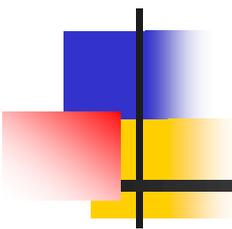


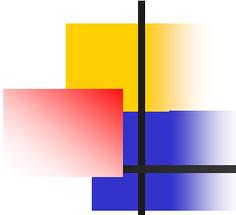
Обзор методов вычисления базовой геометрии сцены



Дмитрий Акимов

Video Group

CS MSU Graphics & Media Lab



Содержание

- **Введение**
- Feature-based алгоритм с обучением
- Интерпретация геометрических примитивов
- Заключение

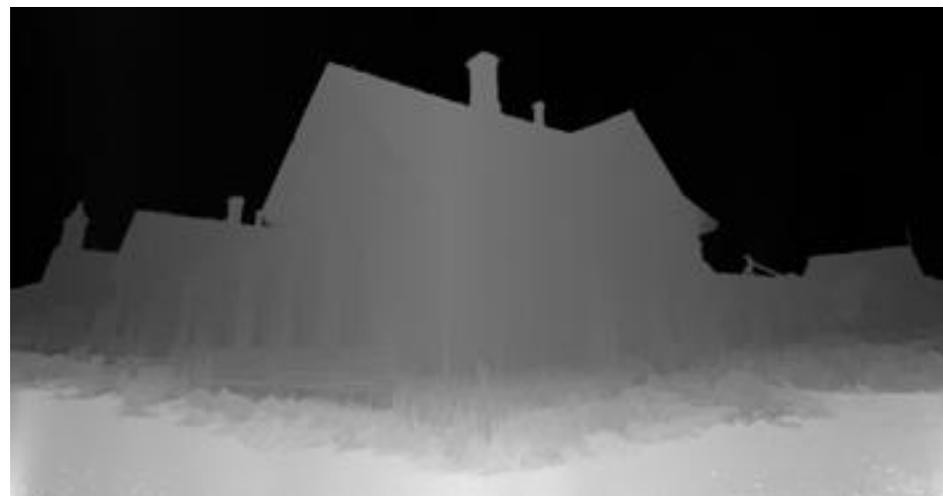
Введение

Карты глубин

Задача автоматического вычисления карты глубины по одному изображению в общем случае неразрешима



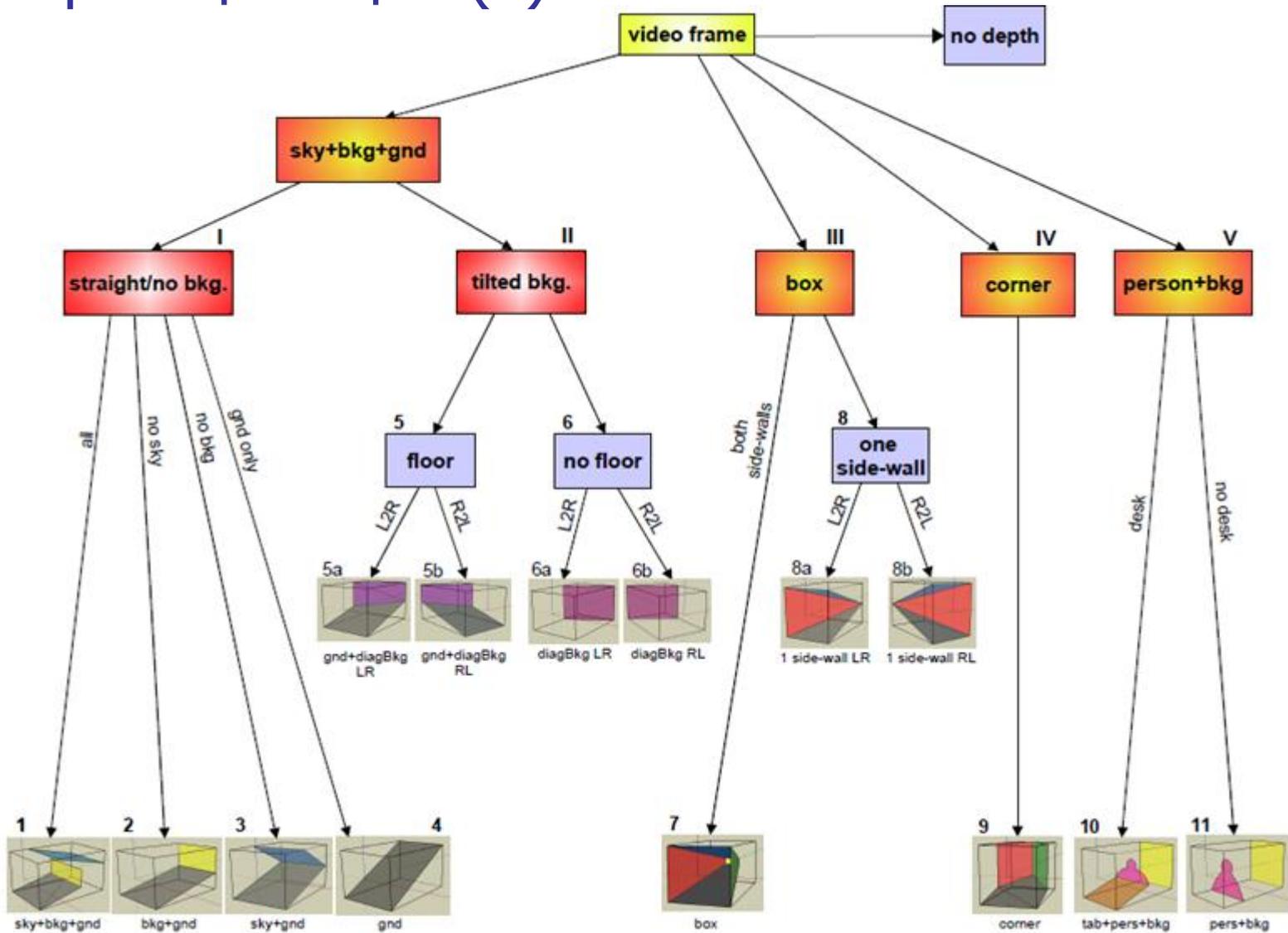
Исходное изображение



Автоматически полученная
карта глубины

Карты глубин

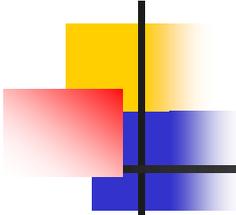
Классификация сцен (1)



Карты глубин

Классификация сцен (2)

- Количество различных классов сцен мало
- Для сцены из каждого конкретного класса с высокой вероятностью возможно автоматически создать базовую карту глубины по шаблону
- Базовая карта глубины значительно упрощает и ускоряет создание финальной карты глубины



Содержание

- Введение
- **Feature-based алгоритмы с обучением**
 - Общие характеристики
 - Geometric context extraction
- Интерпретация геометрических примитивов
- Заключение

Feature-based алгоритмы

Общие черты

- Вычисление большого количества статистик от изображения
- Формализация модели сцены
- Использование размеченной тренировочной базы для обучения и настройки параметров модели на основе вычисляемых статистик
- Получение результата на основе решений обученного классификатора

Geometric context extraction

Работа является частью проекта
Automatic Photo Pop-Up



Исходное изображение



Автоматически полученная
3D-модель сцены

Geometric context extraction

Предположения



- Объекты в сцене можно классифицировать по их пространственной ориентации
- 97% всех объектов принадлежат 3 классам:
 1. Небо
 2. Параллельные земле объекты
 3. Перпендикулярные (в той или иной степени) земле объекты
- Камера направлена параллельно земле

Алгоритм

Исходное изображение

Каждый элемент изображения – часть некоторой плоскости в реальном мире

Требуется извлечь максимум статистик для каждого элемента

- За минимальный элемент принимается суперпиксель изображения
- Вычисляются базовые статистики



Исходное изображение

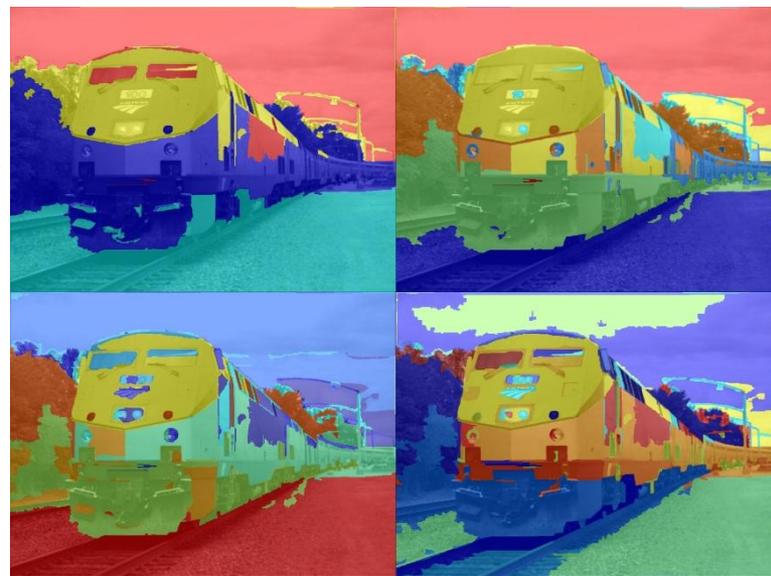


Карта суперпикселей

Алгоритм

Множественная сегментация (1)

- Получение глобальных статистик для суперпикселей
- Для каждой сегментации:
 1. Случайно выбираются $n_r = \{3, 4, 5, 7, 9, 11, 15, 20, 25\}$ исходных суперпикселей
 2. Оставшиеся суперпиксели итеративно добавляются к одному из соседних сегментов на основе абсолютной разности статистик

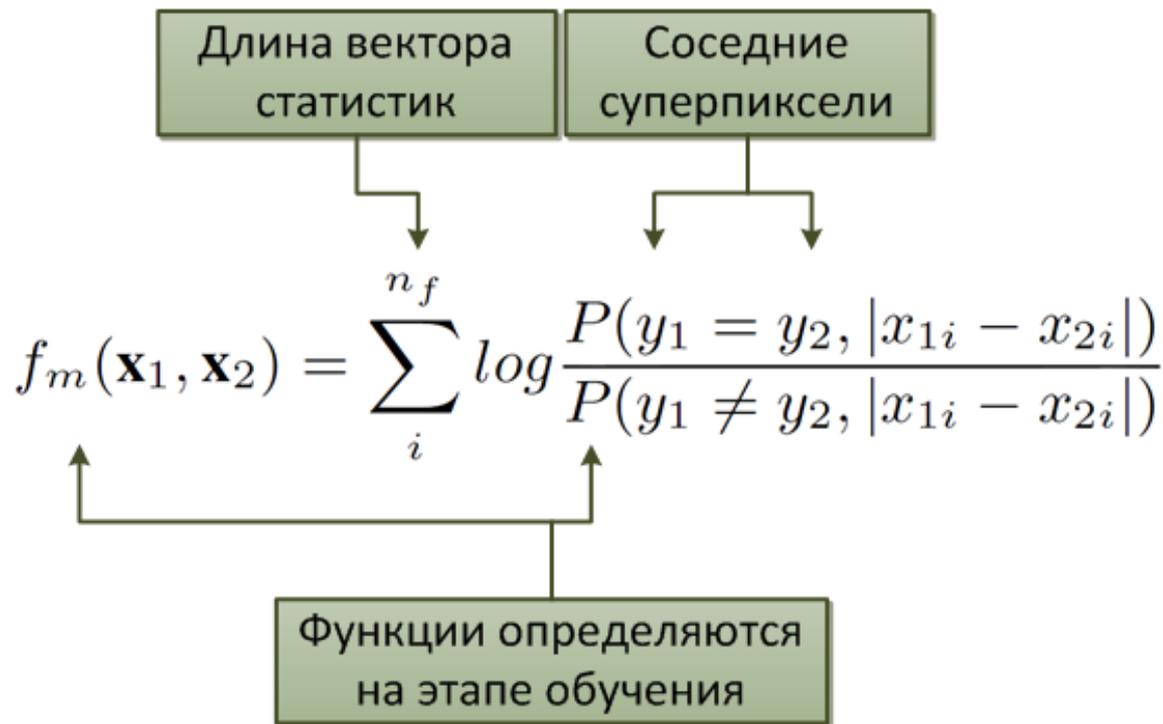


Варианты сегментаций

Алгоритм

Множественная сегментация (2)

Классификатор на основе алгоритма AdaBoost:



Алгоритм

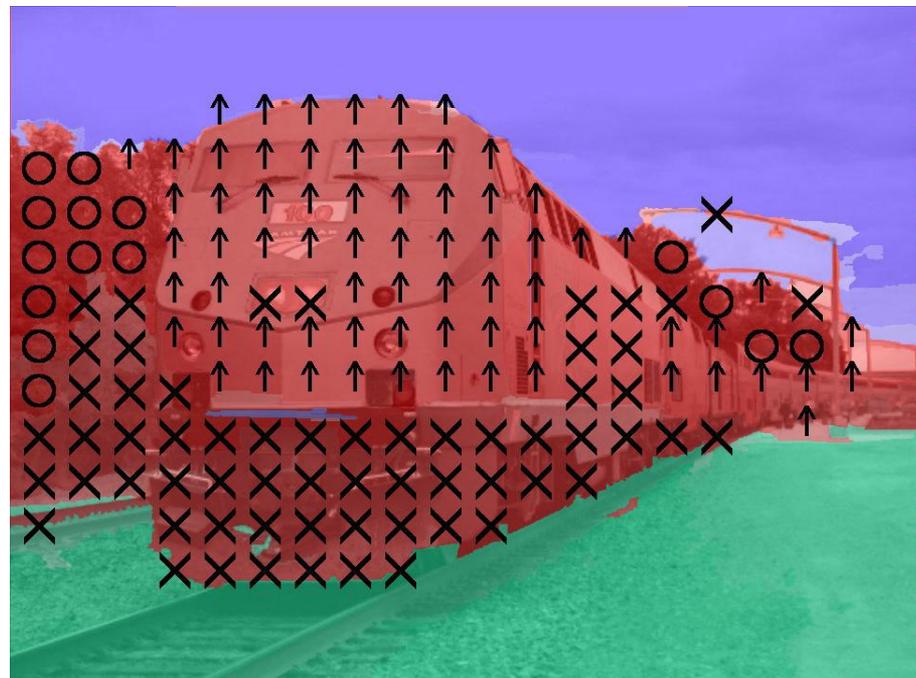
Распределение геометрических меток

Три основных метки:

- Земля
- Небо
- Вертикаль

5 подклассов класса
«вертикаль»:

- Left
- Center
- Right
- Porous
- Solid

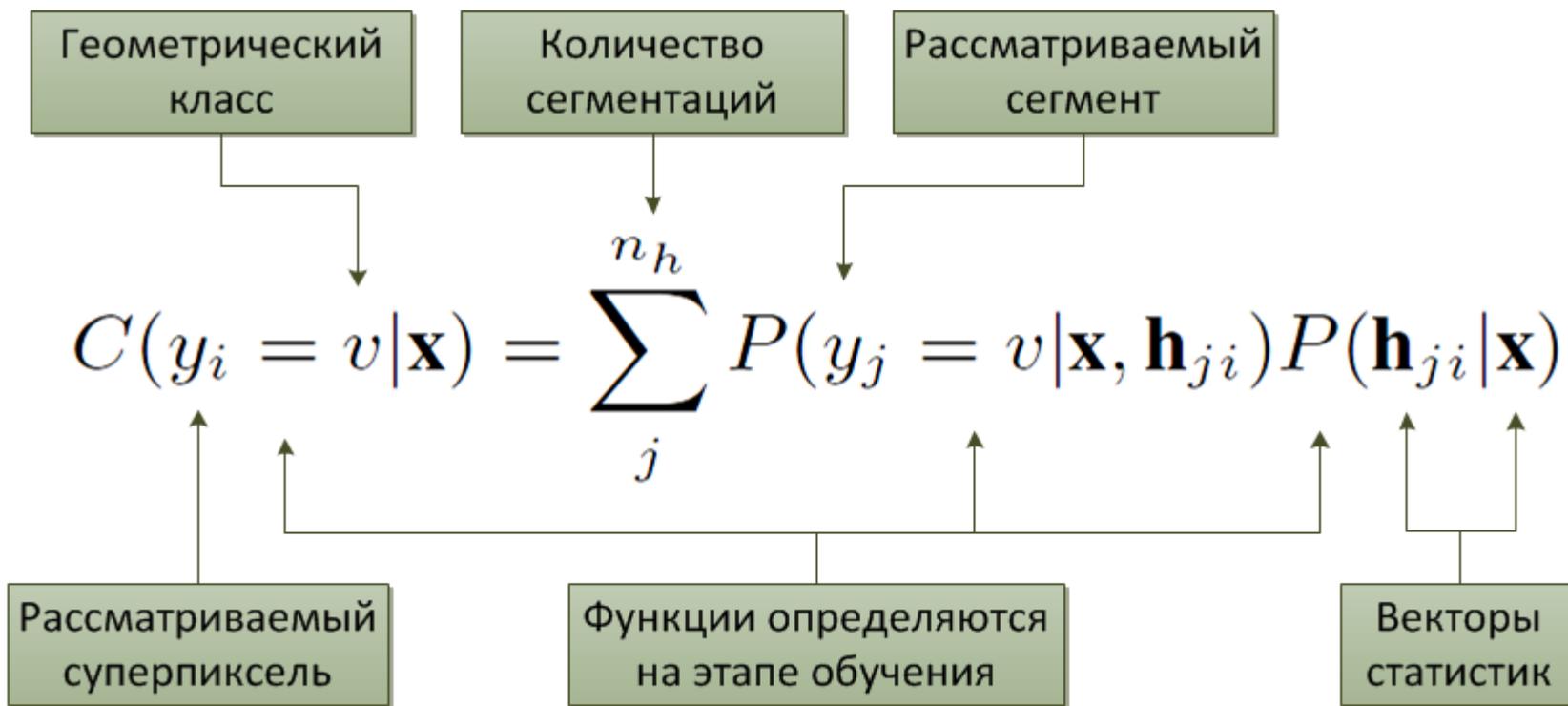


Сегментация по геометрическим
классам

Алгоритм

Распределение геометрических меток (1)

Модель внутри классификатора



Geometric context extraction



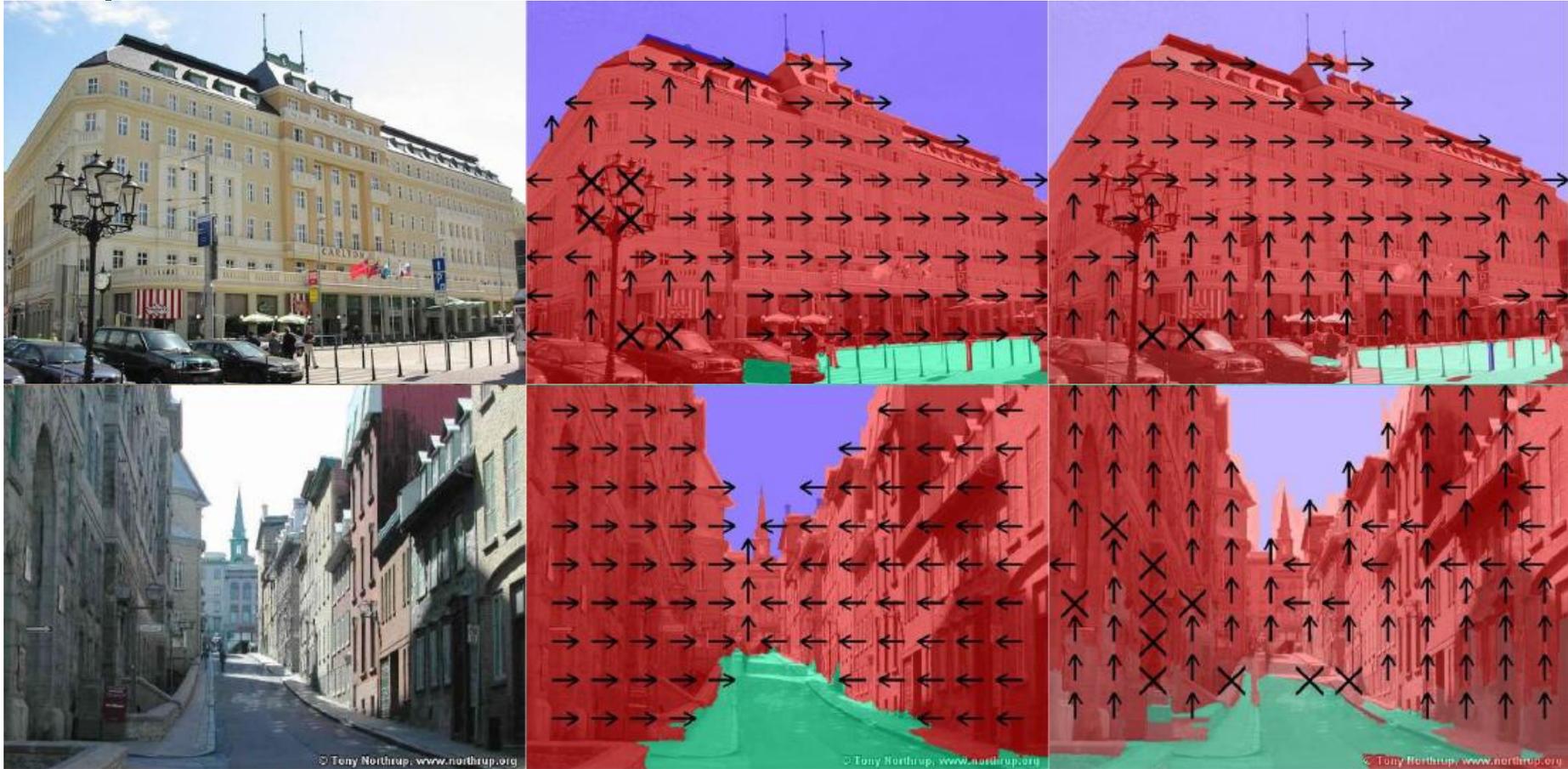
Вычисляемые статистики

Color	16	3D Geometry	35
C1. RGB values: mean	3	G1. Long Lines: total number in region	1
C2. HSV values: C1 in HSV space	3	G2. Long Lines: % of nearly parallel pairs of lines	1
C3. Hue: histogram (5 bins) and entropy	6	G3. Line Intsctn: hist. over 12 orientations, entropy	13
C4. Saturation: histogram (3 bins) and entropy	4	G4. Line Intsctn: % right of center	1
		G5. Line Intsctn: % above center	1
		G6. Line Intsctn: % far from center at 8 orientations	8
		G7. Line Intsctn: % very far from center at 8 orient.	8
		G8. Texture gradient: x and y "edginess" (T2) center	2
Texture	15	Location and Shape	12
T1. DOOG filters: mean abs response of 12 filters	12	L1. Location: normalized x and y, mean	2
T2. DOOG stats: mean of variables in T1	1	L2. Location: norm. x and y, 10 th and 90 th pctl	4
T3. DOOG stats: argmax of variables in T1	1	L3. Location: norm. y wrt horizon, 10 th , 90 th pctl	2
T4. DOOG stats: (max - median) of variables in T1	1	L4. Shape: number of superpixels in region	1
		L5. Shape: number of sides of convex hull	1
		L6. Shape: <i>num pixels/area(convex hull)</i>	1
		L7. Shape: whether the region is contiguous $\in \{0, 1\}$	1

В общей сумме 72 статистики для каждого суперпикселя

Geometric context extraction

Результаты



Source

Ground truth

Result

Geometric context extraction

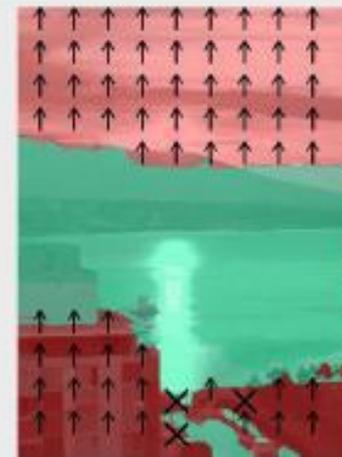
Результаты (failure)



Source



Ground truth



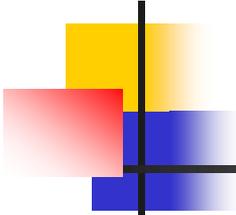
Result

Geometric context Extraction



Выводы

- Достоинства:
 - Уход от семантики объектов в сегментации
 - Возможность использования в качестве confidence map для фильтрации карт глубин
- Недостатки:
 - Требуется объемная тренировочная база
 - Неоднозначность в определении меток даже для ground truth (отмечается в статье)



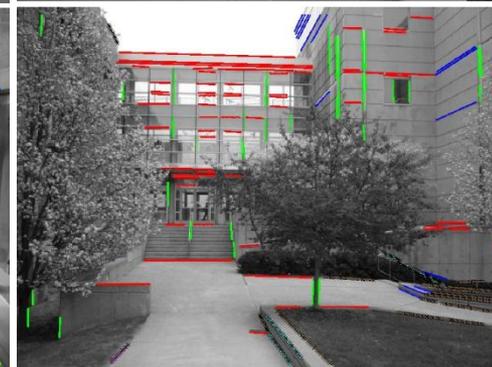
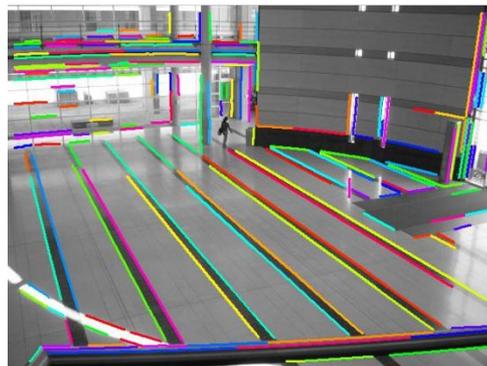
Содержание

- Введение
- Feature-based алгоритмы с обучением
- **Интерпретация геометрических примитивов**
 - Vanishing point detection
 - Geometric reasoning
- Заключение

Vanishing point detection

Алгоритм J-linkage

- На входе – отрезки прямых
- На выходе – классификация линий по основным точкам схода линий



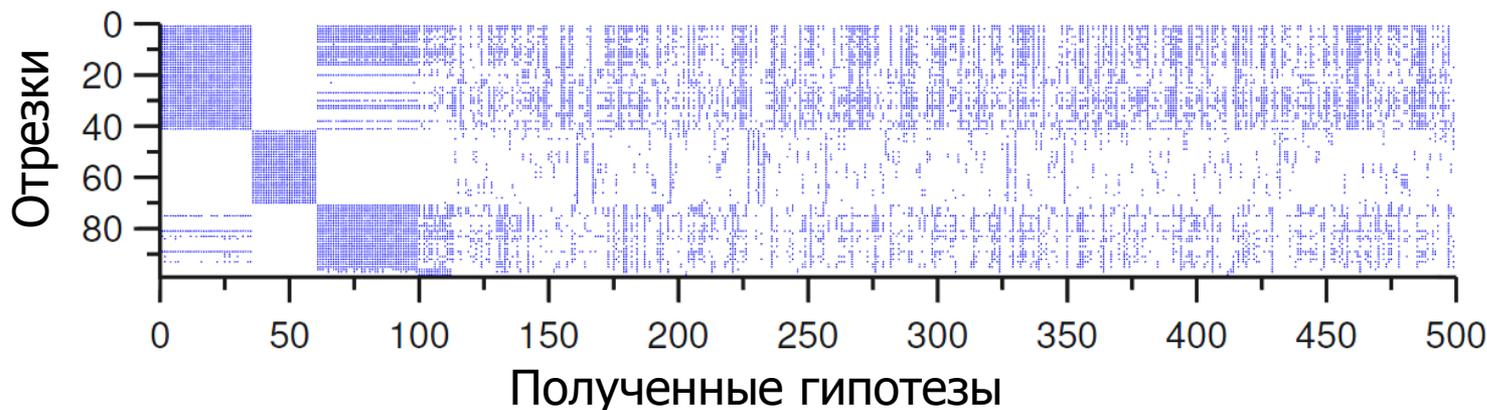
Найденные
отрезки

Классификация
линий
(выделены цветом)

J-linkage

Алгоритм кластеризации

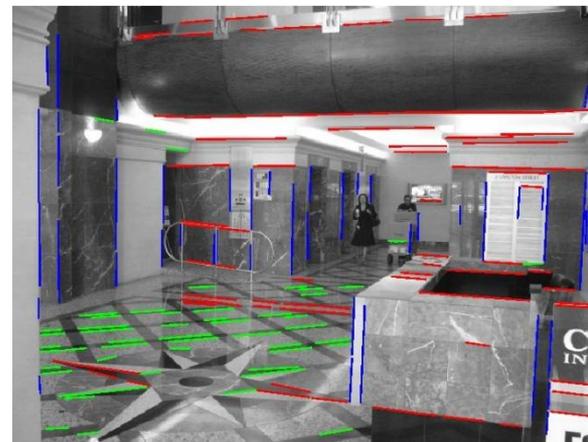
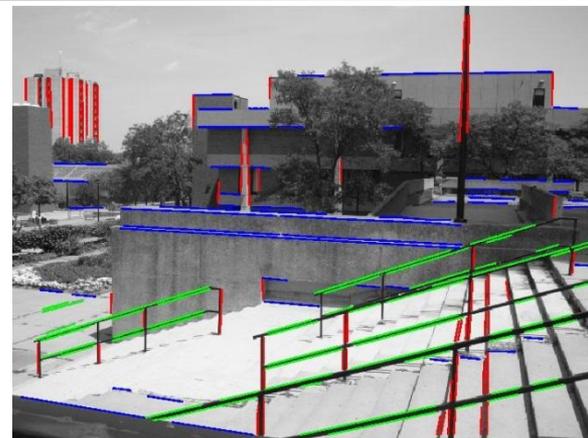
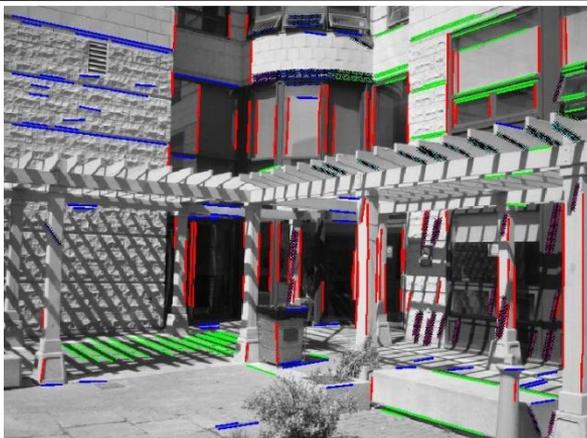
1. Каждая строка матрицы принимается за отдельный кластер
2. Вычисление расстояния между кластерами (количество отличающихся элементов в строках)
3. Слияние наиболее близких кластеров



Характеристическая матрица после кластеризации

J-linkage

Результаты



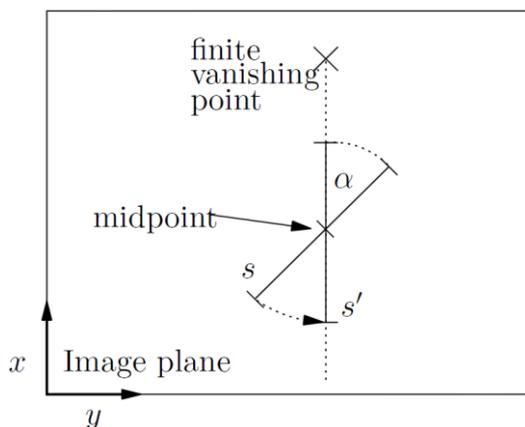
Правильно кластеризованные

Неверно кластеризованные

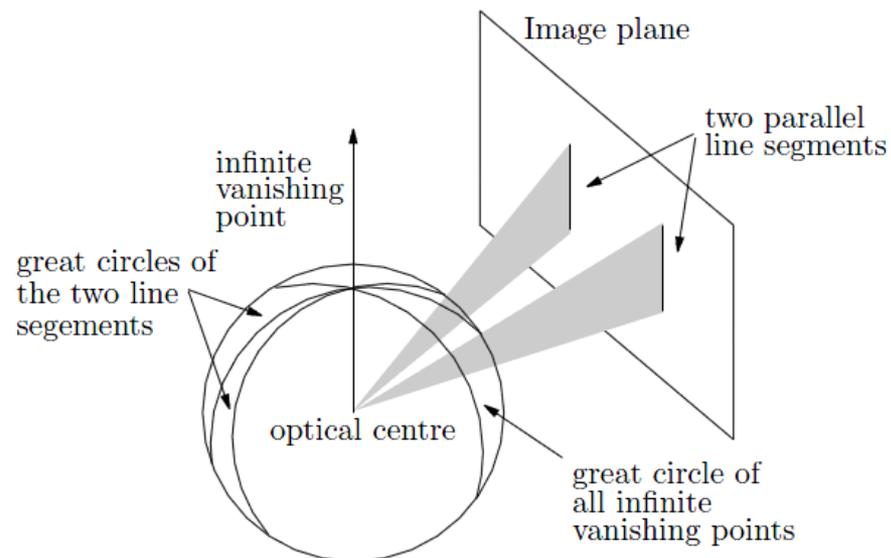
Vanishing point detection

Image plane accumulation

В качестве аккумулирующего пространства может быть выбрана плоскость изображения или гауссова сфера



Отрезок определяется центральной точкой, длиной и углом наклона



Сфера Гаусса

Vanishing point detection

Image plane accumulation

Для каждой точки
аккумулирующей матрицы
каждый отрезок вносит вес:

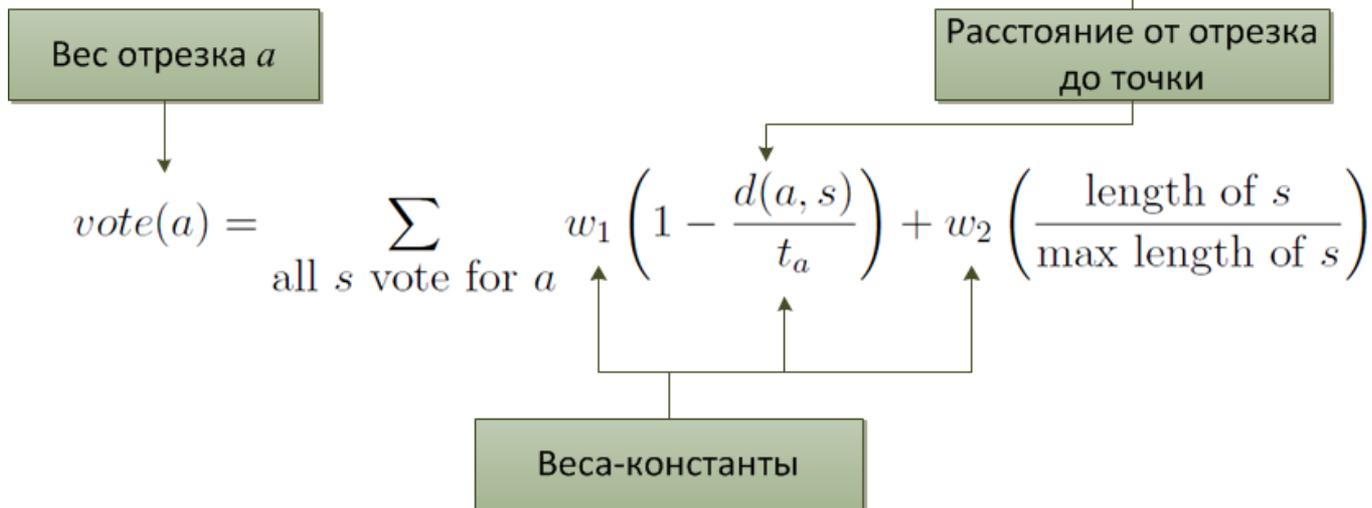
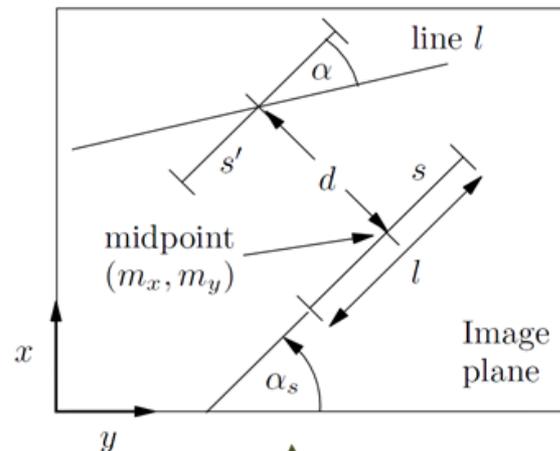
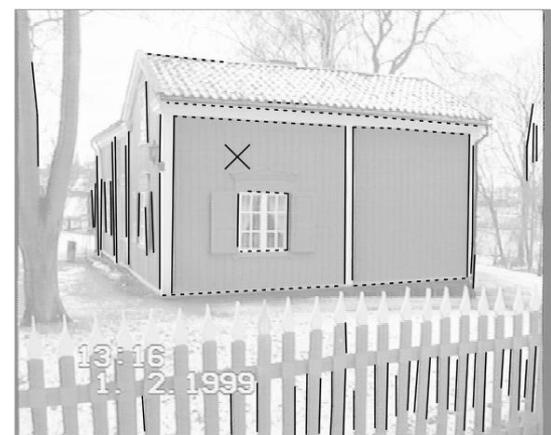
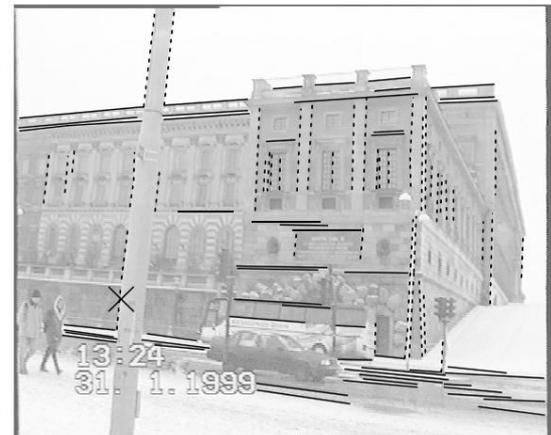
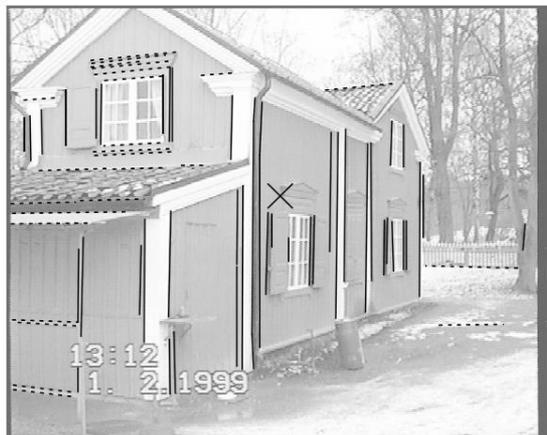
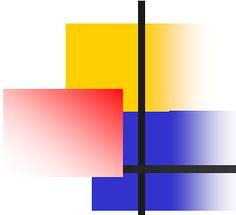


Image plane accumulation

Результаты

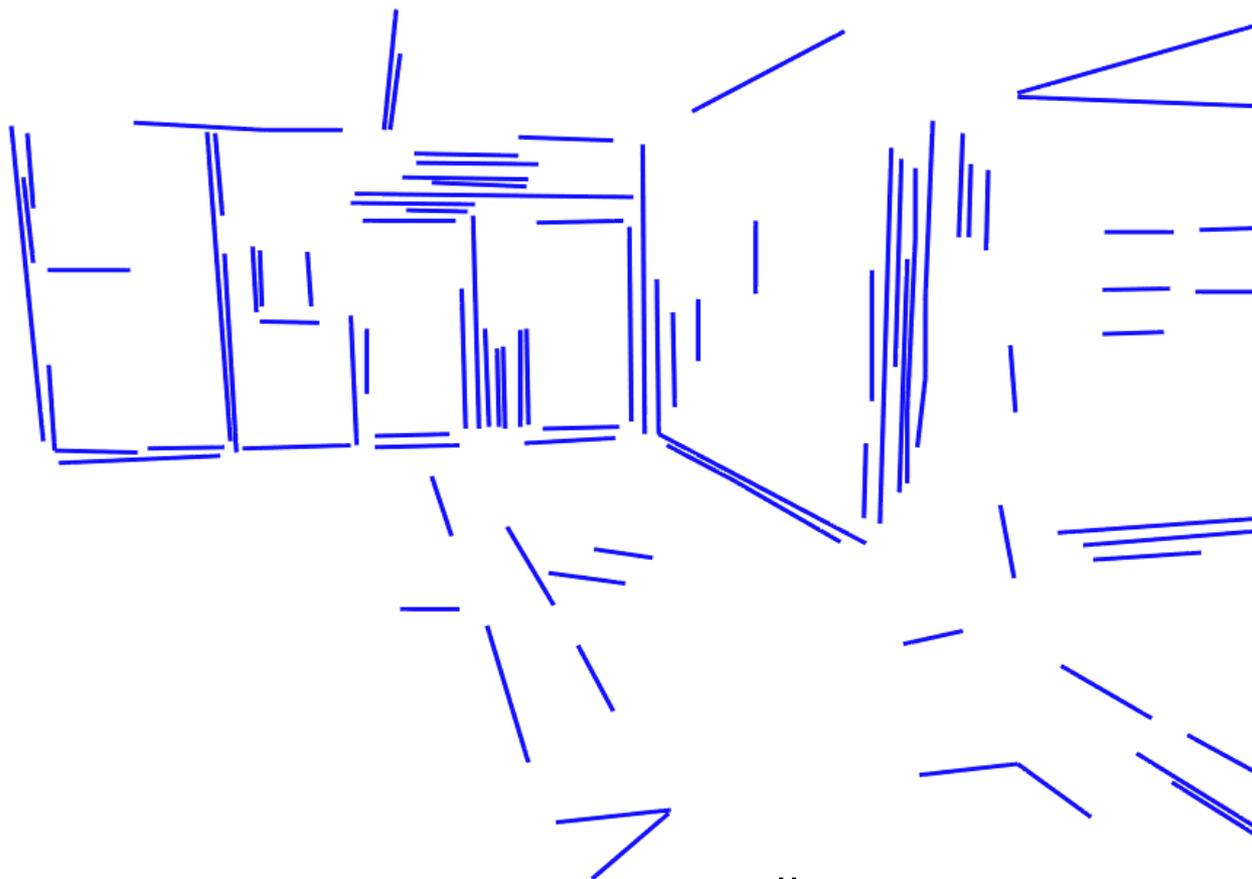




Содержание

- Введение
- Feature-based алгоритмы с обучением
- Интерпретация геометрических примитивов
 - Vanishing point detection
 - **Geometric reasoning**
- Заключение

Geometric reasoning



Отрезки прямых на изображении. Этой информации достаточно, чтобы восстановить геометрию сцены

Geometric reasoning

Алгоритм

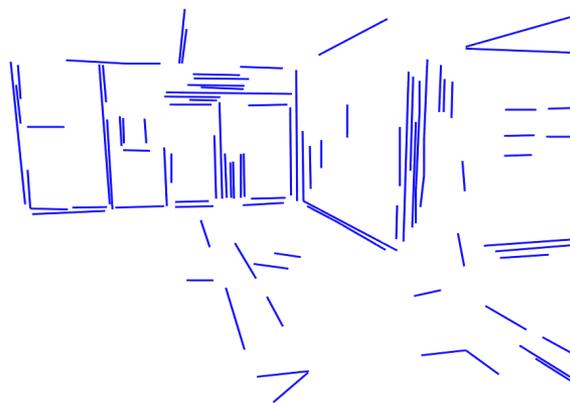
1. Поиск отрезков прямых
2. Вычисление точек схода перспективы и классификация отрезков
3. Построение всех возможных моделей сцены
4. Определение наиболее вероятной модели на основе “supporting line segments”

Line segments and vanishing point detection

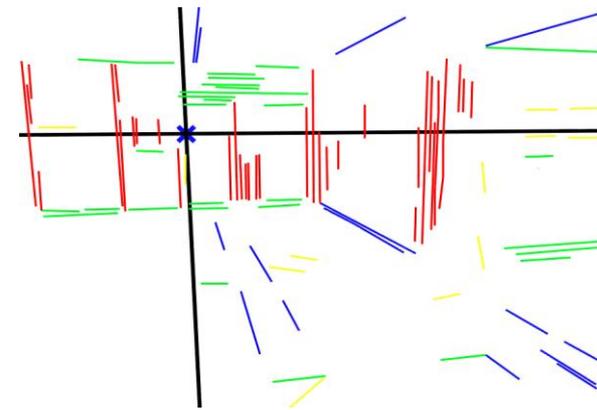
- Используются готовые реализации алгоритмов
- Предполагается наличие трех основных направлений линий (“Manhattan world model”)



Исходное
изображение



Найденные
линии

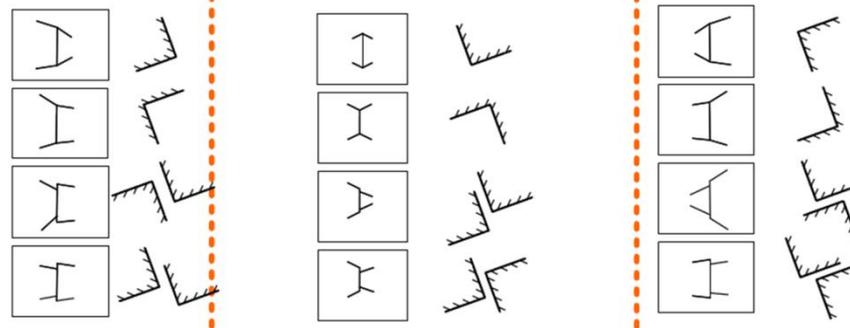
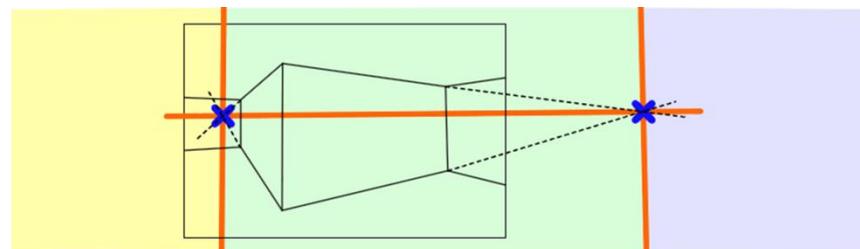


Vanishing point +
три базовых направления

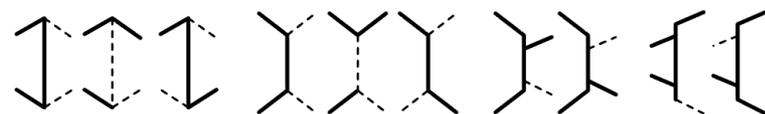
Geometric Reasoning

Построение моделей сцены (1)

- Предполагается существование лишь конечного числа возможных пересечений линий
- Каждое пересечение характеризуется минимальным необходимым количеством линий



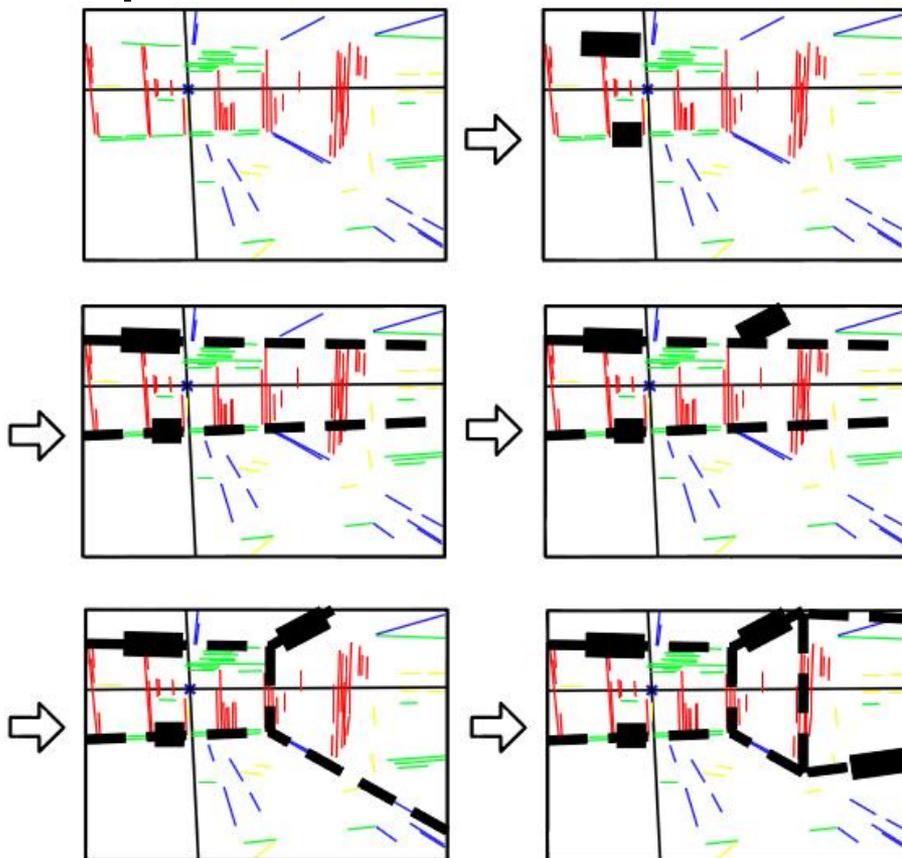
Допустимые гипотезы
о пересечении линий



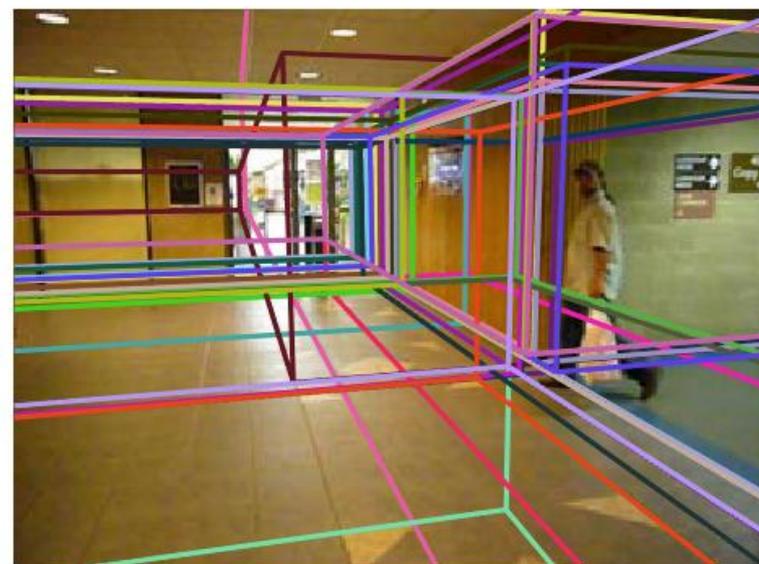
Линии, необходимые
для детектирования пересечения

Geometric Reasoning

Построение моделей сцены (2)



Шаги алгоритма построения
одной из моделей



Множество всех моделей

Построение моделей

Формализация алгоритма

- H – множество моделей сцены
- l_i – отрезок прямой
- H_k – модель с k углами в сцене
- H' – множество моделей,
полученных добавлением одного
угла к исходной модели h
- n – максимально допустимое
количество углов в модели

Algorithm 1 Generating building hypotheses

Set $H_0 \leftarrow \emptyset$, where H_0 is the set of hypotheses with zero corners.

for all pair of line segments (l_i, l_j) **do**
 if l_i above horizon \wedge l_j below horizon \wedge l_i and l_j have
 overlap **then**

 Add scene with no corner (l_i, l_j) to H_0

end if

end for

for $k = 1$ to n , where n is maximum number of corners
in scene **do**

 Set $H_k \leftarrow \emptyset$, where H_k is the set of hypotheses with k
 corners.

for all $h \in H_{k-1}$ **do**

 Find sets of lines that create corners that attaches to
 h and satisfies geometric constraints.

$H' \leftarrow$ Set of all scenes with a new corner attached
 to h

$H_k \leftarrow H_k \cup H'$

end for

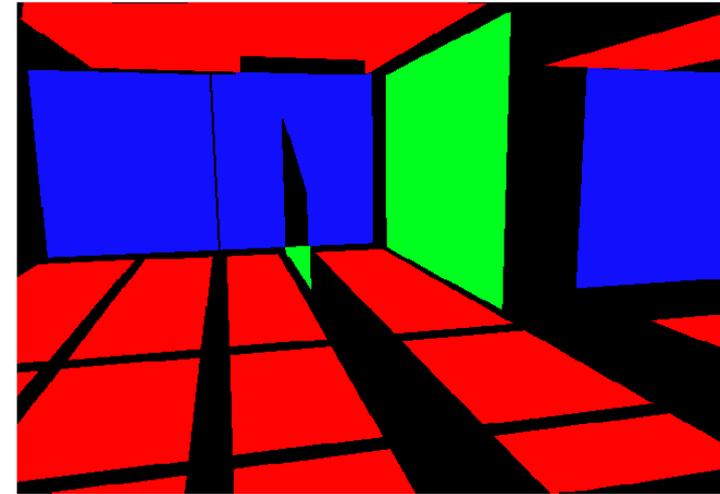
end for

return $H \leftarrow H_0 \cup H_1 \cup \dots \cup H_n$

Определение наиболее вероятной модели

Построение "supporting map":

- Каждая линия в собственной окрестности поддерживает некоторую плоскость
- Две перпендикулярные линии однозначно определяют плоскость в своей окрестности
- В каждой точке определяются вероятности различной ориентации плоскости

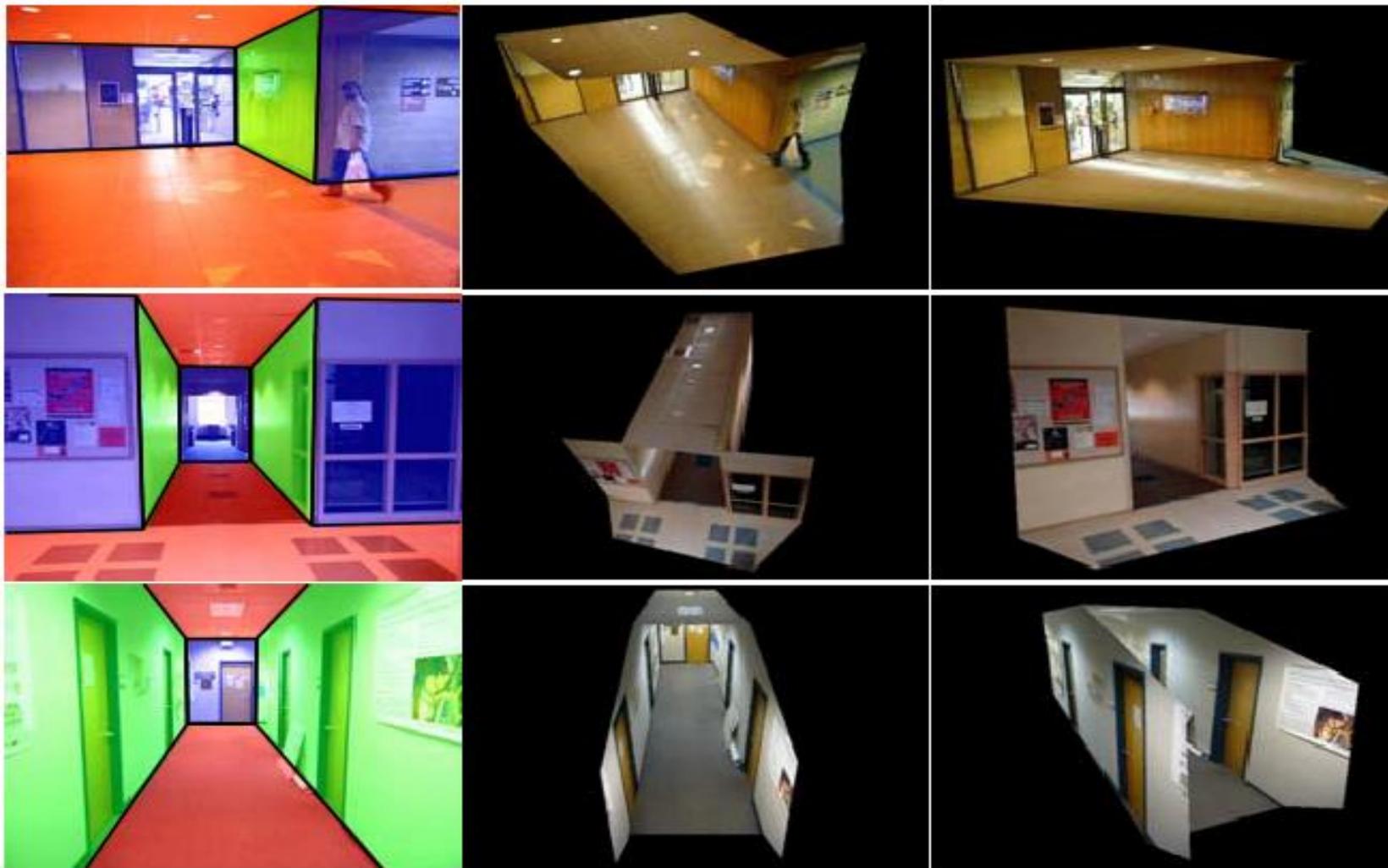


"Supporting map"

Черные области означают конфликт опорных линий

Geometric reasoning

Результаты

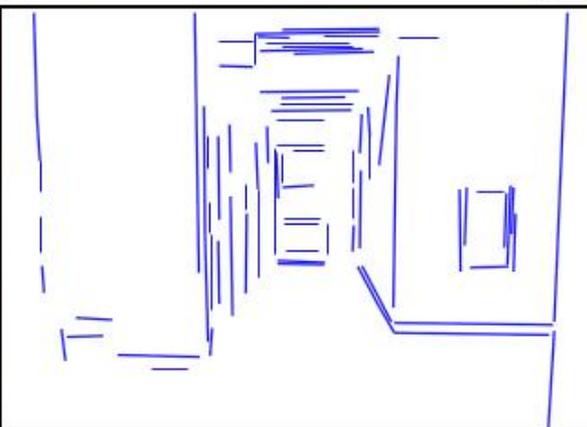
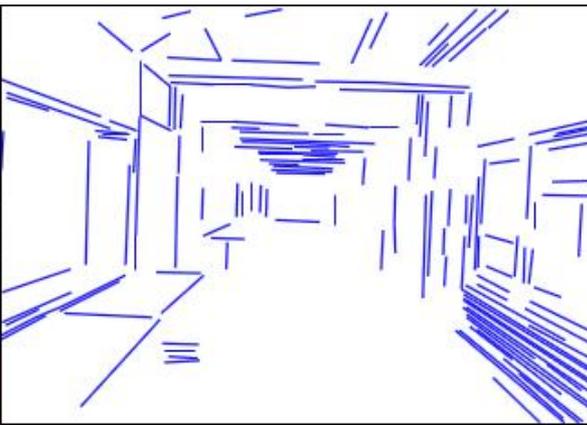


Модель сцены

3D-представление сцены

Результаты

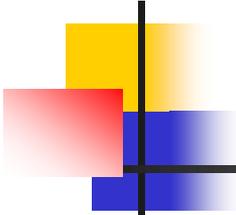
Ошибочные определения моделей



Исходное изображение

Найденные прямые

Модель сцены



Содержание

- Введение
- Feature-based алгоритмы с обучением
- Интерпретация геометрических примитивов
- **Заключение**
 - Собственные наработки

Depth from geometry

Алгоритм

1. Границы – алгоритм Canny edge detection
2. Производные dx, dy – свертка с ядром Собеля
3. Параметры прямых – аккумулялирование голосов точек маски границ

$$y = kx + b$$

$$k(x, y) = \frac{dx}{dy}$$

$$b(x, y) = y - xk(x, y)$$

$$ACC_matr(k, b) ++$$



Найденные прямые

Depth from geometry

Алгоритм (1)

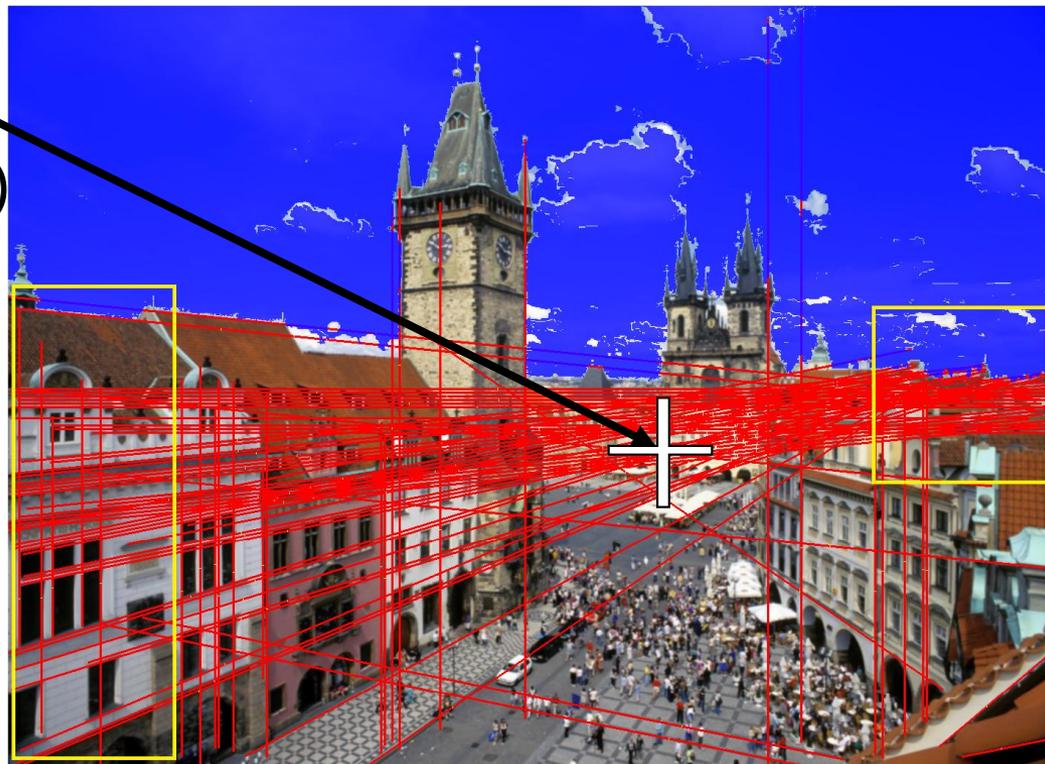
- Сегментация неба по цвету
- Отрезок линии – объединение всех проголосовавших точек данной линии



Depth from geometry

Алгоритм (2)

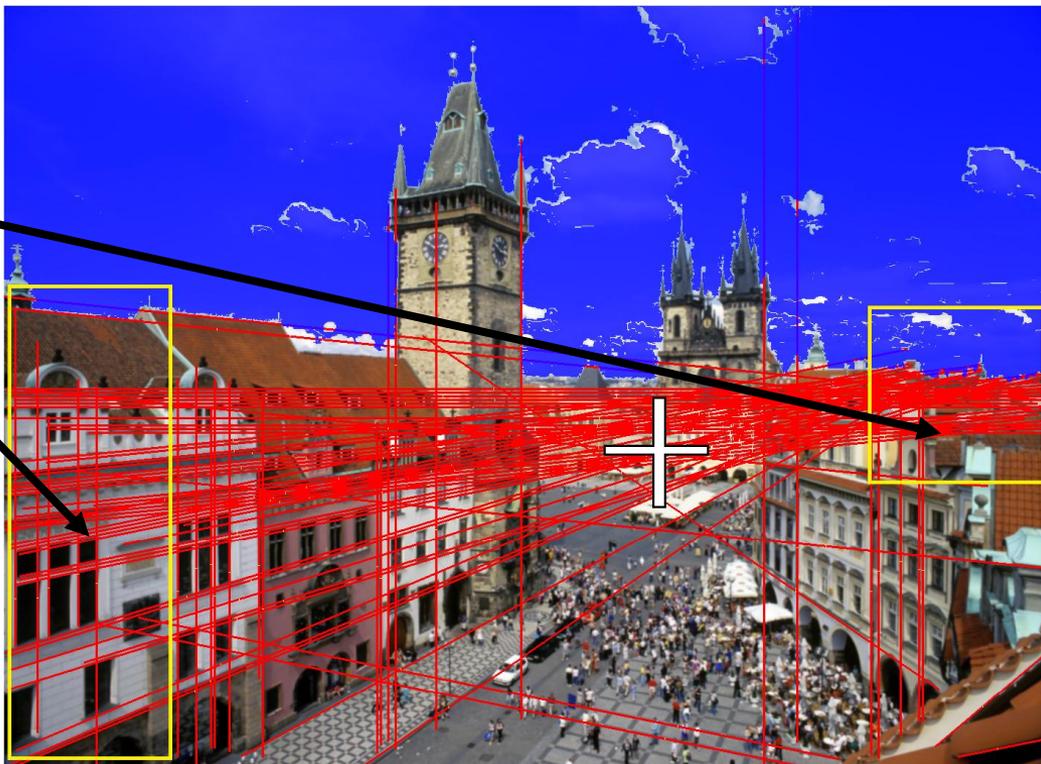
Определение центра
скопления линий
(минимизация невязки)



Depth from geometry

Алгоритм (3)

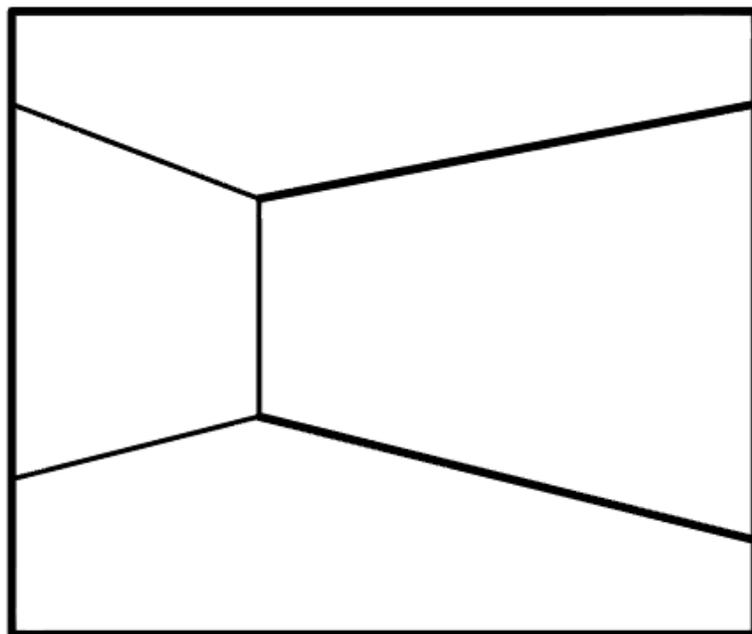
Определение крайних областей пучков



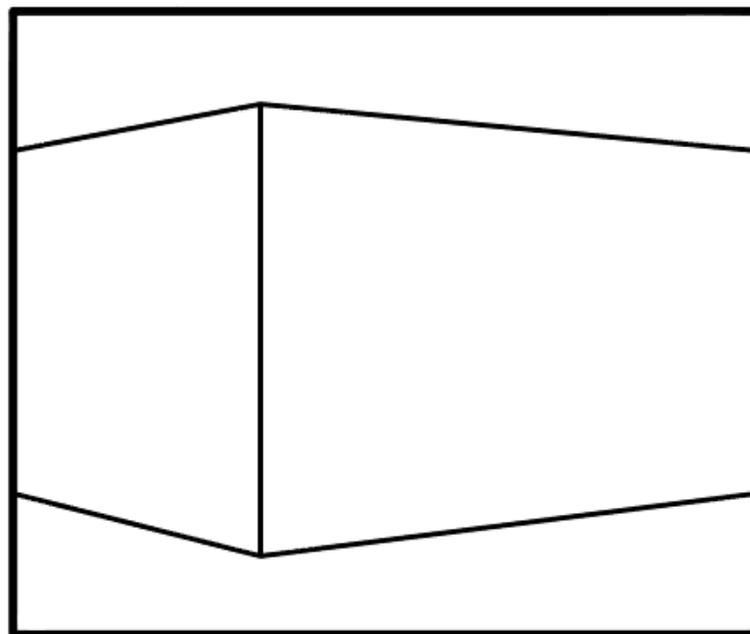
Depth from geometry

Алгоритм (4)

Определение профиля сцены на основе кучности линий в трех основных областях



«Сходящаяся к центру»



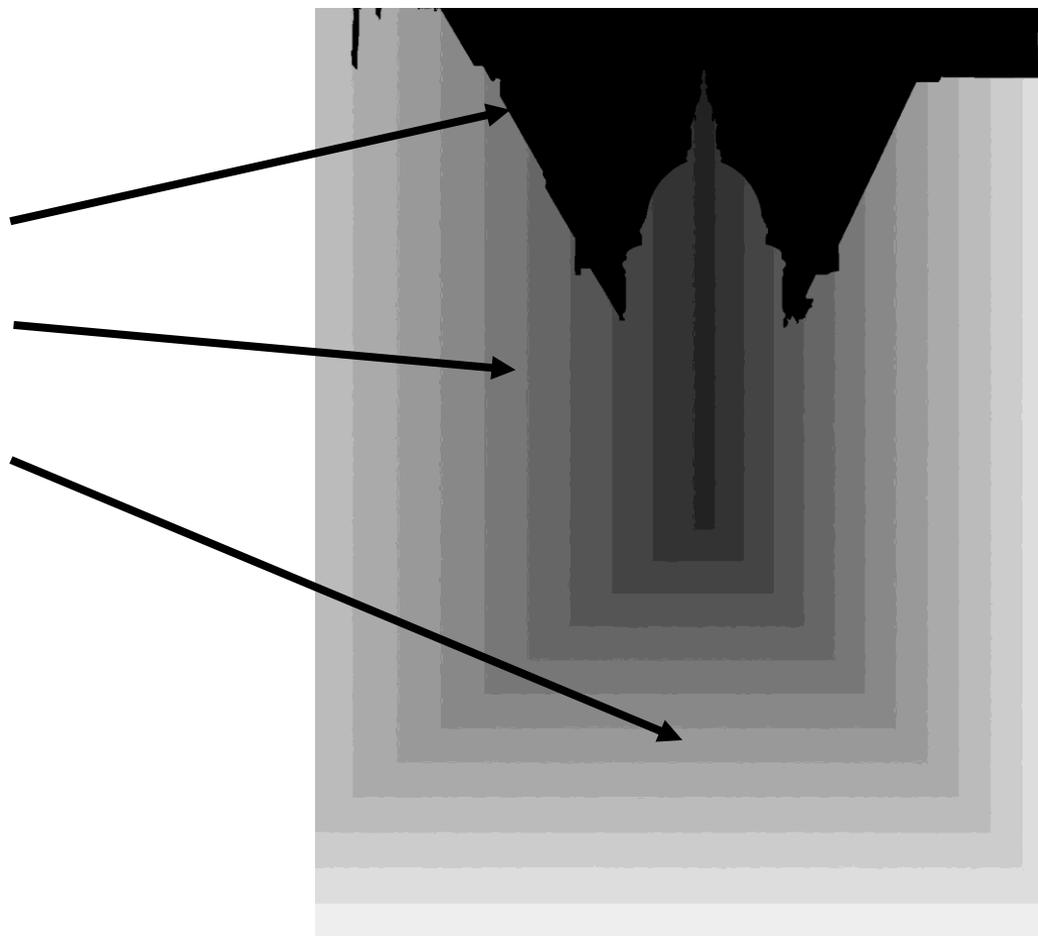
«Расходящаяся от центра»

Depth from geometry

Алгоритм (5)

Объединение:

- Маска неба
- Схема сцены
- Главный градиент плоскости земли



Depth from geometry

Выводы

- **Достоинство:**
Правдоподобная карта глубины на определенной выборке изображений
- **Недостатки:**
 - Излишне высокая дискретизация параметрического пространства
 - Излишнее упрощение отрезков прямых
 - Нахождение только одной точки схода перспективы
 - Простые эвристики для принятия решений

Литература

1. D. Hoiem, A.A. Efros, and M. Hebert, "Recovering Surface Layout from an Image", *IJCV*, Vol. 75, No. 1, October 2007
2. D. Hoiem, A.A. Efros, and M. Hebert, "Geometric Context from a Single Image", *ICCV*, 2005
3. D. C. Lee, M. Hebert, and T. Kanade, "Geometric Reasoning for Single Image Structure Recovery", *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2009
4. J. Tardif, "Non-Iterative Approach for Fast and Accurate Vanishing Point Detection", in *Proc. ICCV*, 2009, pp.1250-1257
5. C. Rother, "A New approach for Vanishing Point Detection in Architectural Environments", in *Proc. Computational Vision and Active Perception Laboratory (CVAP)*, 2000

Лаборатория компьютерной графики и мультимедиа



Видеогруппа — это:

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