

Эпитомы и их применение

Д.А. Елшин
МГУ, ВМК, каф. ММП, 2010.

Доклад на спецсеминаре
«Байесовские методы машинного обучения»

Содержание

Содержание:

Что такое эпитом?

Построение эпитомов.

Построение эпитомов — усложнение модели.

Задача разделения слоёв.

Проблемы.

Применение.

Содержание

Содержание:

Что такое эпитом?

Построение эпитомов.

Построение эпитомов — усложнение модели.

Задача разделения слоёв.

Проблемы.

Применение.

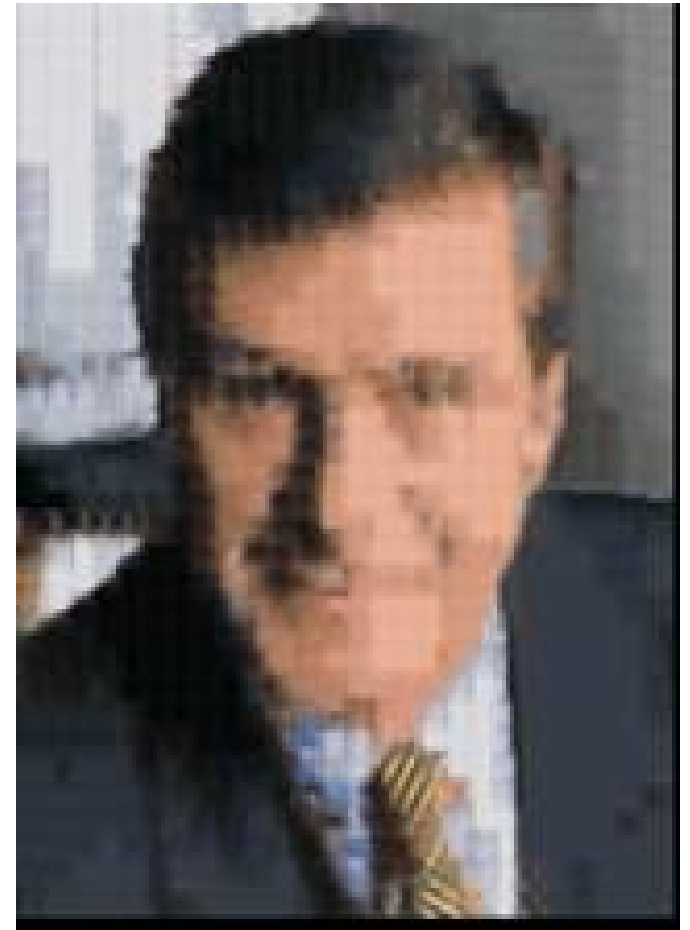
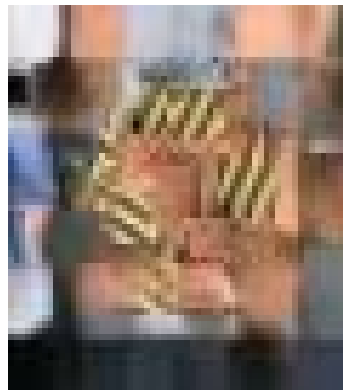
Что это такое?

Эпитом изображения — это его уменьшенная версия, отражающая особенности текстуры и формы объектов, присутствующих на изображении.

Что это такое?

Основная идея: сжать изображение «по-умному», выделив на нём некоторые «похожие» участки, которые в эпитоме будут представлены один раз.

Пример



Пример



Содержание

Содержание:

Что такое эпитом?

Построение эпитомов.

Построение эпитомов — усложнение модели.

Задача разделения слоёв.

Проблемы.

Применение.

Построение эпитома

Пусть размер изображения $N \times M$, эпитома — $N_e \times M_e$.

Исходное изображение разбивается на P блоков (патчей) Z_k . Патчи бывают нескольких видов, могут пересекаться по пикселям изображения.

Преобразованием T_k назовём сопоставление блока изображения блоку (или набору блоков) такого же размера в эпитоме.

Построение эпитома

Вообще говоря, блоки эпитома, формирующие патч изображения, не обязаны быть такого же размера. Для патчей изображения 24×24 можно использовать как цельный 24×24 -блок, так и 4 12×12 -блоков или 9 8×8 -блоков эпитома.

Построение эпитома

Для каждого пикселя эпитома будем хранить матожидание цвета μ и дисперсию φ .

Преобразование T_k переводит пиксели эпитома в S_k — пиксели изображения.

При построенном эпитоме и выбранных преобразованиях патч может быть сгенерирован гауссовским распределением с параметрами μ и φ пикселей эпитома.

Построение эпитома

Тогда имеет место формула:

$$p(Z_k | T_k, e) = \prod_{i \in Sk} N(z_{i,k}; \mu_{T_k(i)}, \varphi_{T_k(i)})$$

Построение эпитома

Для простоты предполагаем, что патчи генерируются из эпитома независимо, а при пересечении блоков соответствующие пиксели совпадают. Тогда:

$$p(\{Z_k, T_k\}_{k=1}^P, e) = p(e) \prod_{k=1}^P p(T_k) \prod_{i \in S_k} N(z_{i,k}; \mu_{T_k(i)}, \varphi_{T_k(i)}).$$

Построение эпитома

Предполагается, что $p(e)$ и $p(T_k)$ (априорные распределения) — константы. Хотя для преобразований можно ввести некоторый приоритет.

Построение эпитома

Теперь можно применить EM, улучшая оценку нижней границы функции правдоподобия:

$$\log p(\{Z_k\}_{k=1}^P) \geq B$$
$$B = \sum_{\{T_k\}_{k=1}^P} \int_e q(\{T_k\}_{k=1}^P, e) \log \frac{p(\{Z_k, T_k\}_{k=1}^P, e)}{q(\{T_k\}_{k=1}^P, e)}$$

Построение эпитома

На $q(T_k)$ используем точечную оценку в предположении о том, что патчи генерируются независимо:

$$q(\{T_k\}_{k=1}^P, e) = \delta(e - \hat{e}) \prod_k q(T_k)$$

Построение эпитома

$$B = \sum_{k=1}^P q(T_k) [\log p(T_k) - \log q(T_k)] \\ + \sum_{k=1}^P q(T_k) \sum_{i \in S_k} \log N(z_i, k; \hat{\mu}_{T_k(i)}, \hat{\varphi}_{T_k(i)})$$

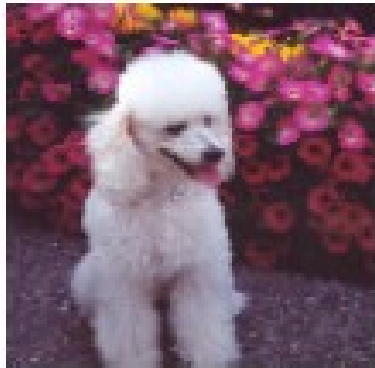
Построение эпитома

$$E\text{-шаг: } q(T_k) \sim p(T_k) \prod_{i \in S_k} N(z_{i,k}; \hat{\mu}_{T_k(i)}, \hat{\varphi}_{T_k(i)})$$

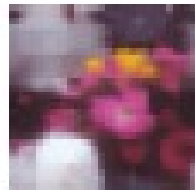
M – шаг:

$$\hat{\mu}_j = \frac{\sum_k \sum_{i \in S_k} \sum_{T_k, T_k(i)=j} q(T_k) z_{i,k}}{\sum_k \sum_{i \in S_k} \sum_{T_k, T_k(i)=j} q(T_k)}$$
$$\hat{\phi}_j = \frac{\sum_k \sum_{i \in S_k} \sum_{T_k, T_k(i)=j} q(T_k) (z_{i,k} - \mu_j)^2}{\sum_k \sum_{i \in S_k} \sum_{T_k, T_k(i)=j} q(T_k)}$$

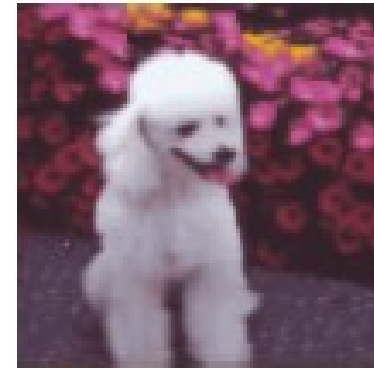
Построение изображения из эпитома



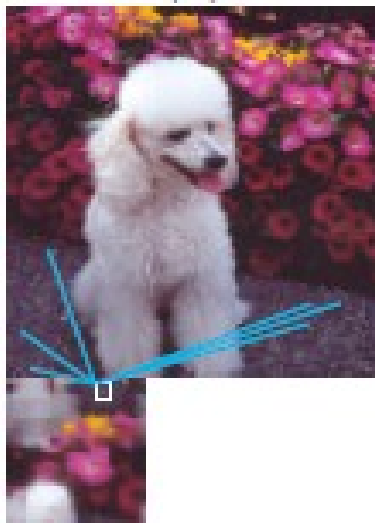
(a)



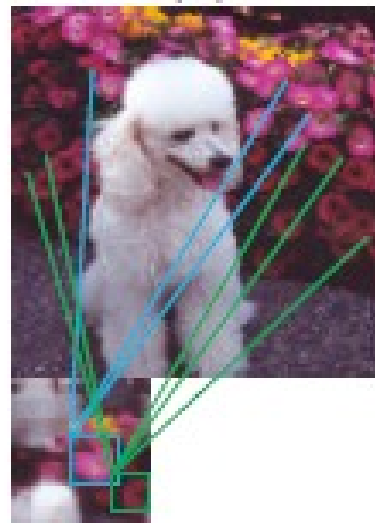
(b)



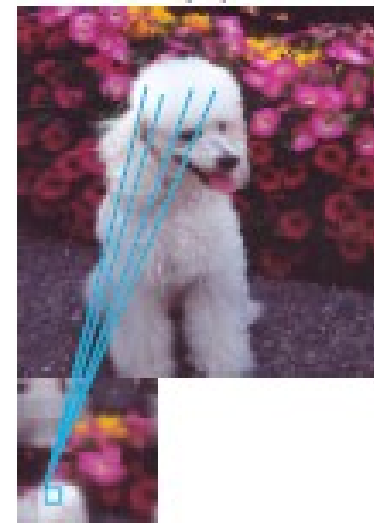
(c)



(d)



(e)



(f)

Содержание

Содержание:

Что такое эпитом?

Построение эпитомов.

Построение эпитомов — усложнение модели.

Задача разделения слоёв.

Проблемы.

Применение.

Построение эпитомов — усложнение модели

На самом деле патчи не генерируются независимо. Мы хотим получить порождающую модель для целого изображения x , а не для патчей по отдельности.

Построение эпитомов — усложнение модели

Для этого в нашей модели будем генерировать пиксель изображения следующим образом:

$$p(x_i | \{Z_k\}_{k=1}^P) = N(x_i; \frac{1}{N_{i k, i \in S_k}} \sum z_{i,k}, \psi_i),$$

где N_i — число патчей, пересекающихся
на i — ом пикселе

Тогда:

$$\begin{aligned} & p(x, \{Z_k, T_k\}_{k=1}^P, e) \\ &= p(\{Z_k, T_k\}_{k=1}^P, e) \prod_i p(x_i | \{Z_k\}_{k=1}^P) \end{aligned}$$

Построение эпитомов — усложнение модели

Но нас будет интересовать такая генерация,
при которой накладываемые патчи
совпадают в общих пикселях:

$$q(\{z_{i,k}\}, \{T_k\}, e) = \delta(e - \hat{e}) \prod_k q(T_k) \prod_{i \in S_k} \delta(z_{i,k} - \zeta_i)$$

ζ_i — цвет пикселей патчей, имеющих
общий i — ый пиксель

Построение эпитома — усложнение модели

$$\log p(\mathbf{x}) \geq B =$$

$$\int_{\{z_{i,k}\}} \sum_{\{\mathcal{T}_k\}} \int_{\mathbf{e}} q(\{z_{i,k}\}, \{\mathcal{T}_k\}, \mathbf{e}) \log \frac{p(\mathbf{x}, \{\mathbf{Z}_k, \mathcal{T}_k\}, \mathbf{e})}{q(\{z_{i,k}\}, \{\mathcal{T}_k\}_{k=1}^P, \mathbf{e})}$$

$$B = \sum_{k=1}^P q(\mathcal{T}_k) [\log p(\mathcal{T}_k) - \log q(\mathcal{T}_k)] +$$

$$+ \sum_{k=1}^P q(\mathcal{T}_k) \sum_{i \in S_k} \log \mathcal{N}(\zeta_i; \hat{\mu}_{\mathcal{T}_k(i)}, \hat{\phi}_{\mathcal{T}_k(i)})$$

$$+ \sum_i \log \mathcal{N}(x_i; \zeta_i, \psi_i),$$

Формулы пересчёта

Используем в итерационном алгоритме следующие формулы пересчёта :

$$\zeta_i = \frac{x_i/\psi_i + \sum_{k,i \in S_k} \sum_{T_k} q(T_k) \mu_{T_k} / \phi_{T_k}}{1/\psi_i + \sum_{k,i \in S_k} \sum_{T_k} q(T_k) / \phi_{T_k}}$$

$$q(T_k) \sim p(T_k) \prod_{i \in S_k} \mathcal{N}(\zeta_i; \hat{\mu}_{T_k(i)}, \hat{\phi}_{T_k(i)}),$$

$$\hat{\mu}_j = \frac{\sum_k \sum_{i \in S_k} \sum_{T_k, T_k(i)=j} q(T_k) \zeta_i}{\sum_k \sum_{i \in S_k} \sum_{T_k, T_k(i)=j} q(T_k)}$$

$$\hat{\phi}_j = \frac{\sum_k \sum_{i \in S_k} \sum_{T_k, T_k(i)=j} q(T_k) (\zeta_i - \mu_j)^2}{\sum_k \sum_{i \in S_k} \sum_{T_k, T_k(i)=j} q(T_k)}$$

Пример



(a)



(b)

Содержание

Содержание:

Что такое эпитом?

Построение эпитомов.

Построение эпитомов — усложнение модели.

Задача разделения слоёв.

Проблемы.

Применение.

Задача разделения слоёв

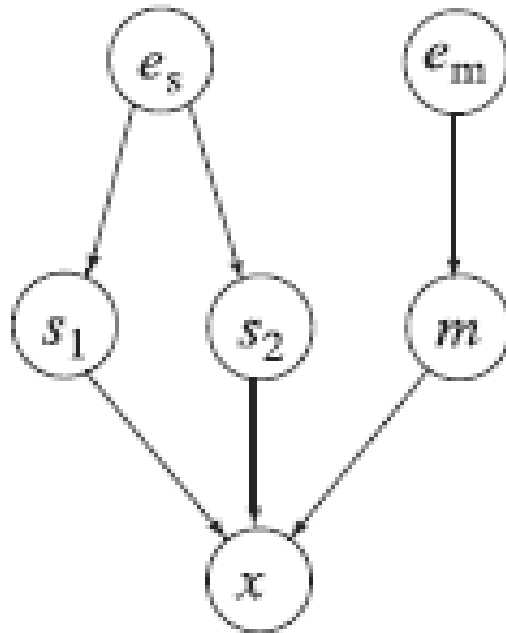
Рассмотрим задачу порождения изображения с несколькими слоями. Эпитомы e_s порождают слои s_1, s_2 , а эпитом формы e_m (маска) отвечает за порождение суммы слоёв.

Разделение слоёв

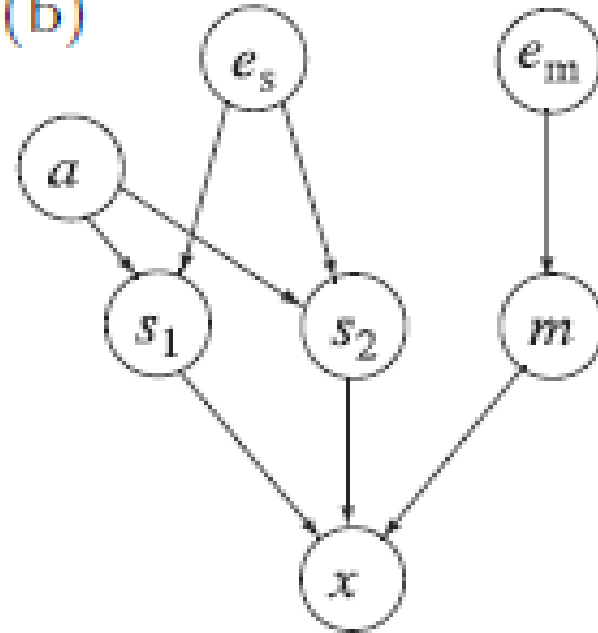
Значения маски лежат в интервале $[0, 1]$ — фактически, это коэффициент сложения значений первого и второго изображений.

Разделение слоёв

(a)



(b)



$$x = m \circ s_1 + (1 - m) \circ s_2 + \text{noise}.$$

Разделение слоёв

Тогда совместное распределение представимо
в виде:

$$\begin{aligned} & p(x, s_1, s_2, m, e_s, e_m) = \\ & = p(x|s_1, s_2, m) p(s_1|e_s) p(s_2|e_s) p(m|e_m) p(e_s) p(e_m) \end{aligned}$$

$$p(x|s_1, s_2, m) = N(x; m \circ s_1 + (1 - m) \circ s_2, \Psi_x)$$

Разделение слоёв — обоснование EM

$$q = \delta(s_1 - \hat{s}_1) \delta(s_2 - \hat{s}_2) \delta(m - \hat{m}) \delta(e_m - \hat{e}_m) \delta(e_s - \hat{e}_s)$$

$$\log p(\mathbf{x}) > B =$$

$$\begin{aligned} &= \int_{\mathbf{s}_1, \mathbf{s}_2, \mathbf{m}} q \log \frac{p(\mathbf{x}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{m}, \mathbf{e}_s, \mathbf{e}_m)}{q(\mathbf{s}_1)q(\mathbf{s}_2)q(\mathbf{m})} \\ &= \log p(\mathbf{x} | \hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \hat{\mathbf{m}}) + \log p(\hat{\mathbf{s}}_1 | \hat{\mathbf{e}}_s) + \log p(\hat{\mathbf{s}}_2 | \hat{\mathbf{e}}_s) \\ &\quad + \log p(\hat{\mathbf{m}} | \hat{\mathbf{e}}_m) + \text{const}, \end{aligned} \tag{17}$$

Разделение слоёв — обоснование EM

$$\log p(\hat{x} | \hat{s}_1, \hat{s}_2, \hat{m}) = B_x = \sum \log N(x_i; \zeta_i^m \zeta_i^{s_1} + (1 - \zeta_i^m) \zeta_i^{s_2}, \psi_i)$$

$$\log p(\mathbf{s}_\ell | \mathbf{e}_s) \geq B_\ell = \sum_{k=1}^P q(\mathcal{T}_k^{s_\ell}) [\log p(\mathcal{T}_k^{s_\ell}) - \log q(\mathcal{T}_k^{s_\ell})] + \sum_{k=1}^P q(\mathcal{T}_k^{s_\ell}) \sum_{i \in S_k} \log \mathcal{N}(\zeta_i^{s_\ell}; \hat{\mu}_{\mathcal{T}_k^{s_\ell}(i)}^s, \hat{\phi}_{\mathcal{T}_k^{s_\ell}(i)}^s)$$

$$\log p(\mathbf{m} | \mathbf{e}_m) \geq B_m = \sum_{k=1}^P q(\mathcal{T}_k^m) [\log p(\mathcal{T}_k^m) - \log q(\mathcal{T}_k^m)] + \sum_{k=1}^P q(\mathcal{T}_k^m) \sum_{i \in S_k} \log \mathcal{N}(\zeta_i^m; \hat{\mu}_{\mathcal{T}_k^m(i)}^m, \hat{\phi}_{\mathcal{T}_k^m(i)}^m)$$

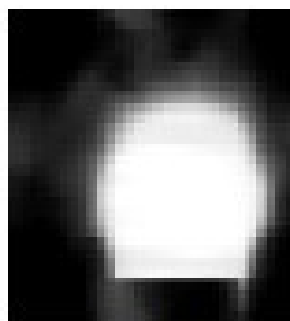
Разделение слоёв — обоснование EM

$$B = B_x + B_1 + B_2 + B_m$$

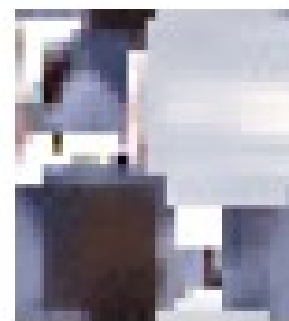
Хорошее разделение



(a)



(b)



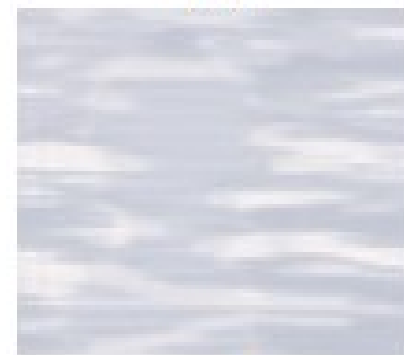
(c)



(d)

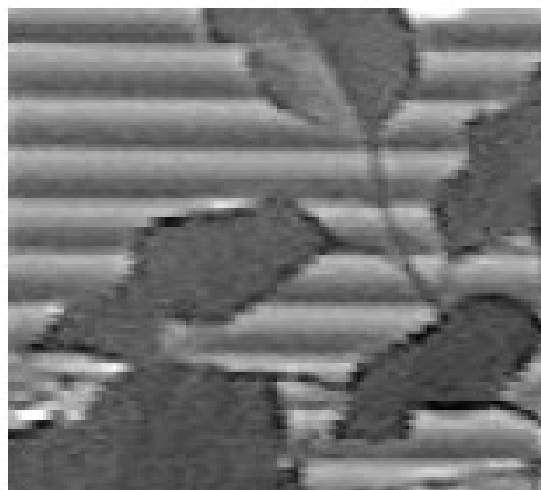


(e)

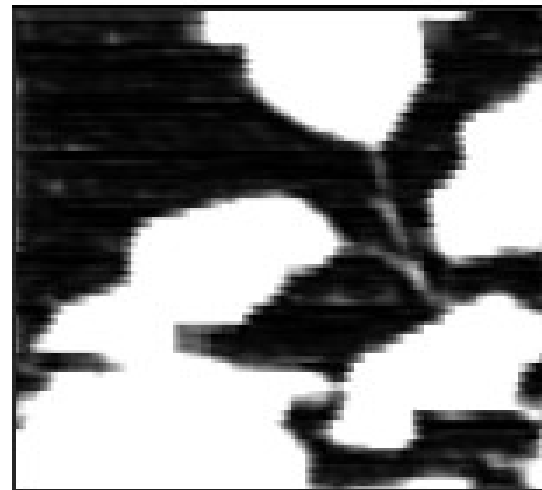


(f)

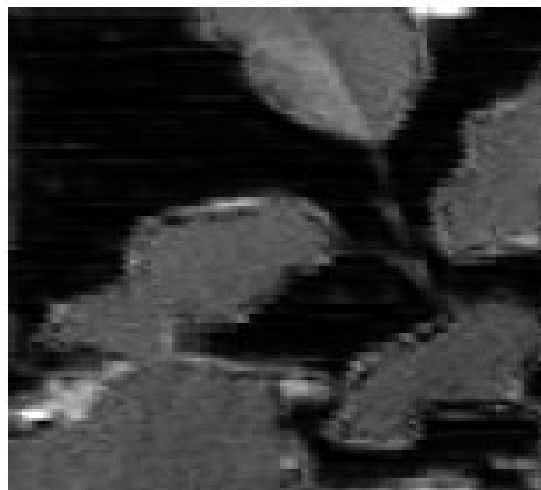
Хорошее разделение



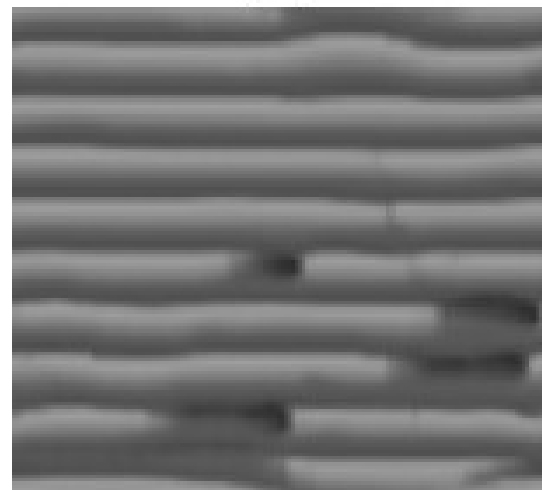
(a)



(b)



(d)

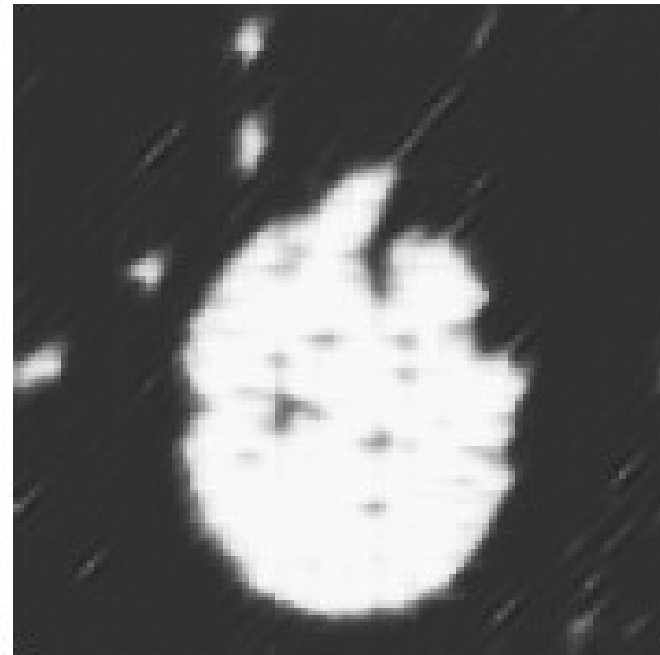


(e)

Плохое разделение



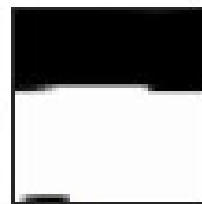
(a)



(b)



(c)



(b)



(d)

Содержание

Содержание:

Что такое эпитом?

Построение эпитомов.

Построение эпитомов — усложнение модели.

Задача разделения слоёв.

Проблемы.

Применение.

Проблемы

1. Выбор оптимального размера блоков.
2. Комбинаторный взрыв для блоков разного вида.
3. Проблемы «похожести» разнородных структур (решается гораздо эффективнее, чем в других подходах).
4. Корректное разделение слоёв.

Содержание

Содержание:

Что такое эпитом?

Построение эпитомов.

Построение эпитомов — усложнение модели.

Задача разделения слоёв.

Применение.

Применение

1. Сжатие нескольких последовательных кадров, в том числе анализ видеоряда.
2. Задача восстановления изображения.
3. Выделение слоёв на изображении.
4. Удаление шума с изображений.
5. Сегментирование изображений.

Восстановление изображений

$$I^* = \operatorname{argmax}_I E[\log p(\{Z_k(I)\}_{k=1}^P, \{T_k\}_{k=1}^P | e(I_g))]$$

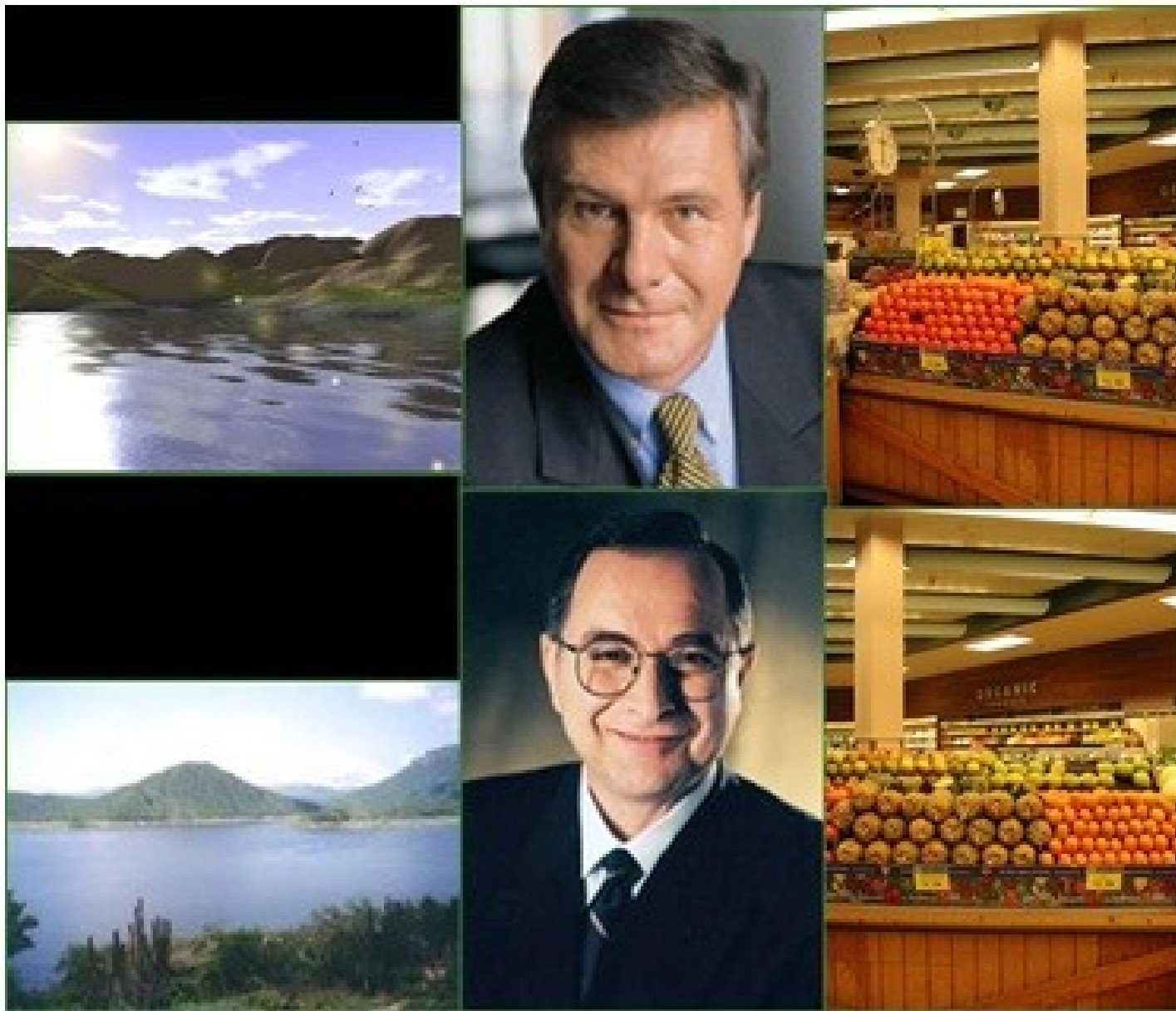
Восстановление изображений



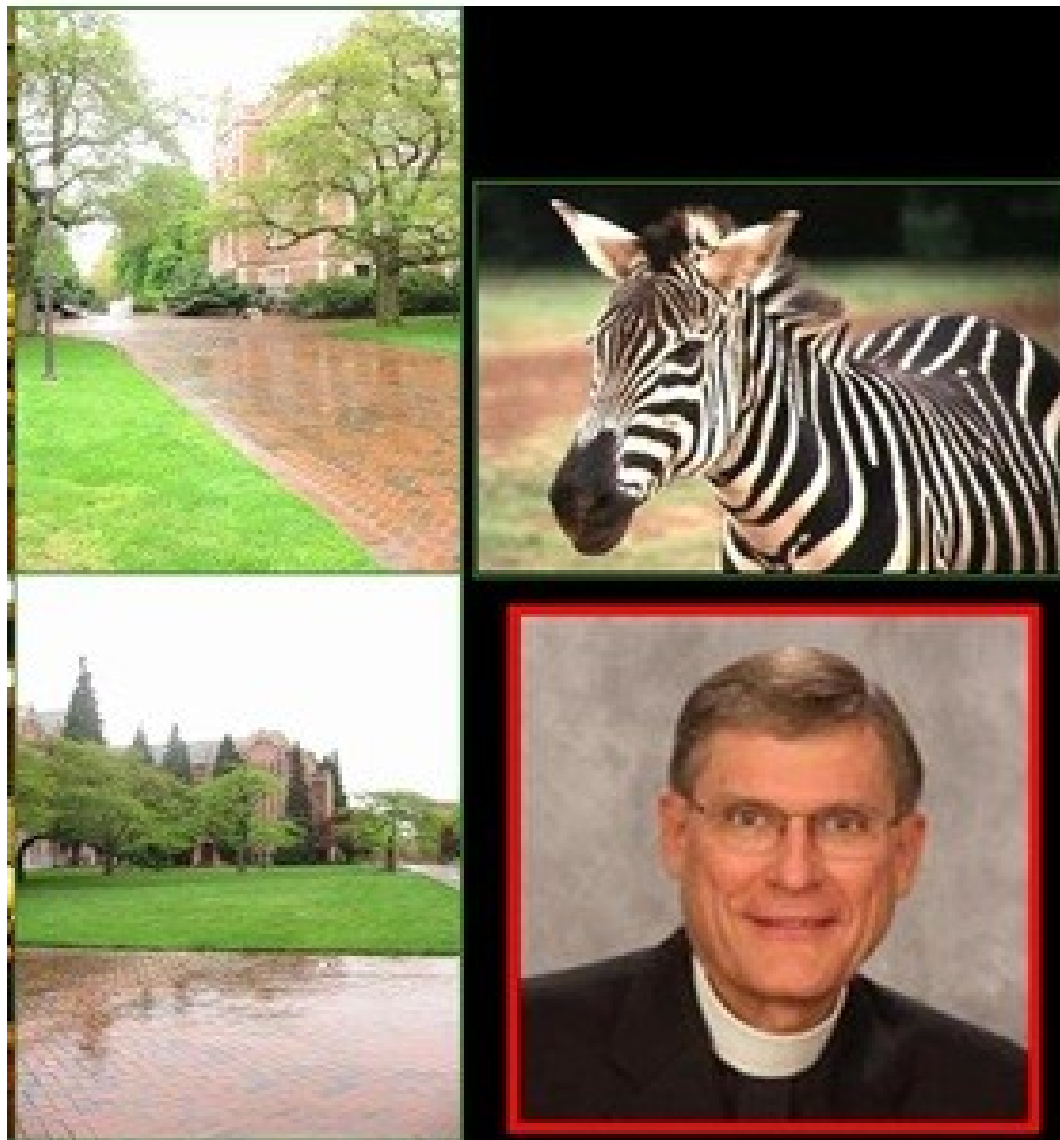
Восстановление изображений



Восстановление изображений



Восстановление изображений



Литература

Epitomic analysis of appearance and shape

Nebojsa Jojic Brendan J. Frey Anitha Kannan

Microsoft Research University of Toronto

University of Toronto

www.research.microsoft.com/~jojic

www.psi.toronto.edu www.psi.toronto.edu

<http://ssli.ee.washington.edu/~lixiao/epitome/paper.html>