

MRF + ShapeVM для сегментации изображений

Екатерина Лобачева, Александр Кириллов, Дмитрий Ветров

lobacheva.tjulja@gmail.com

2014

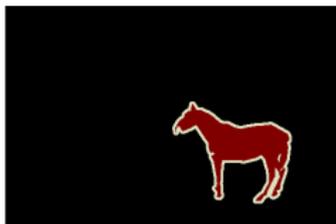
Оглавление

- 1 **Постановка задачи**
- 2 Основы
 - MRF
 - ShapeBM
- 3 MRF + ShapeBM
 - Строение
 - Вывод: EM
 - Эксперименты
- 4 Применение выхода детектора
- 5 Future work

Глобальная задача

Вход: изображение и информация о типе находящегося на нем объекта.

Выход: попиксельная разметка изображения на два класса: объект и фон.



Локальная задача

Вход: центрированное и масштабированное изображение и информация о типе находящегося на нем объекта.

Выход: попиксельная разметка изображения на два класса: объект и фон.



Оглавление

- 1 Постановка задачи
- 2 Основы
 - MRF
 - ShapeBM
- 3 MRF + ShapeBM
 - Строение
 - Вывод: EM
 - Эксперименты
- 4 Применение выхода детектора
- 5 Future work

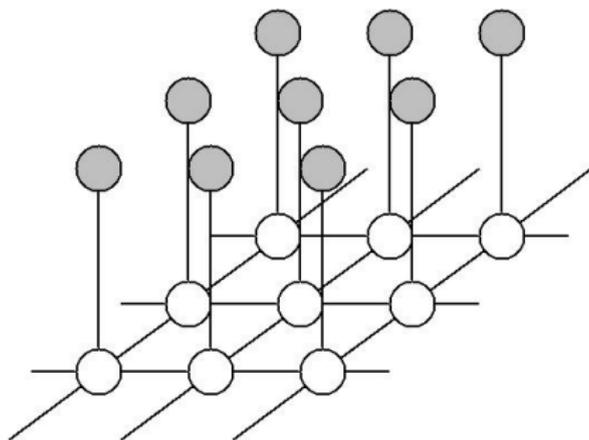
Markov Random Fields

Пусть дан граф $G(\mathbf{v}, \mathbf{e})$ и вектор меток для вершин этого графа $\mathbf{y} \in \{0, 1\}^{|\mathbf{v}|}$ (0 – фон, 1 – объект).

Энергия MRF:

$$E(\mathbf{y}) = (1 - C) \sum_{v \in \mathbf{v}} \phi_v(y_v) + C \sum_{(v,u) \in \mathbf{e}} \phi_{vu}(y_v, y_u). \quad (1)$$

Цель: $\mathbf{y}^* = \operatorname{argmin}_{\mathbf{y}} E(\mathbf{y})$.



MRF

Markov Random Fields

- Унарные потенциалы $\phi_i(v_i)$ определяются с помощью модели цвета, которая обучается на размеченной обучающей выборке. Они определяют вероятность принадлежности каждого пикселя к объекту или фону.
- Парные потенциалы $\phi_{i,j}(v_{i,j})$ задаются как взвешенные потенциалы Поттса и штрафуют разметку за прохождение границы по участку изображения со слабым градиентом.
- Существует эффективный алгоритм вывода — graph cuts.



MRF

Недостатки MRF

- MRF не принимает во внимание форму объекта
- не умеет эффективно работать с глобальными ограничениями, например формой объекта



ShapeBM

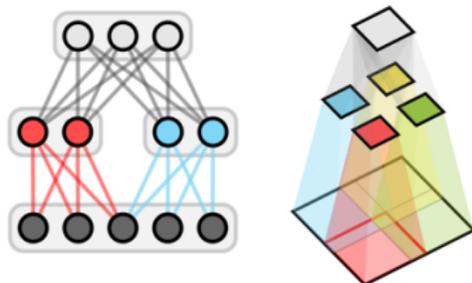
ShapeBM [S. M. Ali Eslami et al. 2012]

Shape Boltzmann Machine (ShapeBM) — вариант Deep Boltzmann Machine с обнулением (и, возможно, отождествлением) некоторых весов.

Энергия:

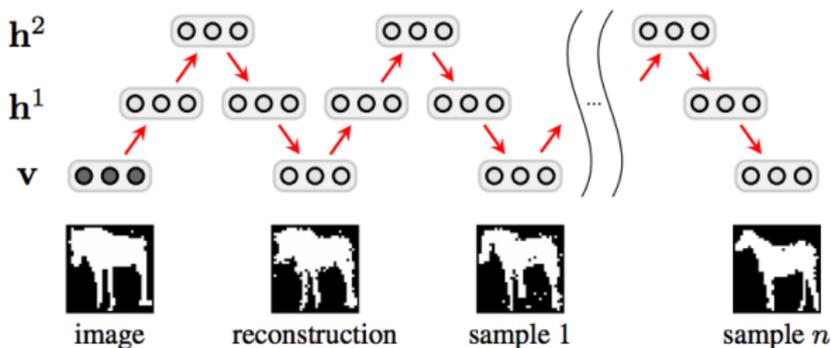
$$E(v, h^1, h^2) = -v^T W^1 h^1 - (h^1)^T W^2 h^2 - b_v^T v - b_{h^1}^T h^1 - b_{h^2}^T h^2$$

v — наблюдаемые переменные (разметка), h^1, h^2 — слои скрытых переменных.



ShapeBM

ShapeBM является генеративной моделью формы, она позволяющей восстановить форму объекта из неполного или поврежденного бинарного изображения.



ShapeBM также дает оценку вероятности того, что на рисунке изображена лошадь.

ShapeVM — плюсы и минусы

- + Задает распределение вероятностей на форму объекта, которое настраивается по обучающей выборке.
- + Генерирует новые объекты, отличные от объектов обучающей выборки.
- Не применима для изображений большого разрешения. Для изображения 32×32 уже необходимо примерно 3000 вершин в графе.
- Не учитывает локальные ограничения: цвета пикселей и градиенты.

Оглавление

- 1 Постановка задачи
- 2 Основы
 - MRF
 - ShapeBM
- 3 MRF + ShapeBM**
 - Строение
 - Вывод: EM
 - Эксперименты
- 4 Применение выхода детектора
- 5 Future work

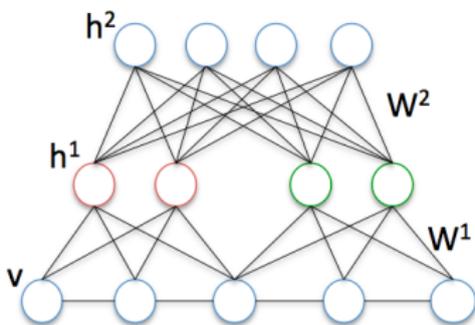
MRF + ShapeBM

Основная цель: построение модели сегментации, которая будет учитывать как локальные ограничения, так и глобальное ограничение на форму объекта.

Энергия:

$$E(v, h^1, h^2) = (1 - \gamma)E_{MRF}(v) + \gamma E_{ShapeBM}(v, h^1, h^2) \quad (2)$$

Модель объединение MRF и ShapeBM:



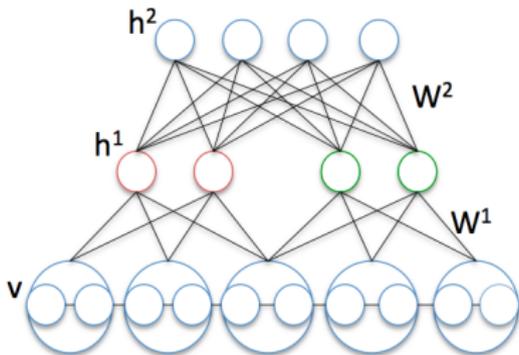
Разные разрешения

Мы используем ShapeBM для задания примерной формы, поэтому нам достаточно ShapeBM с низким разрешением. Энергия:

$$E(v, h^1, h^2) = (1 - \gamma)E_{MRF}(v) + \gamma E_{ShapeBM}(v^s, h^1, h^2) \quad (3)$$

где $v_i^s = \frac{\sum_{j \in fields(i)} v_j}{|fields(i)|}$.

Модель объединения MRF высокого разрешения и ShapeBM низкого разрешения:



MRF + ShapeBM

- В ходе экспериментов использовалась ShapeBM, уже обученная на центрированных и масштабированных изображениях 32×32 из Weismann horse dataset [S. M. Ali Eslami et al. 2012].
- Унарные потенциалы $\phi_i(\theta, v_i)$ определялись с помощью модели цвета θ – смеси гауссиан в пространствах RGB и LUV. Модель цвета обучалась на подмножестве той же выборки.
- Парные потенциалы $\phi_{i,j}(v_i, v_j) = \exp(-\|v_i - v_j\|_{RGB}^2)$.

Обучение

Обучение объединенной модели явным образом распадается на:

- Стандартное обучение ShapeVM.
- Обучение начальной модели цвета θ для определения унарных потенциалов MRF.

Обучать и то, и то можно на одних и тех же данных: попиксельно размеченной выборке объектов.

Вывод: EM

EM для задачи вывода

Задача: максимизировать $p(v|\theta) = \sum_{h_1, h_2} \exp(-E(\theta, v, h^1, h^2))$.

- E-step:

Вариационное приближение $q(h^1, h^2) = \prod_i q_i(h_i^1) \cdot \prod_i q_i(h_i^2)$

при фиксированных v и θ :

$$KL(q(h^1, h^2) \| p(h^1, h^2 | v, \theta)) \rightarrow \min_{q(h^1, h^2)}.$$

Пересчет проводится по следующим формулам:

$$h_j^1 = \frac{1}{1 + \exp(-\sum_i W_{ij}^1 v_i^s - \sum_k W_{jk}^2 h_k^2)},$$

$$h_k^2 = \frac{1}{1 + \exp(-\sum_j W_{jk}^2 h_j^1)},$$

$$h_j^1 = \frac{1}{1 + \exp(-\sum_i W_{ij}^1 v_i^s - \sum_k W_{jk}^2 h_k^2)}$$

Вывод: EM

EM для задачи вывода

- M_1 -step:
Graph cuts:

$$\mathbb{E}_{q(h^1, h^2)} E(\theta, v, h^1, h^2) \rightarrow \min_v$$

- M_2 -step:
Цветовая модель настраивается на конкретное изображение:

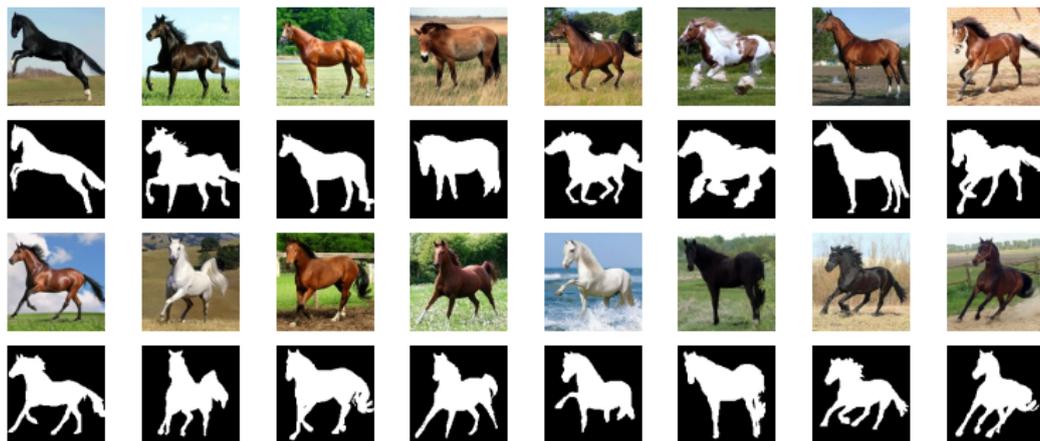
$$\mathbb{E}_{q(h^1, h^2)} E(\theta, v^*, h^1, h^2) \rightarrow \min_{\theta}$$

v^* — результирующая разметка.

Тестовая выборка

Тестовая выборка из 16 изображений размера 128×128 размечена вручную.

Функционал качества — среднее взвешенное расстояние Хэмминга на тестовой выборке.



Эксперименты

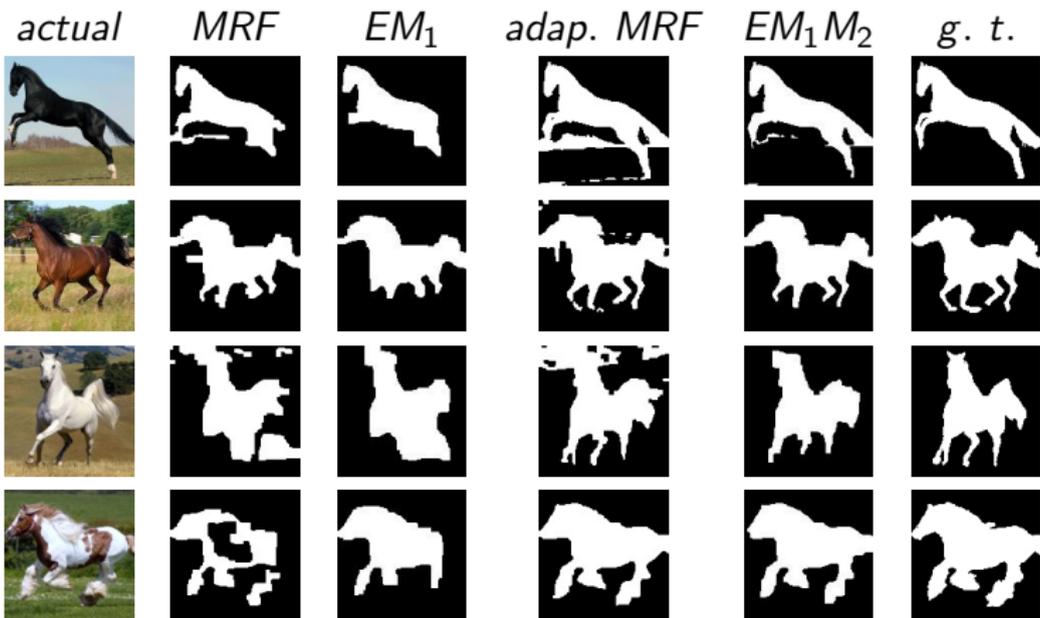
Сравним следующие схемы вывода:

- *MRF* с фиксированной моделью цвета
- *EM*: шаги E и M_1
- адаптивный *MRF*: модель цвета пересчитывается после каждой итерации на основе текущей сегментации
- *EM*: шаги E , M_1 and M_2

<i>MRF</i>	<i>EM</i> ₁	adaptive <i>MRF</i>	<i>EM</i> ₁ <i>M</i> ₂
0.3247	0.2883	0.1790	0.1585

Эксперименты

Эксперименты

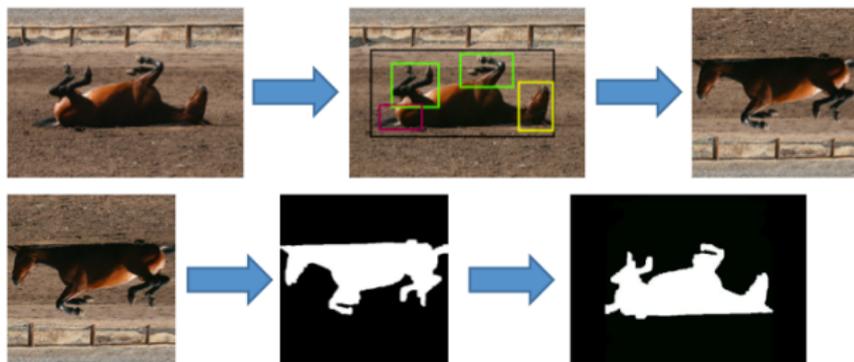


Оглавление

- 1 Постановка задачи
- 2 Основы
 - MRF
 - ShapeBM
- 3 MRF + ShapeBM
 - Строение
 - Вывод: EM
 - Эксперименты
- 4 Применение выхода детектора**
- 5 Future work

Алгоритм решения глобальной задачи

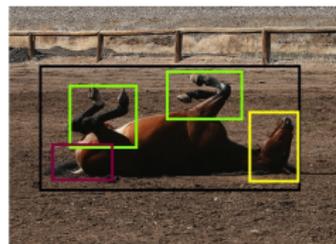
Три основных этапа: детекция объекта и частей, прикладывание шейпа, сегментация.



Part based object detection [Felzenszwalb P. et al. 2008]

Обучение: требует выборки с bounding box-ами объектов. Обучаются несколько моделей, для каждой: bounding box-ы объекта и частей. Части определяются автоматически.

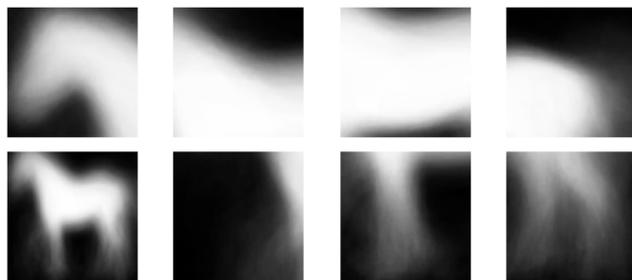
Детекция: находит bounding box-ы объекта и частей и определяет к какой модели данный объект относится.



Улучшение начальной цветовой модели

Раньше: начальная цветовая модель общая для всех лошадей.

Хотим: настраивать ее под конкретную лошадь. **Решение:** для каждой модели детектора построим средние маски для всех bounding box-ов:



Будем обучать цветовую модель для конкретного изображения на пикселях этого же изображения. Метки для пикселей будем брать их результатов детекции.

Эксперименты

Эксперименты проводились только для одной модели, полученной из детектора. Под нее подошли только 11 изображений из нашей тестовой выборки.

Color model	<i>MRF</i>	adaptive <i>MRF</i>	EM_1M_2
general	0.3118	0.1542	0.0943
from detection	0.2237	0.0899	0.0743

Оглавление

- 1 Постановка задачи
- 2 Основы
 - MRF
 - ShapeBM
- 3 MRF + ShapeBM
 - Строение
 - Вывод: EM
 - Эксперименты
- 4 Применение выхода детектора
- 5 Future work

OneCut [Yuri Boykov et al., 2013]

- Позволяет сегментировать изображение с помощью MRF без использования цветовой модели.
- Основана на переходе от идеи цветowych моделей к идее уменьшения энтропии цветов внутри класса.
- Оптимизирует функционал более простого вида сразу по v и θ .
- Работает за один разрез графа.
- Требуется дополнительное ограничение на расположение объекта на изображении для предотвращения схлопывания сегментации в один класс.

Future work

- Замена шагов $M1$ и $M2$ на OneCut.
- Изменение процедуры обучения ShapeVM: попытаться обучить ShapeVM без попиксельно размеченной выборки. Вместо попиксельной разметки использовать результаты OneCut с ограничениями на расположение объекта на основе выхода детектора.
- Изменение структуры ShapeVM: использование получаемых при детекции bounding box-ы в качестве «плавающих» паттернов в ShapeVM.