

The Shape Boltzmann Machine

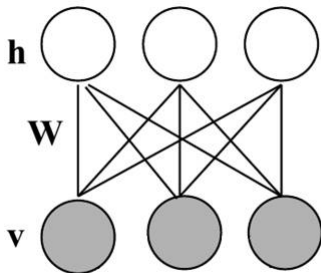
Лобачева Екатерина

MSU

2013

Restricted Boltzmann Machine

$V \in \{0, 1\}^n$ — наблюдаемые переменные, $H \in \{0, 1\}^d$ — скрытые переменные. RBM представляет из себя следующую графическую модель:



Энергия

$$\mathbf{E}(v, h|\Theta) = -v^T W h - v^T b^v - h^T b^h,$$

где $\Theta = \{b^v, b^h, W\}$

Вероятность

$$\mathbf{P}(v, h|\Theta) = \frac{1}{Z(\Theta)} \exp(-\mathbf{E}(v, h|\Theta))$$

Структура RBM



$$P(v_i = 1|h, \Theta) = \frac{1}{1 + \exp(-\sum_j W_{ij}h_j - b_i^v)}$$

$$P(h_j = 1|v, \Theta) = \frac{1}{1 + \exp(-\sum_i W_{ij}v_i - b_j^h)}$$



Схема Гиббса

Суть

Настройка Θ по обучающей выборке \hat{V} путем максимизации неполного правдоподобия $\mathbf{P}(\hat{V}|\Theta) \rightarrow \max_{\Theta}$.

Процедура

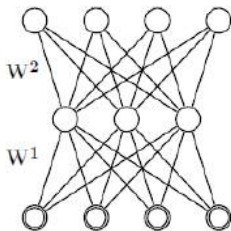
Стохастический градиентный спуск.

Градиент с помощью Contrastive Divergences:

$$\frac{\partial \log \mathbf{P}(v|\Theta)}{\partial \Theta_k} \approx \frac{\partial \mathbf{E}(v|\Theta)}{\partial \Theta_k} - \frac{\partial \mathbf{E}(v^*|\Theta)}{\partial \Theta_k}$$

Deep Boltzmann Machine

$V \in \{0, 1\}^n$ — наблюдаемые переменные,
 $H_1 \in \{0, 1\}_1^d, H_2 \in \{0, 1\}_2^d$ — скрытые переменные. DBM
представляет из себя следующую графическую модель:



Энергия

$$\mathbf{E}(v, h^1, h^2 | \Theta) = -v^T W^1 h^1 - h^1 T W^2 h^2 - v^T b^v - h^1 T b^{h^1} - h^2 T b^{h^2},$$

где $\Theta = \{b^v, b^{h^1}, b^{h^2}, W^1, W^2\}$

Вероятность

$$\mathbf{P}(v, h^1, h^2 | \Theta) = \frac{1}{Z(\Theta)} \exp(-\mathbf{E}(v, h^1, h^2 | \Theta))$$

Структура DBM



Аналитически выражаются $\mathbf{P}(v_i = 1|h^1, \Theta)$,
 $\mathbf{P}(h_j^1 = 1|v, h^2, \Theta)$, $\mathbf{P}(h_k^2 = 1|h^1, \Theta)$



Схема Гиббса

Суть

Настройка Θ по обучающей выборке V путем максимизации неполного правдоподобия $\mathbf{P}(\hat{V}|\Theta) \rightarrow \max_{\Theta}$.

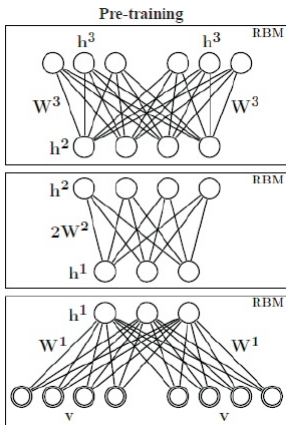
Процедура

Предобучение.

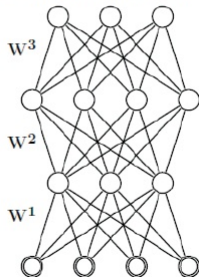
Стохастический градиентный спуск.

Градиент с помощью Contrastive Divergences:

$$\frac{\partial \log \mathbf{P}(v|\Theta)}{\partial \Theta_k} \approx \frac{\partial \mathbf{E}(v|\Theta)}{\partial \Theta_k} - \frac{\partial \mathbf{E}(v^*|\Theta)}{\partial \Theta_k}$$



Deep Boltzmann Machine



Strong probabilistic model of binary object shapes:

- 1 Realism
- 2 Generalization

	Realism		Generalization
	Globally	Locally	
Mean	✓	-	-
Factor Analysis	✓	-	✓
Fragments	-	✓	✓
Grid MRFs/CRFs	-	✓	✓
High-order potentials	limited	✓	✓
Database	✓	✓	-
ShapeBM	✓	✓	✓

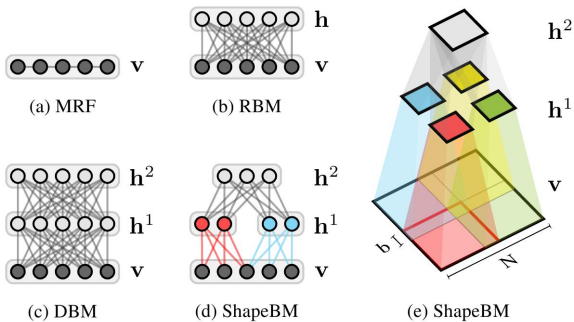
Realism + Generalization



DBM: много параметров,
а объем обучающей
выборки ограничен



ShapeBM



Веса W^1 общие для всех групп пикселей.

Смысл структуры

- Отбрасывание связей: наиболее сильные зависимости имеют локальный характер.
- Общие веса W^1 : многие свойства формы (н-р, гладкость) не зависят от места на рисунке.
- H^1 : локальные ограничения.
- H^2 : глобальные ограничения.
- $d_2 \ll d_1$.

Обучение

Аналогично обучению DBM, с учетом структуры ShapeBM при подсчете градиентов.

Параметры ShapeBM

Вход: картинки 32×32 .

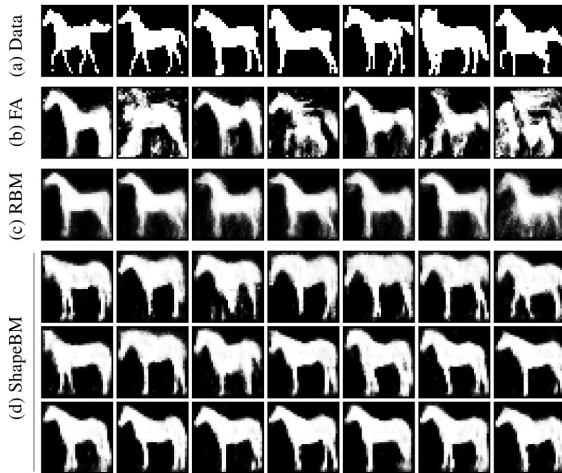
Overlap $b = 4$, $d_1 = 2000$, $d_2 = 100$.

Обучение ShapeBM

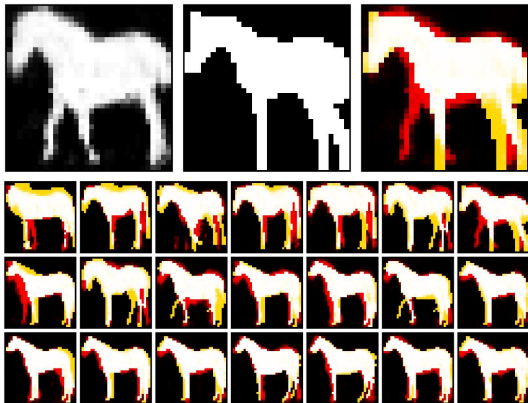
Предобучение: 1 уровень - 3000 итераций, 2 - 1000 итераций.

Общее обучение: 1000 итераций.

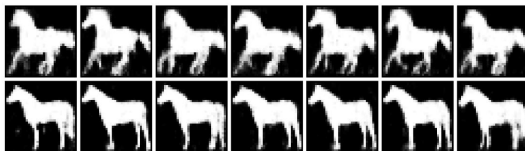
Weizmann horses: Realism



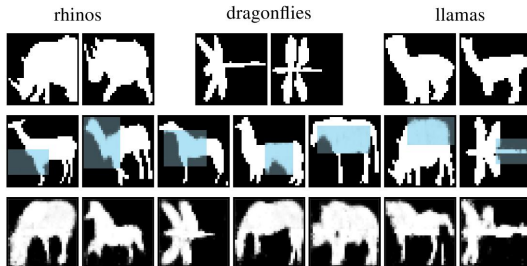
Weizmann horses: Generalization



Weizmann horses: суть H^2



Multiple object categories



При обучении не давалась информация о классах!

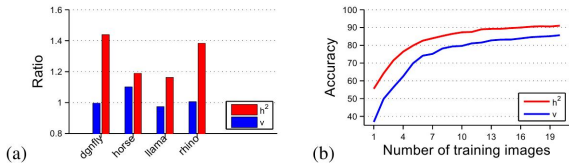


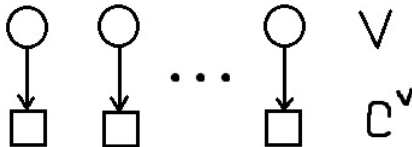
Figure 13. (a) Inter- and intra-class distance ratios (values > 1 indicate that inter-class distances are larger). (b) GLM classification.

Segmentation with ShapeBM: same size

Вводим дополнительные унарные веса на видимые переменные c^v . Вводим коэффициент α , отвечающий за соотношение цвета и формы. Сложение цвета и формы на уровне v .

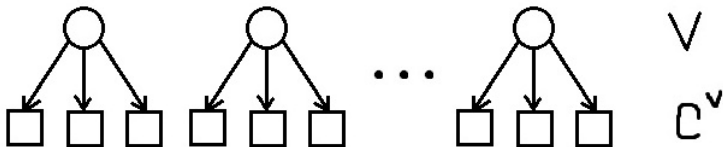
Энергия

$$\mathbf{E}(v, h^1, h^2 | \Theta) = -v^T W^1 h^1 - h^{1T} W^2 h^2 - v^T b^v - v^T c^v - h^{1T} b^{h^1} - h^{2T} b^{h^2}$$



Segmentation with ShapeBM: larger size

Сложение цвета и формы на уровне c^v , то есть не модель цвета уменьшаем, а модель формы увеличиваем.



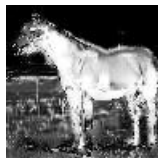
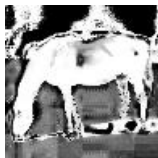
Segmentation with ShapeBM: Color model

Обучаются 4 модели цвета: по две для объектов и для фона. Они представляют из себя смесь гауссиан в пространстве RGB и LUV.

Композиция моделей для RGB, LUV \rightarrow для изображения получаем: p_obj, p_not_obj .

$$c^v = \log(p_obj) - \log(p_not_obj)$$

Segmentation with ShapeBM: Results (train)



Segmentation with ShapeBM: Results (test)

