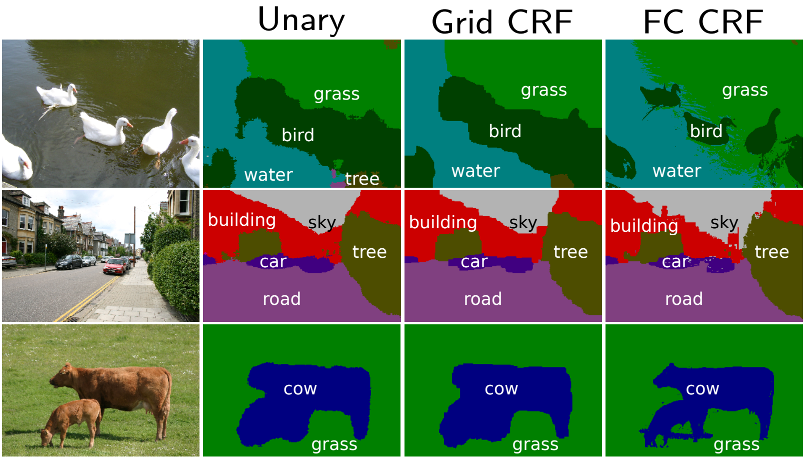
$O(NMd)$ Суммирование сообщений по ядрам:
$Q'_i(l) = \sum_{m=1}^{M} w_m Q'^m_i(l) \quad \forall i,\ l$

$O(Nd)$ Суммирование сообщений с функцией совместимости:
$\overline{Q}_i(l) = \sum_{l' \in \mathcal{L}} \mu(l, l') Q'_i(l') \quad \forall i,\ l$

$O(Nd)$ Добавляем унарный потенциал:
$Q_i(l) = \exp(-\varphi_i(x_i) - \overline{Q}_i(l)) \quad \forall i,\ l$

$O(Nd)$ Нормализуем $Q_i(l), \quad \forall i,\ l$

Примеры работы на датасете MSRC



Krähenbühl et al. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials. NIPS 2011

Резюме

Используем полносвязное условное случайное поле, которое не склонно сильно сглаживать границы

Аппроксимируем распределение mean-field приближением, для которого существует сходящийся итерационный процесс, называемый передачей сообщений

Если парные потенциалы — гауссовские, можем быстро пересчитывать сообщения с помощью гауссовских сверток в пространстве высокой размерности