

Некоторые методы сопоставления и трекинга границ во времени

Александр Новиков

Video Group
CS MSU Graphics & Media Lab



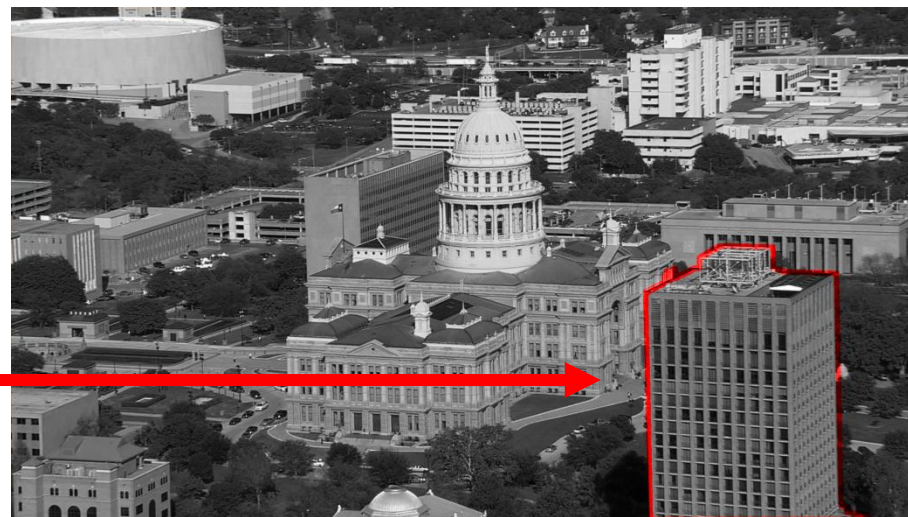
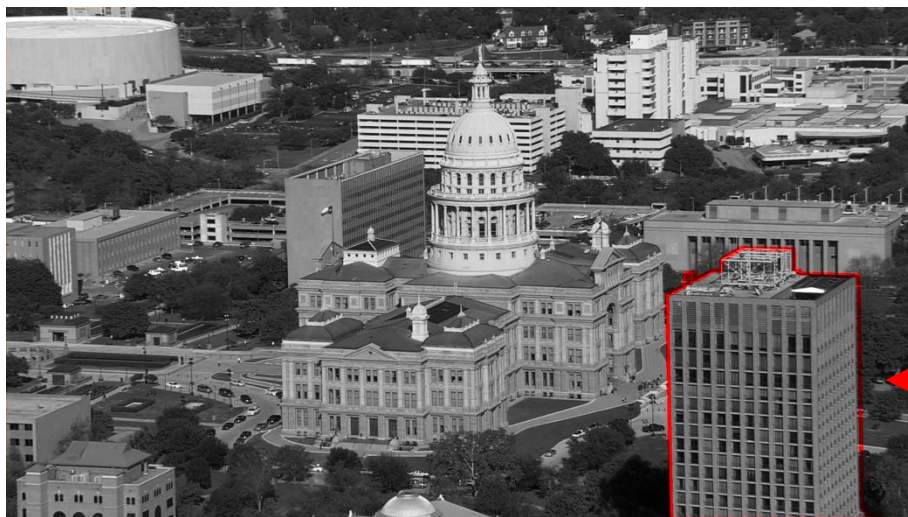
Содержание

- **Введение**
- Joint Pixel Features Based Tracking
- Generalized Hough Transform Based Tracking
- Tracking via Edgel Templates
- Заключение

Постановка задач

Задача сопоставления границ

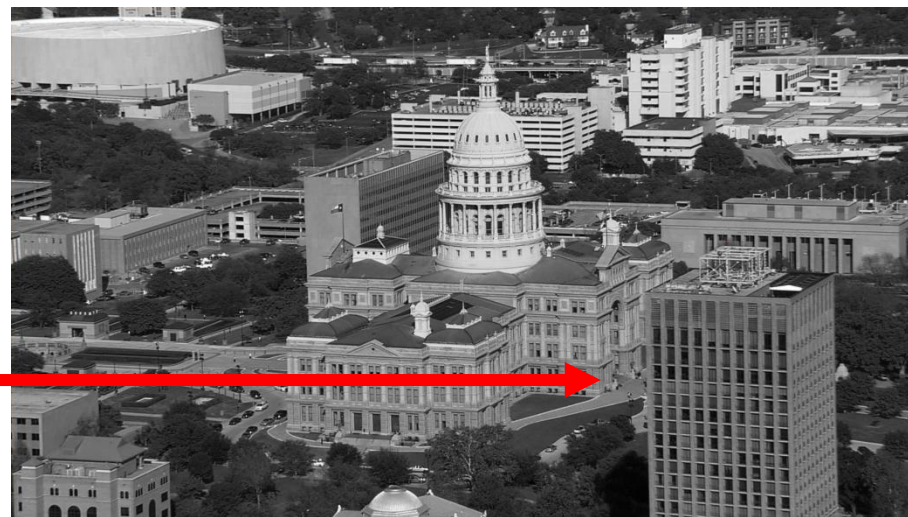
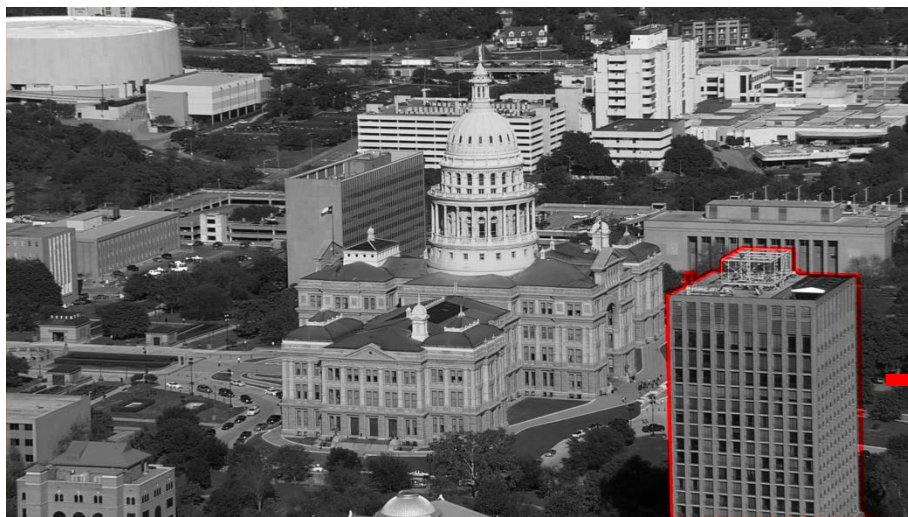
Для пары кадров произвести сопоставление их карт границ



Постановка задач

Задача трекинга границ

Проследить движение границ, выделенных на первом кадре видеопоследовательности



Применение (1)

1. Трекинг объектов



Frame 0



Frame 15



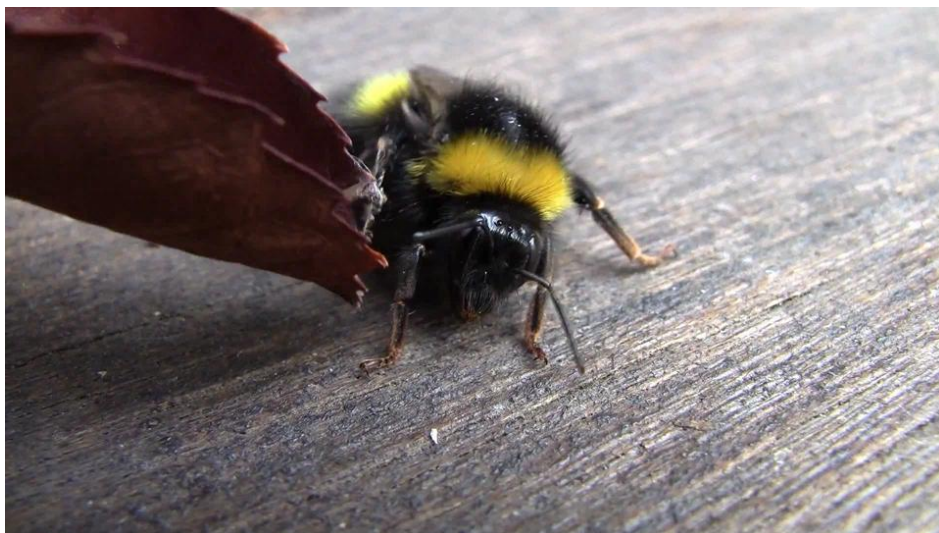
Frame 30

2. Надежный поиск и обработка областей открытия

M.-C. Roh, T.-Y. Kim, J. Park, S.-W. Lee, "Accurate object contour tracking based on boundary edge selection," in *Pattern Recognition*, Volume 40, Issue 3, 2007

Применение (2)

3. Избавление от «затеканий», т.е. попадания значений каких-либо карт за границы соответствующих объектов



Исходное видео



Карта салиентности



Тривиальный подход

Для решения задачи можно использовать алгоритмы Motion Estimation

Проблемы:

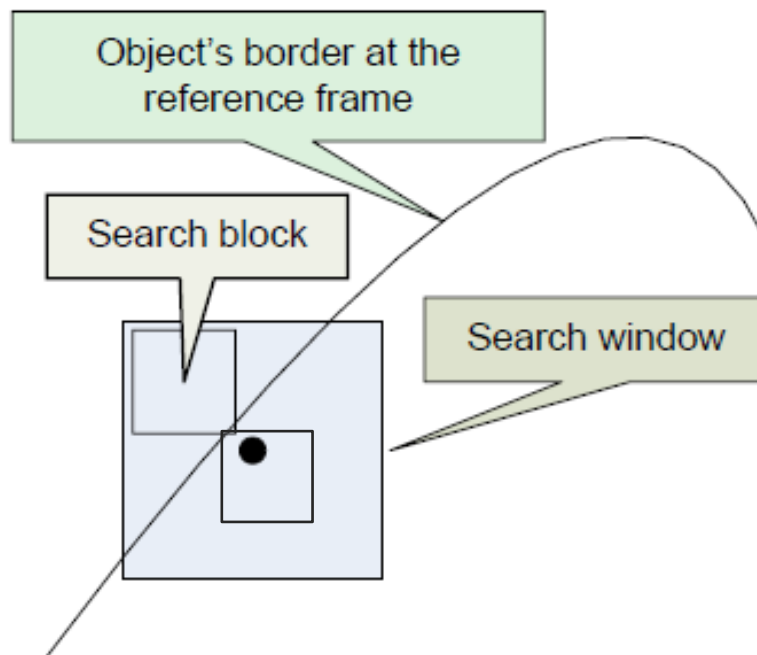
- Блочные методы с простыми метриками неточны на границах
- Неверная работа на областях открытия



Содержание

- Введение
- **Joint Pixel Features Based Tracking**
- Generalized Hough Transform Based Tracking
- Tracking via Edgel Templates
- Заключение

Идея метода



Поиск для каждого пикселя границы на первом кадре его положения на втором сравнением блоков по некоторой метрике



Варианты реализации

Поблочное сравнение – самый распространенный подход

Часто используемые метрики – разница:

- значений цветочных каналов
- модулей/направлений градиентов яркости



Joint Feature Vector

В данной статье используется евклидово расстояние между Joint Feature Vector'ами пикселей, состоящими из:

1. выбранных коэффициентов Undecimated Wavelet Packet Transform (UWPT)
2. значения яркости пикселя

Undecimated Wavelet Packet Transform (1)



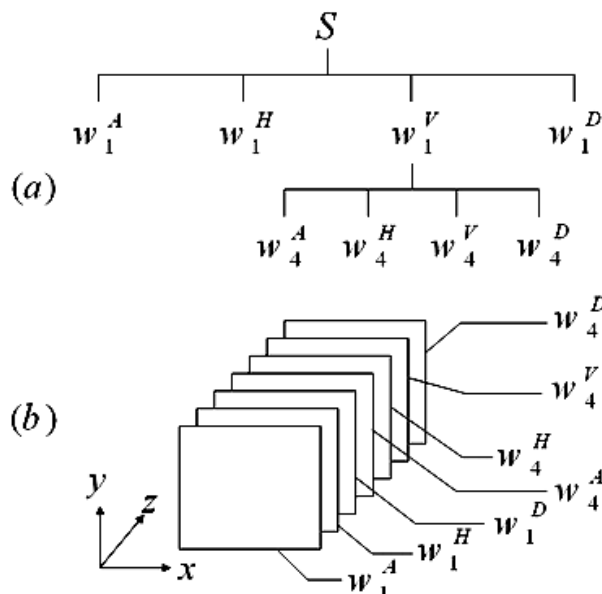
UWPT – один из видов вейвлет-преобразований

Используется авторами в связи с:

- достаточно подробным описанием окрестности пикселя
- его инвариантностью к сдвигам
- устойчивостью к различным видам шума

Undecimated Wavelet Packet Transform (2)

1. Строится дерево UWPT и выполняется его прореживание
2. Задается Feature Vector для каждого пикселя кадров



Пример прореженного дерева и формула построения Feature Vector по нему

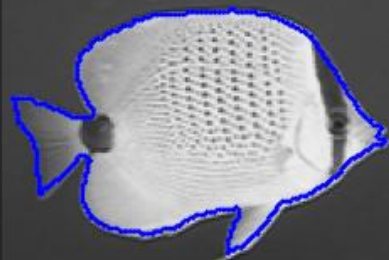
w – матрицы коэффициентов вейвлет-преобразования

$$FV(x, y) = \{w_1^A(x, y), w_1^H(x, y), w_1^D(x, y), w_4^A(x, y), w_4^H(x, y), w_4^V(x, y), w_4^D(x, y)\}$$

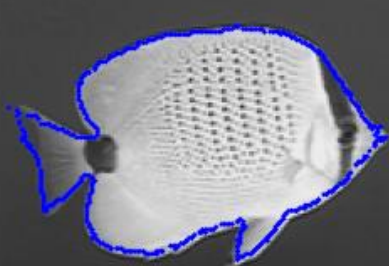
M. Khansari, H. R. Rabiee, et al., "An Adaptive Semi-Automatic Video Object Extraction Algorithm Based on Joint Transform and Spatial Domains Features," in *CBMI*, 2005

Результаты

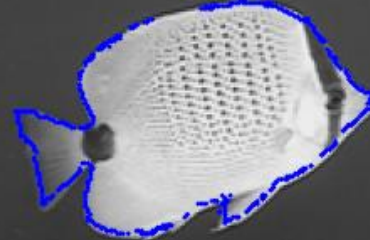
Кадр №0



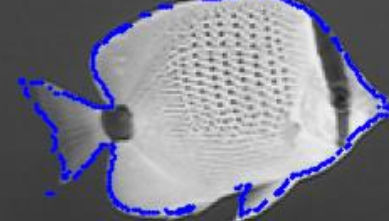
Кадр №20



Кадр №47



Кадр №76



Кадр №169



Кадр №176



Кадр №194



Кадр №195



Референсный кадр

Результаты трекинга границы на некоторых кадрах

M. Khansari, H. R. Rabiee, et al., "An Adaptive Semi-Automatic Video Object Extraction Algorithm Based on Joint Transform and Spatial Domains Features," in *CBMI*, 2005



Выводы

Достоинство:

- Использование привычного поблочного поиска – возможность использовать МЕ

Недостаток:

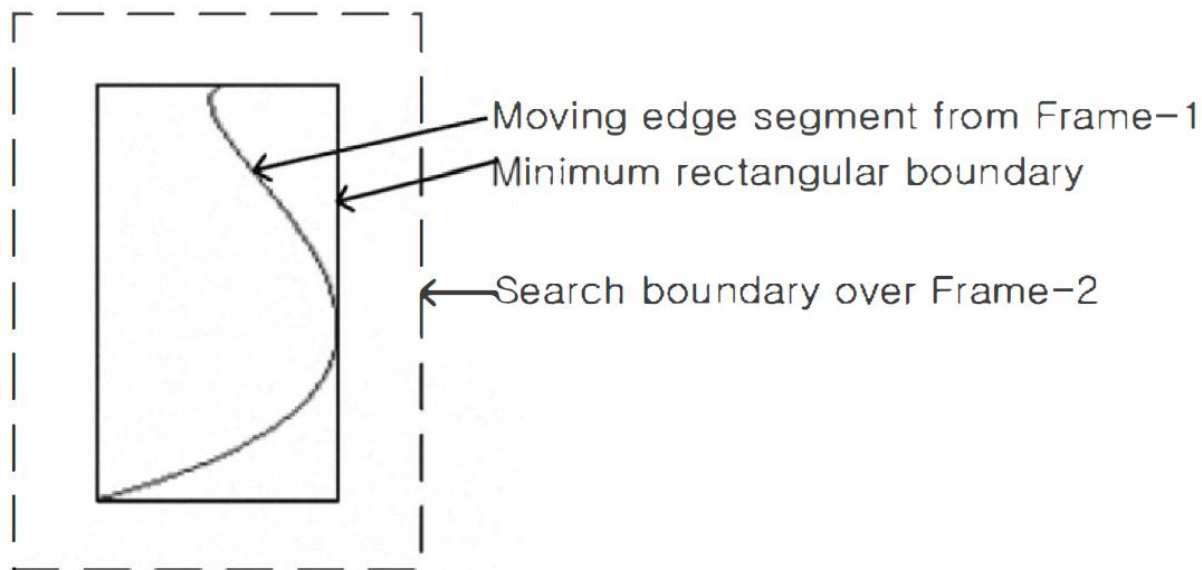
- Относительная сложность реализации метрики



Содержание

- Введение
- Joint Pixel Features Based Tracking
- **Generalized Hough Transform Based Tracking**
- Tracking via Edgel Templates
- Заключение

Идея метода



По построенным картам границ двух кадров, для каждого сегмента границы с первого кадра ищется ближайший по форме со второго

Generalized Hough Transform

Generalized Hough Transform (GHT) – это метод представления границы объекта в виде набора параметров, характеризующих её форму (R-таблицы)

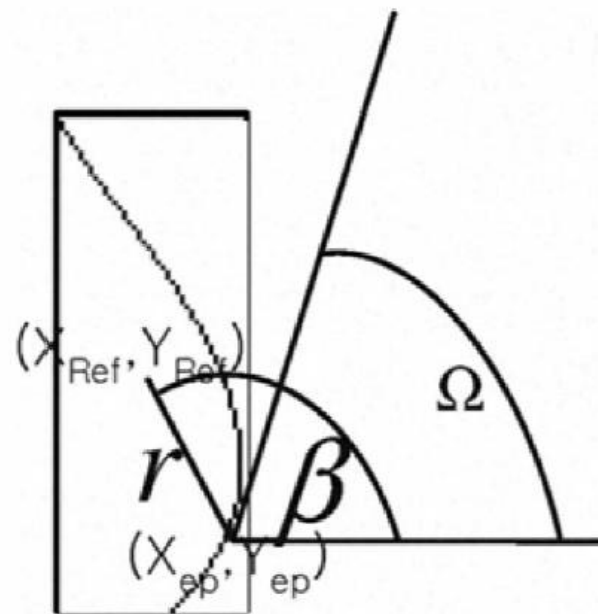
- Позволяет быстро проверять объекты на схожесть с референсным
- Устойчиво к шуму и искажениям границ

Generalized Hough Transform

Построение R-таблицы (1)

Для каждой точки исходного сегмента определяется три параметра: r , β , Ω

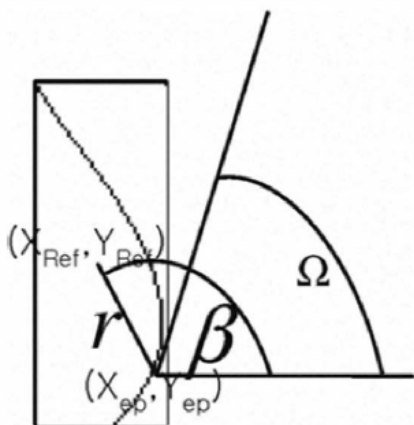
X_{Ref} , Y_{Ref} – центр ограничивающего прямоугольника
 X_{ep} , Y_{ep} – текущая точка
 r – длина отрезка между точками (X_{Ref}, Y_{Ref}) и (X_{ep}, Y_{ep})
 β – полярный угол точки (X_{Ref}, Y_{Ref}) относительно (X_{ep}, Y_{ep})
 Ω – угол наклона касательной к сегменту в точке (X_{ep}, Y_{ep})



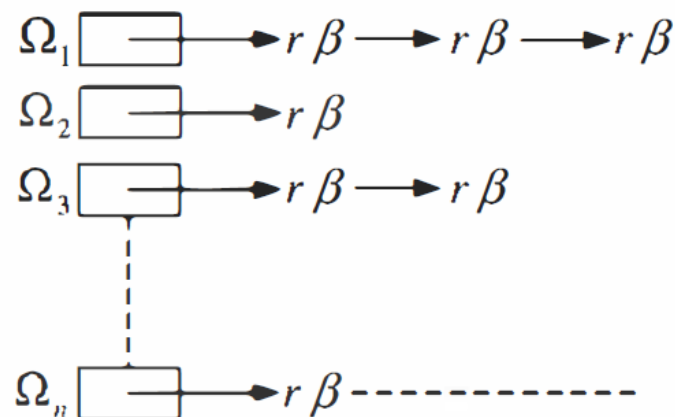
Generalized Hough Transform

Построение R-таблицы (2)

Параметры каждой точки заносятся в R-таблицу исходного сегмента



Параметры точки сегмента

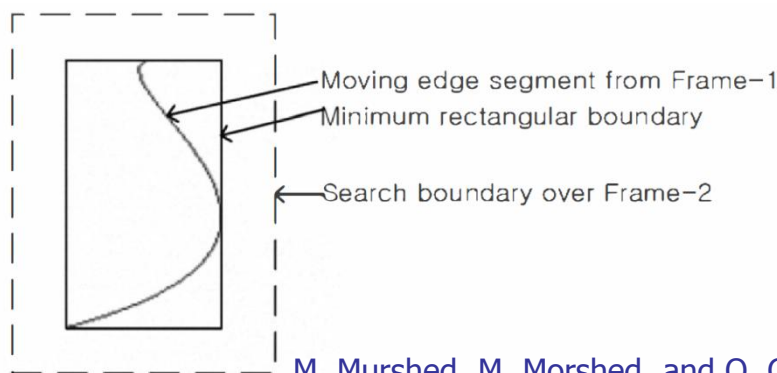


Общий вид R-таблицы

Алгоритм сопоставления

Для каждого сегмента границ с первого кадра:

1. Строится R-таблица
2. Находятся сегменты-кандидаты в окрестности поиска на втором кадре
3. С помощью R-таблицы выбирается сегмент-кандидат, наиболее похожий на исходный

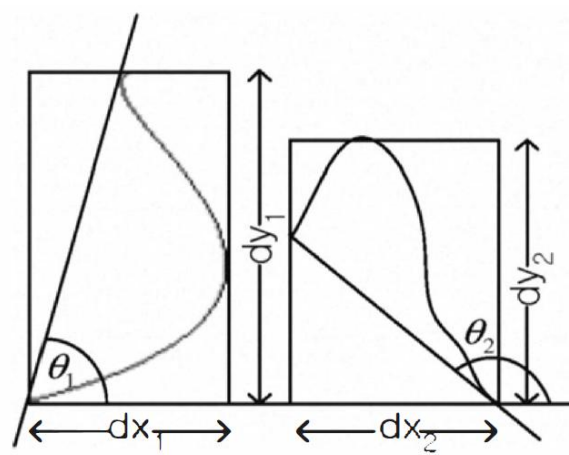
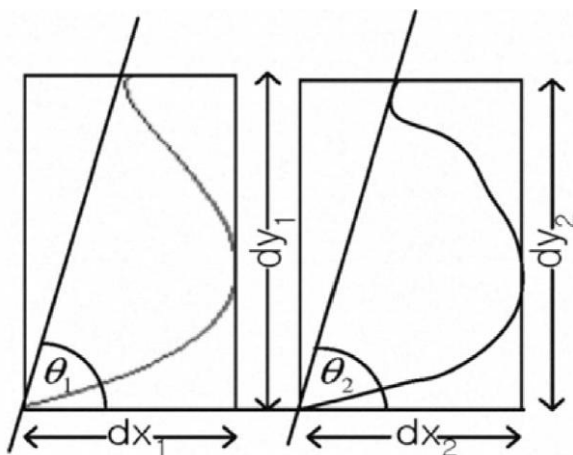


M. Murshed, M. Morshed, and O. Chae, "Moving Edge Matching for Moving Object Tracking," in *Proc. of 14th International Conference on Computer and Information Technology*, 2011

Выбор сегментов-кандидатов

Для исходного сегмента границы на втором кадре выбираются сегменты-кандидаты, близкие по:

1. длине и ширине ограничивающего прямоугольника
2. углу между прямой, соединяющей начало и конец сегмента и осью x

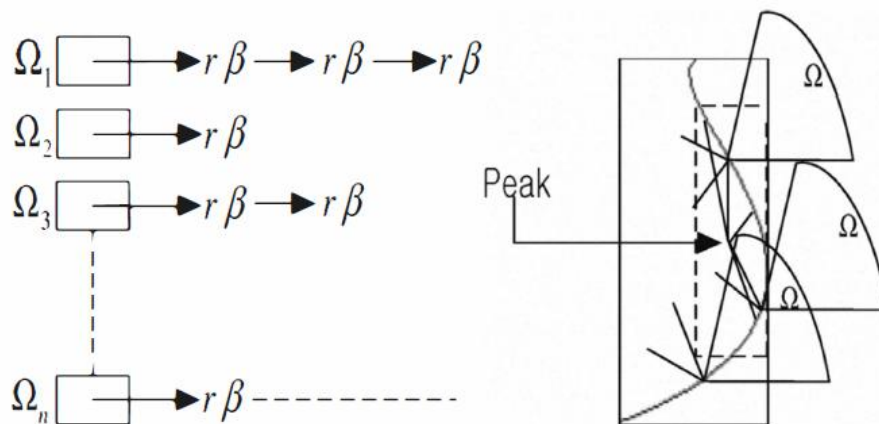


M. Murshed, M. Morshed, and O. Chae, "Moving Edge Matching for Moving Object Tracking," in *Proc. of 14th International Conference on Computer and Information Technology*, 2011

Сопоставление кандидатов с исходным сегментом (1)

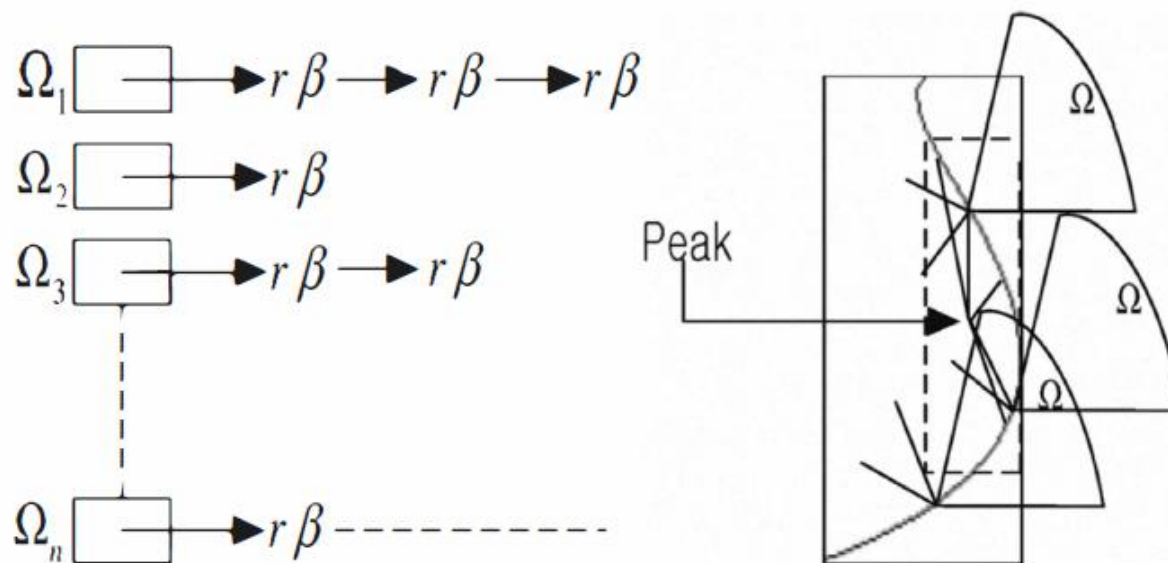
Для каждой точки сегмента-кандидата:

1. Вычисляется угол Ω между осью x и касательной к сегменту в ней
2. Для Ω из R-таблицы извлекаются все пары (r, β) , и от точки откладываются соответствующие векторы



Сопоставление кандидатов с исходным сегментом (2)

Точка Peak, в которой заканчивается наибольшее число векторов, считается новым положением центральной точки (X_{Ref} , Y_{Ref}) исходного сегмента



Определение движения сегмента границы



- Поиск точки Peak выполняется для каждого сегмента-кандидата
- Количество векторов, заканчивающихся в ней, служит мерой надежности сегмента
- Выбирается сегмент с наивысшим значением надежности, определяется сдвиг

$$dx = X_{\text{MaxPeak}} - X_{\text{Ref}}$$
$$dy = Y_{\text{MaxPeak}} - Y_{\text{Ref}}$$

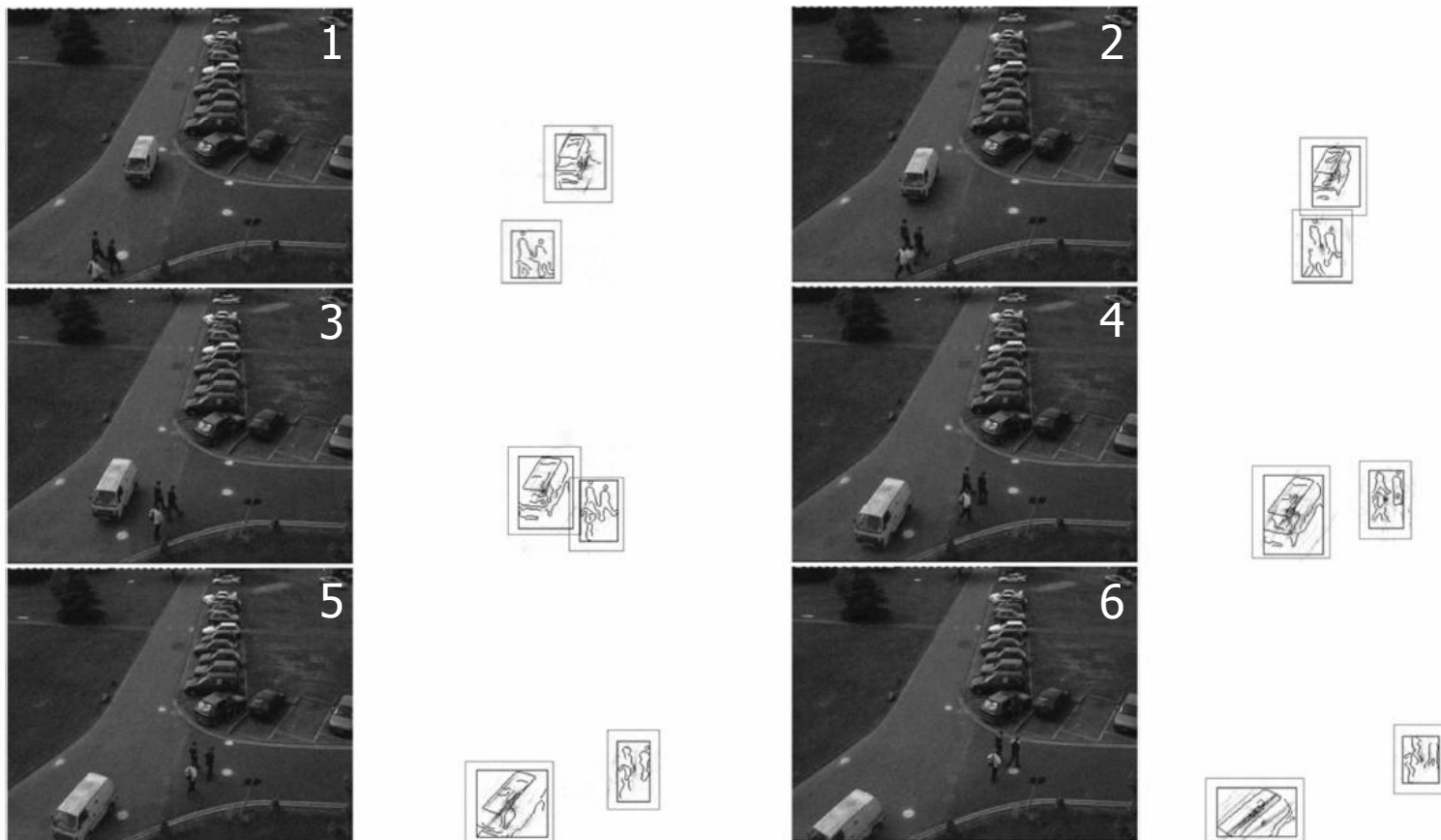
Тестирование метода

Авторы использовали свой метод сопоставления границ для трекинга объектов



Маленький прямоугольник – граница объекта, большой – окно поиска
Видео 640×520 пикселей, скорость работы реализации авторов – 7 fps

Результаты



Результаты трекинга объектов при помощи сопоставления границ

M. Murshed, M. Morshed, and O. Chae, "Moving Edge Matching for Moving Object Tracking," in *Proc. of 14th International Conference on Computer and Information Technology*, 2011



Выводы

Достоинство:

- Устойчивость к шуму и искажениям формы границ

Недостаток:

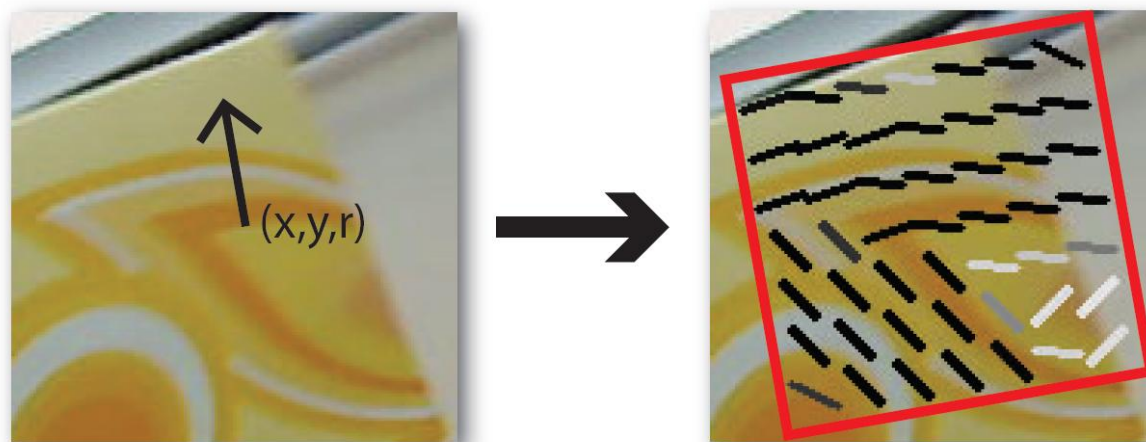
- Низкая устойчивость к поворотам границ



Содержание

- Введение
- Joint Pixel Features Based Tracking
- Generalized Hough Transform Based Tracking
- **Tracking via Edgel Templates**
- Заключение

Идея метода



Оценка совпадения направлений градиентов яркости пикселей в окрестностях сопоставляемых границ

Что такое Edgel?

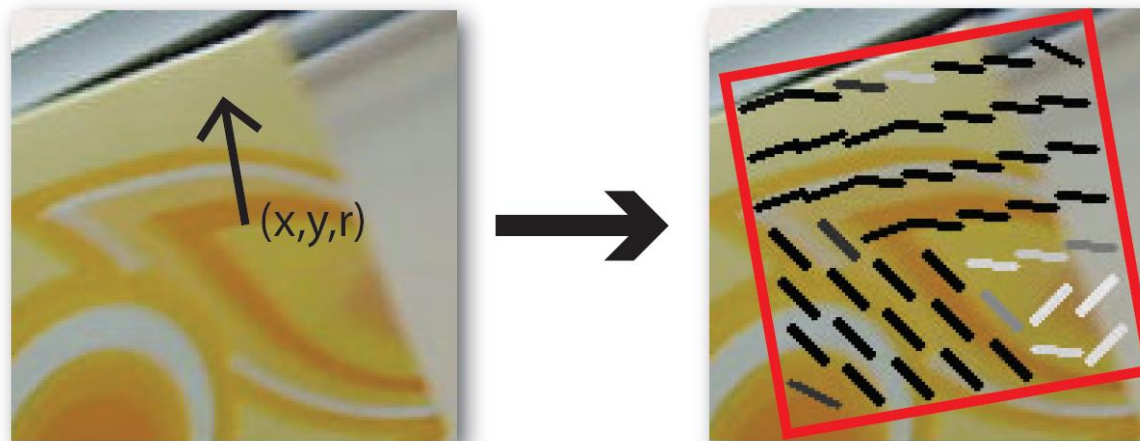
Edgel – пиксель границы, характеризующийся координатами и направлением градиента яркости



Edgel Template (1)

Для сопоставляемых Edgel'ов вычисляются Edgel Template'ы (*ET*) – матрицы, характеризующая их окрестности

Шаг 1: разбиение окрестности на $M \times M$ блоков и вычисление градиентов яркости в них

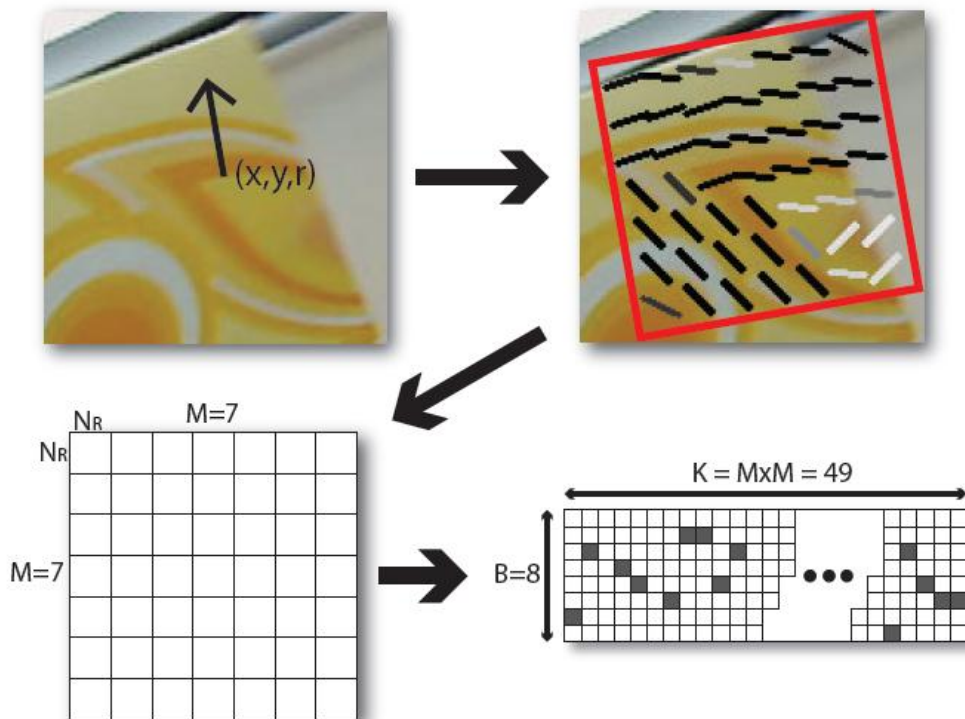


Edgel Template (2)

Шаг 2: дискретизация направлений градиентов, заполнение матрицы

$$ET(i, j) = \begin{cases} 1, & \frac{(j-1) * \pi}{B} \leq \varphi_i < \frac{j\pi}{B} \\ 0, & otherwise \end{cases}$$

φ_i – направление градиента яркости в i -м блоке
 B – параметр дискретизации



Edgel Templates Matching

Метрика степени совпадения двух Edgel Template'ов имеет вид:

$$F(\phi_1, \phi_2) \doteq \frac{1}{K} \|\phi_1 \circ \phi_2\|$$

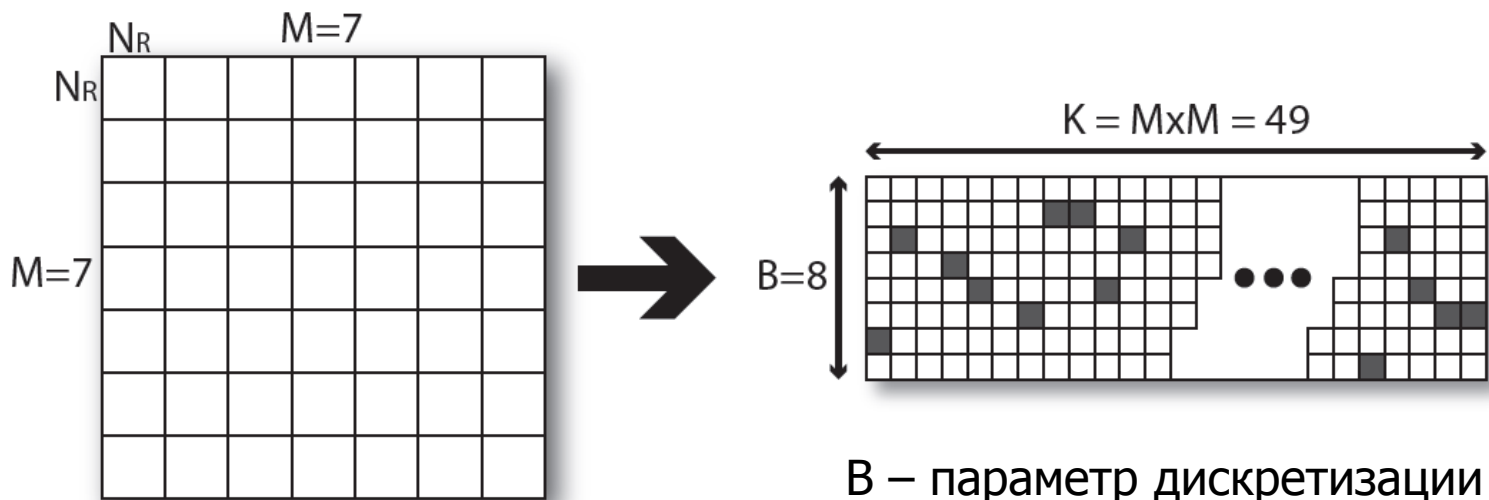
$\|\ \ \|$ – количество единиц в поэлементном произведении двух Edgel Template'ов

K – константа

Edgel Templates Matching

Быстрая реализация

Хранение Template'ов в виде матриц из 0 и 1 позволяет реализовать их сравнение быстрыми битовыми операциями



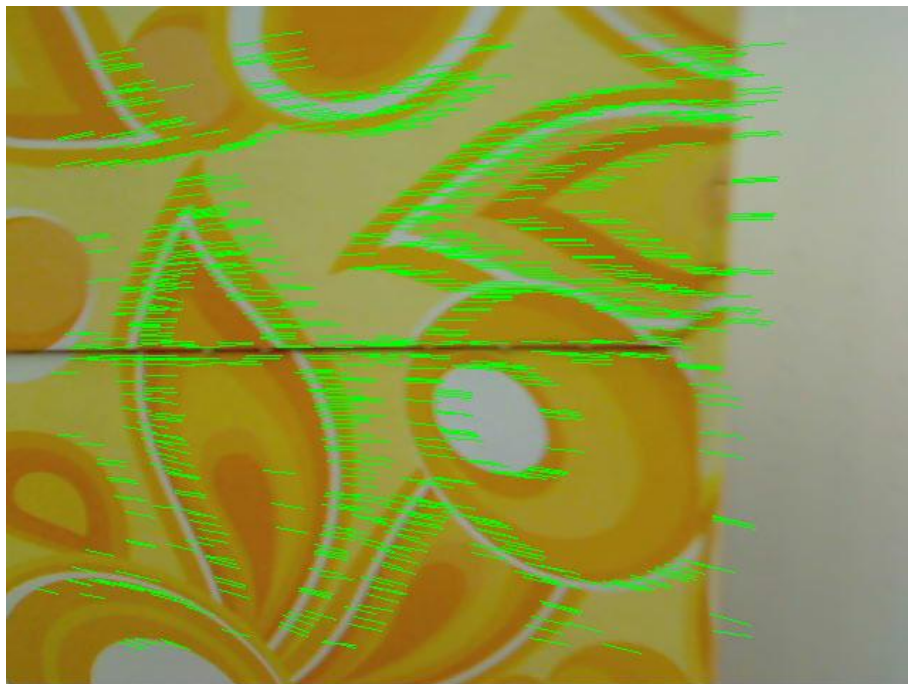
Оценка скорости работы

Реализация авторов работает в реальном времени на видео с разрешением 640×480

tasks	edgel+tracking time (ms)	no tracking time (ms)
image pyramids	2.3	2.3
tracking edgels	40.7	-
image gradients	1.4	1.6
extracting edgels	12.7	14.7
matching edgels	5.2	7.9
RANSAC	4.9	6.2
total	67.2	32.7

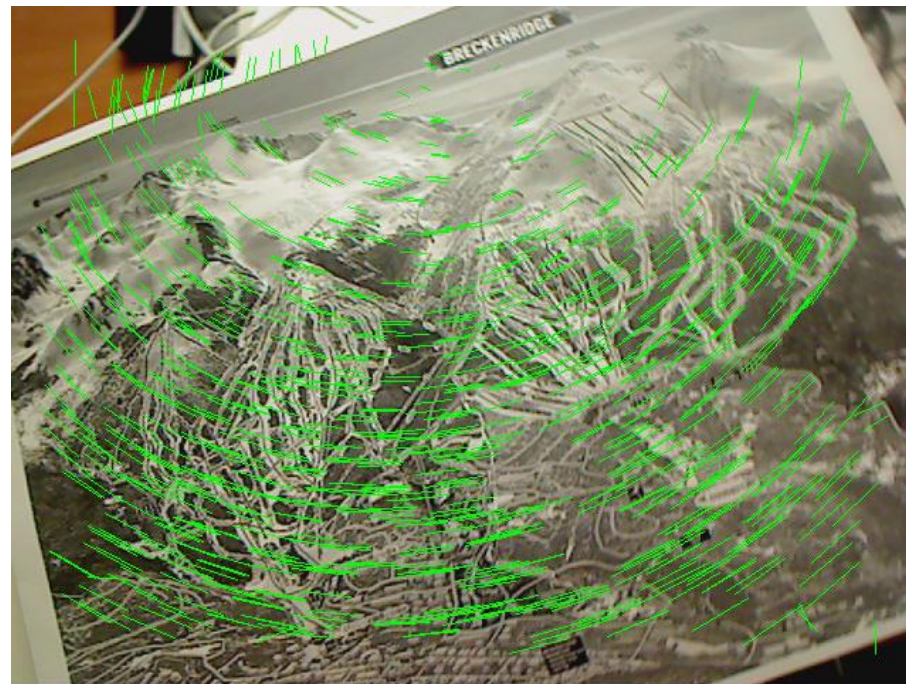
Время выполнения основных этапов для пары кадров

Результаты (1)



Результаты трекинга
Векторы сдвига границ между соседними кадрами

Результаты (2)



Результаты трекинга
Векторы сдвига границ между соседними кадрами



Выводы

Достоинства:

- Высокая скорость работы
- Инвариантность к изменению яркости
- Устойчивость к поворотам

Недостаток:

- Вероятность низкого качества работы в случаях более сложного движения (авторы привели простые примеры)



Содержание

- Введение
- Joint Pixel Features Based Tracking
- Generalized Hough Transform Based Tracking
- Tracking via Edgel Templates
- **Заключение**

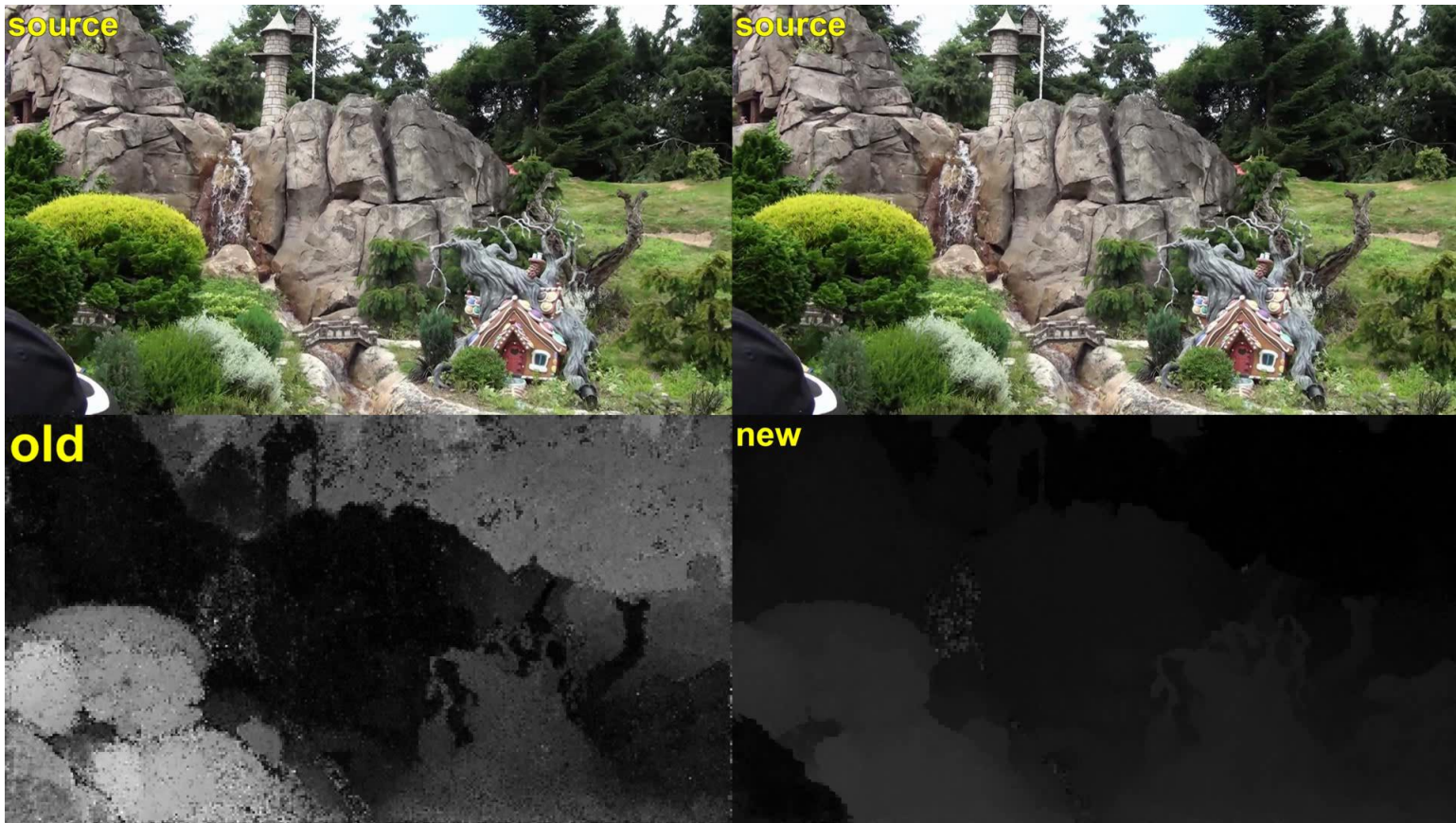
Depth from Motion

Текущие результаты (1)



Depth from Motion

Текущие результаты (2)



Depth from Motion

Текущие результаты (3)



Depth from Motion

Текущие результаты (3)



Depth from Motion

Возможные улучшения

- Уточнение границ при помощи трекинга
- Выделение областей открытия дает
ВОЗМОЖНОСТЬ:
 - полной автоматизации построения карт глубины
 - корректной обработки движущихся объектов

Литература

1. T. Lee and S. Soatto, "Fast Planar Object Detection and Tracking via Edgel Templates," in *IEEE Workshop on the Applications of Computer Vision*, Breckenridge, Colorado, 2012, pp. 473–480.
2. M. Khansari, H. R. Rabiee, and M. Asadi, M. Ghanbari, M. Nosrati, and M. Amiri, "An Adaptive Semi-Automatic Video Object Extraction Algorithm Based on Joint Transform and Spatial Domains Features," in *International Workshop on Content-Based Multimedia Indexing*, Riga, Latvia, June 21–23, 2005.
3. M. Murshed, M. Morshed, and O. Chae, "Moving Edge Matching for Moving Object Tracking," in *Proceedings of 14th International Conference on Computer and Information Technology*, Dhaka, Bangladesh, December 22–24, 2011, pp. 355–359.
4. M.-C. Roh, T.-Y. Kim, J. Park, and S.-W. Lee, "Accurate object contour tracking based on boundary edge selection," in *Pattern Recognition*, Volume 40, Issue 3, 2007, pp. 931–943.

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