

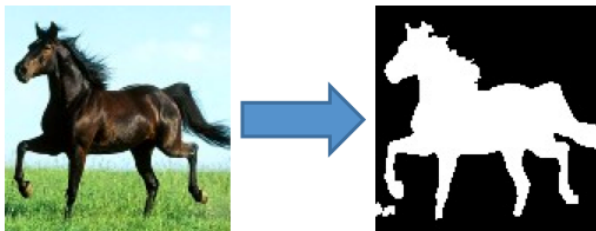
# Многоклассовая модель формы со скрытыми переменными

Кириллов Александр<sup>1</sup>, Гавриков Михаил<sup>2</sup>, Лобачева Екатерина<sup>3</sup>,  
Осокин Антон<sup>4</sup>, Ветров Дмитрий<sup>5</sup>

<sup>1</sup>ВМК МГУ, <sup>2</sup>ООО "Яндекс", <sup>3</sup>ФКН НИУ ВШЭ,  
<sup>4</sup>SIERRA team, INRIA, <sup>5</sup>Сколтех

Математические методы распознавания образов  
2015

# Сегментация изображений

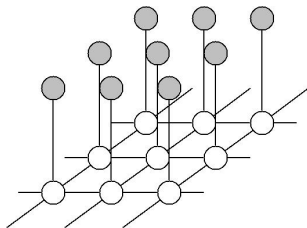


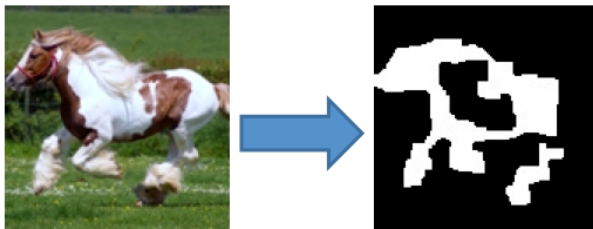
Марковское случайное поле:

$G(\mathbf{v}, \mathbf{e})$  — граф,  $\mathbf{y} \in \{0, 1\}^{|\mathbf{v}|}$  — вектор меток.

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

Только локальные ограничения!

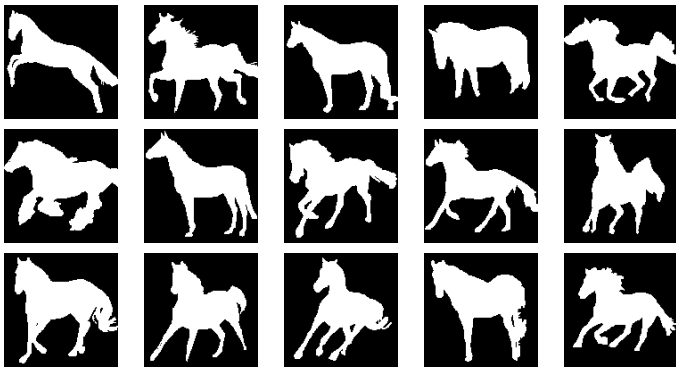




Если известен тип объекта на изображении, то можно использовать глобальные ограничения на форму объекта.

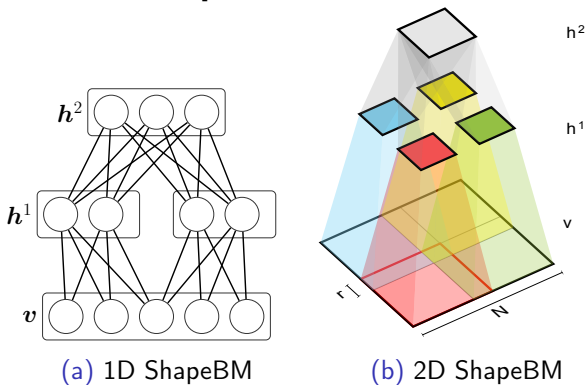
# Модель формы

**Модель формы** — вероятностное распределение над изображениями, где различные формы интересующего нас объекта имеют высокую вероятность, а любые другие изображения имеют низкую вероятность.



# Модель формы Больцмана (ShapeBM)

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



Энергия:

$$E(v, h^1, h^2) = -v^T W^1 h^1 - (h^1)^T W^2 h^2 - a^T v - (b^1)^T h^1 - (b^2)^T h^2$$

$$\Theta = \{W^1, W^2, a, b^1, b^2\}$$

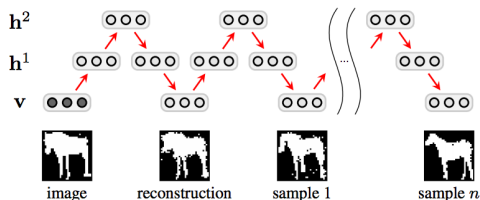
# Модель формы Больцмана (ShapeBM)

Обучение:

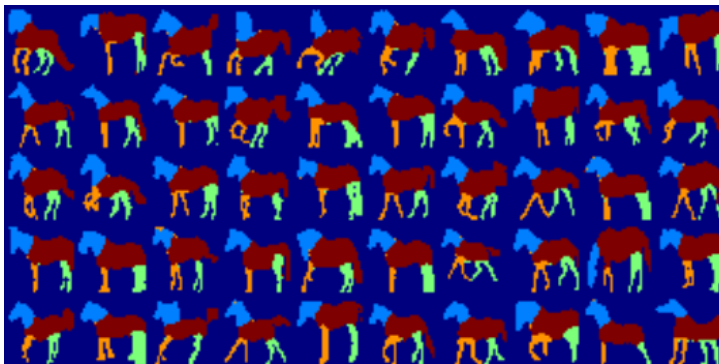
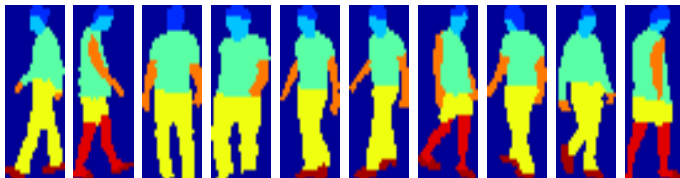
$$P(\hat{V}|\theta) \rightarrow \max_{\theta}, \quad \hat{V} \text{ — обучающая выборка.}$$

Стохастический градиентный спуск.

Генерация (Gibbs sampling):



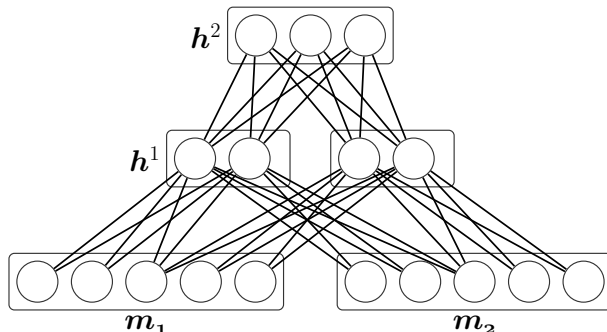
# Многоклассовая модель формы





# Многоклассовая модель формы Больцмана (MSBM)

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



$m_1, m_2$  соответствуют двум частям объекта.

- Модели формы Больцмана успешно применяют в задаче сегментации изображений
- Многоклассовая модель обладает большей выразительной способностью по сравнению с бинарной
- Многоклассовая модель требует данных с многоклассовой разметкой для обучения

**Задача:** построение алгоритма обучения многоклассовой модели формы Больцмана, не требующего полной многоклассовой разметки обучающих данных

# Многоклассовая модель с семенами



(a)  
Бинарная  
разметка —  $b$

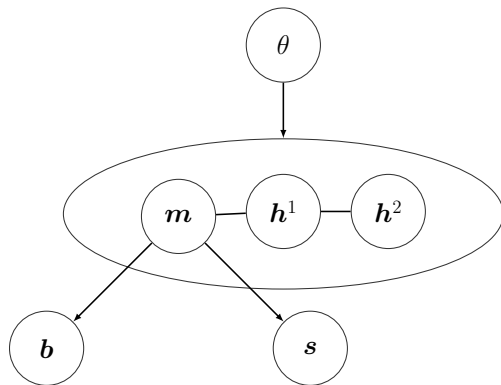


(b)  
Многоклассовая  
разметка —  $m$



(c) Семена  
частей объекта —  
 $s$

$$p(\mathbf{b}, \mathbf{s}, \mathbf{m}, \mathbf{h}^1, \mathbf{h}^2 | \theta) = p(\mathbf{b} | \mathbf{m}) p(\mathbf{s} | \mathbf{m}) p(\mathbf{m}, \mathbf{h}^1, \mathbf{h}^2 | \theta)$$



$$\log P(B, S | \theta) = \sum_{d=1}^D \log p(\mathbf{b}^d, \mathbf{s}^d | \theta) \rightarrow \max_{\theta}$$

Обучение основано на вариационном EM-алгоритме:

**E-шаг:**

$$q(\mathbf{m}, \mathbf{h}^1, \mathbf{h}^2) = \arg \min_q \text{KL} (q(\mathbf{m}, \mathbf{h}^1, \mathbf{h}^2) \| p(\mathbf{m}, \mathbf{h}^1, \mathbf{h}^2 | \mathbf{b}, \mathbf{s}, \theta_{old}))$$

$q$  — семейство полностью факторизованных распределений.

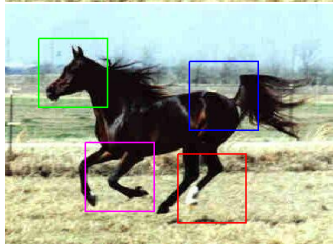
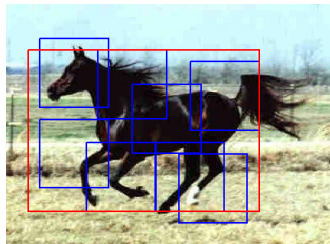
**M-шаг:**

$$\max_{\theta} \sum_{d=1}^D \left[ \mathbb{E}_{\mathbf{m}, \mathbf{h}^1, \mathbf{h}^2 \sim q(\mathbf{m}, \mathbf{h}^1, \mathbf{h}^2)} \log p(\mathbf{b}^d, \mathbf{s}^d, \mathbf{m}^d, \mathbf{h}^1, \mathbf{h}^2 | \theta) \right]$$

Задача аналогична стандартной задаче обучения MSBM.

# Автоматическая разметка семян

Part-Based Object Detector [P. F. Felzenszwalb et al. 2010]



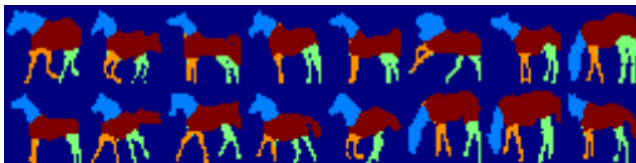
## Weizmann horse dataset:

327 изображений лошадей и с их бинарными масками.

Сделана многоклассовая разметка.

Обрезаны и масштабированы в  $32 \times 32$ .

Обучение — 277 изображений, тест — 50 изображений.

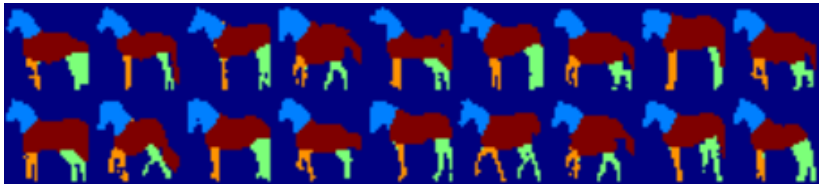


- Бинарная модель формы Больцмана.
- Многоклассовая модель формы Больцмана, обученная по данным с полной разметкой.
- Многоклассовая модель формы Больцмана, обученная с помощью предложенного алгоритма. Для нахождения семян использовался детектор частей.

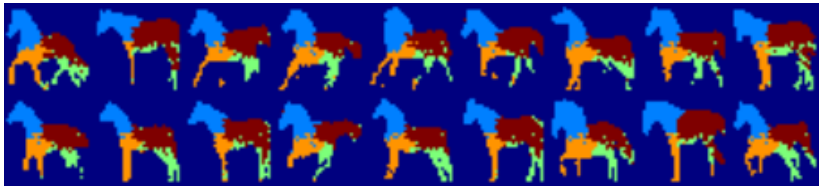


# Генерация из модели

Из MSBM, обученной по данным с полной разметкой:

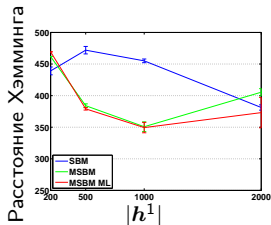


Из MSBM, обученной нашим алгоритмом:

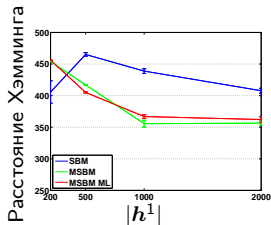


# Порождение формы из семян

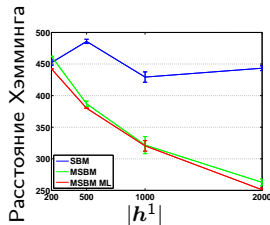
- Возьмем истинные семена для тестового объекта и сгенерируем бинарную форму.
- Подсчитаем расстояние Хэмминга между получаемой формой и исходной для тестового объекта.



(a)  $|h^2| = 50$



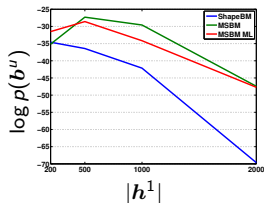
(b)  $|h^2| = 100$



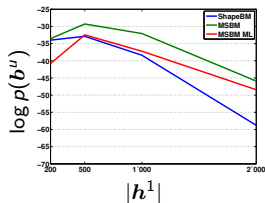
(c)  $|h^2| = 200$

# Восстановление формы

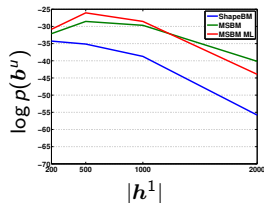
- Делим бинарную маску тестового изображения на 9 частей (сеткой 3 на 3)
- Каждый раз считаем одну из частей скрытой и приближенно вычисляем вероятность пронаблюдать эту скрытую часть.



(a)  $|h^2| = 50$



(b)  $|h^2| = 100$



(c)  $|h^2| = 200$

- Построена вероятностная модель, объединяющая бинарную разметку, многоклассовую разметку и семена частей.
- На основе этой модели предложен алгоритм обучения многоклассовой модели Больцмана по бинарным данным и семенам частей объектов, которые можно получать автоматически с помощью детектора частей.
- Результаты экспериментов показывают, что многоклассовая модель, обученная предложенным алгоритмом, превосходит по качеству бинарную модель и сопоставима с многоклассовой моделью, обученной по данным с полной разметкой.