



# Метод полуавтоматической контекстной разметки сцены

Егоров А. И., Вишняков Б. В.

ФГУП ГосНИИАС, Москва

ИОИ 2014, Москва

# Задача контекстной разметки сцены



Под контекстной разметкой обычно понимается сегментация изображения, которая заключается в присвоении каждому пикселю некой категории или метки, которая однозначно связывала бы его с каким-либо объектом или областью сцены.

1. Предобработка начального изображения
2. Сегментация изображения – на основе цветовых и контурных характеристик объектов сцены
3. Маркировка сегментированного изображения – на основании каких-либо априорных данных (разметка оператора, дополнительная информация о сцене и т.д.)
4. Составление списка описаний объектов – на основании маркировки сегментированного изображения, а также пространственных и характеристических описаний сегментов

# Алгоритм метода

В работе предлагается метод полуавтоматической контекстной разметки сцены, организованной по итерационному принципу:

1. На вход принимается предварительная разметка (например, ручная разметка оператором)
2. Выполняется сегментация сцены с последующей маркировкой сегментов
3. Оператор проверяет и при желании исправляет возможные ошибки разметки и маркировки
4. Далее разметка может быть снова отправлена на уточнение предложенным методом, либо объявлена окончательной

# Предобработка исходного изображения

Предобработка изображения позволяет не только повысить эффективность сегментации, но и выявить текстурные характеристики объектов, которые в дальнейшем будут использоваться в сегментации

1. Размытие исходного изображения низкочастотным фильтром
2. Получение контуров полутонового изображения с использованием оператора Собеля
3. Подавление не максимумов контуров полутонового изображения
4. Пороговая фильтрация контуров полутонового изображения
5. Морфологическое восстановление разорванных контуров



# Размытие исходного изображения и получение контуров полутонового изображения

$$G_1 = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} * G_0$$

Целочисленный дискретный  
Гауссовский фильтр

---

$$G_x^s = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * G_1$$

$$G_y^s = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * G_1$$

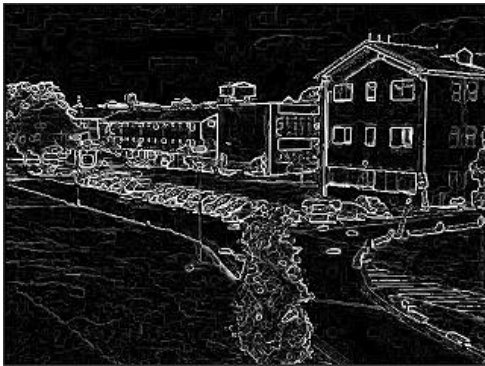
Дискретный оператор Собеля:

- Функции приближённых производных яркостей по осям  $x$  и  $y$
- Функция модулей и направлений градиентов яркости

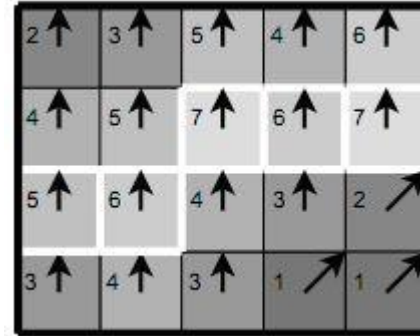
$$G^s(x, y) = \sqrt{G_x^{s2}(x, y) + G_y^{s2}(x, y)},$$

$$\Theta^s(x, y) = \arctan\left(\frac{G_y^s(x, y)}{G_x^s(x, y)}\right)$$

# Обработка контуров полутонового изображения



1. Подавление не максимумов производится для того, чтобы каждая линия маски контуров имела ширину в один пиксель по следующему принципу:



2. Пороговая фильтрация для удаления слабых градиентов, наиболее подверженных шумам или являющимися таковыми
3. Морфологическое восстановление разорванных контуров:

$$G^m = G^f \bullet B = (G^f \oplus B) \ominus B$$

где:

- – морфологическая операция закрытия
- $\oplus$  – морфологическая операция расширения
- $\ominus$  – морфологическая операция сжатия

# Сегментация сцены (первый этап)



Сегментация производится на основе цветовых и текстурных характеристик изображения в два этапа:

## 1. Первоначальная сегментация

$$d(x, y, S_i) = |\bar{P}_1(x, y) - \bar{S}_i^{color}| + \bar{G}^m(x, y) \\ i = 1..N_s$$

$$\bar{P}_1(x, y) = \begin{pmatrix} H(G_1(x, y)) \\ S(G_1(x, y)) \end{pmatrix}$$

$$\bar{S}_i^{color} = \begin{pmatrix} H(S_i^{color}) \\ S(S_i^{color}) \end{pmatrix}$$



На данном примере исходное изображение размером 640x480 пикселей было разбито на около 10000 первоначальных сегментов

# Сегментация сцены (второй этап)



## 2. Удаление малых сегментов

Большое количество сегментов плохо поддаётся анализу, поэтому малые сегменты объединяются с соседом, с которым у него будет минимальное характеристическое расстояние





# Маркировка сегментированной сцены



- Исходная неточная разметка, поступившая от оператора



Маркировка производится по мажоритарному принципу на сравнении количества попавших пикселей разметки от оператора в каждый сегмент сцены

# Результаты работы алгоритма



- |   |        |   |        |
|---|--------|---|--------|
|  | Трава  |  | Здание |
|  | Дерево |  | Небо   |
|  | Дорога |  | Другое |

# Выводы

В работе рассмотрен метод полуавтоматической контекстной разметки сцены, основанный на итерационном принципе:

- Реализация алгоритма на C++, без использования сторонних библиотек. Реализация программного обеспечения оператора на WPF
- Время работы на изображении 640x480 на персональном компьютере ~300мс

Дальнейшие пути развития:

- Использование априорных данных о характеристиках объектов (плавность дорог, небо ниже дороги)
- Интеллектуализация алгоритма – подсказка ошибок оператора
- Сравнение с существующими методами

Спасибо за внимание