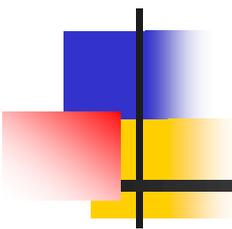
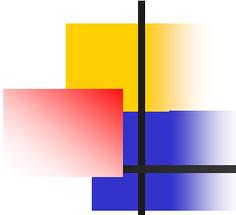


Обработка областей открытия при распространении карт глубины



Сергей Матюнин

CS MSU Graphics & Media Lab
Video group



Содержание

- **Введение**
- Совместный поиск областей открытия и OF
- Использование машинного обучения
- Выделение границ объектов
- Заключение

Распространение глубины

Исходное видео



Распространение глубины

Карта глубины



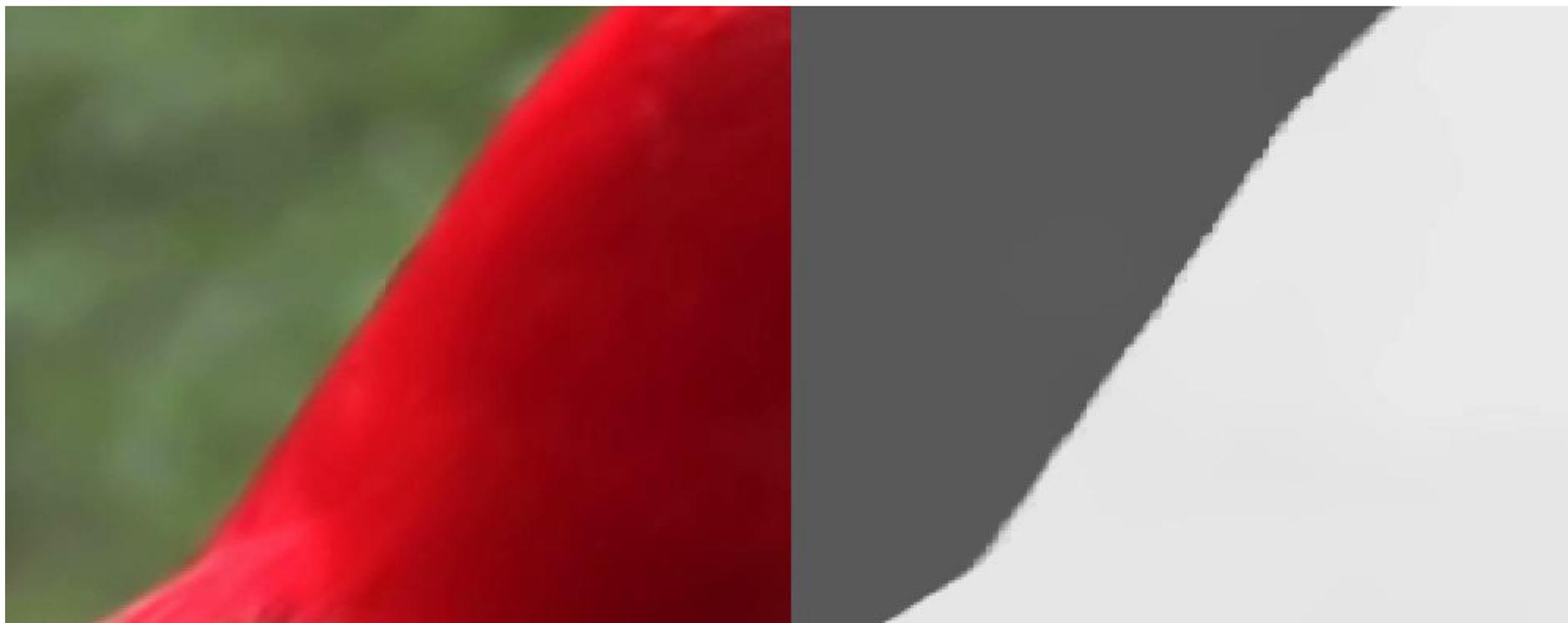
Распространение глубины

Пример работы



Распространение глубины

Основная проблема – «затекания»

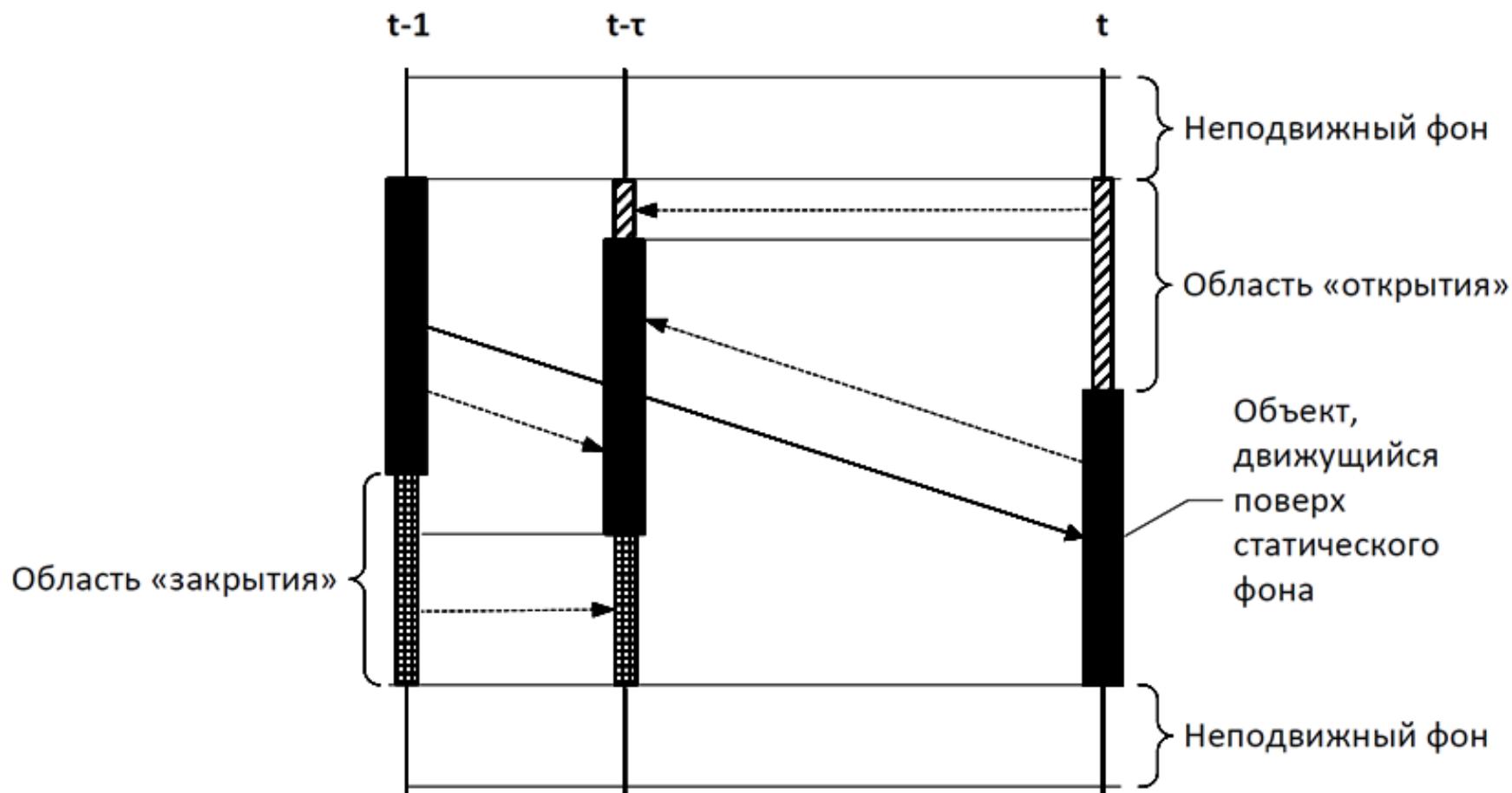


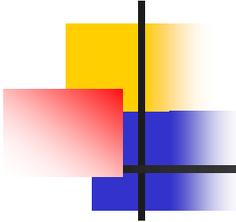
Исходное видео

Полученная глубина

Распространение глубины

Причина «затеканий»





Распространение глубины

Направления улучшения



- Двухстороннее распространение с учетом меры доверия
- Обработка областей открытия
- Отслеживание объектов

Области открытия

Опорный кадр



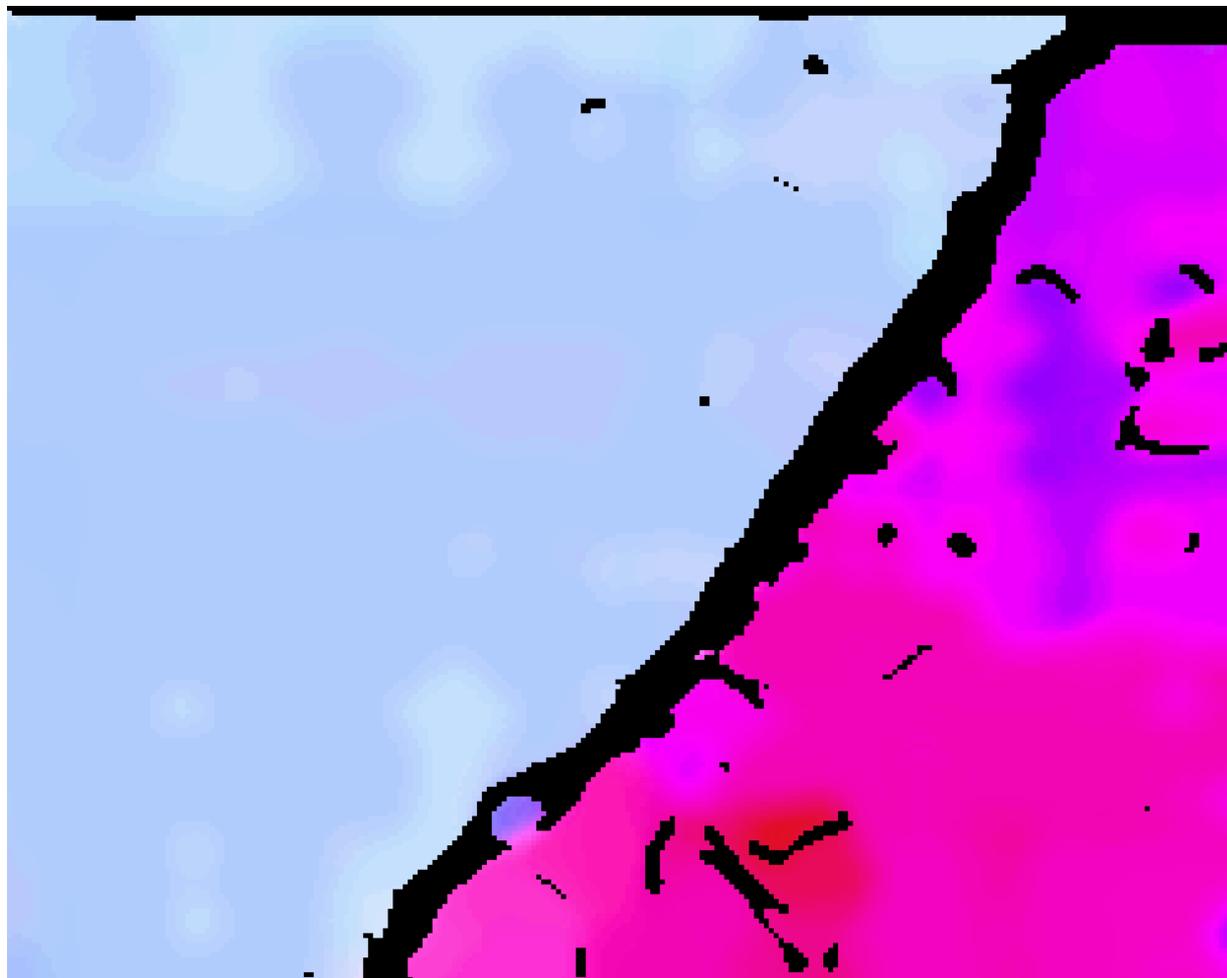
Области открытия

Текущий кадр



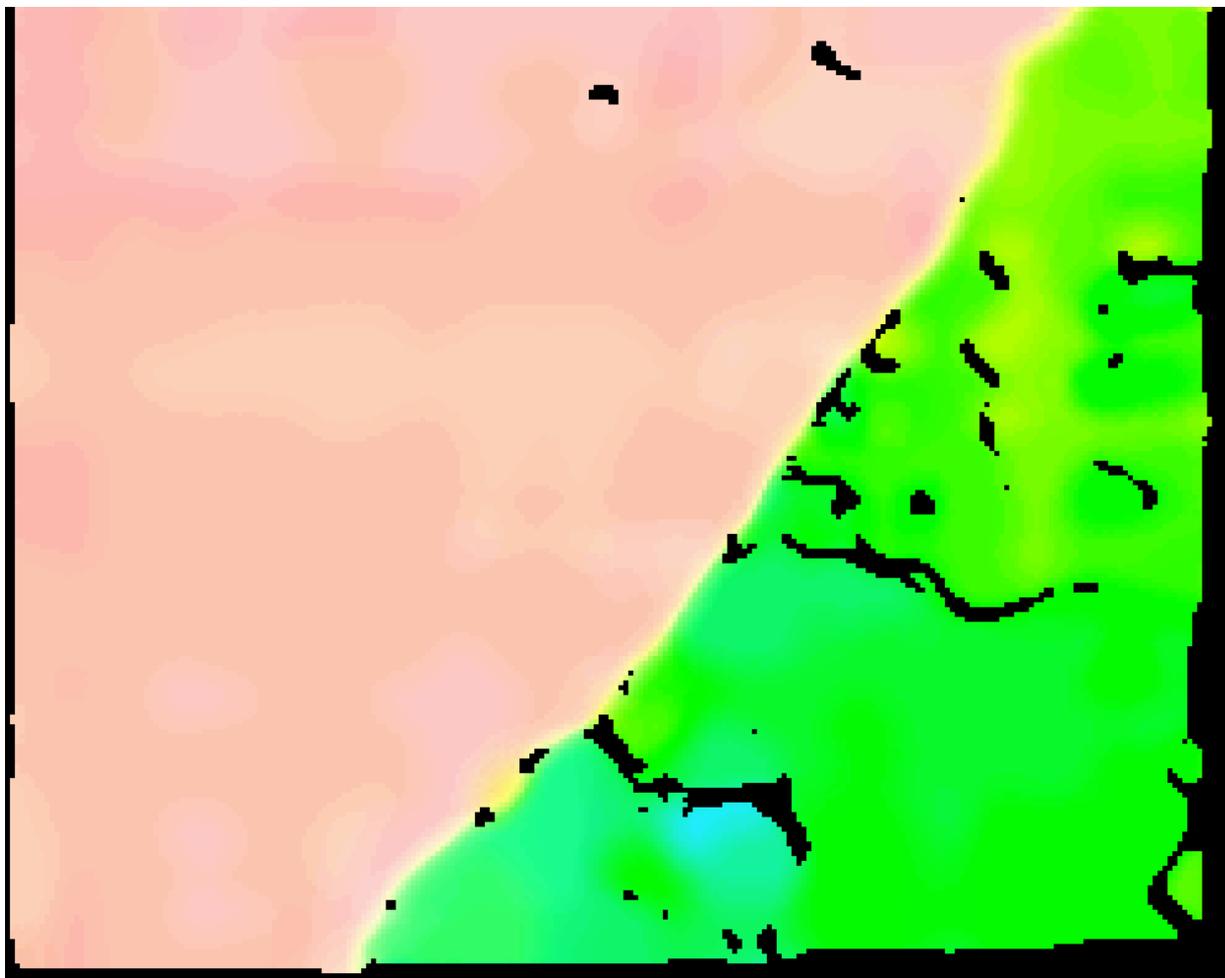
Области открытия

Поток с областями открытия



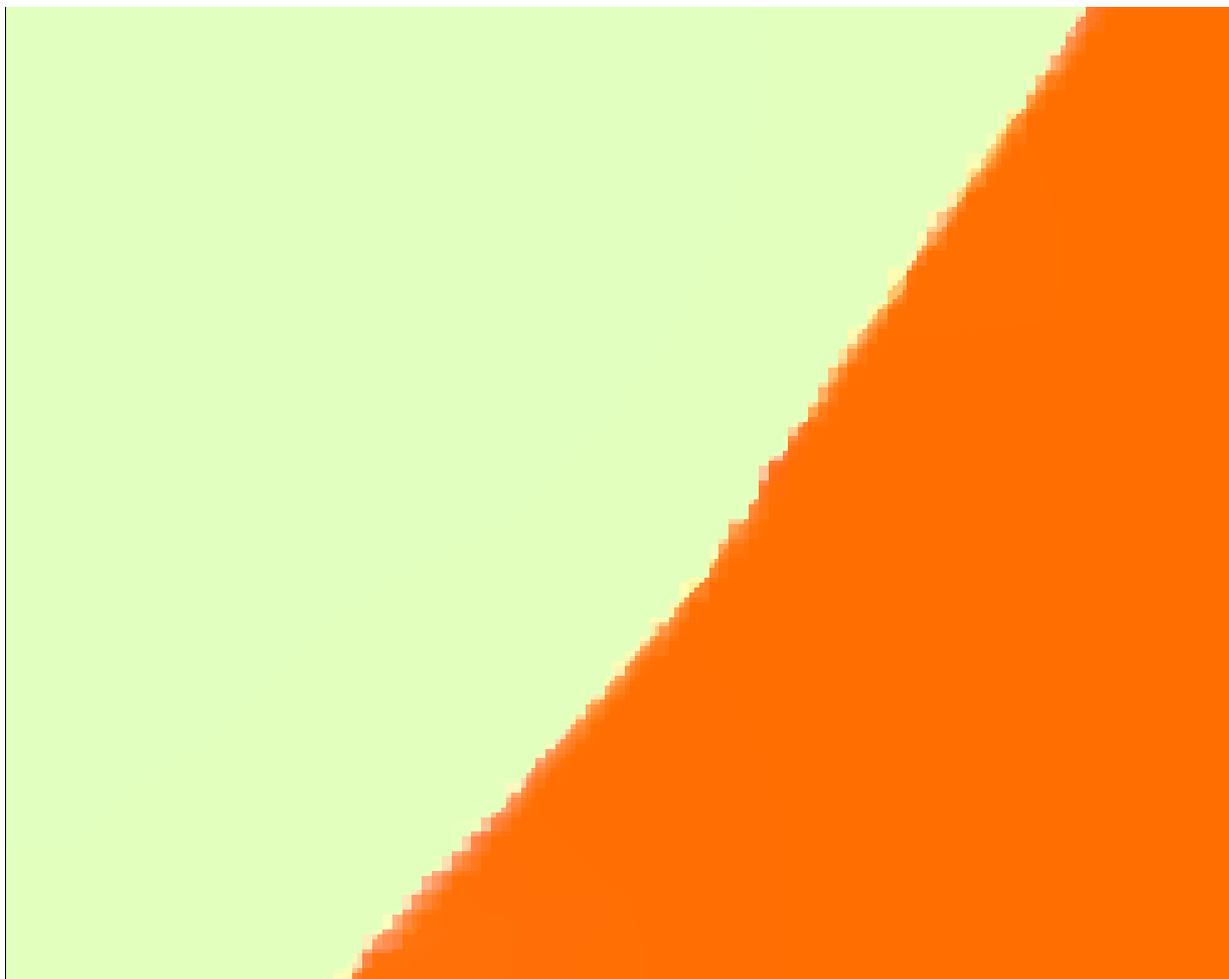
Области открытия

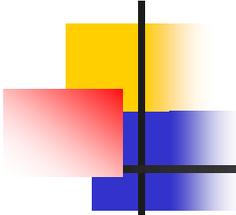
Обратный поток с областями открытия



Оптический поток

В ~ 60 раз медленнее





Содержание

- Введение
- **Совместный поиск областей открытия и OF**
- Использование машинного обучения
- Выделение границ объектов
- Заключение

Встроенное обнаружение областей открытия



- Стандартные подходы:
 - ограничение на гладкость
 - постобработка
- Предлагаемая схема:
 - устойчивый OF (пирамидальный алгоритм)
 - поиск областей открытия (закрытия)
 - улучшение границ OF билатеральным сглаживанием с учетом областей открытия

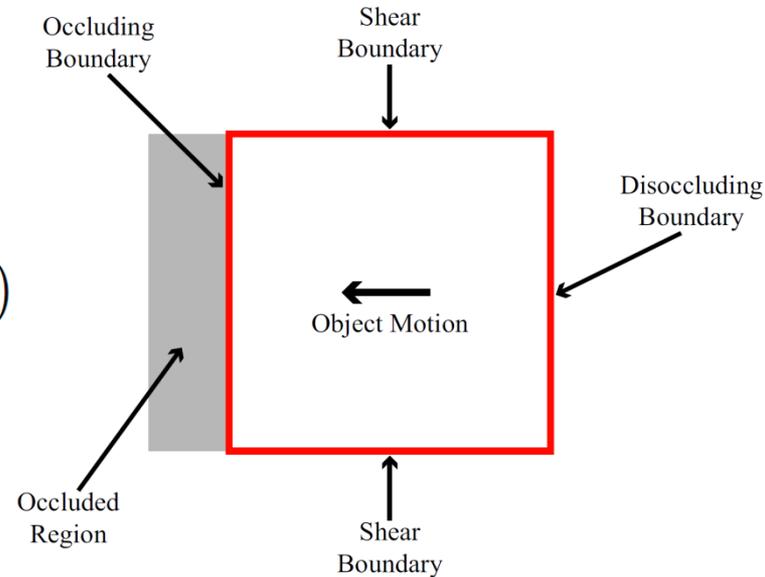
Поиск областей открытия (1/2)

Дивергенция потока:

$$\text{div}(x, y, t) = \frac{\partial}{\partial x} u(x, y, t) + \frac{\partial}{\partial y} v(x, y, t)$$

Выбираем области закрытия:

$$d(x, y, t) = \begin{cases} \text{div}(x, y, t) & \text{div}(x, y, t) < 0 \\ 0 & \text{otherwise.} \end{cases}$$



Поиск областей открытия (2/2)



Межкадровая разность:

$$e(x, y, t) = I(x, y, t) - I(x + u(x, y, t), y + v(x, y, t), t + 1)$$

Объединение критериев:

$$r(x, y, t) = N(d(x, y, t); \sigma_d) \cdot N(e(x, y, t); \sigma_e)$$

N – функция плотности нормального
распределения

Из экспериментов: $\sigma_d = 0.3$

$$\sigma_e = 20$$

Модификация оценки OF

$$E_{Flow}(u, v, t) = E_{FlowData}(u, v, t) + E_{FlowSmooth}(u, v, t)$$

$$E_{FlowData}(u, v, t) = \sum_{x,y,k} r(x, y, t) \Psi([I^{[k]}(x + u(x, y, t), y + v(x, y, t), t + 1) - I^{[k]}(x, y, t)]^2)$$

$$\Psi(s^2) = \sqrt{s^2 + \varepsilon^2}; \quad \varepsilon = 0.001$$

$$E_{FlowSmooth}(u, v, t) =$$

$$\sum_{x,y} (\alpha_g + \alpha_l \cdot b(x, y, t)) \cdot \Psi(u_x(x, y, t)^2 + u_y(x, y, t)^2 + v_x(x, y, t)^2 + v_y(x, y, t)^2)$$

$$b(x, y, t) = N\left(\sqrt{\frac{\partial}{\partial x} I(x, y, t)^2 + \frac{\partial}{\partial y} I(x, y, t)^2}; \sigma_b\right)$$

Базовый метод: T. Brox et al., "High accuracy optical flow estimation based on a theory for warping," ECCV, 2004.

Фильтрация OF

$$u'(x, y, t) = \frac{\sum_{x_1, y_1} u(x_1, y_1, t) w(x, y, x_1, y_1, t)}{\sum_{x_1, y_1} w(x, y, x_1, y_1, t)}$$

$$w(x, y, x_1, y_1, t) = N(\sqrt{(x - x_1)^2 + (y - y_1)^2}; \sigma_x) \\ \cdot N(I(x, y, t) - I(x_1, y_1, t); \sigma_i) \\ \cdot N(\sqrt{(u - u_1)^2 + (v - v_1)^2}; \sigma_m) \\ \cdot r(x_1, y_1, t)$$

$$\sigma_x = 4, \sigma_i = 7.5, \sigma_m = 0.5$$

Радиус фильтрации = 10

Фильтрация OF

Ускорение

- Модуль градиента потока:

$$g(x, y, t) = \sqrt{u_x^2(x, y, t) + u_y^2(x, y, t) + v_x^2(x, y, t) + v_y^2(x, y, t)}$$

- Сглаживание по Гауссу, $\sigma_g = 3$
- Сглаживание потока применяется там, где усредненное значение $\hat{g}(x, y, t) > 0.25$

Результаты

Кадр 1



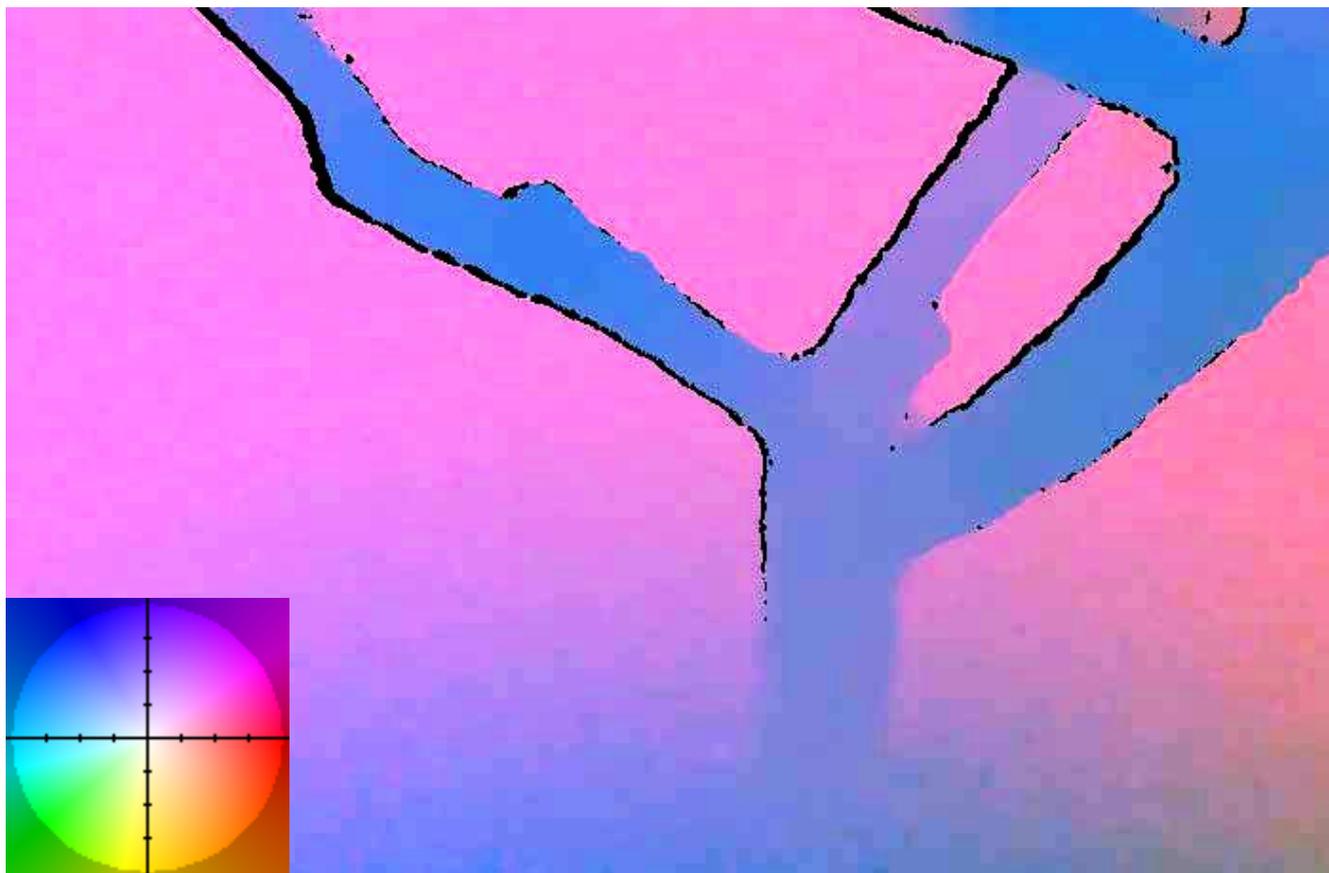
Результаты

Кадр 2

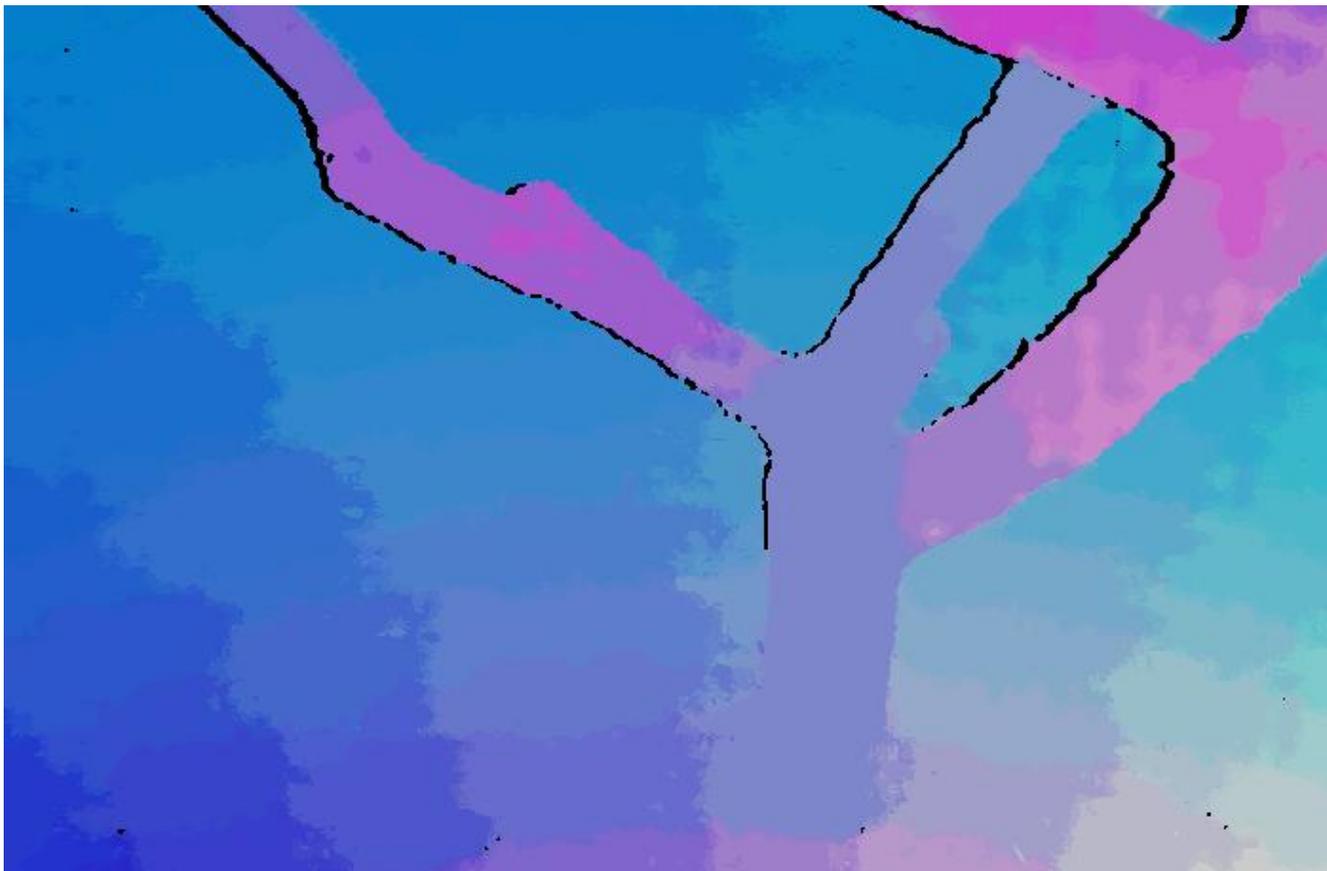


Результаты

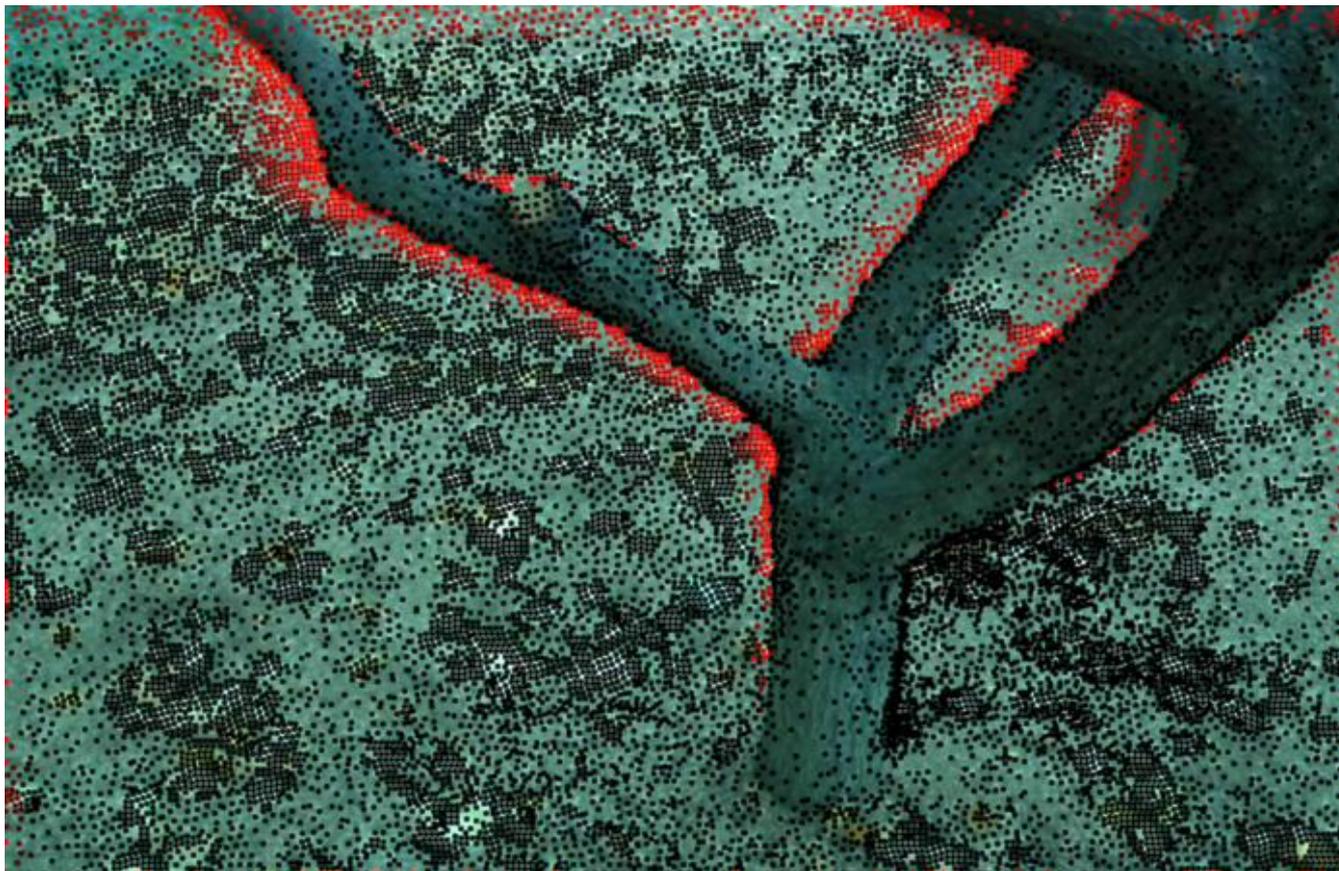
Из статьи (40,53 с)

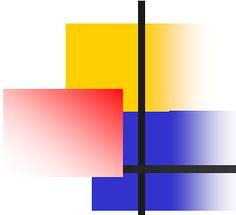


Результаты Наши (13,74 с)



Тема статьи – построение траекторий частиц





Содержание

- Введение
- Совместный поиск областей открытия и OF
- **Использование машинного обучения**
- Выделение границ объектов
- Заключение

Классификатор для областей открытия

- Задача классификации

$$\mathcal{O} := \left\{ (f_i \rightarrow L_i) \mid f_i \in \mathbb{R}^d, L_i \in [0, 1] \right\}_{i=1}^t$$

$L = 1$ для областей открытия

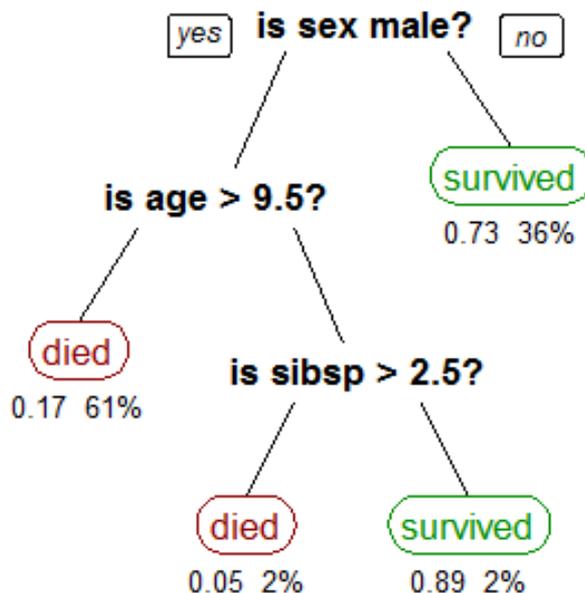
d – длина вектора признаков

t – размер обучающей базы

- Random forest (решающий лес)

Решающий лес

Набор решающих деревьев



Итоговое решение – результат голосования

Признак по границам объектов



Области открытия расположены вблизи границ объектов

- Canny edge detector
- $f_{ED}(\mathbf{x}, z) = \text{distTrans}(\|\nabla I_1(z)\| > \tau_{ED})$
 \mathbf{x} – текущий пиксель
 z – уровень разрешения изображения

Признак по постоянству цвета



Межкадровая разность

по скомпенсированным кадрам:

$$f_{PC,a}(\mathbf{x}, z) = |I_1(\mathbf{x}, z) - \text{bicubic}(I_2(\mathbf{x} + u_a(\mathbf{x}, z), z))|$$

bicubic – интерполяция для дробных сдвигов

За границами кадра – большие значения

Области открытия

Постоянство текстуры (1/3)

$$f_{ST,a}^n(\mathbf{x}) = \frac{1}{D} \sum_{d=1}^D \left(\frac{\mu_n(\mathbf{T}_{1,d}(\mathbf{x})) - \text{bicubic}(\mu_n(\gamma_d^a(\mathbf{x})))}{\sigma_n(\mathbf{T}_{1,d}(\mathbf{x})) - \text{bicubic}(\sigma_n(\gamma_d^a(\mathbf{x})))} \right)^2$$

$$D = 5$$

$$\gamma_d^a(\mathbf{x}) \equiv \mathbf{T}_{2,d}(\mathbf{x} + u_a(\mathbf{x}))$$

$\mathbf{T}_{1,d}(\mathbf{x})$ и $\mathbf{T}_{2,d}(\mathbf{x})$ – результат действия текстурных фильтров

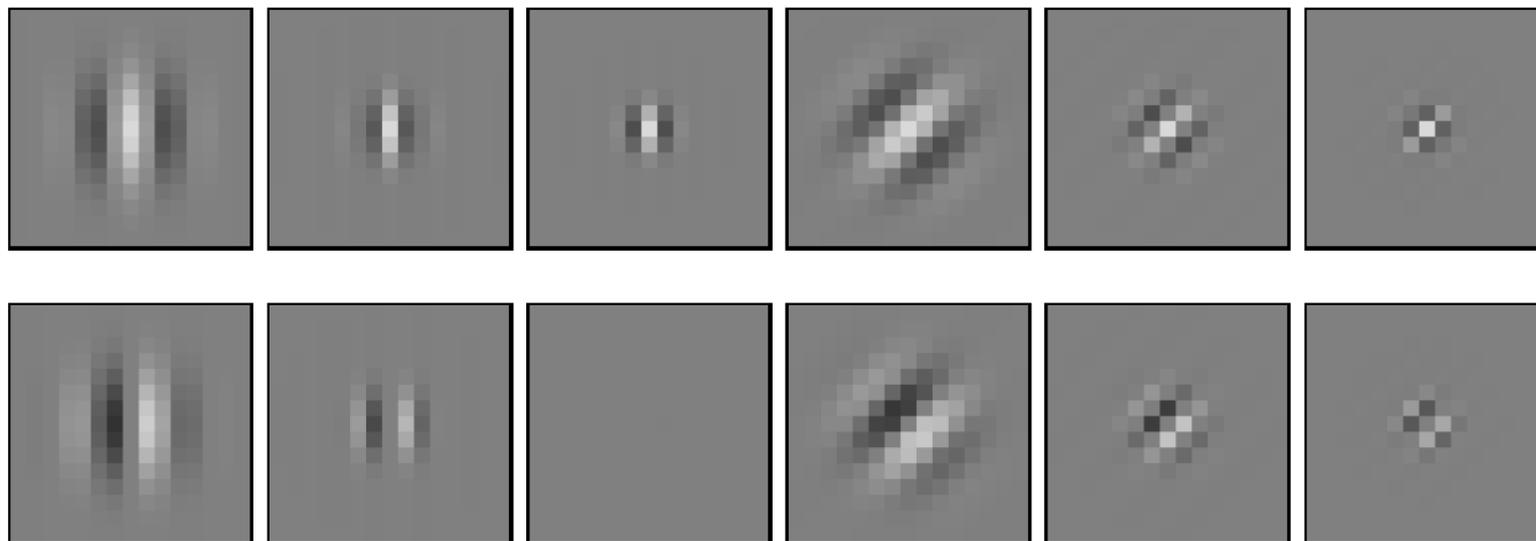
$\mu_n(\cdot)$ – среднее, $\sigma_n(\cdot)$ – дисперсия
в окрестности $n \times n$

Области открытия

Постоянство текстуры (2/3)

$$f_{ST,a}^n(\mathbf{x}) = \frac{1}{D} \sum_{d=1}^D \left(\frac{\mu_n(\mathbf{T}_{1,d}(\mathbf{x})) - \text{bicubic}(\mu_n(\gamma_d^a(\mathbf{x})))}{\sigma_n(\mathbf{T}_{1,d}(\mathbf{x})) - \text{bicubic}(\sigma_n(\gamma_d^a(\mathbf{x})))} \right)^2$$

Пример текстурных фильтров:



Области открытия

Постоянство текстуры (3/3)

Расстояние Махаланобиса:

$$f_{\text{STM},a}(\mathbf{x}) = \sqrt{\sum_{d=1}^D [\mathbf{T}_{1,d}(\mathbf{x}) - \text{bicubic}(\gamma_d^a(\mathbf{x}))]^2 / \sigma_d^2}$$

σ_d^2 – дисперсия по $\mathbf{T}_{1,d}$ и $\mathbf{T}_{2,d}$

Признаки по потоку

Направление векторов

- $f_{\text{TG},x}(\mathbf{x}, z) = \|\nabla \bar{u}_x\|$, $f_{\text{TG},y}(\mathbf{x}, z) = \|\nabla \bar{u}_y\|$

\bar{u}_x и \bar{u}_y – медианы компонент потока по набору алгоритмов OF (4 алгоритма)

- Дисперсия направления потока в окне $n \times n$

$$f_{\text{AV},a}^n(\mathbf{x}, z) = \mathbf{E} \left[(\theta_a(\mathbf{x}_i, z) - \mathbf{E}[\theta_a(\mathbf{x}, z)])^2 \right]$$

$$\theta_a(\mathbf{x}, z) = \arctan [u_{x,a}(\mathbf{x}, z) / u_{y,a}(\mathbf{x}, z)]$$

Признаки по потоку

Длина векторов

$$f_{LV,a}^n(\mathbf{x}, z) = \mathbf{E} \left[(M_a(\mathbf{x}_i, z) - \mathbf{E} [M_a(\mathbf{x}, z)])^2 \right]$$

$$M_a(\mathbf{x}, z) = \|u_a(\mathbf{x}, z)\|$$

Среднее оценивается в окне $n \times n$

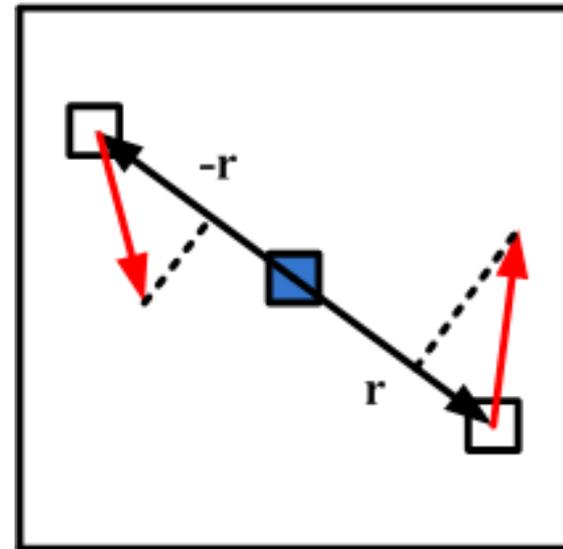
Признаки на основе дисперсии длины и направления векторов считаются также по каждому из алгоритмов

Относительное движение

$$f_{CS,a}^n(\mathbf{x}, z) = \max \{ \Psi_a(\mathbf{x}, \mathbf{r}_i, z) : \mathbf{r}_i \in \mathcal{N}_{CS} \}$$

$$\Psi_a(\mathbf{x}, \mathbf{r}, z) = \frac{2\|\mathbf{r}\|}{\text{proj}_{\vec{v}}(u_a(\mathbf{x}+\mathbf{r}, z)) + \text{proj}_{\vec{v}}(u_a(\mathbf{x}-\mathbf{r}, z))}$$

Время «столкновения»
противоположных пикселей



Другие признаки

- Медиана потока по суперпикселям (сегментация по яркости)

- LRC (Left-Right Consistency)

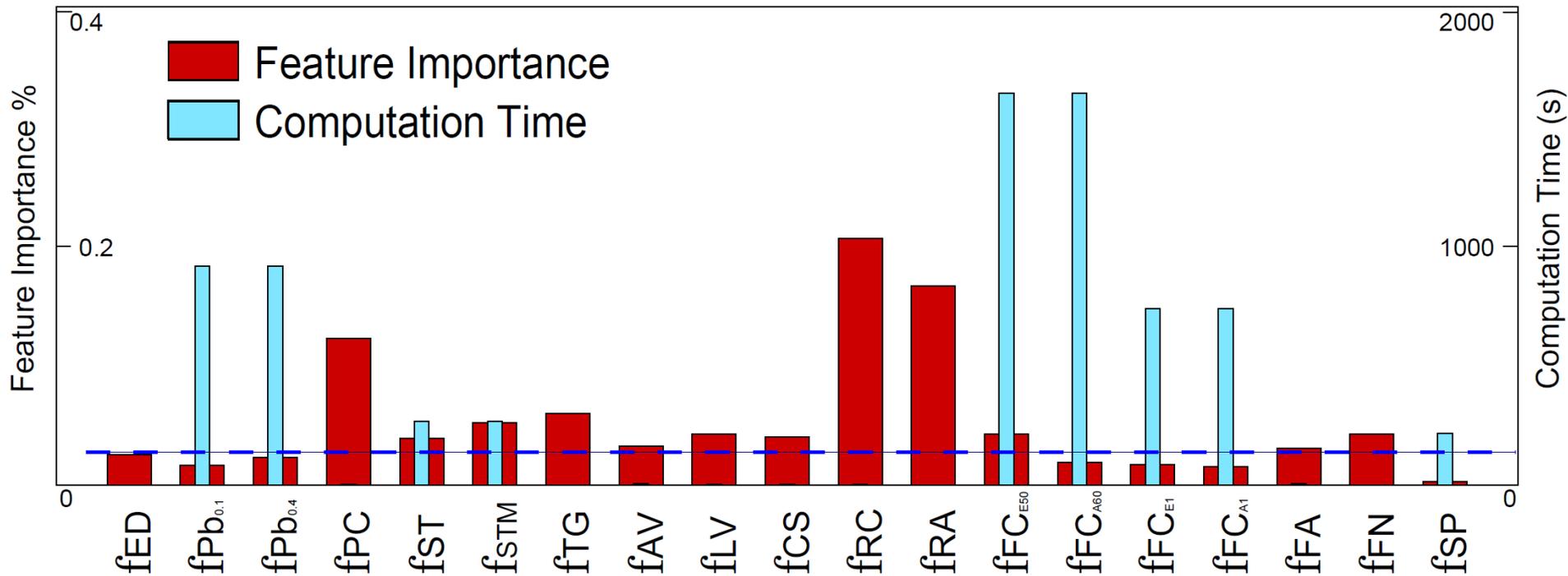
$$f_{RC,a}(\mathbf{x}, z) = \|\mathbf{x} - (\mathbf{x}'_a + u'_a(\mathbf{x}'_a, z))\|$$

$$\mathbf{x}'_a = \text{round}(\mathbf{x} + u_a(\mathbf{x}, z))$$

$u'_a(\mathbf{x}'_a, z)$ – поток в обратном направлении

- LRC по углу

$$f_{RA,a}(\mathbf{x}, z) = |\pi - \arccos[u_a(\mathbf{x}, z) \cdot u'_a(\mathbf{x}'_a, z)]|$$



ED – близость к границам

Pb – близость к границам (другие детекторы границ)

PC – межкадровая разность

ST – пространственные текстуры

STM – пространственные текстуры (Махаланобис)

TG – градиенты медианы потока по алгоритмам

AV – дисперсия угла

LV – дисперсия длины векторов

SC – скорость «сближения»

RC – LRC

RA – LRC по углу

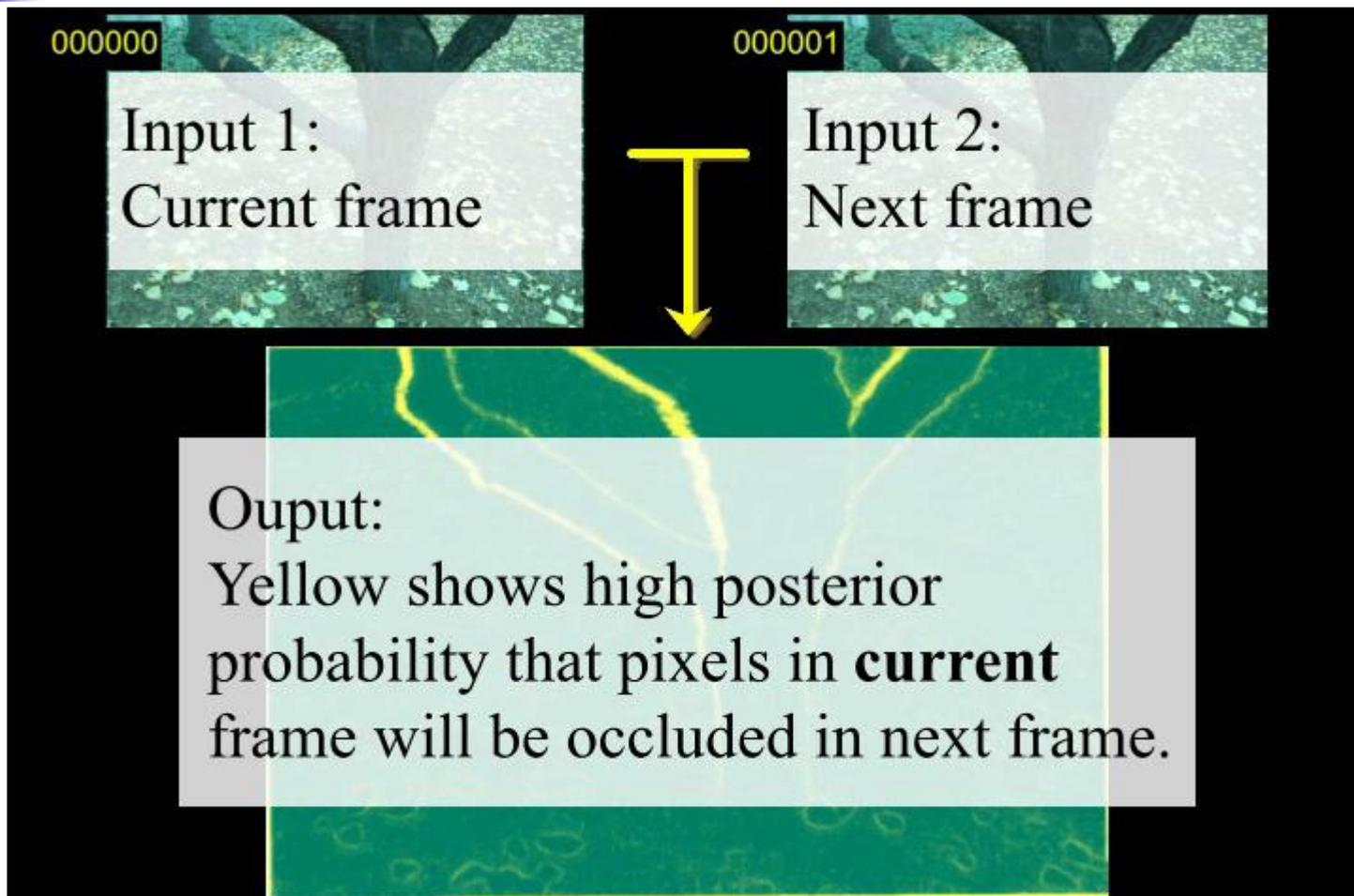
FC_i – набор метрик оптического потока

FA – дисперсия угла по алгоритмам

FN – дисперсия длины векторов по алгоритмам

SP – градиенты потока между суперпикселями

Результаты (1/2)

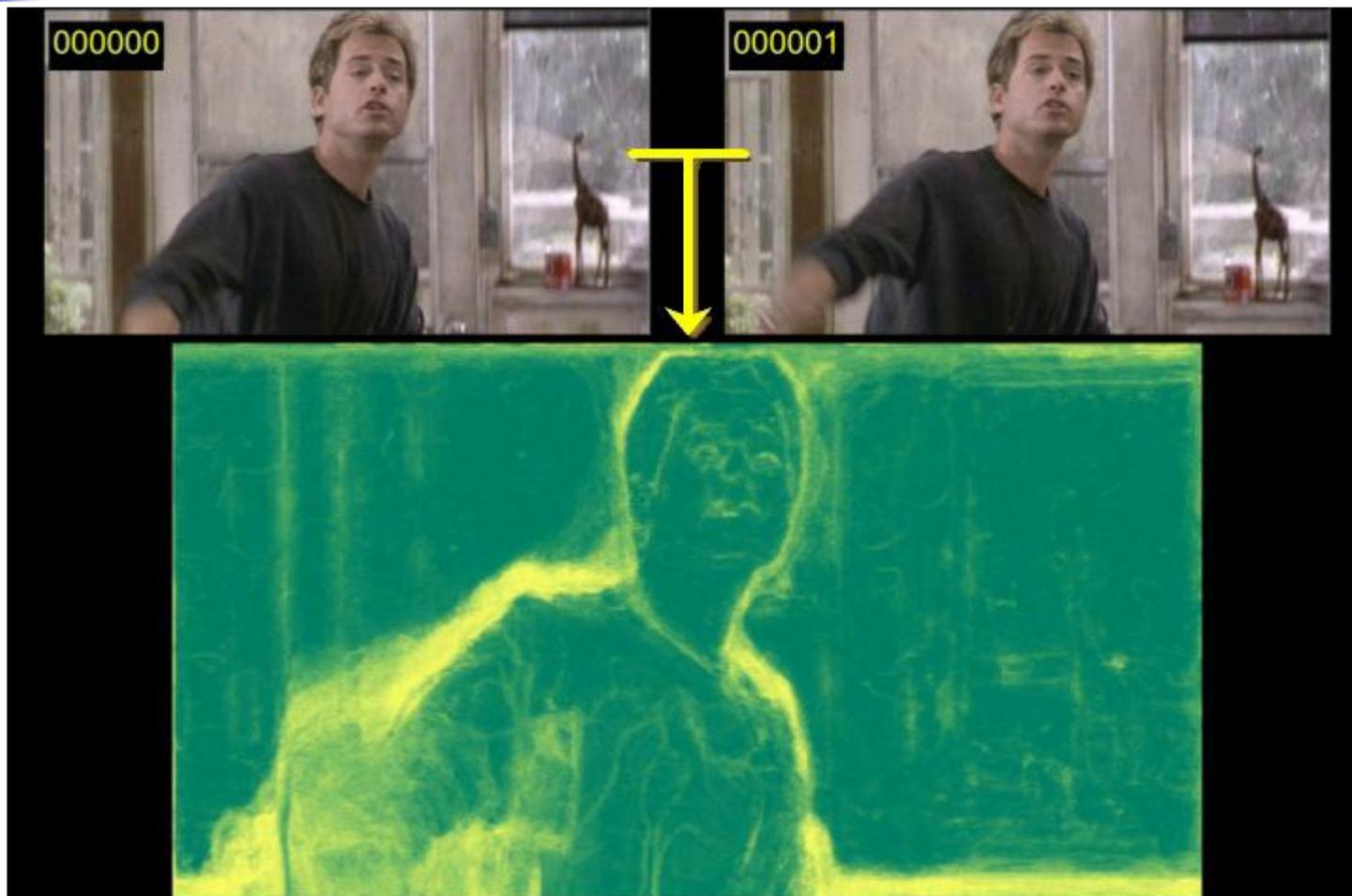


Результаты (1/2)



Current frame

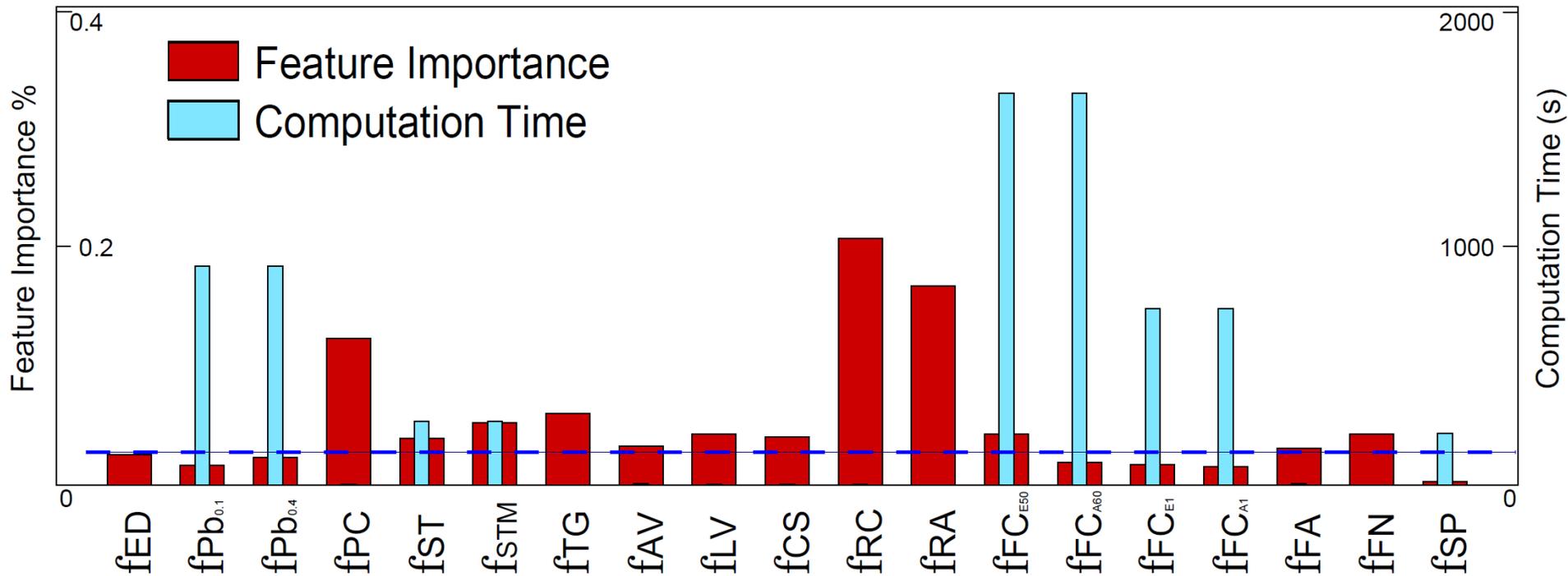
Результаты (2/2)



Результаты (2/2)



Current frame



ED – близость к границам

Pb – близость к границам (другие детекторы границ)

PC – межкадровая разность

ST – пространственные текстуры

STM – пространственные текстуры (Махалонобис)

TG – градиенты медианы потока по алгоритмам

AV – дисперсия угла

LV – дисперсия длины векторов

SC – скорость «сближения»

RC – LRC

RA – LRC по углу

FC_i – набор метрик оптического потока

FA – дисперсия угла по алгоритмам

FN – дисперсия длины векторов по алгоритмам

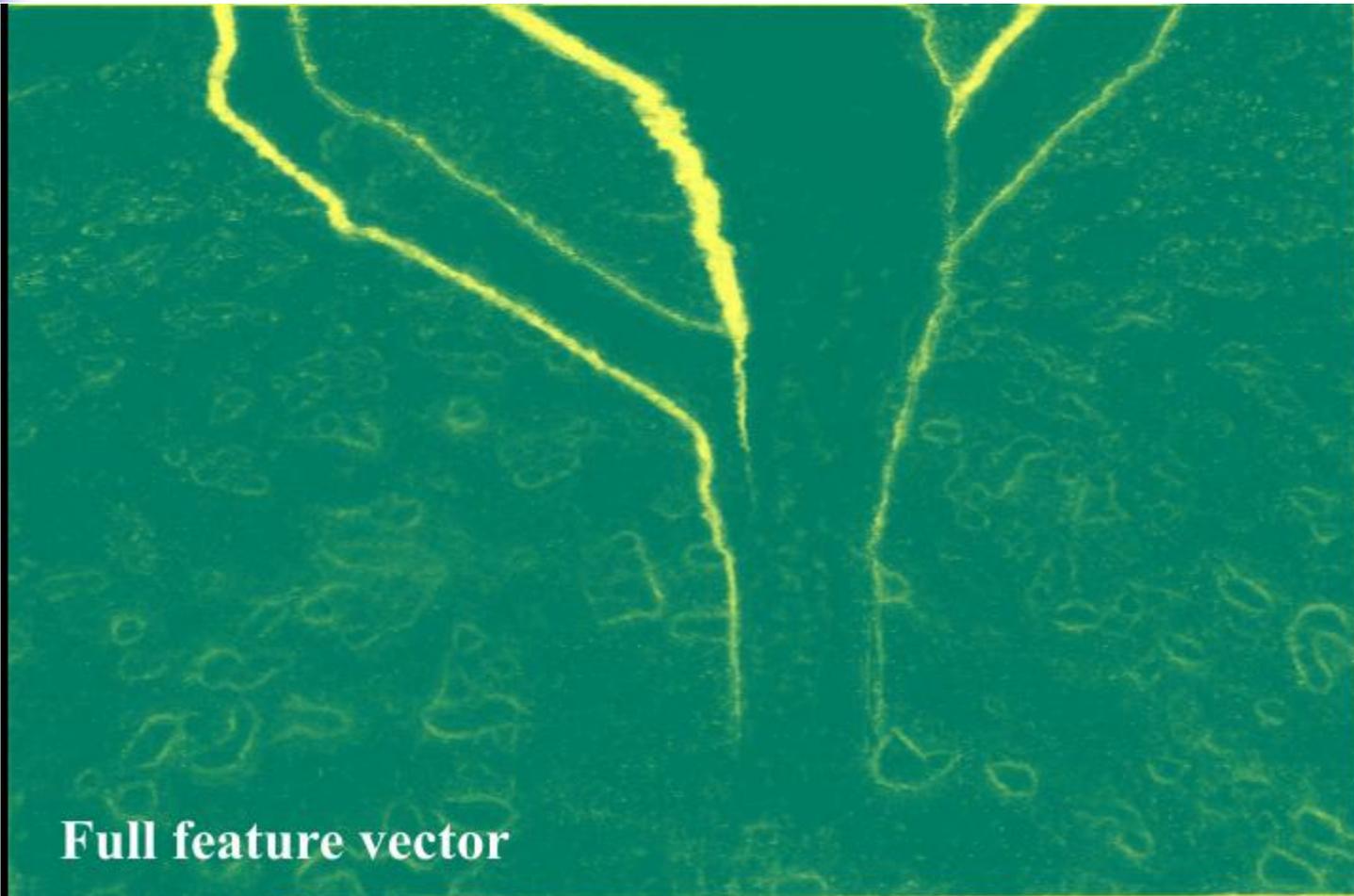
SP – градиенты потока между суперпикселями

Сокращенный набор признаков

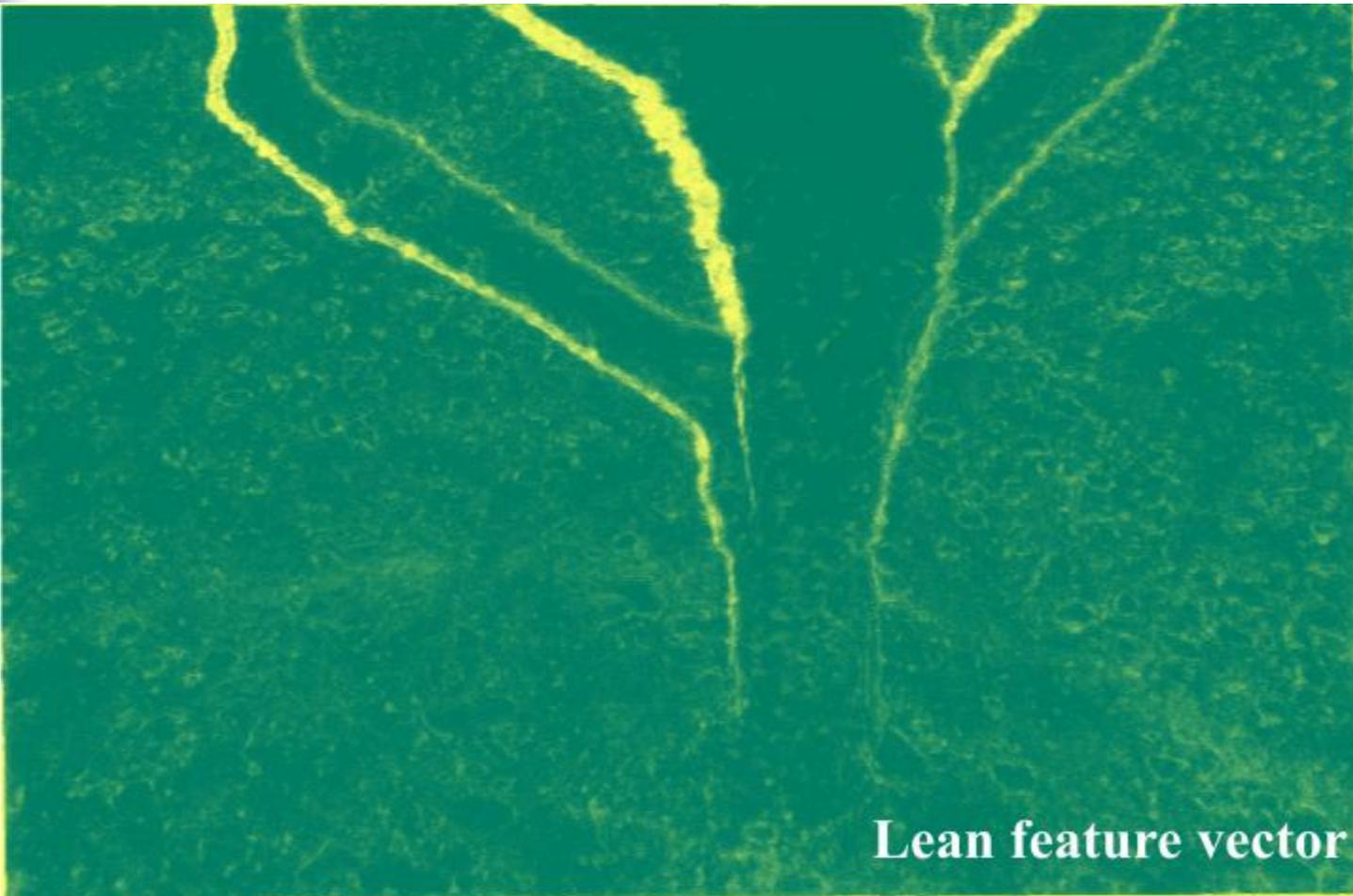


- Только «быстрые» признаки (<2 минут)
- Время обучения: 54 → 24 минуты (для 13-ти последовательностей)
- Обработка: 124 → 1.3 минуты

Результаты сокращения набора признаков (1/2)



Результаты сокращения набора признаков (1/2)



Результаты сокращения набора признаков (2/2)

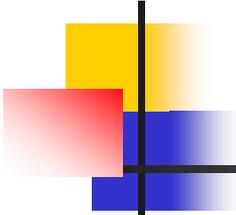


Full feature vector

Результаты сокращения набора признаков (2/2)



Lean feature vector



Содержание

- Введение
- Совместный поиск областей открытия и OF
- Использование машинного обучения
- **Выделение границ объектов**
- Заключение

Поиск границ объектов

Исходное изображение



Поиск границ объектов

Желаемая карта границ



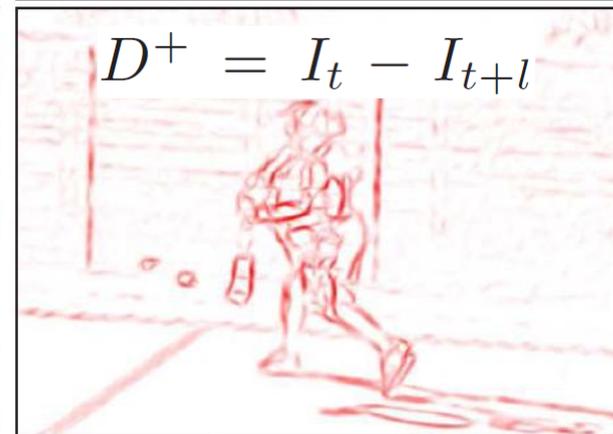
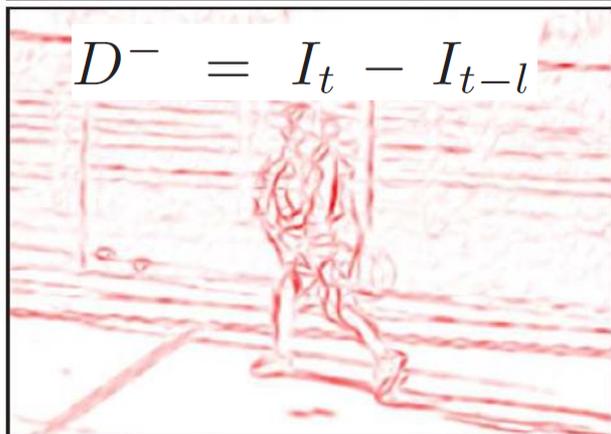
Поиск границ объектов

Разность с соседними кадрами

t-1

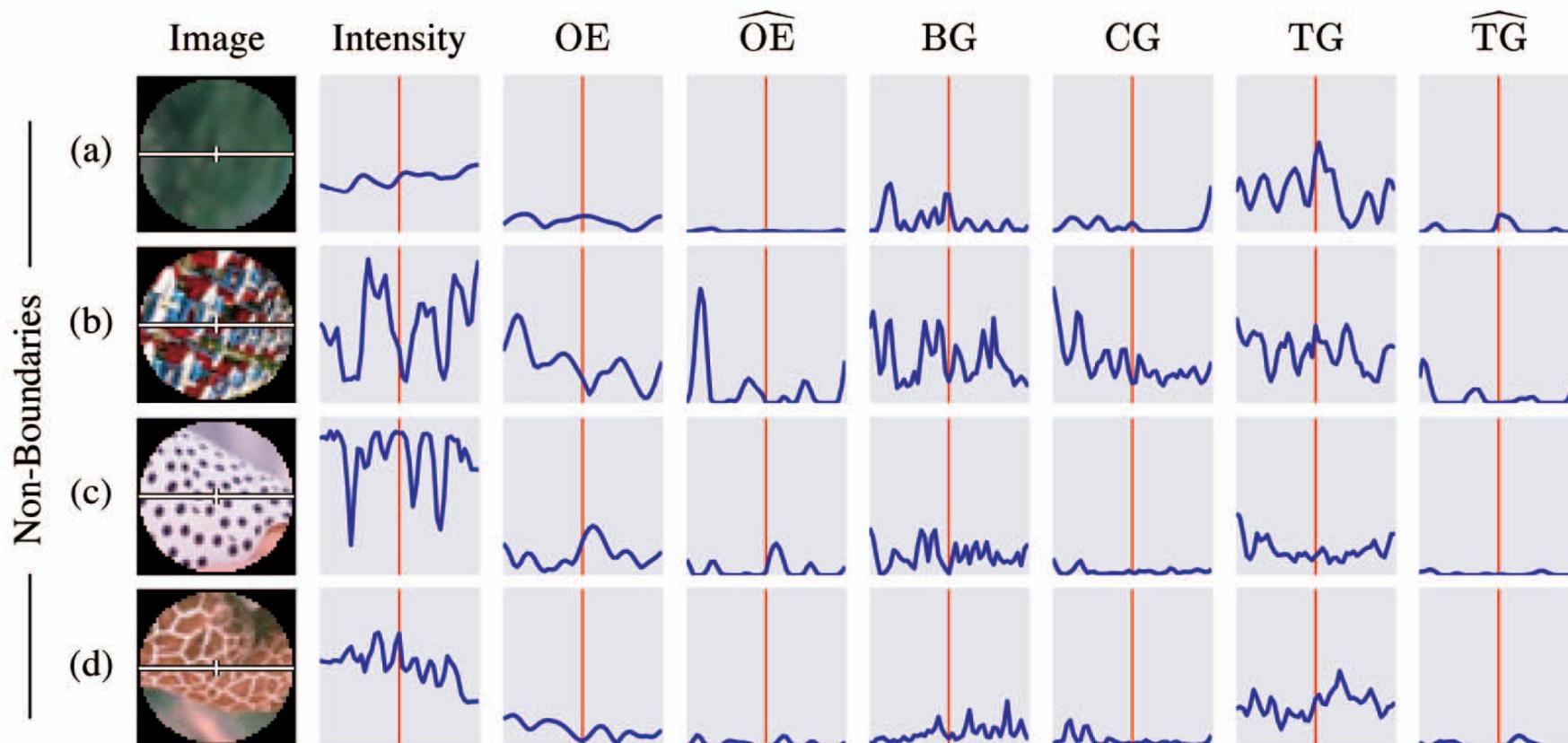
t

t+1



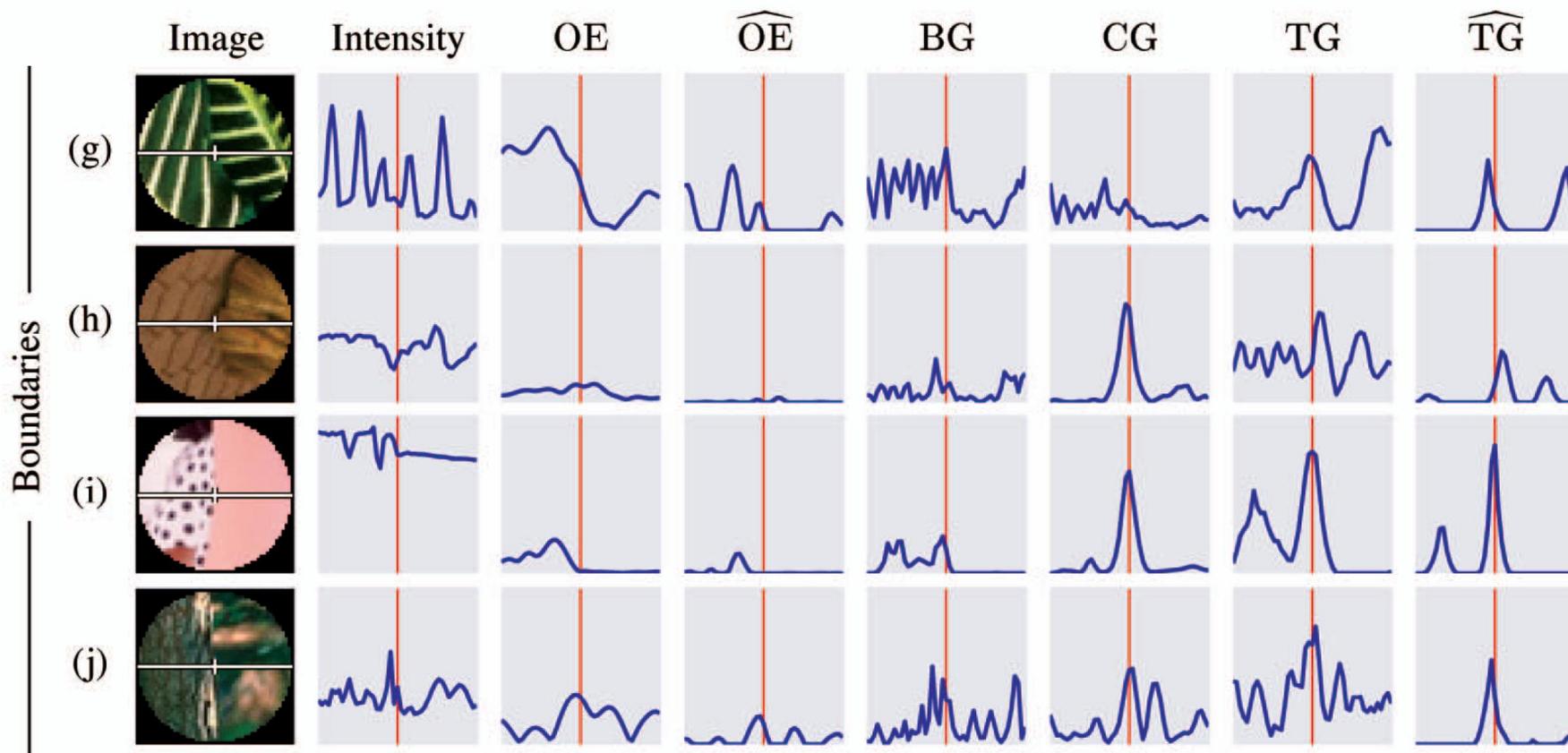
Поиск границ объектов

Градиент по текстуре (1/2)



Поиск границ объектов

Градиент по текстуре (2/2)



Поиск границ объектов

Градиент по движению

$$G_r(x, y, \theta) = \chi^2(g, h) = \frac{1}{2} \sum \frac{(g_i - h_i)^2}{g_i + h_i}$$

MG^- и MG^+ – применение $G_r(x, y, \theta)$
к D^- и D^+

Удаляем повторные границы:

$$MG(x, y, \theta) = \sqrt{MG^-(x, y, \theta) \cdot MG^+(x, y, \theta)}$$

Используем как дополнительную
информацию $MG(x, y) = \max_{\theta} \{MG(x, y, \theta)\}$

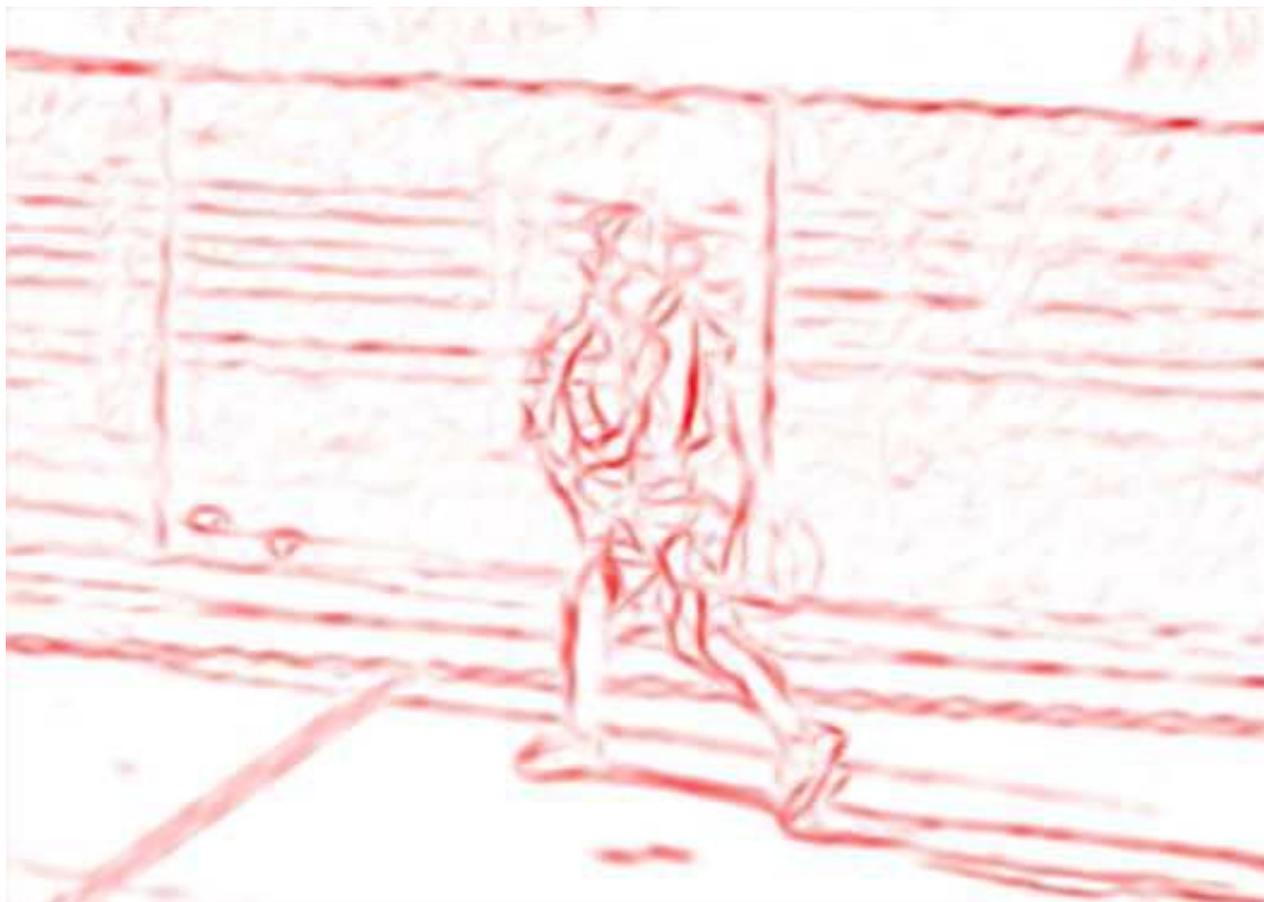
Градиент по движению

Исходный кадр



Градиент по движению

Межкадровая разность



Градиент по движению

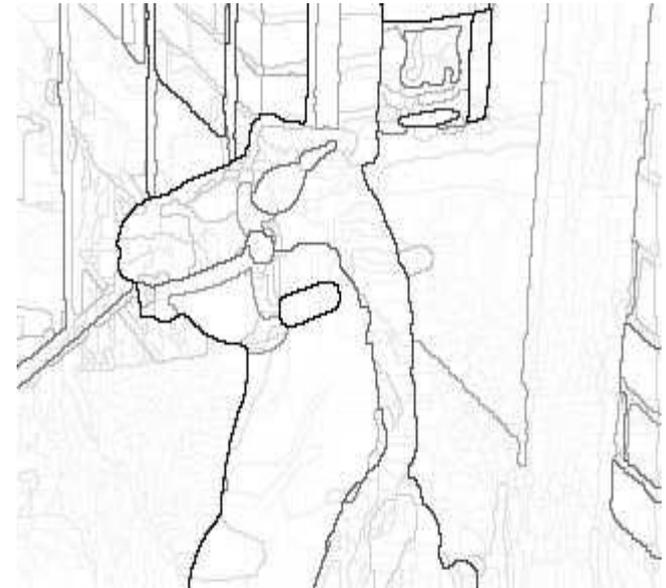
Результат



Поиск границ объектов

Ключевые компоненты

- Предположения о границах
 - пространственные
 - цвет
 - яркость
 - текстура
 - из движения
- Данные оптического потока
 - сравнение движения по разные стороны от границ



Взвешенная карта границ

Использование оптического потока (1/2)

$x_i = (x_i, y_i)$ – точка границы

R – соседняя область

$$w_i^R(x, y) = \frac{1}{Z} \exp \left[-\frac{(x-x_i)^2 + (y-y_i)^2}{2\sigma_w^2} \right] \delta(r(x, y), R) i_d(x, y)$$

$r(x, y)$ – область текущей точки

$i_d(x, y) = 1$, если текущая точка внутренняя
(вблизи нет границ), иначе – 0

Использование оптического потока (2/2)



Наилучшее приближение потока:

$$e_u^R = \sum_{(x,y)} w_i^R(x,y) (A_u^R x + B_u^R y + C_u^R - u(x,y))^2$$

$$e_v^R = \sum_{(x,y)} w_i^R(x,y) (A_v^R x + B_v^R y + C_v^R - v(x,y))^2$$

Поток по разные стороны от границы:

$$\mathbf{w}_i^+ := (u_i^+, v_i^+) \quad \mathbf{w}_i^- := (u_i^-, v_i^-)$$

Мера границы по оптическому потоку:

$$\delta = |(u_i^+ - u_i^-, v_i^+ - v_i^-)|$$

Поиск границ объектов

Итоговая формула

$$f = \rho * \delta + (1 - \rho) * gPb$$

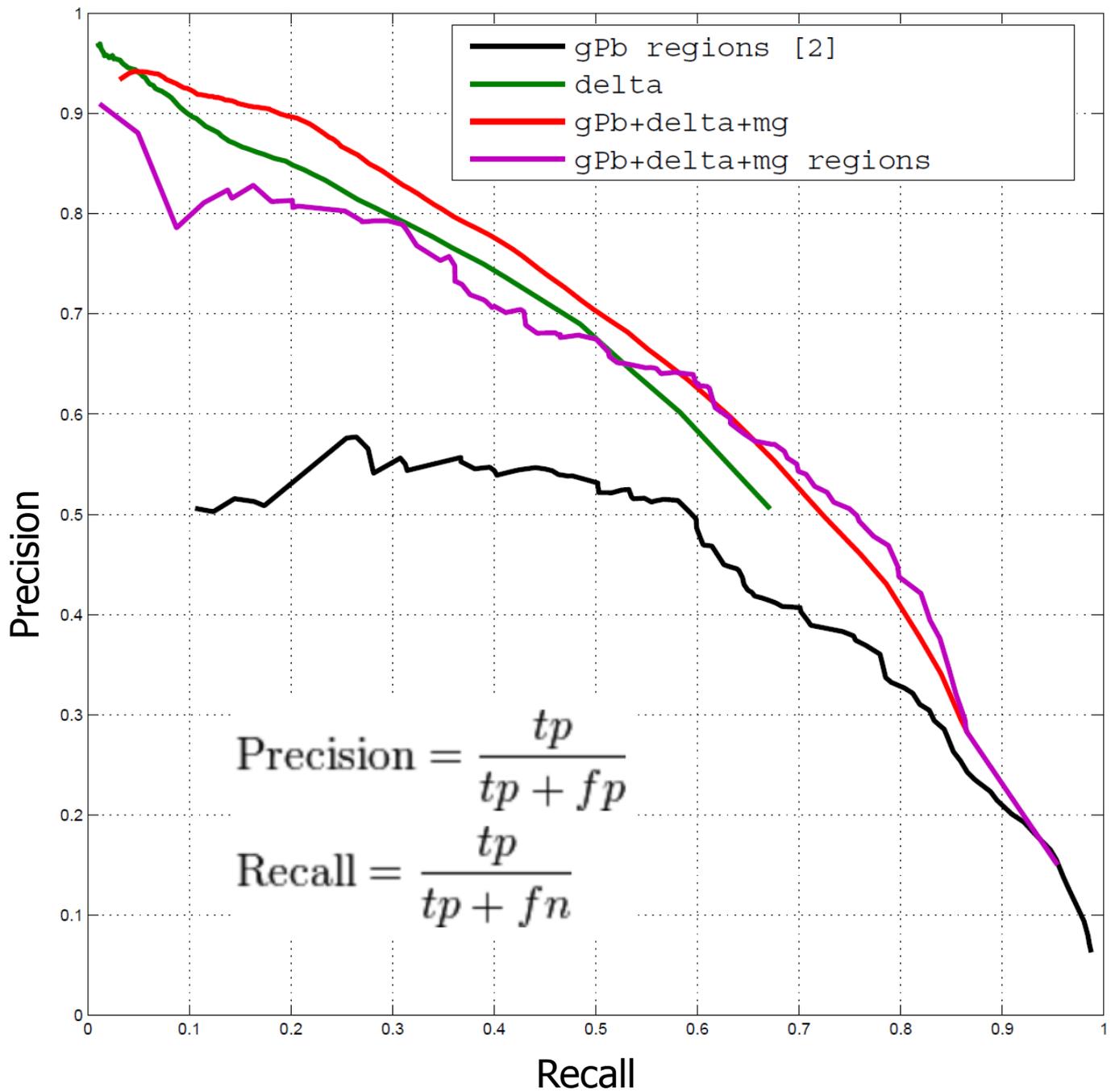
Предположения
о потоке

Предположения
о границах

$$\rho = 0.7 \text{ (подобран с помощью SVM)}$$

Дополнительно предлагаются:

- разделение на связные объекты
- сегментация объект/фон



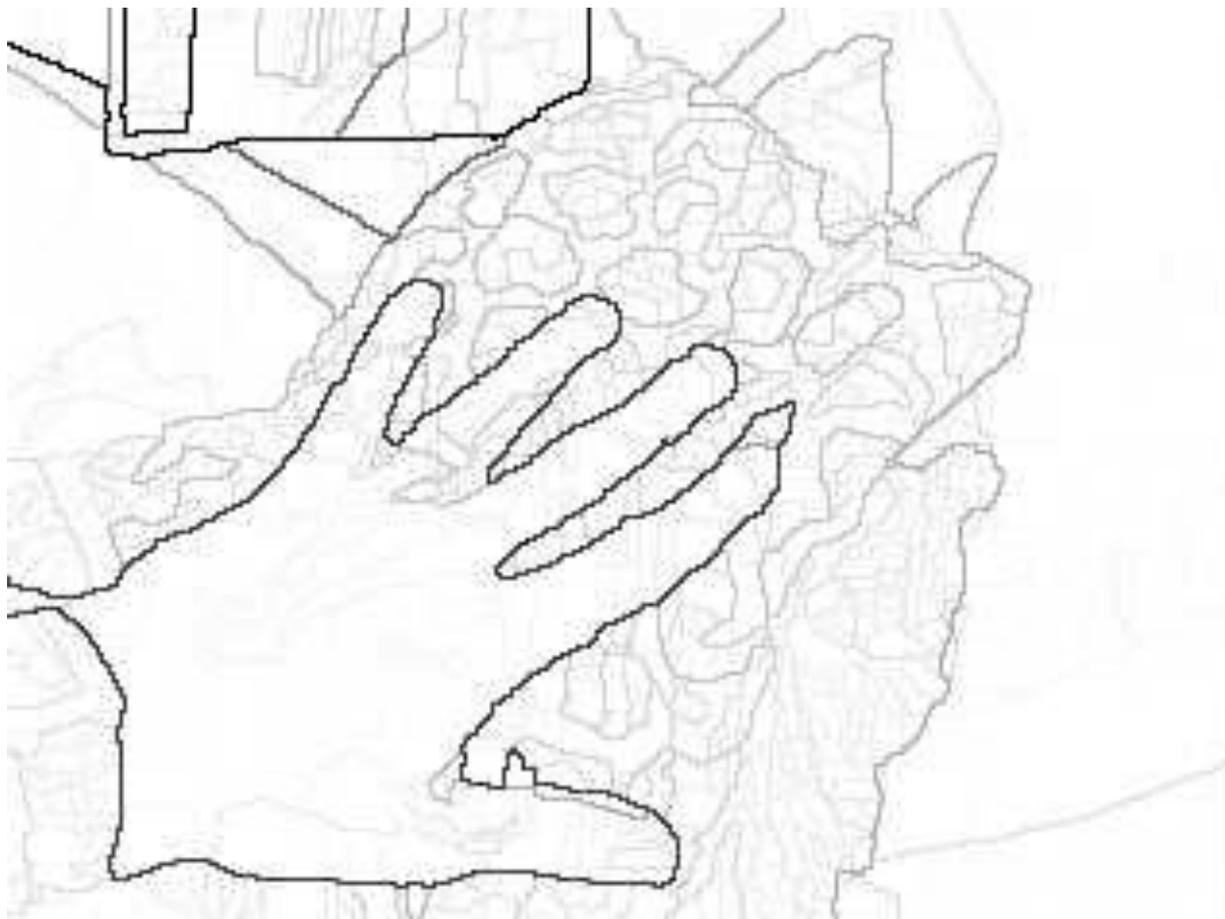
Результаты (1/2)

Исходное изображение



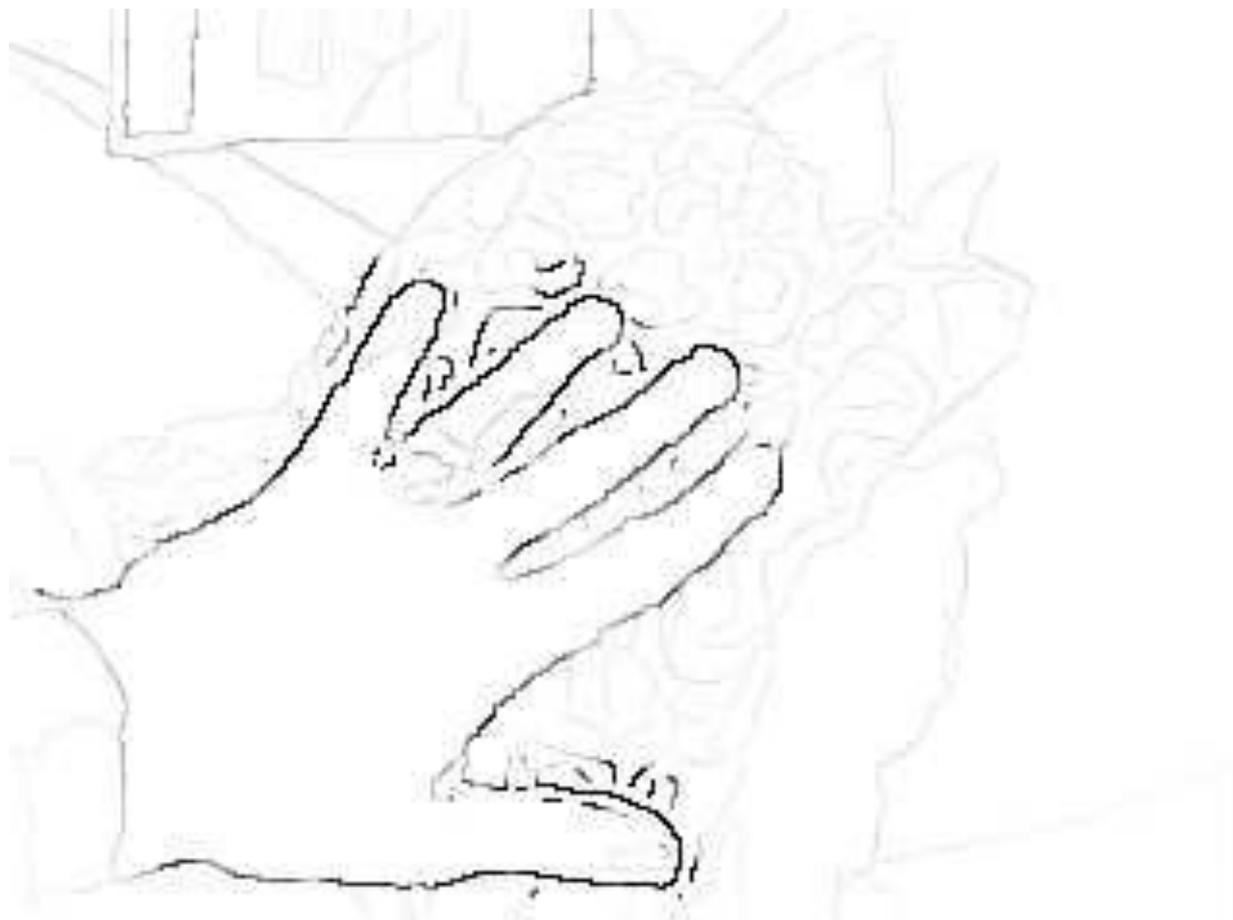
Результаты (1/2)

Без оптического потока



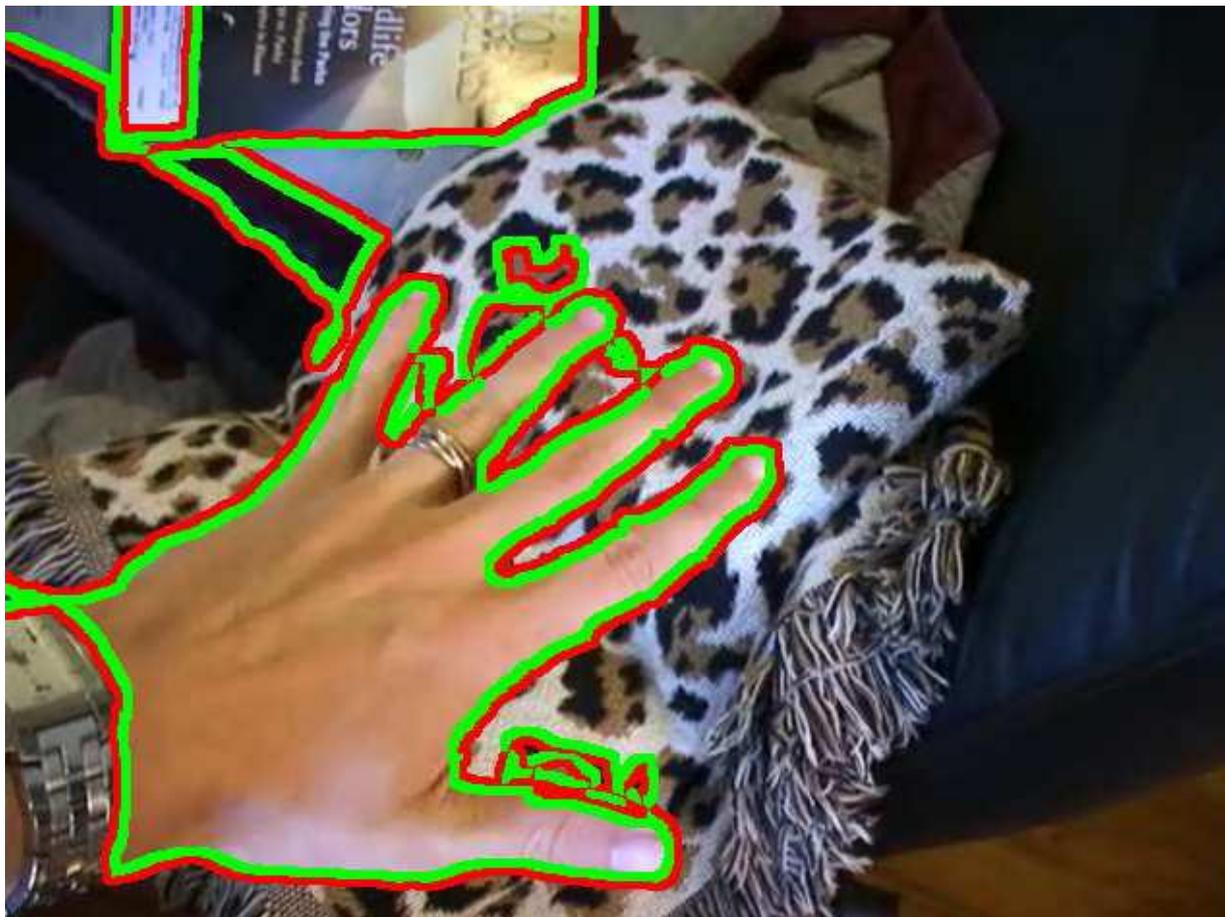
Результаты (1/2)

С оптическим потоком



Результаты (1/2)

Сегментация объект/фон



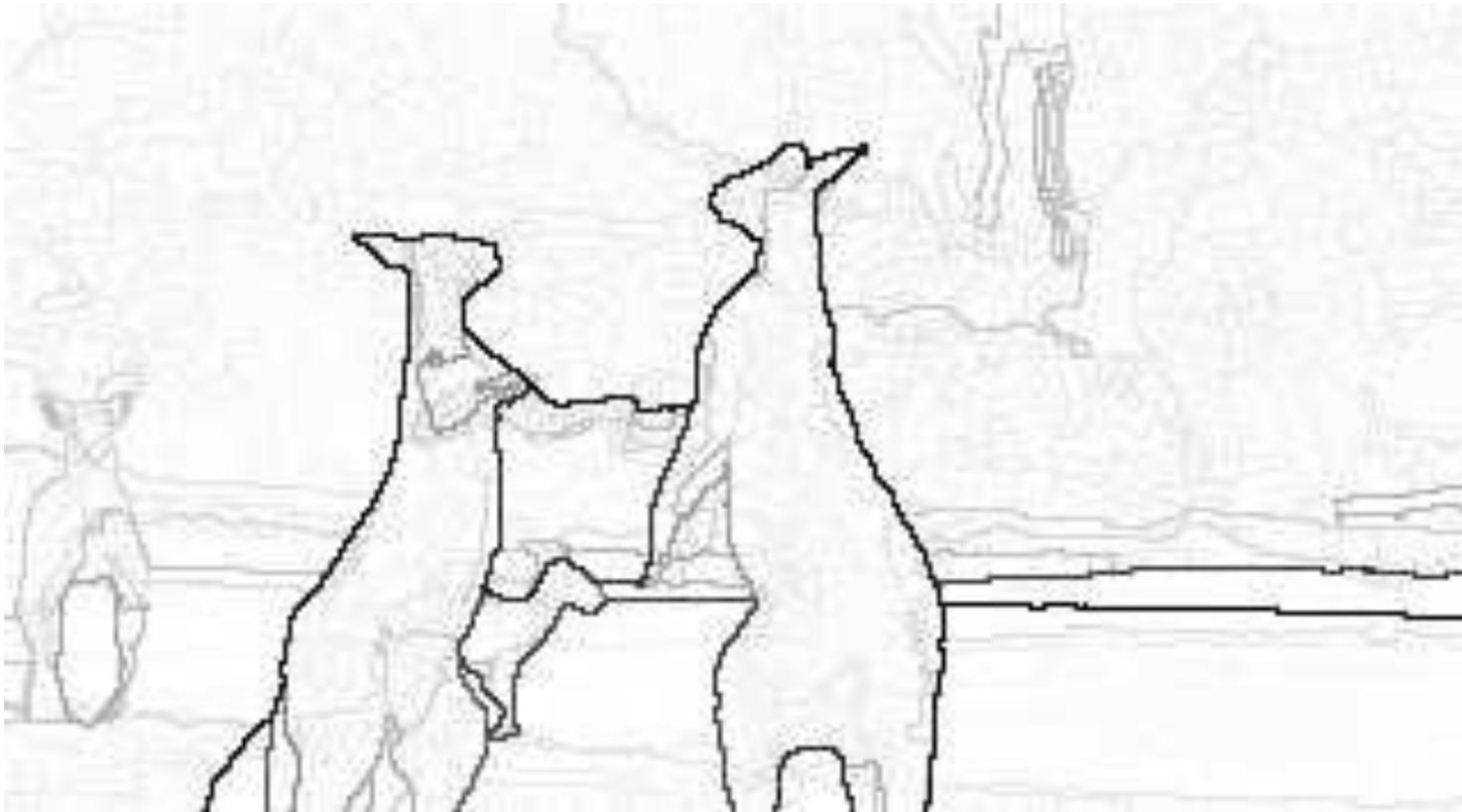
Результаты (2/2)

Исходное изображение



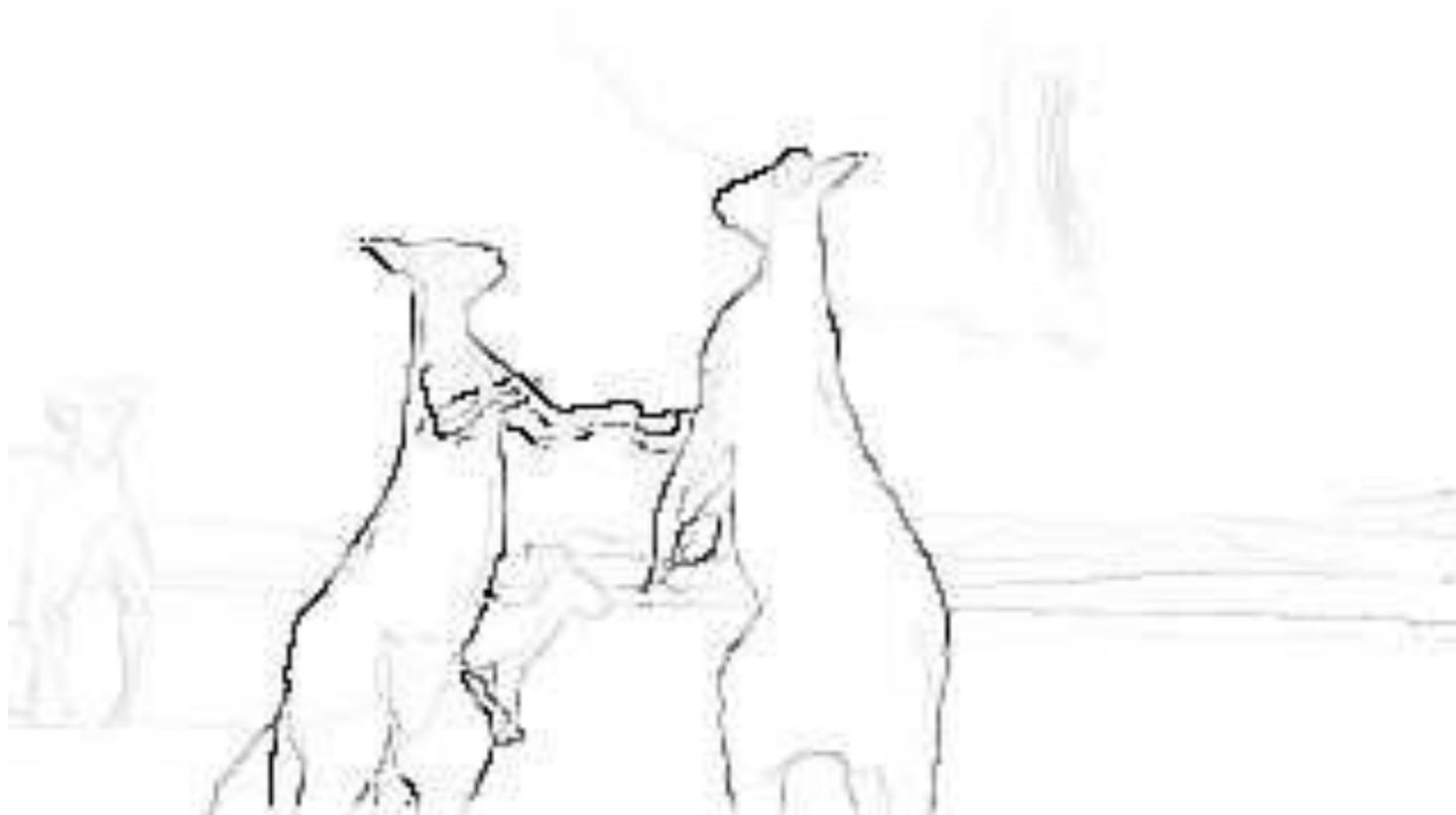
Результаты (2/2)

Без оптического потока



Результаты (2/2)

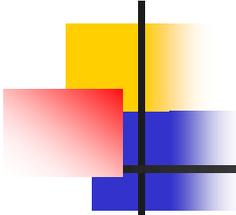
С оптическим потоком



Результаты (2/2)

Сегментация объект/фон





Содержание

- Введение
- Совместный поиск областей открытия и OF
- Использование машинного обучения
- Выделение границ объектов
- **Заключение**

Заключение

- Поиск областей открытия только по потоку ненадежен
- Нужно
 - искать поток одновременно с областями открытия
 - учитывать пространственные признаки
 - использовать несколько признаков (машинное обучение?)
 - тестировать сегментацию объект/фон на стандартных базах
- Глобальные оптимизационные методы поиска OF не такие медленные, как казалось (?)

Применение к распространению глубины

Улучшение детектора областей открытия

- Избавиться от ошибок типа False Positive
Заполнять области открытия глубиной фона
- Добавить пространственные признаки
Заполнять приблизительные области открытия
только из пространственной информации

Вопросы?

Литература

1. T. Brox, "From pixels to regions: partial differential equations in image analysis," PhD thesis, Faculty of Mathematics and CS, Saarland University, 2005
2. A. Humayun, O. M. Aodha, G. J. Brostow, "Learning to Find Occlusion Regions," *CVPR*, 2011
3. P. Sand, S. Teller, "Particle Video: Long-Range Motion Estimation using Point Trajectories," *CVPR*, 2006
4. T. Brox, A. Bruhn, N. Papenberg, J. Weickert, "High accuracy optical flow estimation based on a theory for warping," *ECCV*, pages 25–36, 2004
5. P. Sundberg, T. Brox, M. Maire, P. Arbelaez, J. Malik, "Occlusion boundary detection and figure/ground assignment from optical flow," *CVPR*, 2011
6. С. Гришин, "Программная система для преобразования частоты кадров цифровых видео сигналов," 2009

Лаборатория компьютерной графики и мультимедиа



Видеогруппа — это:

- Выпускники в аспирантурах Англии, Франции, Швейцарии (в России в МГУ и ИПМ им. Келдыша)
- Выпускниками защищены 5 диссертаций
- Наиболее популярные в мире сравнения видеокодеков
- Более 3 миллионов скачанных фильтров обработки видео