

Д.А. ЕЛШИН¹, Д.А. КРОПОТОВ², Д.П. ВЕТРОВ¹

¹ Московский государственный университет им. М.В. Ломоносова

² Вычислительный центр РАН, Москва

elshin.den@yandex.ru, dmitry.kropotov@gmail.com, vetrovd@yandex.ru

ВЕРОЯТНОСТНАЯ МОДЕЛЬ ЭПИТОМОВ ДЛЯ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ И ЕЕ ПРИМЕНЕНИЕ ДЛЯ АНАЛИЗА ГИСТОЛОГИЧЕСКИХ ИЗОБРАЖЕНИЙ СРЕЗОВ МОЗГА МЫШИ*

Рассматривается задача автоматической сегментации изображений на основе подхода минимизации энергии марковского случайного поля. В работе вводится текстурный потенциал, основанный на модификации вероятностной модели эпитомного подхода [1]. Эффективность предложенного текстурного потенциала демонстрируется на модельной задаче сегментации текстур, а также на практической задаче сегментации гистологических изображений срезов головного мозга на анатомические структуры. Произведено сравнение результатов с алгоритмом textonboost [2].

Ключевые слова: сегментация изображений, марковское случайное поле, вероятностные модели, текстурные дескрипторы, эпитомы, экспрессия генов в мозге

Введение

Одним из перспективных направлений в современной нейробиологии для понимания когнитивных процессов, происходящих в головном мозге животных, является анализ экспрессии генов в мозге. Стандартной процедурой здесь является декапитация головного мозга животного, замораживание, нарезка на тонкие слои и последующая окраска на гистологию (общую структуру) и экспрессию определенных генов [3]. Проведение статистического анализа уровня экспрессии генов в различных анатомических структурах мозга требует, в частности, процедуру автоматической сегментации гистологических изображений срезов мозга на анатомические структуры.

* Работа выполнена при поддержке РФФИ (номера проектов 11-04-12174 и 10-01-00419).

В данной работе рассматривается подход к подобной сегментации для данных из Алленовского атласа мозга мыши [4]. Этот атлас содержит 132 коронарных среза мозга, для каждого из которых доступно его гистологическое изображение (рис. 1a) и ручная экспертная разметка на анатомические структуры (рис. 1b). Особенностью данных гистологических изображений является отсутствие видимых границ между многими парами соседних анатомических структур. Поэтому основной задачей при построении процедуры сегментации таких изображений является выделение текстурных особенностей каждой анатомической структуры.

В данной работе для текстурного описания изображения предлагается адаптировать т.н. эпитомный подход [1]. Для обучающего изображения и его разметки строится их уменьшенная версия (эпитом), содержащая текстурные особенности изображения, а также вероятностное распределение на классы каждого пикселя эпитома. Затем тестовое изображение проектируется на известный эпитом, а вероятностное распределение на классы отдельных пикселей тестового изображения вычисляется путем обратного преобразования вероятностного распределения для эпитома. Итоговая сегментация тестового изображения получается путем минимизации энергии соответствующего марковского случайного поля, в энергии которого к текстурному потенциалу, полученному на основе эпитома, добавляются стандартные компоненты, включающие потенциалы цвета, положения и Поттса [2].

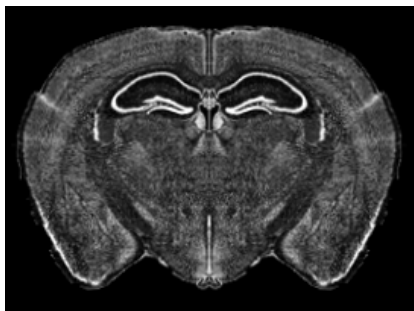


Рис. 1a. Изображение среза.



Рис. 1b. Разметка на структуры

Для решения задачи сегментации изображения на классы существует множество подходов (SIFT-дескрипторы [5], texton boost [2] и др.). В нашей работе мы проводили сравнение предложенного метода и подхода Texton boost [2].

Постановка задачи и описание алгоритма

Задача состоит в следующем: данному изображению $Z = \{Z(i)\}_{i=1}^R$ поставить в соответствие маску $Q = \{Q(i)\}_{i=1}^R$:

$$Z(i) \longrightarrow Q(i), Q(i) \in \{1 \dots K\},$$

где K – число классов. Для её решения используется модель марковского случайного поля. Изображение рассматривается как марковская сеть (V, E) , где каждому пикселю изображения ставится в соответствие вершина сети $i \in V$, а ребра $(i, j) \in E$ определяют соседние пиксели. Здесь и далее используется четырехсвязная система соседства. Переменная в вершине i принимает значения $Q(i) \in \{1 \dots K\}$. Тогда энергия разметки Q при известном изображении Z определяется следующим образом:

$$E(Q|Z) = \sum_{i \in V} \theta_i(Q(i), Z(i)) + \sum_{(i,j) \in E} \theta_{ij}(Q(i), Q(j), Z(i), Z(j)) \quad (1)$$

где $\theta_i(Q(i), Z(i))$ – унарные потенциалы, отвечающие за степень принадлежности пикселя i классу $Q(i)$, а $\theta_{ij}(Q(i), Q(j), Z(i), Z(j))$ – парные потенциалы, штрафующие ситуацию разных классов у соседних пикселей.

Оптимальная разметка получается как

$$\bar{Q} = \arg \min_Q E(Q|Z)$$

В работе [2] в качестве унарных потенциалов θ_i предлагается взять линейную комбинацию текстурных потенциалов на основе т.н. текстонов, потенциалов цвета и потенциалов положения. В качестве парных потенциалов θ_{ij} использовалась модель Поттса. При этом энергия $E(Q|Z)$ минимизировалась алгоритмом α -расширения [6]. В нашей работе используется аналогичный подход, где вместо текстурных потенциалов на основе текстонов предлагается выбрать другой текстурный потенциал на основе т.н. эпитомов [1]. Опишем подробнее каждый из потенциалов.

Парные потенциалы

Здесь используются простейшие потенциалы, штрафующие длину границы между классами:

$\theta_{ij}(Q(i), Q(j), Z(i), Z(j)) = C[Q(i) \neq Q(j)], C > 0$ – штраф за различие классов у соседних пикселей.

Потенциалы положения и потенциалы цвета

В качестве потенциалов положения учитывается доля пикселей класса l в окрестности рассматриваемого пикселя.

Цветовая модель представляет собой гауссовскую смесь для каждого класса. Восстановление смеси осуществляется стандартным EM-алгоритмом. Условное распределение на цвет пикселя задаётся как

$$P(z | q) = \sum_{k=1}^{n(q)} w_k^q N(z | \mu_k^q, \sigma_k^q), q \in \{1 \dots K\}, \forall q \in \{1 \dots K\} : \sum_{k=1}^q w_k^q = 1.$$

Значением потенциала является величина $-\log P(Z(i) | Q(i))$.

Эпитомный подход

Рассмотрим вероятностную модель эпитомов, предложенную в работе [1]. Эпитомом $e = (\mu, \varphi)$ для изображения размера $N \times M$ называется изображение меньшего размера $N_e \times M_e$, в каждом пикселе t которого помимо цветовой характеристики μ_t хранится её дисперсия φ_t . Пример изображения и его эпитома (компонента μ) показаны на рис. 2. Исходное изображение разбивается на P пересекающихся блоков $\{Z_k\}_{k=1}^P$. Обозначим через S_k множество индексов пикселей блока Z_k , $Z_k = \{z_{k,i} | i \in S_k\}$. Рассмотрим взаимно однозначное преобразование T_k блока Z_k в некоторый блок эпитома. Далее рассматриваются только квадратные блоки Z_k размера $D \times D$ и преобразования, сопоставляющие блоку Z_k блок такого же размера в эпитоме. Тогда при известном эпитоме $e = (\mu, \varphi)$ и преобразовании T_k блок Z_k генерируется путём копирования блока из μ с добавлением гауссовского шума с дисперсией φ (Рис. 2):

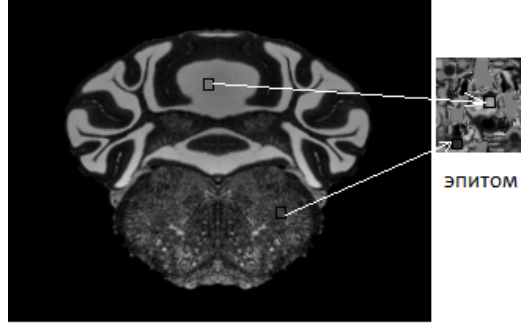
$$P(Z_k | T_k, e) = \prod_{i \in S_k} N(z_{k,i} | \mu_{T_k(i)}, \varphi_{T_k(i)}).$$

Полагая, что блоки генерируются независимо, получаем совместное распределение:

$$P(Z, T, e) = \prod_{k=1}^P P(Z_k, T_k, e) = P(e) \prod_{k=1}^P P(T_k) P(Z_k | T_k, e).$$

Априорные распределения $P(e)$ и $P(T_k)$ в дальнейшем будем полагать равномерными.

Представленная вероятностная модель не может быть использована для решения задачи сегментации изображений напрямую, т.к. в ней не учитывается разметка изображения на классы. В нашей работе предлагается модификация вероятностной модели эпитомов, которая учитывает помимо самого изображения и его разметку на классы.



обучающее изображение

Рис. 2. Пример эпитома для изображения. Стрелками обозначены T_k .

Рассмотрим данную модель. Для этого введём новые понятия: $Q = \{Q_k\}_{k=1}^P$ – набор разметок на классы для каждого из блоков. Кроме того, расширим понятие эпитома – введём в него третье поле $\psi = \{\psi_{t,l}\}$ – вероятности класса l для элемента эпитома t , $l \in \{1 \dots K\}$.

Новая модель имеет такой вид:

$$P(Z, Q, T, e, \psi | \alpha) = \prod_k P(Z_k, Q_k, T_k, e, \psi | \alpha) = P(e) P(\psi | \alpha) \prod_k P(Z_k, Q_k | T_k, e, \psi) P(T_k)$$

$$P(Z_k, Q_k | T_k, e, \psi) = \prod_{i \in S_k} N(z_{k,i} | \mu_{T_k(i)}, \sigma_{T_k(i)}) \psi_{T_k(i), Q_k(i)},$$

$$P(\psi | \alpha) = \prod_t Dir(\psi_t | \alpha) \text{ – априорное распределение Дирихле.}$$

Теперь модель необходимо обучить на изображениях с известной разметкой. Обучение производится путём максимизации апостериорной вероятности с использованием EM-алгоритма:

$$P(e, \psi | Z, Q, \alpha) \longrightarrow \max_{e, \psi}.$$

E-шаг:

$$q(T_k) = P(T_k | Z_k, Q_k, e, \psi) \propto \prod_{i \in S_k} N(z_{k,i} | \mu_{T_k(i)}, \varphi_{T_k(i)}) \psi_{T_k(i), Q_k(i)}, \quad (2a)$$

M-шаг:

$$\mu_t = \frac{\sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k) z_{k,i}}{\sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k)}, \quad \varphi_t = \frac{\sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k) (z_{k,i} - \mu_t)^2}{\sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k)}, \quad (2b)$$

$$\psi_{t,l} = \frac{\alpha_l - 1 + \sum_k \sum_{i: Q_k(i)=l} \sum_{T_k: T_k(i)=t} q(T_k)}{\sum_j \alpha_j - K + \sum_k \sum_i \sum_{T_k: T_k(i)=t} q(T_k)}.$$

Для ускорения вычислений на E-шаге можно воспользоваться представлением операции (2a) через свертку [7].

Сегментация тестового изображения

Рассмотрим процедуру сегментации тестового изображения Z с помощью обученной эпитомной модели (e, ψ) . Для этого в энергию E введём глобальный текстурный потенциал, определяемый с помощью блоков:

$$E(Q | Z) = - \sum_k \log P(Z_k, Q_k | e, \psi) + \sum_{i \in V} \theta_i(Q(i), Z(i)) + \sum_{(i,j) \in E} \theta_{ij}(Q(i), Q(j), Z(i), Z(j))$$

Приведём эту энергию к виду (1), т.е. к функционалу только с унарными и парными потенциалами. Из неравенства Йенсена следует, что для любого распределения $q(T_k)$ верно:

$$- \sum_k \log P(Z_k, Q_k | e, \psi) \leq - \sum_k \sum_{i \in S_k} \sum_{T_k} q(T_k) [\log \psi_{T_k(i), Q_k(i)} + \log N(z_{k,i} | \mu_{T_k(i)}, \varphi_{T_k(i)})] + \sum_k E_{q(T_k)} \log q(T_k) - \sum_k E_{q(T_k)} \log P(T_k). \quad (3)$$

Заметим, что подчёркнутые члены в оценке (3) не зависят от разметки Q . Таким образом, при фиксированном распределении $q(T_k)$ оценка свёрху на глобальный потенциал распадается на сумму унарных потенци-

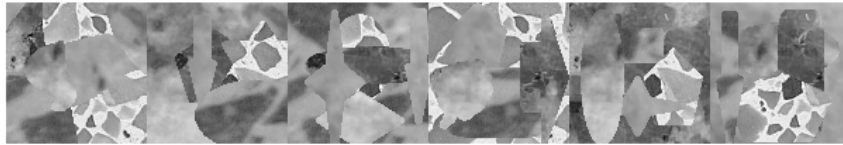
алов по отдельным пикселям. Тогда, зафиксировав это распределение, разметку можно найти алгоритмом α -расширения. Получаем итерационный алгоритм: сначала при фиксированной разметке Q находим распределение $q(T_k)$ по формуле (2a), затем при фиксированном $q(T_k)$ находим разметку Q . Итерационный процесс продолжается до сходимости.

Эксперименты

В первой части экспериментов рассматривается модельная задача сегментации текстур. Эпитом обучается на наборе изображений выборки, содержащих равную долю каждого из классов (рис.3), после чего производится сегментация тестового изображения (рис.4), где в качестве текстурного потенциала выступает результат работы textonboost (рис.5) и предложенного алгоритма (рис. 6). В экспериментах помимо текстурного потенциала задействованы потенциалы цвета и парные потенциалы.

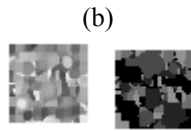
Вторая часть экспериментов посвящена изучению поведения алгоритма на реальной задаче сегментации гистологических изображений срезов мозга мыши на анатомические структуры. Алленовский атлас мозга разбивается на тройки изображений, по двум из трёх изображений строится эпитом, а третье изображение выступает в качестве тестового (рис. 7). Производится его сегментация на основе потенциалов textonboost (рис. 8), эпитомных потенциалов (рис. 9). Во всех экспериментах помимо текстурных потенциалов задействованы потенциалы положения, цвета и парные потенциалы.

Для обеих задач стандартные потенциалы считаются по обучающей выборке (в модельной задаче – по рис.3a-b, в реальной – по рис.7a-b).



(a)





(c) (d)

Рис. 3 (a) – обучающее изображение, (b) – его истинная сегментация, (c) – эпитом, (d) – сегментация эпитома (наиболее вероятные классы)

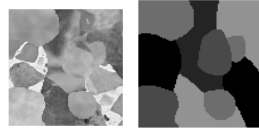


Рис. 4. Тестовое изображение (слева) и его истинная сегментация (справа).



Рис.5. Результат применения текстонов в качестве текстурных потенциалов. Используемые в энергии потенциалы: слева – текстоны (точность 55%), в середине – текстоны + парные потенциалы (56,5%), справа – текстоны, потенциалы цвета и парные потенциалы (59%).

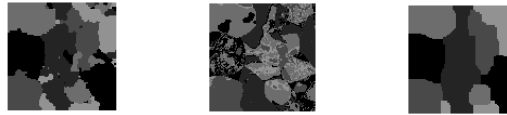
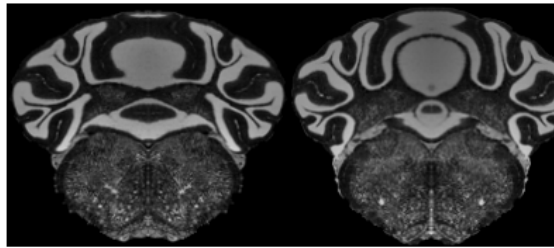


Рис. 6. Результат работы предложенного алгоритма. Используемые потенциалы: слева – эпитомный потенциал (74%), в середине – цветовой потенциал (точность 45%), справа – эпитомный, цветовой и парный потенциалы (78%).



(a)



(c) (d)

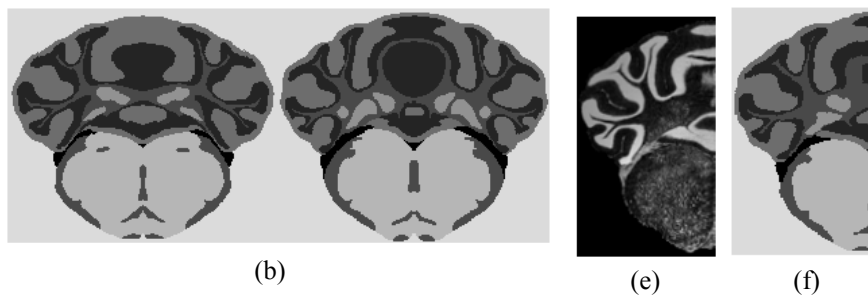


Рис. 7. (a) – обучающее изображение, (b) – его сегментация, (c) – эпитом, (d) – сегментация эпитома (наиболее вероятные классы), (e) - тестовое изображение, (f) - его разметка на анатомические структуры.

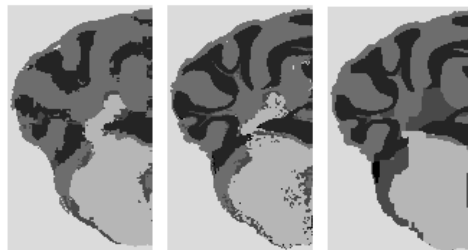


Рис. 8. Результат работы textonboost. Используемые в энергии потенциалы: слева – текстоны (точность 75%), в середине – текстоны и цветовые потенциалы (80%), справа – текстоны, потенциалы цвета, положения и парные потенциалы (84%)

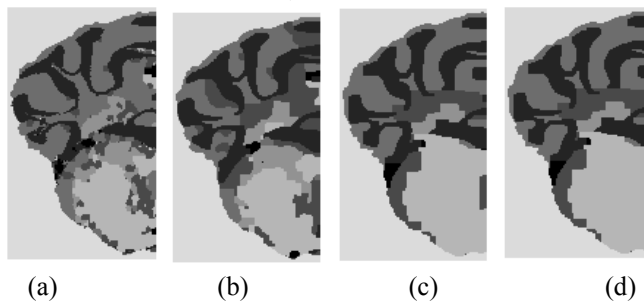


Рис. 9. Результаты работы предложенного алгоритма. (a) – эпитомный потенциал (76%), (b) – цветовые и парные потенциалы (74%), (c) – цветовые, положения и парные потенциалов (86%), (d) – эпитомный потенциал, потенциалы цвета, положения и парные потенциалы (88%).

Выводы

Предложен и реализован алгоритм сегментации изображений на основе эпитомного подхода. Алгоритм выделяет текстурные особенности каждого из классов обучающего изображения, что позволяет повысить эффективность сегментации.

Одним из важных аспектов сегментации изображений в рамках рассматриваемого подхода является выбор коэффициентов линейной комбинации унарных и бинарных потенциалов в функционале энергии (1). В данной работе эти коэффициенты настраивались вручную. Одним из направлений дальнейших исследований является решение задачи автоматической настройки коэффициентов потенциалов с помощью структурных методов обучения, например, структурного метода опорных векторов [8]. Другим направлением дальнейших исследований является усовершенствование потенциалов положения.

Список литературы

1. Jojic.N., Frey.B., Kannan.A. Epitomic analysis of appearance and shape // International Conference on Computer Vision. 2003.
2. Shotton.J., Winn.J., Rother.C., Criminisi.A. TextonBoost for Image Understanding: Multi-Class Object Recognition and Segmentation by Jointly Modeling Texture, Layout, and Context // International Journal on Computer Vision. 2009.
3. Ng L., Pathak S.D., Cuan C., Lau C., Dong H., Sodt A., Dang C., Avants B., Yushkevich P., Gee J.C., Haynor D., Lein E., Jones A., Hawlyrycz M. Neuroinformatics for Genome-Wide 3D Gene Expression Mapping in the Mouse Brain // IEEE Trans. Comp. Biol. and Bioinf. 2007. Vol.4. No.3.
4. Allen Brain Atlas (Internet). Allen Institute for Brain Science, Seattle 2008, <http://www.brain-map.org>
5. Lowe D.G. Object recognition from local scale-invariant features // International Conference on Computer Vision. 1999.
6. Boykov.Y., Veksler.O., Zabih.R. Fast Approximate Energy Minimization via Graph Cuts // IEEE Trans. Pat. Anal. Mach. Intell. 2001.
7. Елшин Д.А., Кропотов Д.А. Использование эпитомного подхода в задаче автоматической сегментации гистологических изображений срезов мозга мыши // Доклады 15-й Всероссийской конференции «Математические методы распознавания образов», 2011.
8. Tsochantaridis I., Joachims T., Hofmann T., Altun Y. Large Margin Methods for Structured and Interdependent Output Variables // Journal of Machine Learning Research. 2005. Vol. 6. P. 1453-1484.