

Точное выделение границ объектов



Михаил Ерофеев

Video Group

CS MSU Graphics & Media Lab



Содержание

- **Введение**
- Robust Matting
- Video SnapCut

Введение

Постановка задачи

- $$I(x) = \alpha(x)F(x) + (1 - \alpha(x))B(x)$$

x – точка на плоскости

I – исходное изображение

F – объект для выделения

B – фон

α – коэффициент прозрачности

Задача: по данному изображению найти остальные неизвестные

Введение

Некорректно поставленная задача

Три неизвестных (F, B, α), одна известная (I).
Чтобы решить задачу нужны дополнительные условия.

Возможные дополнительные условия:

- информация от пользователя
- изображения с разных камер
- освещение объектов вспышкой
- данные о движении



Содержание

- Введение
- **Robust Matting**
- Video SnapCut

Robust Matting

Описание метода

Метод с использованием входного изображения и trimap строит α -маску.

Основная идея: использовать для определения коэффициента прозрачности не все точки фона и объекта в окрестности, а лишь наиболее «надежные»

Robust Matting

Выбор точек

Для пары точек фона и объекта
из окрестности неизвестной точки найдем:

- Предполагаемый коэффициент прозрачности

$$\hat{\alpha} = \frac{(C - B^j)(F^i - B^j)}{\|F^i - B^j\|^2} \quad C, B^j, F^i - \text{наблюдаемый цвет, цвет фона и объекта}$$

- Разница между наблюдаемым цветом и предсказанным

$$R_d(F^i, B^j) = \frac{\|C - (\hat{\alpha}F^i + (1 - \hat{\alpha})B^j)\|}{\|F^i - B^j\|}$$

Robust Matting

Выбор точек (2)

- Введем два веса для поощрения точек, сильно похожих на фон или объект

$$w(X^i) = \exp\left(-\frac{\|X^i - C\|^2}{\min_i (\|X^i - C\|)^2}\right) \quad X \in \{F, B\}$$

- Определим надежность пары точек

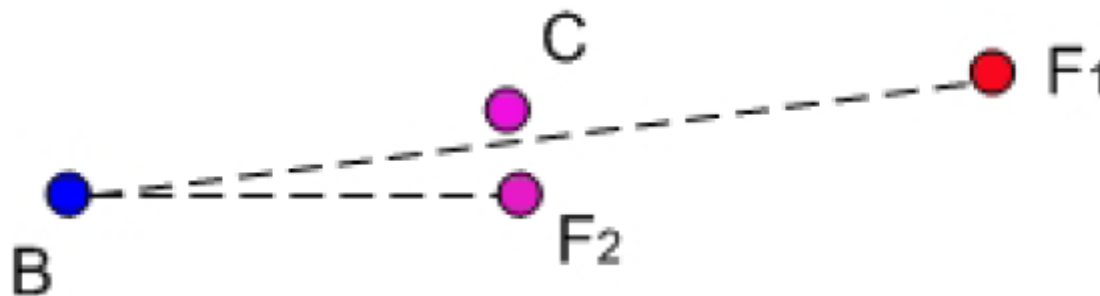
$$f(F^i, B^j) = \exp\left\{-\frac{R_d(F^i, B^j)^2 \cdot w(F^i) \cdot w(B^j)}{\sigma^2}\right\}$$

$$\sigma = 0.1$$

Robust Matting

Пояснение смысла надежности

Введенной функцией надежности будет признана предпочтительнее пара точек B, F_2



Robust Matting

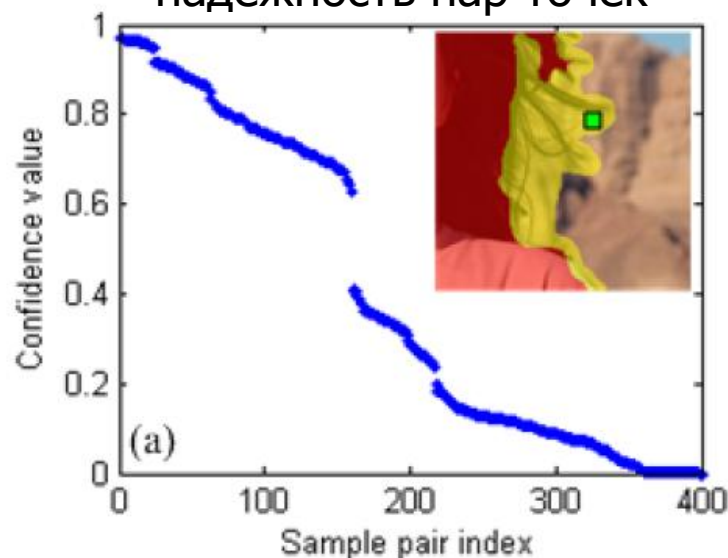
Карты надежности и прозрачности

- Для каждой неопределенной точки выберем три пары точек фона и объекта с наибольшей надежностью
- Усредним надежность и прозрачность этих пар
- Занесем полученные значения в карту прозрачности и надежности

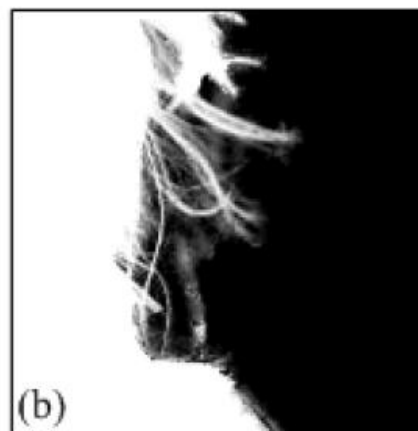
Robust Matting

Карты надежности и прозрачности

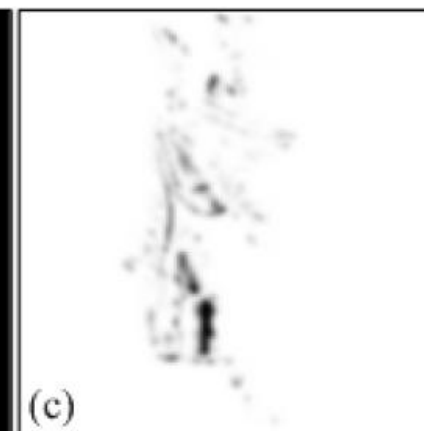
Исходное
изображение и
надежность пар точек



Карта
прозрачности

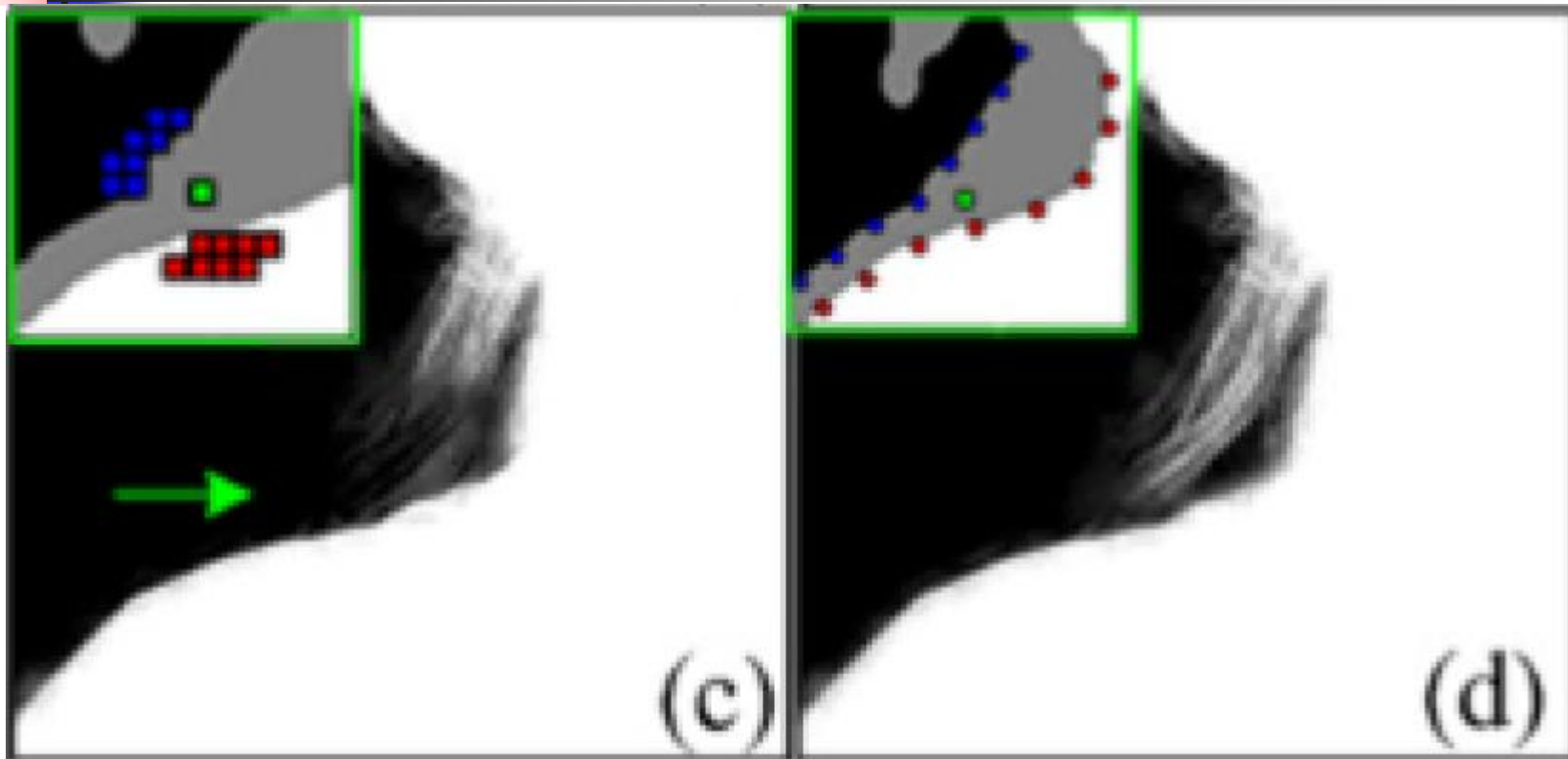


Карта
надежности



Robust Matting

Форма окрестности



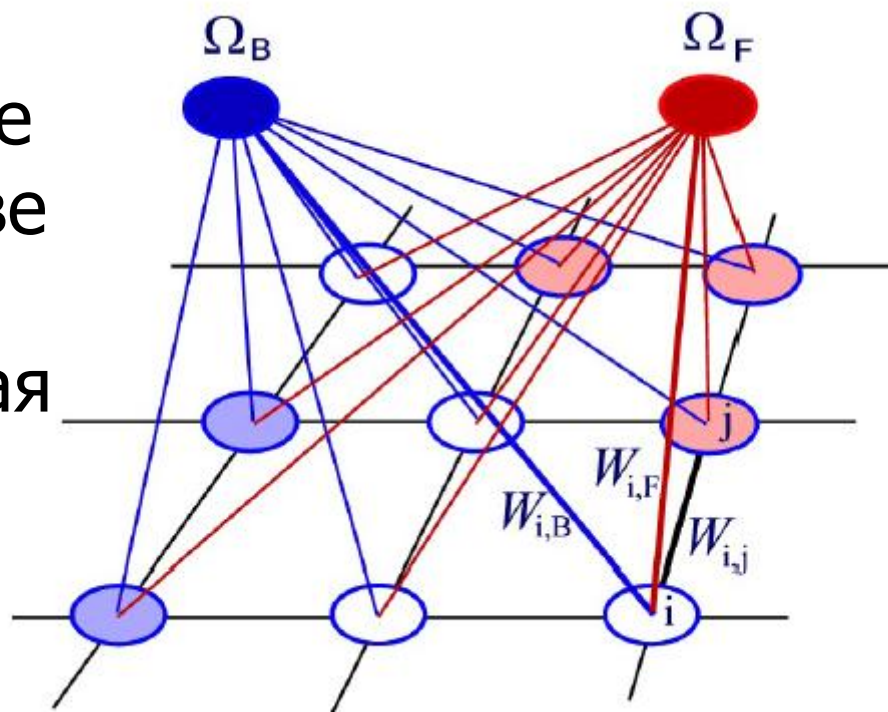
Традиционная
окрестность

Предложенная
окрестность

Robust Matting

Оптимизация маски

Рассмотрим граф.
 Множество вершин – все точки изображения и две специальных вершины фон B и объект F . Каждая вершина соединена со своими соседями и специальными вершинами



Robust Matting

Оптимизация маски

Веса ребер:

$$W(i, F) = \gamma \cdot [\hat{f}_i \hat{\alpha}_i + (1 - \hat{f}_i) \delta(\hat{\alpha}_i > 0.5)]$$

$$W(i, B) = \gamma \cdot [\hat{f}_i (1 - \hat{\alpha}_i) + (1 - \hat{f}_i) \delta(\hat{\alpha}_i < 0.5)]$$

$$W_{ij} = \sum_{(i,j) \in w_k} \frac{1}{9} (1 + (C_i - \mu_k)(\Sigma_k + \frac{\epsilon}{9}I)^{-1}(C_j - \mu_k))$$

w_k – окно 3×3 , C_i – наблюдаемый цвет,

μ_k – математическое ожидание внутри окна,

Σ_k – матрица ковариации, $\gamma = 0.1$, $\epsilon = 10^{-5}$

WANG, J., AND COHEN, M. 2007.

Optimized color sampling for robust matting.

In Proc. of IEEE CVPR, 2007

Robust Matting

Построение конечной маски

К построенному графу применим метод Random Walk, т.е. найдем вероятность прийти из вершины F в каждую точку изображения. Составим из этих вероятностей конечную маску.

Robust Matting

Примеры работы и сравнение

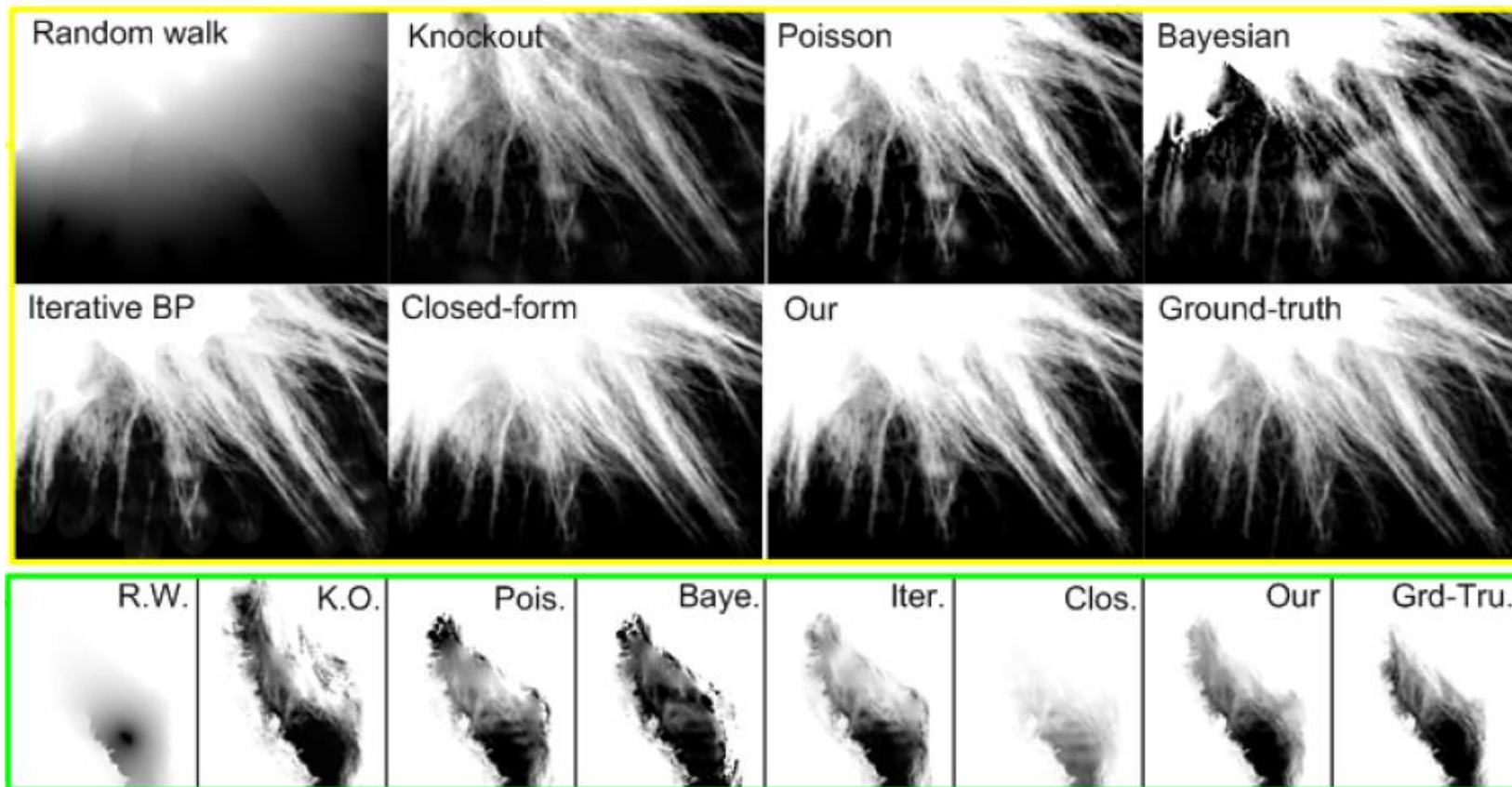
Объективное сравнение алгоритмов
на основе среднеквадратичной ошибки

Ground-truth:

- Три изображения на основе blue screen matting
- Три на основе подмены фона
- Два натуральных изображения

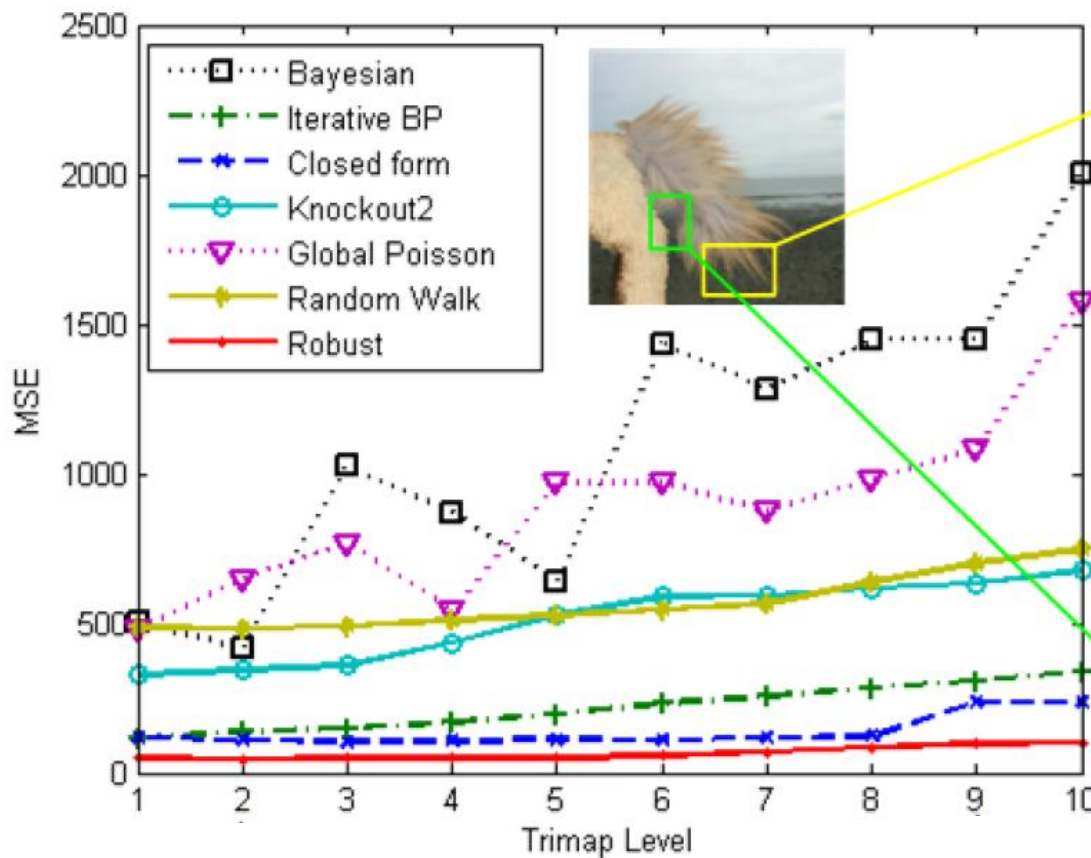
Robust Matting

Примеры работы и сравнение (2)



Robust Matting

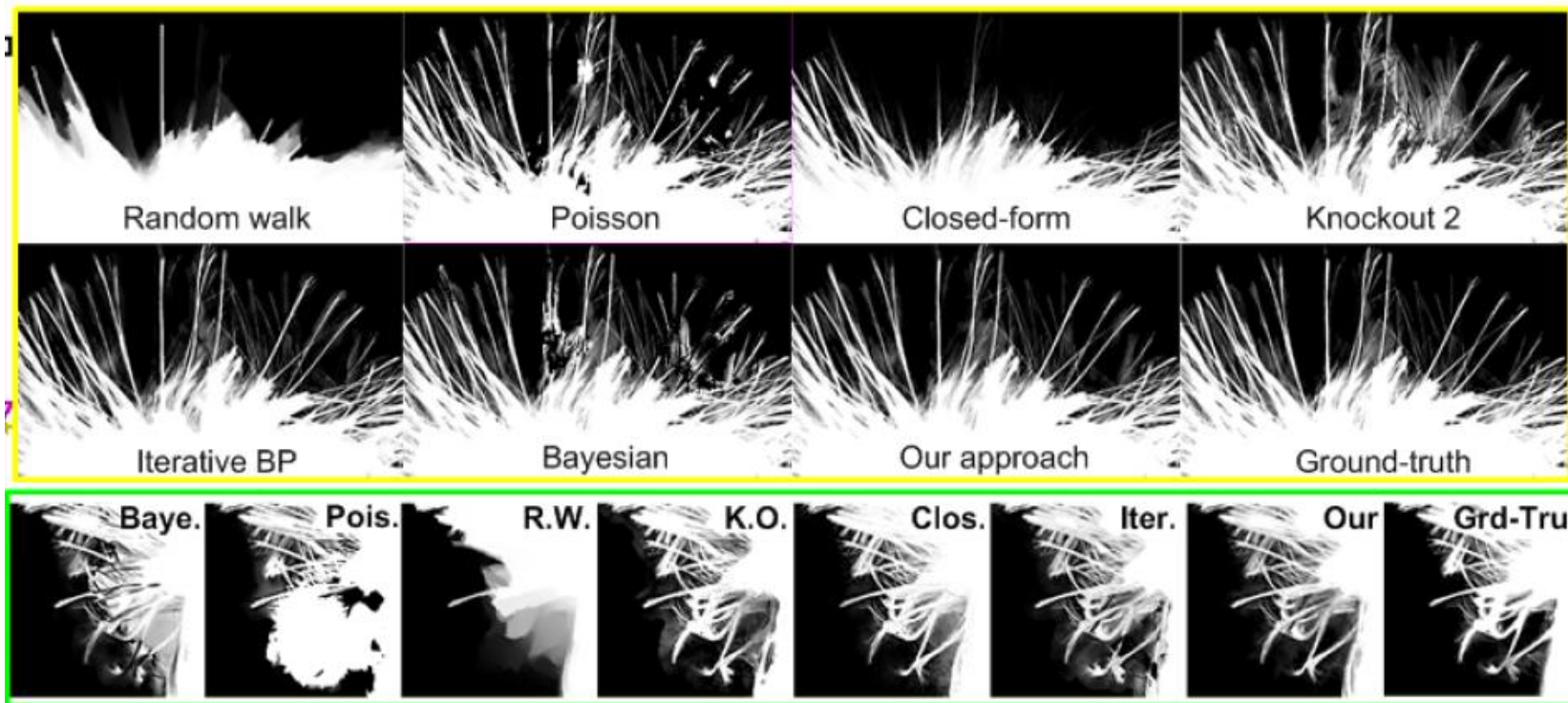
Примеры работы и сравнение (3)



WANG, J., AND COHEN, M. 2007.
Optimized color sampling for robust matting.
In Proc. of IEEE CVPR, 2007

Robust Matting

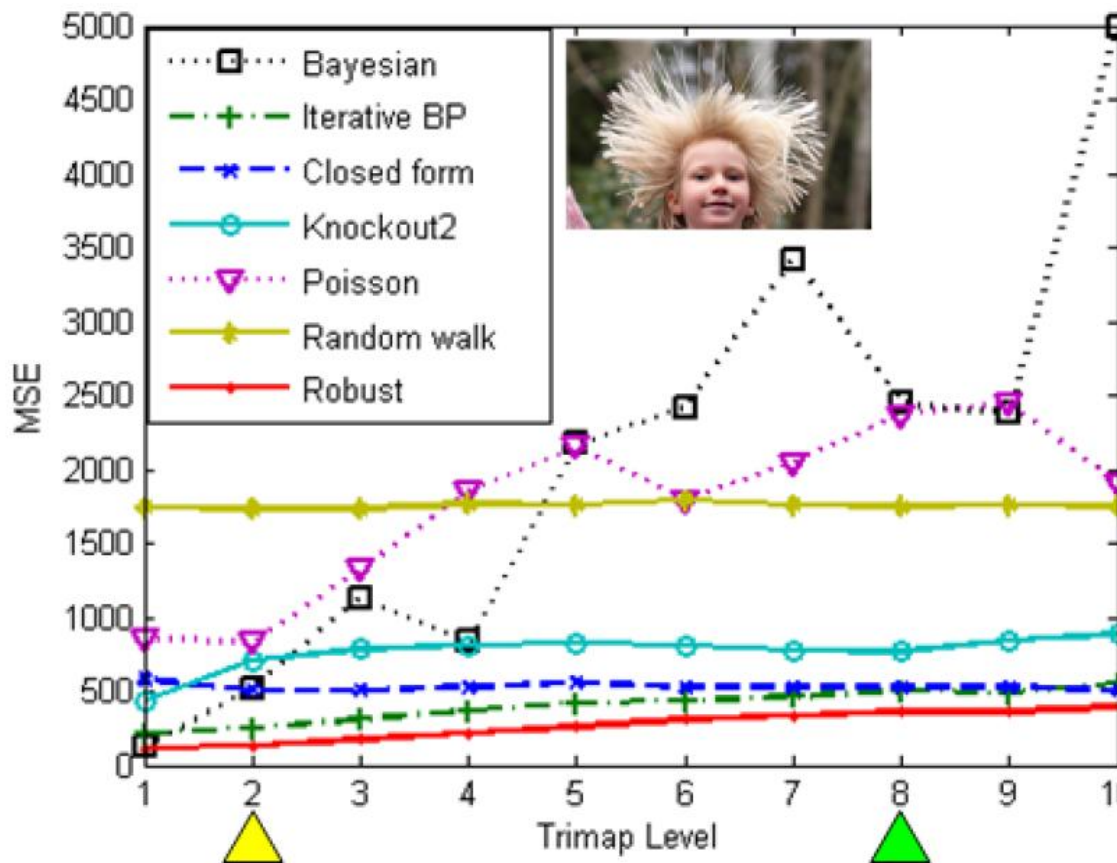
Примеры работы и сравнение (4)



WANG, J., AND COHEN, M. 2007.
Optimized color sampling for robust matting.
In Proc. of IEEE CVPR, 2007

Robust Matting

Примеры работы и сравнение (5)



WANG, J., AND COHEN, M. 2007.
Optimized color sampling for robust matting.
In Proc. of IEEE CVPR, 2007

Robust Matting

Анализ метода

- Достоинства
 - Лучший результат в сравнительном тесте
 - Мало деградирует при снижении качества trimap
- Недостатки
 - Уступает методам, основанным на распространении маски, при сравнении на грубых масках



Содержание

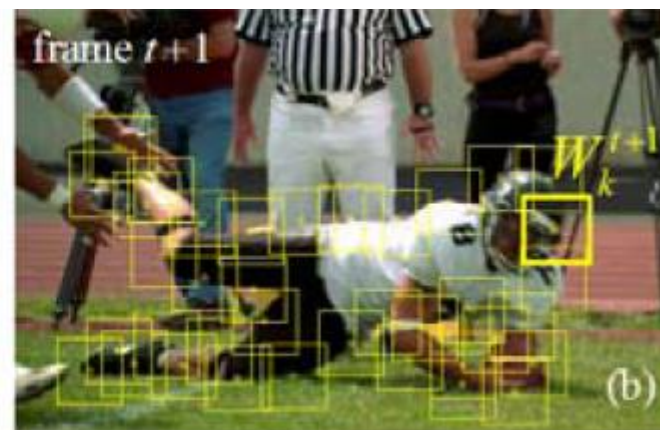
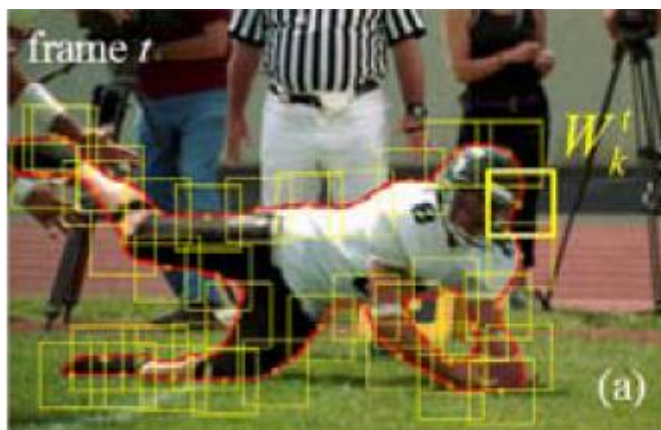
- Введение
- Robust Matting
- **Video SnapCut**

Video SnapCut

Описание метода

В качестве дополнительных данных используется грубая маска объектов для ключевых кадров.

Основная идея: решать задачу в маленьких окнах, расположенных вдоль границы объекта и перемещать их вслед за объектом.



Video SnapCut

Схема работы метода

1. Для первого кадра получить грубую маску
2. Расположить вдоль границы объекта небольшие перекрывающиеся друг друга окна
3. Инициализировать модель цвета и формы внутри каждого окна
4. Построить бинарную маску
5. Построить α -маску

Video SnapCut

Схема работы метода (2)

6. Перейти к следующему кадру
7. На основе данных optical flow переместить окна и обновить маску
8. Обновить модель цвета и формы внутри каждого окна
9. Построить бинарную и α -маску
10. Перейти к шагу 6

Video SnapCut

Инициализация окон

Для каждого окна необходимо инициализировать:

- Модель цвета M_c
- Надежность модели цвета f_c
- Модель формы M_s

Video SnapCut

Модель цвета

Распределение цвета моделируется с помощью Gaussian mixture model (GMM)

Алгоритм инициализации:

- Выбрать внутри окна точки фона на расстоянии не менее пяти пикселей от границы
- Построить GMM из трех компонент
- Повторить аналогичные действия для объектов

Video SnapCut

Gaussian mixture model (GMM)

Вектор \mathbf{z} принадлежит объекту с вероятностью

$$p_{fg}(\mathbf{z}) = \sum_k \pi_{fg}(k) G(\mathbf{z}; \vec{\mu}_{fg}(k), R_{fg}(k))$$

Вектор \mathbf{z} принадлежит фону с вероятностью

$$p_{bg}(\mathbf{z}) = \sum_k \pi_{bg}(k) G(\mathbf{z}; \vec{\mu}_{bg}(k), R_{bg}(k))$$

Плотность многомерного распределения Гаусса

$$G(\mathbf{z}; \vec{\mu}, R) = \|R\|^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}(\mathbf{z} - \vec{\mu})^T R^{-1}(\mathbf{z} - \vec{\mu})\right\}$$

Video SnapCut

Построение GMM

$\mu_{bg}(k), \mu_{fg}(k)$ – среднее значение векторов в k -ом кластере фона или объекта

$R_{bg}(k), R_{fg}(k)$ – матрица ковариаций векторов в k -ом кластере

$\pi_{bg}(k), \pi_{fg}(k)$ – вес кластера пропорционален количеству векторов в нем

Вероятность, что точка принадлежит объекту

$$p_c(x) = \frac{p_{fg}(x)}{p_{fg}(x) + p_{bg}(x)}$$

Video SnapCut

Надежность цветовой модели

Отражает, насколько фон и объекты различимы с точки зрения построенной модели

$$f_c = 1 - \frac{\int_{W_k} |L^t(x) - p_c(x)| \cdot \omega_c(x) dx}{\int_{W_k} \omega_c(x) dx}$$

$$\omega_c(x) = \exp\left(-\frac{d^2(x)}{\sigma_c^2}\right)$$

W_k – окно, L^t – маска, d – расстояние до края,

σ_c – половина размера окна

Video SnapCut

Модель формы

Состоит из текущей маски и надежности модели

$$f_s(x) = 1 - \exp(-d^2(x)/\sigma_s^2)$$

d – расстояние до края

$$\sigma_s = \begin{cases} \sigma_{min} + a(f_c - f_{cutoff})^r & f_{cutoff} < f_c \leq 1 \\ \sigma_{min} & 0 \leq f_c \leq f_{cutoff} \end{cases}$$

$$a = (\sigma_{max} - \sigma_{min}) / (1 - f_{cutoff})^r$$

$$f_{cutoff} = 0.85, \sigma_{min} = 2, r = 2 \quad \sigma_{max} - \text{ширина окна}$$

Video SnapCut

Построение бинарной маски

Вероятность признать пиксель объектом
внутри окна

$$p_{\mathcal{F}}^k(x) = f_s(x)L^{t+1}(x) + (1 - f_s(x))p_c(x)$$

Итоговая вероятность

$$p_{\mathcal{F}}(x) = \frac{\sum_k p_{\mathcal{F}}^k(x) (|x - c_k| + \epsilon)^{-1}}{\sum_k (|x - c_k| + \epsilon)^{-1}}$$

$\epsilon = 0.1$, c_k – центр окна

Video SnapCut

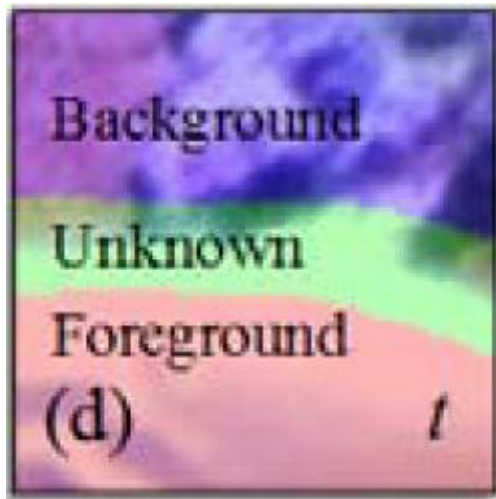
Построение бинарной маски (2)

- Построим на основе вероятностей функцию энергии и применим Graph Cut, получим маску
- Обновим модель цвета и формы в каждом окне
- Построим новую маску
- Будем повторять, пока маска не перестанет меняться

Video SnapCut

Построение α -маски

Построим trimap, применив к бинарной маске последовательно сужение и расширение



Video SnapCut

Построение α -маски (2)

$$\alpha^t = \arg \min_{\alpha^t} \sum_{x \in U} [\lambda_x^T (\alpha_x^t - \hat{\alpha}_x^{t-1})^2 + \lambda_x^C (\alpha_x^t - \tilde{\alpha}_x^t)^2] + \lambda^S \alpha^{tT} L_m \alpha^t,$$

U – неизвестная область trimap

Первое слагаемое – требование непрерывности во времени

Второе – насколько хорошо данный цвет можно описать смешением фона и объекта

Третье – непрерывность в пространстве

Video SnapCut

Построение α -маски (3)

$$\alpha^t = \arg \min_{\alpha^t} \sum_{x \in U} \left[\lambda_x^T (\alpha_x^t - \hat{\alpha}_x^{t-1})^2 + \lambda_x^C (\alpha_x^t - \tilde{\alpha}_x^t)^2 \right] + \lambda^S \alpha^{tT} L_m \alpha^t,$$

$$\hat{\alpha}_x^{t-1} = \frac{\sum_k \beta_{W_k} \cdot \alpha^{t-1}(x_{W_k}^{t-1})}{\sum_k \beta_{W_k}} \quad x_{W_k}^{t-1} - \text{точка, соответствующая данной из прошлого кадра}$$

$$\beta_{W_k} = \exp \left(- \frac{1}{|W_k|} \sum_{i \in W_k} |L_{W_k}^t(i) - L_{W_k}^{t-1}(i)| / \sigma_w \right)$$

Video SnapCut

Построение α -маски (4)

$$\alpha^t = \arg \min_{\alpha^t} \sum_{x \in U} \left[\lambda_x^T (\alpha_x^t - \hat{\alpha}_x^{t-1})^2 + \lambda_x^C (\alpha_x^t - \tilde{\alpha}_x^t)^2 \right] + \lambda^S \alpha^{tT} L_m \alpha^t,$$

$$L_m(i, j) = \sum_{k | (i, j) \in w_k} \left(\delta_{ij} - \frac{1}{|w_k|} \left(1 + \frac{1}{\frac{\epsilon}{|w_k|} + \sigma_k^2} (I_i - \mu_k)(I_j - \mu_k) \right) \right)$$

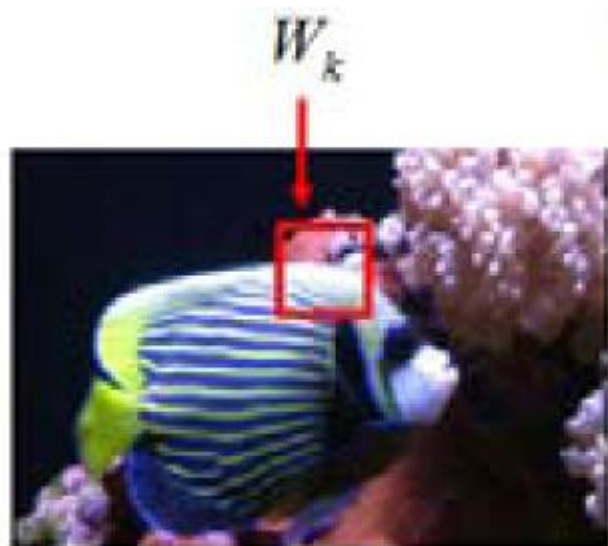
δ_{ij} – символ Кронекера

μ_k, σ_k – математическое ожидание и дисперсия интенсивности внутри окна

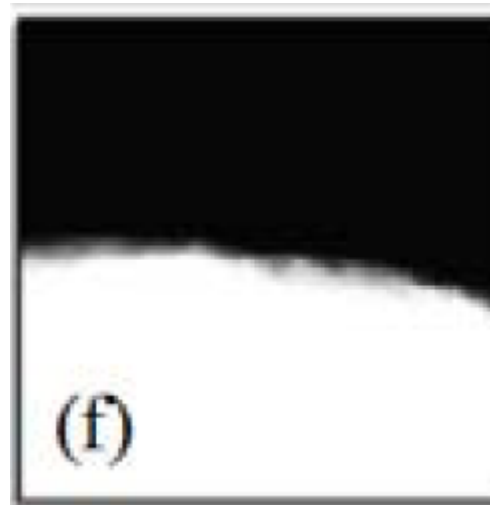
Video SnapCut

Построение α -маски (5)

Получение α -маски теперь сводится к решению СЛАУ.



Исходный фрагмент



α -маска

Video SnapCut

Перемещение окон

- Найдем особые точки внутри объекта для пары кадров
- Найдем аффинное преобразование между этими кадрами
- Между новым кадром и преобразованным найдем optical flow
- Сдвинем окно на среднее значение векторов optical flow в этом окне

Video SnapCut

Обновление моделей

- Модель формы обновляется на основании новой маски
- Модель цвета
Две модели
 - Историческую – модель с прошлого кадра
 - Текущую – модель, составленную из модели для данного кадра и историческойЕсли использование исторической модели приводит к уменьшению области фона – возьмем текущую модель.

Video SnapCut

Изменение топологии объекта

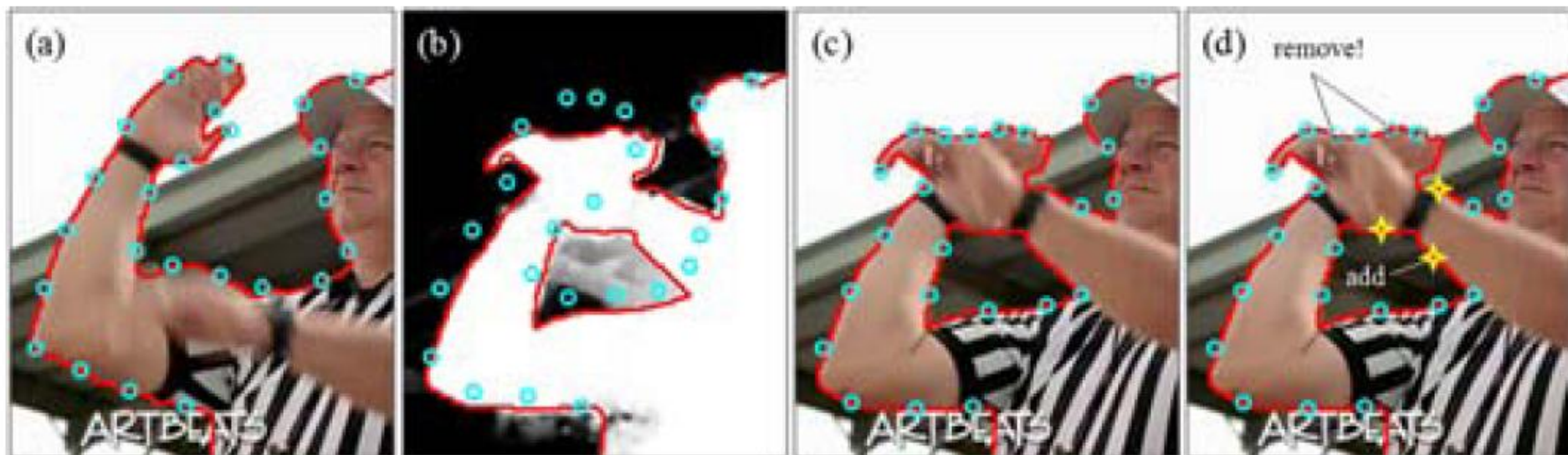
Форма объекта может существенно изменяться, необходимо удалять окна и добавлять новые:

- Сместим центры окон на границу объекта
- Удалим окно, если плотность окон в окрестности высока
- Добавим окно, если есть точки, не покрытые ни одним окном

Video SnapCut

Изменение топологии объекта (2)

Пример добавления новых окон



Video SnapCut

Распространение в двух направлениях

■ Применим алгоритм в обратном направлении.

Новая карта вероятностей:

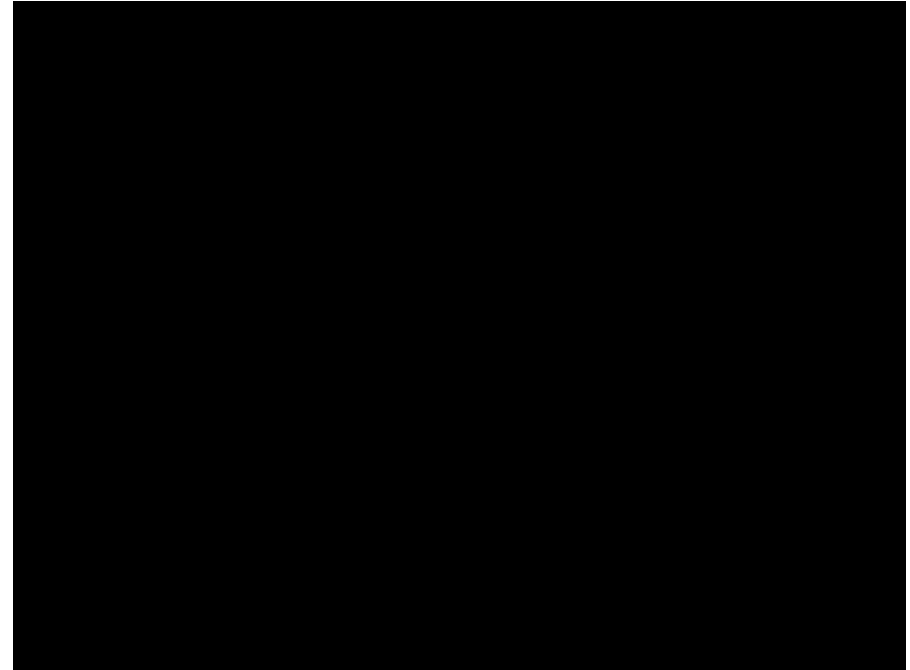
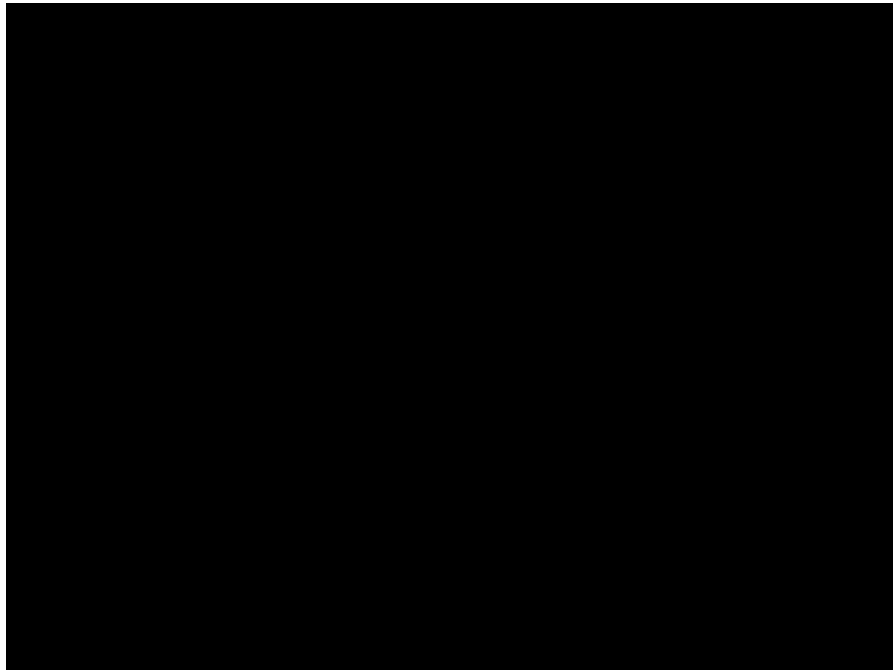
$$p_{\mathcal{F}}^t(x) = \frac{(t-t^1) \cdot p_{\mathcal{F}}^2(x) + (t^2-t) \cdot p_{\mathcal{F}}^1(x)}{t^2-t^1}$$

$p_{\mathcal{F}}^1, p_{\mathcal{F}}^2$ – карта вероятностей для прямого и обратного хода

t^1, t^2 – номера ключевых кадров

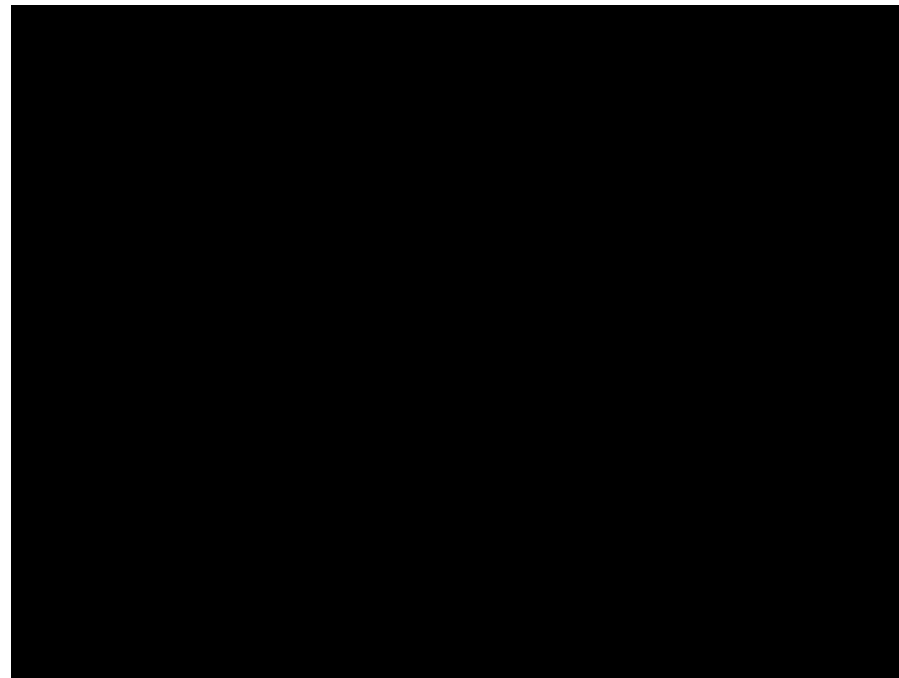
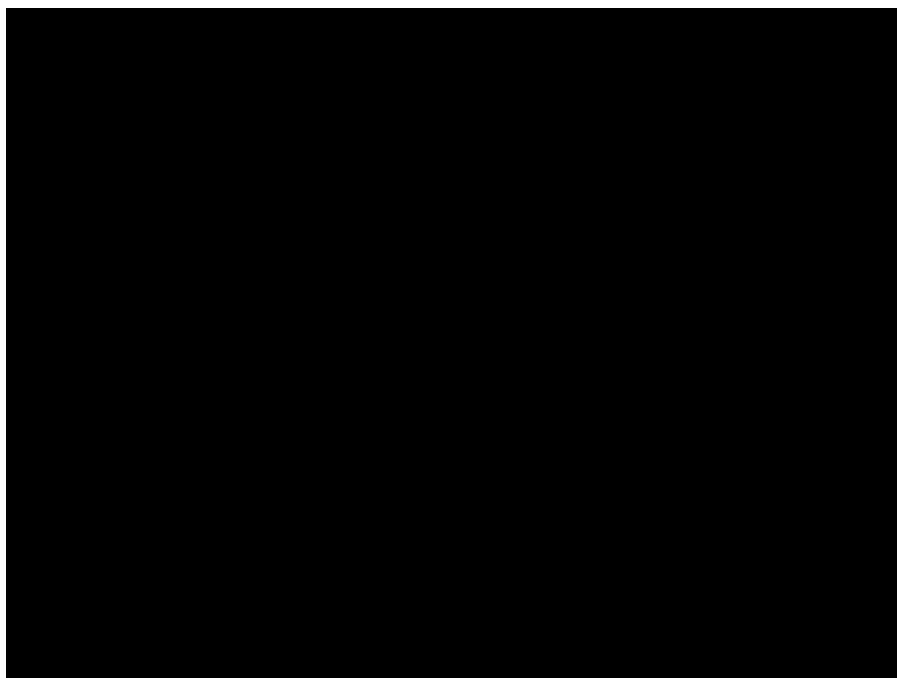
Video SnapCut

Пример работы



Video SnapCut

Пример работы (2)



Video SnapCut

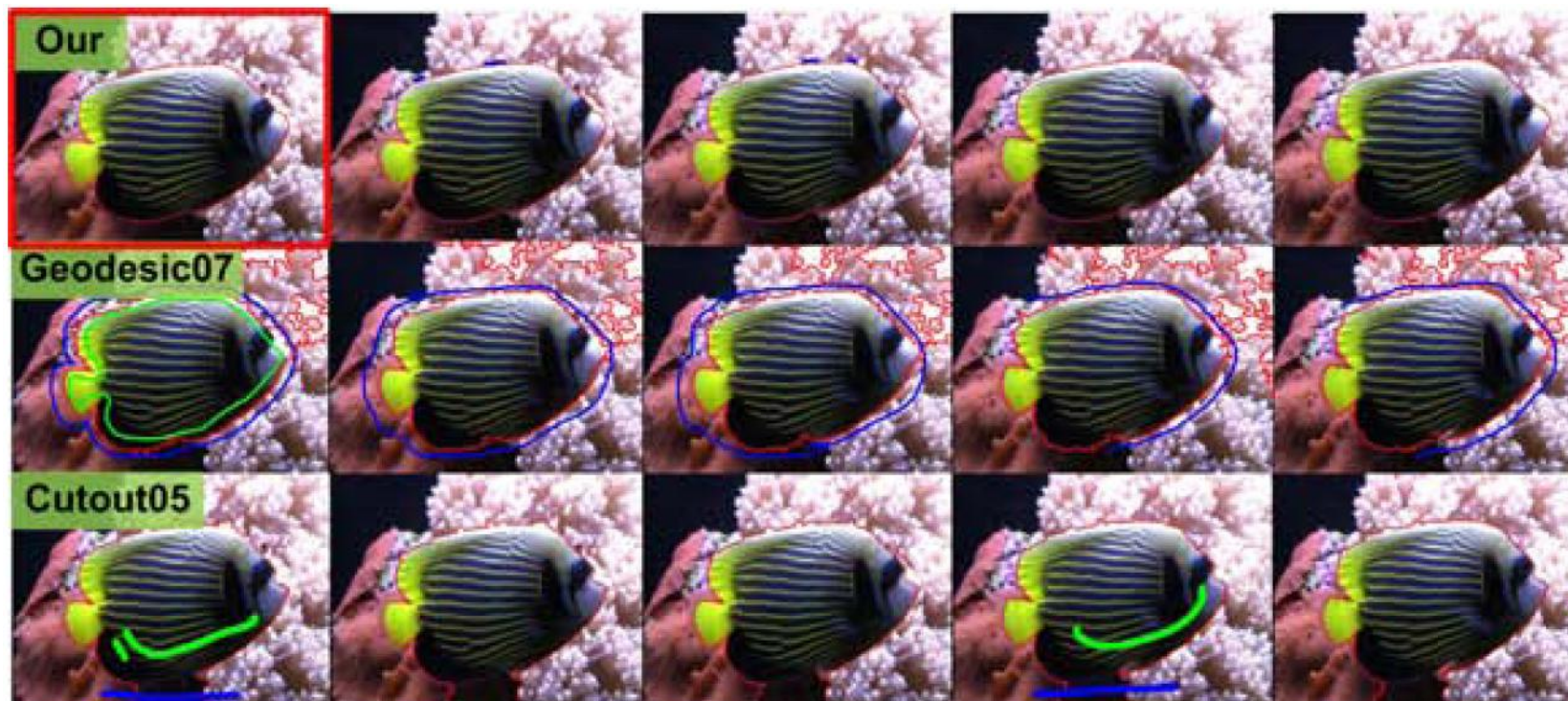
Сравнение с другими методами

Цветом выделена дополнительная информация, которую вносил пользователь для получения желаемого результата



Video SnapCut

Сравнение с другими методами (2)



Video SnapCut

Анализ метода

- **Достоинства**
 - Алгоритм включен в состав Adobe After Effects CS5
 - Локальность окон позволяет вносить небольшие правки без повторной обработки всего видео
- **Недостатки**
 - Не обрабатывает появление новых цветов на границе объекта
 - Не обрабатывает объекты неопределенной формы (огонь, дым)



Литература

1. WANG, J., AND COHEN, M. 2007. Optimized color sampling for robust matting. In Proc. of IEEE CVPR, 2007
2. LEVIN, A., LISCHINSKI, D., AND WEISS, Y. 2008. A closedform solution to natural image matting. IEEE Trans. Pattern Analysis and Machine Intelligence, 2008
3. X. Bai, J. Wang, D. Simons, and G. Sapiro. Video SnapCut: robust video object cutout using localized classifiers. ACM Transactions on Graphics, 2009