

Share prior на основе упрощенного циркулярного графа

Борис Янгель

ВМиК МГУ

24 марта 2011 г.

- 1 Введение
- 2 Некоторые известные методы введения shape prior
 - Жестко заданная форма
 - Star shape prior
 - Tightness prior
- 3 Shape prior на основе упрощенного циркулярного графа
 - Итерационная сегментация с моделью формы
 - Модель формы
 - Связь с EM-алгоритмом
- 4 Картинки
 - Итерационный процесс
 - Некоторые результаты

Задача сегментации изображения может быть выражена в терминах минимизации MRF-энергии:

$$L^* = \arg \min_L \left\{ \sum_i \phi_i(L_i) + \sum_{(i,j) \in \mathcal{N}} \phi_{ij}(L_i, L_j) \right\}.$$

- L_i — метки пикселей изображения;
- $\phi_i(L_i)$ — безусловный штраф за назначение пикселу данной метки:
 - $\phi_i(L_i) = -\log P(I_i | L_i)$ — штраф за несоответствие цвета;
- $\phi_{ij}(L_i, L_j)$ накладывает ограничения на соседние метки:
 - $\phi_{ij}(L_i, L_j) = \alpha I[L_i \neq L_j]$ — соседние пиксели скорее всего имеют одинаковые метки;
 - $\phi_{ij}(L_i, L_j) = \beta I[L_i \neq L_j] e^{-\gamma |I_i - I_j|}$ — граница проходит там, где сильный перепад цвета.

Потребность в модели формы

К сожалению, локальных ограничений на метки не всегда достаточно.



- В практических задачах зачастую существуют ограничения на форму сегментируемого объекта;
- Однако такие ограничения на форму, как правило, не выражаются через локальные ограничения на метки;
- Поэтому использование моделей формы для регуляризации сегментации представляется интересной и перспективной областью исследований.

С этого момента будем рассматривать только задачу бинарной сегментации, т.е. $L_i \in \{0, 1\}$.

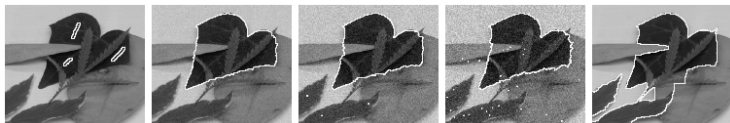
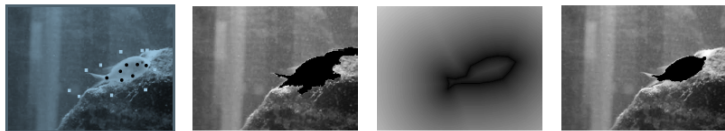
- 1 Введение
- 2 **Некоторые известные методы введения shape prior**
 - Жестко заданная форма
 - Star shape prior
 - Tightness prior
- 3 **Shape prior на основе упрощенного циркулярного графа**
 - Итерационная сегментация с моделью формы
 - Модель формы
 - Связь с EM-алгоритмом
- 4 **Картинки**
 - Итерационный процесс
 - Некоторые результаты

В ряде работ предлагается считать, что форма известна с точностью до аффинного преобразования и задана при помощи двоичной маски или distance transform. Если преобразование известно, существует несколько способов учесть форму при сегментации:

- $\phi_i^S(L_i) = \phi_i(L_i) + \eta I[L_i \neq S_i]$: к каждому унарному потенциалу добавляется штраф за несоответствие метки бинарной маске формы [VM08].
- $\phi_{ij}^S(L_i, L_j) = \phi_{ij}(L_i, L_j) + D_i I[L_i \neq L_j]$: для каждой пары соседних пикселей с разными метками добавляется штраф, возрастающий с увеличением расстояния до границы формы D_i [Fre05].

- Поскольку преобразование формы нам неизвестно, предлагается сделать сегментацию итерационной:
 - 1 Сегментируем изображение без учета модели формы;
 - 2 Ищем “лучшее” аффинное преобразование, которое переводит форму в маску, являющуюся результатом сегментации;
 - 3 Сегментируем изображение с учетом модели формы;
 - 4 Если нет сходимости, то переходим на пункт 2.
- Поскольку метод задает весьма жесткие ограничения на форму, он плохо работает, если в форме присутствует высокая степень вариабельности. Пример: стоящая и скачущая лошадь.

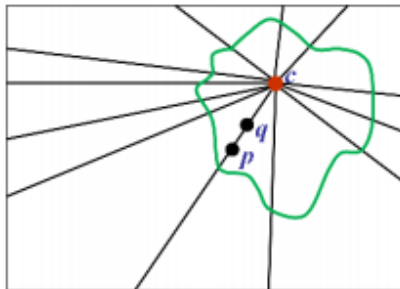
Жестко заданная форма — результаты



Можно наоборот накладывать самые общие ограничения на форму объекта, например требовать *звездности* [Vek08].

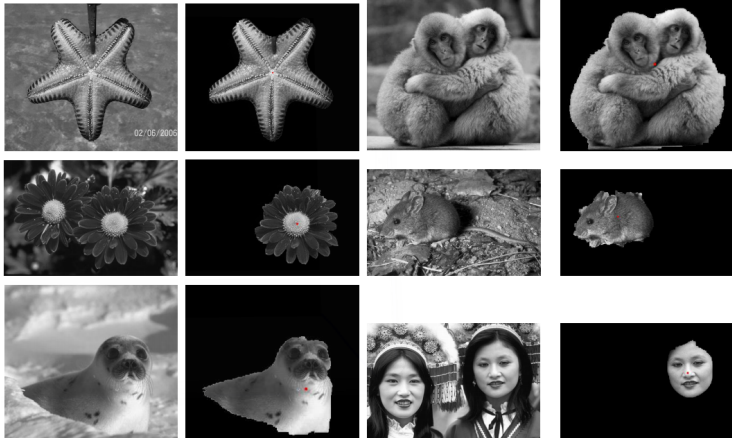


- Назовем область *звездной относительно ее внутренней точки c* , если любой отрезок, соединяющий точку c с некоторой внутренней точкой, целиком лежит внутри области;
- Чтобы гарантировать звездность сегментации, для всех соседних пикселей p и q , лежащих на выходящем из c луче так, что p от c дальше, чем q , положим $\phi_{pq}^S(1, 0) = +\infty$.



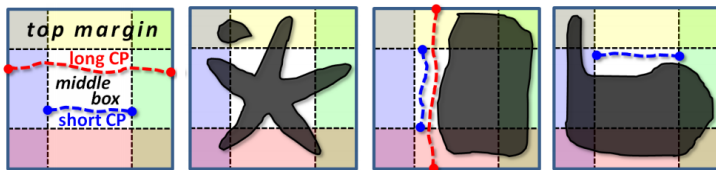
- Если при этом положить $\phi_{pq}^S(0, 1) = \lambda < 0$, начнем штрафовать сегментации с небольшим периметром объекта;
- Для инициализации требуется одна точка;
- Метод быстрый, поскольку не является итерационным;
- К сожалению, далеко не все потенциально интересные объекты имеют звездную форму.

Star shape prior — результаты



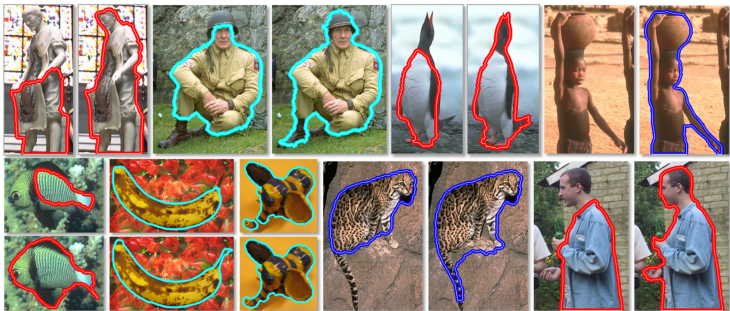
Еще одно ограничение общего характера — требование близости объекта к границе bounding box, который задает пользователь [LKRS09].

- Объект называется *сильно (слабо) натянутым*, если он пересекается со всеми *короткими (длинными) путями*;



- Условие на пересечение объекта с некоторым путем можно записать как ограничение в виде линейного неравенства на сегментацию;
- Можно минимизировать энергию итерационно, на каждой итерации добавляя некоторое подмножество невыполненных ограничений к текущему набору и решая задачу линейного программирования.

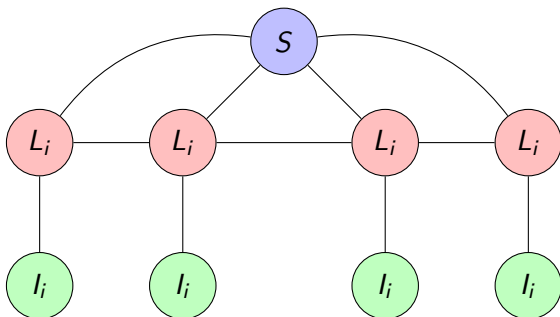
Tightness prior — результаты



- 1 Введение
- 2 Некоторые известные методы введения shape prior
 - Жестко заданная форма
 - Star shape prior
 - Tightness prior
- 3 **Shape prior на основе упрощенного циркулярного графа**
 - Итерационная сегментация с моделью формы
 - Модель формы
 - Связь с EM-алгоритмом
- 4 **Картинки**
 - Итерационный процесс
 - Некоторые результаты

Сперва рассмотрим следующую порождающую модель для сегментации:

- Форма S порождает метки пикселей L_i , причем метки соседних пикселей согласованы друг с другом;
- Каждый пиксель изображения I_i независимо порождается на основе метки L_i .



Запишем совместную вероятность всех переменных модели:

$$\begin{aligned}
 P(S, L, I) &= P(I | S, L)P(S, L) = P(I | L)P(L | S)P(S) = \\
 &= \frac{1}{Z}P(S) \prod_i P(I_i | L_i) \prod_{(i,j) \in \mathcal{N}} \tilde{\phi}_{ij}(L_i, L_j, S) = \\
 &= \frac{1}{Z}P(S) \prod_i P(I_i | L_i) \prod_{(i,j) \in \mathcal{N}} \phi_{ij}(L_i, L_j) \prod_i \phi_i(L_i, S).
 \end{aligned}$$

Тогда можно сформулировать задачу сегментации с моделью формы так: найти

$$\begin{aligned}
 \langle S^*, L^* \rangle &= \arg \max_{S, L} P(S, L | I) = \arg \min_{S, L} \left[-\log P(S) - \right. \\
 &\left. - \sum_i (\log P(I_i | L_i) + \log \phi_i(L_i, S)) - \sum_{(i,j) \in \mathcal{N}} \log \phi_{ij}(L_i, L_j) \right].
 \end{aligned}$$

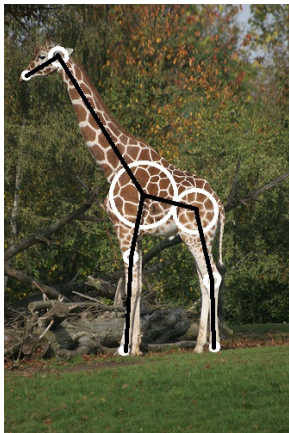
Такую энергию можно минимизировать покоординатным спуском:

$$S^{new} = \arg \min_S \left[-\log P(S) - \sum_i \log \phi_i(L_i^{old}, S) \right],$$

$$L^{new} = \arg \min_L \left[-\sum_i (\log P(l_i | L_i) + \log \phi_i(L_i, S^{new})) - \sum_{(i,j) \in \mathcal{N}} \log \phi_{ij}(L_i, L_j) \right].$$

- Шаг по L представляет из себя задачу сегментации с модифицированными унарными потенциалами;
- Шаг по S — более сложная оптимизационная задача, метод стоит выбирать по ситуации.

- Можно построить модель формы на основе графа, в каждой вершине которого дополнительно задан радиус:



- Такой граф является упрощенным аналогом циркулярного графа, представляющего из себя скелет фигуры вместе с максимальными вписанными кругами.

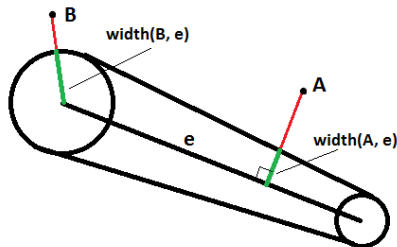
- $P(S)$ задает вероятность различных конфигураций графа
 - Для модели формы жирафа задавались мягкие ограничения на углы между частями и соотношения длин соседних частей;
 - Может также задавать глобальные ограничения на форму, например на поворот и масштаб;
- По набору графов можно выбирать параметры модели автоматически, например методом максимального правдоподобия;
- На практике $P(S)$ может не иметь вероятностной природы (выступить в роли потенциала).

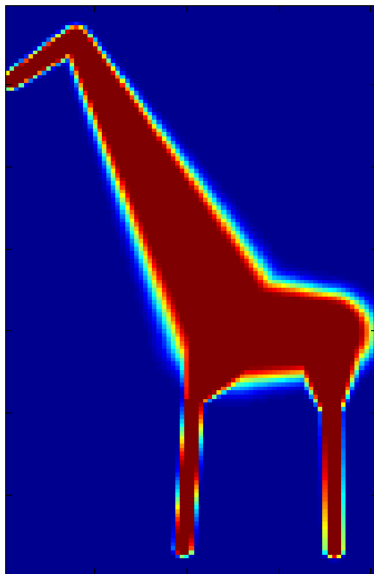
Унарные потенциалы можно задать в виде

$$\phi_i(L_i, S) = L_i \max_{e \in E_S} W_e(i) + (1 - L_i)(1 - \max_{e \in E_S} W_e(i)),$$

где

- E_S — множество ребер графа формы S ;
- $W_e(i) = \exp(-\omega D^P(i, e))$ — функция, убывающая с ростом $D(i, e)$;
- $D(i, e) = \frac{\text{dist}(i, e)}{\text{width}(i, e)}$ — нормированное расстояние от пикселя i до ребра e .





- Оказалось, что если начальное приближение для формы находится “недалеко” от оптимального значения, для оптимизации неплохо подходит метод имитации отжига;
- Для генерации новых решений можно использовать случайную модификацию длин ребер, радиусов вершин и углов между ребрами;
- На тестовых изображениях имитация отжига сходилась за 2000-4000 итераций;
- Вопрос о существовании более эффективного метода оптимизации пока открыт.

- EM-алгоритм служит для максимизации неполного правдоподобия

$$P(X | \theta) = \sum_Z P(X, Z | \theta);$$

- Представляет из себя итерационный процесс:

E-шаг Найти распределение на скрытые переменные

$$P^*(Z) = P(Z | X, \theta_{old});$$

M-шаг Обновить текущее решение

$$\theta_{new} = \max_{\theta} \mathbb{E}_{Z \sim P^*(Z)} \log P(X, Z | \theta);$$

- Можно также максимизировать неполную апостериорную вероятность. Для этого к выражению, максимизируемому на M-шаге, нужно прибавить $\log P(\theta)$.

Будем считать S скрытой переменной и максимизировать $P(L | I)$. Тогда на M-шаге мы получим оптимизационную задачу

$$\begin{aligned} L_{new} &= \max_L \left[\mathbb{E}_{S \sim P^*(S)} \log P(I, S | L) + \log P(L) \right] = \\ &= \max_L \mathbb{E}_{S \sim P(S|I, L_{old})} \log P(I | L) P(L | S) P(S). \end{aligned}$$

При этом

$$P^*(S) = P(S | I, L_{old}) = \frac{1}{Z} P(I | L_{old}) P(L_{old} | S) P(S).$$

- Если в EM-алгоритме на E-шаге приближать распределение $P(S | I, L_{old})$ дельта-функцией в точке максимума, получим в точности алгоритм сегментации методом покоординатной оптимизации.
- Вместо этого можно попробовать применить схему Monte Carlo EM.

- 1 Введение
- 2 Некоторые известные методы введения shape prior
 - Жестко заданная форма
 - Star shape prior
 - Tightness prior
- 3 Shape prior на основе упрощенного циркулярного графа
 - Итерационная сегментация с моделью формы
 - Модель формы
 - Связь с EM-алгоритмом
- 4 Картинки
 - Итерационный процесс
 - Некоторые результаты

- Сегментация инициализировалась ограничивающим прямоугольником;
- В качестве цветowych моделей объекта и фона использовались смеси гауссиан;
- В качестве L^{old} на первой итерации использовались метки, полученные в результате сегментации без модели формы;
- Вес shape prior линейно увеличивался в течение первых 10 итераций покоординатного спуска;
- На всех картинках процесс сошелся за 10-15 итераций.







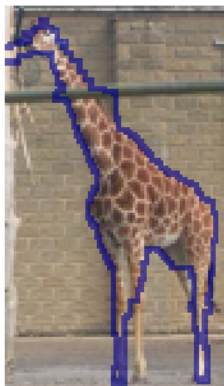






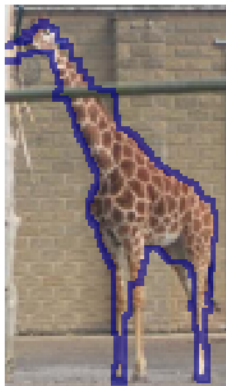




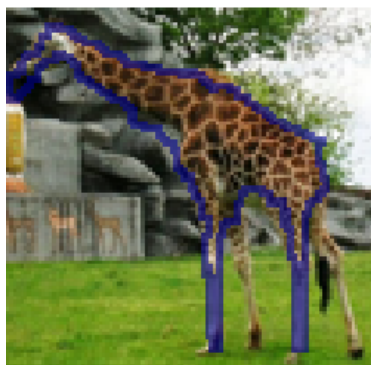
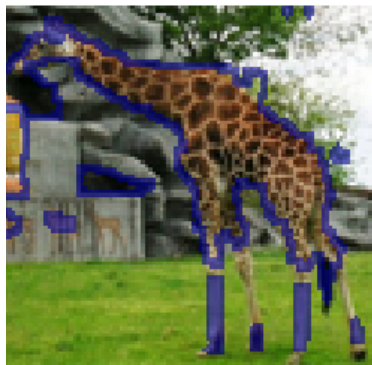




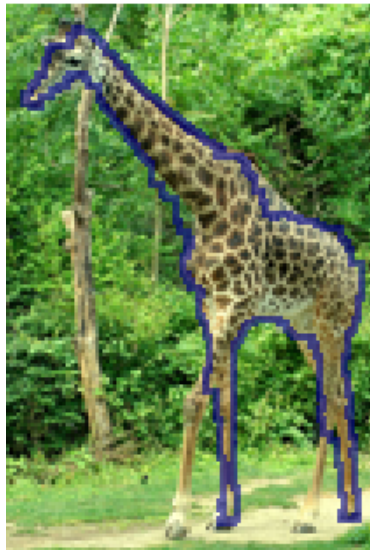
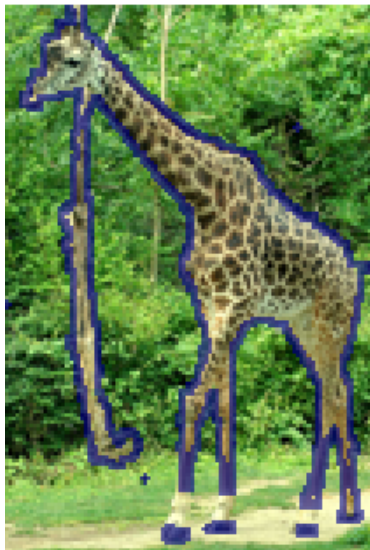




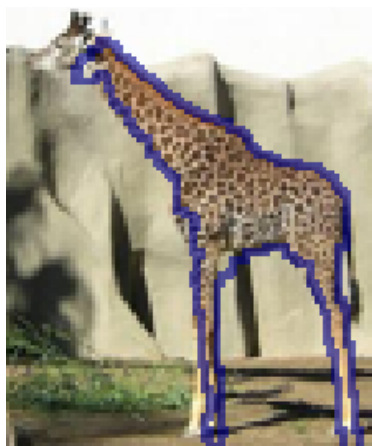
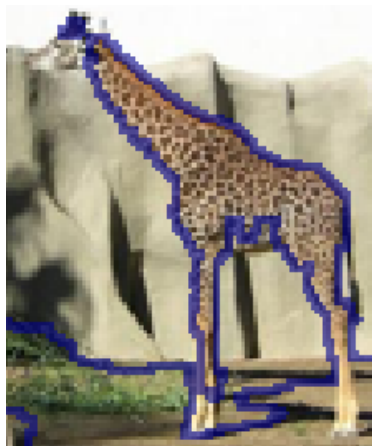
Некоторые результаты 1

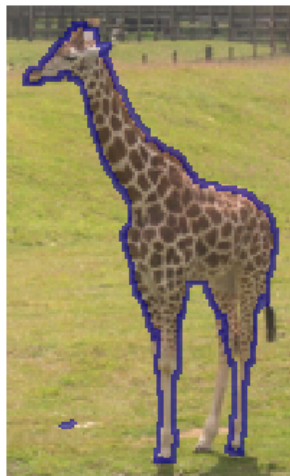


Некоторые результаты 2

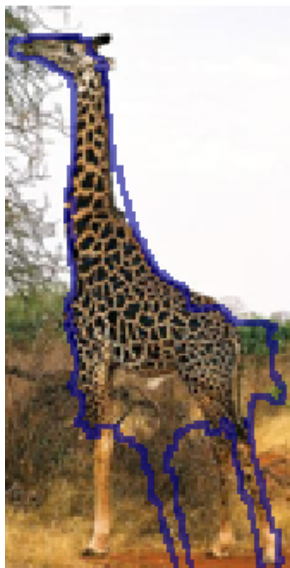
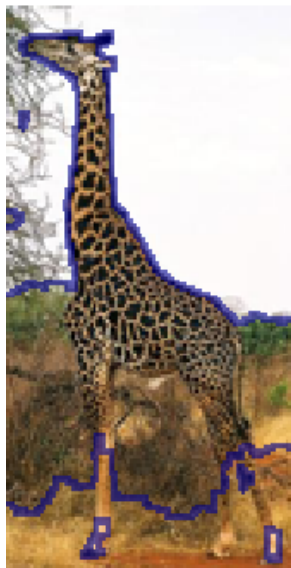


Некоторые результаты 3

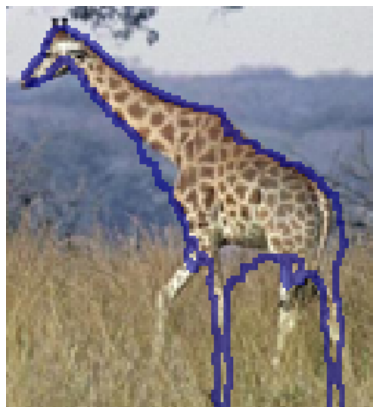
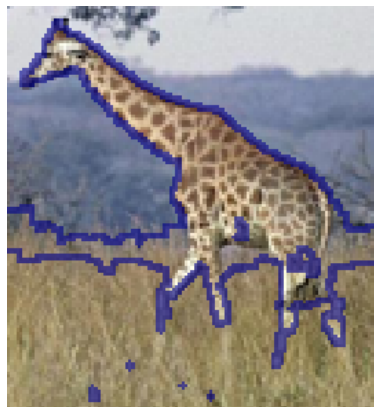




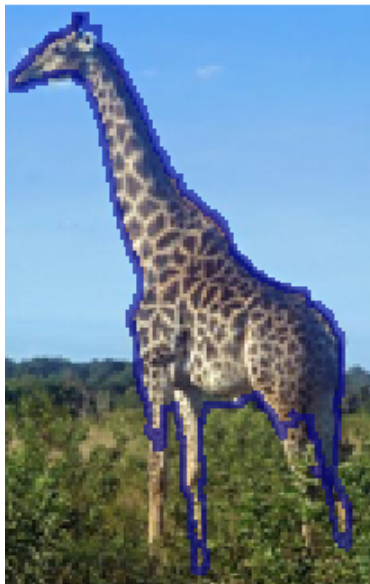
Некоторые результаты 5

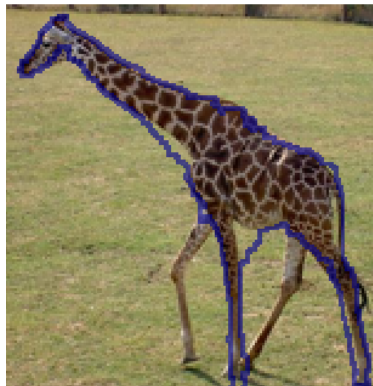
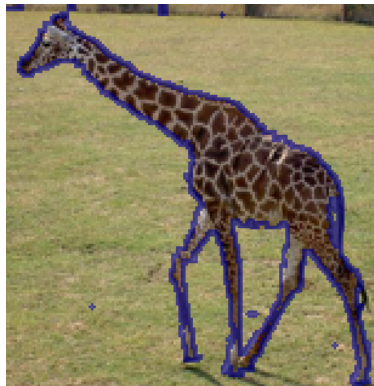


Некоторые результаты 6



Некоторые результаты 7







D. Freedman.

Interactive Graph Cut Based Segmentation with Shape Priors.
2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pages 755–762, 2005.



Victor Lempitsky, Pushmeet Kohli, Carsten Rother, and Toby Sharp.

Image segmentation with a bounding box prior.

In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 277–284. IEEE, September 2009.



Olga Veksler.

Star Shape Prior for Graph-Cut Image Segmentation.

In *Proceedings of the 10th European Conference on Computer Vision*, 2008.



Nhat Vu and B.S. Manjunath.

Shape prior segmentation of multiple objects with graph cuts.
2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008.