LSD: a Line Segment Detector

Тарасов Михаил Николаевич
tarasov.mn@phystech.edu

Московский физико-технический институт

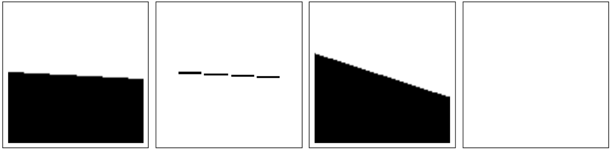
23 апреля 2023 г.

Содержание

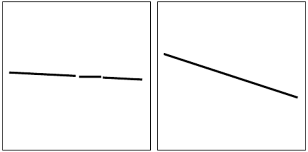
Алгоритм

1. Масштабирование изображения
2. Вычисление градиента
3. Псевдосортировка градиентов
4. Порог градиента
5. Разрастание региона
6. Прямоугольное приближение
7. Расчет NFA
8. Плотность выровненных точек
9. Улучшение прямоугольника
10. Вычислительная сложность

Масштабирование изображения



Обнаружения без масштабирования.



Обнаружения с использованием масштабирования.

Вычисление градиента

Градиент вычисляется для каждого пикселя с помощью маски 2x2:

$i(x,y)$	$i(x+1,y)$
$i(x,y+1)$	$i(x+1,y+1)$

где $i(x,y)$ – значение серого на пикселе (x,y), градиент пикселя вычисляется как

$$g_x(x,y) = \frac{i(x+1,y) + i(x+1,y+1) - i(x,y) - i(x,y+1)}{2},$$

$$g_y(x,y) = \frac{i(x,y+1) + i(x+1,y+1) - i(x,y) - i(x+1,y)}{2}.$$

Угол линии уровня вычисляется как

$$\arctan\left(\frac{g_x(x,y)}{-g_y(x,y)}\right)$$

и величина градиента как

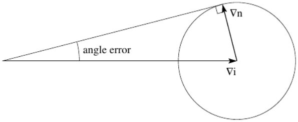
$$G(x,y) = \sqrt{g_x^2(x,y) + g_y^2(x,y)}$$

Псевдосортировка градиентов

Возможна псевдосортировка пикселей за линейное время. Пиксели классифицируются в 1024 контейнера в соответствии со значением их градиента. Сначала LSD использует пиксели из контейнера с наибольшим значением градиента; затем берет пиксели из второго контейнера и так далее. 1024 контейнеров достаточно для почти полной сортировки значений градиента, когда значения серого являются целыми числами из отрезка $[0, 255]$.

Порог градиента

В LSD рассматриваются только пиксели с величиной градиента больше ρ. Предполагая шум квантования n и идеальное изображение i, получаем: $\tilde{i} = i + n \quad \nabla\tilde{i} = \nabla i + \nabla n$.

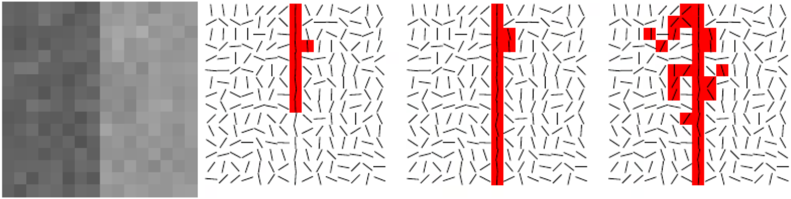


Оценка угловой ошибки из-за шума квантования.

Имеем $|\text{ angle error }| \leq \arcsin\left(\frac{q}{|\nabla i|}\right)$, где q – граница $|\nabla n|$.
Критерий состоит в том, чтобы отбрасывать пиксели, в которых угловая ошибка больше, чем угол допуска τ. То есть накладываем $|\text{angle error}| \leq \tau$ и получаем

$$\rho = \frac{q}{\sin\tau}$$

Разрастание региона



Примеры роста области, начиная со среднего верхнего пикселя для трех значений допуска угла.
Слева направо: изображение; $\tau = 11.25$; $\tau = 22.5$; $\tau = 45$.

Прямоугольное приближение

Центр прямоугольника (c_x, c_y):

$$c_x = \frac{\sum_{j \in \text{ Region}} G(j) \cdot x(j)}{\sum_{j \in \text{ Region}} G(j)}$$

$$c_y = \frac{\sum_{j \in \text{ Region}} G(j) \cdot y(j)}{\sum_{j \in \text{ Region}} G(j)}$$

где $G(j)$ – величина градиента пикселя j. Угол прямоугольника устанавливается равным углу собственного вектора, связанного с наименьшим собственным значением матрицы $M = \begin{pmatrix} m^{xx} & m^{xy} \\ m^{xy} & m^{yy} \end{pmatrix}$.

$$m^{xx} = \frac{\sum_{j \in \text{ Region}} G(j) \cdot (x(j) - c_x)^2}{\sum_{j \in \text{ Region}} G(j)}$$

$$m^{xy} = \frac{\sum_{j \in \text{ Region}} G(j) \cdot (x(j) - c_x)(y(j) - c_y)}{\sum_{j \in \text{ Region}} G(j)}.$$

Расчет NFA

Количество ложных срабатываний (NFA), связанных с прямоугольником r:

$$\mathrm{NFA}(r) = (NM)^{5/2}\gamma \cdot B(n,k,p)$$

где N и M – количество столбцов и строк изображения,

$$B(n,k,p) = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j}$$

Для каждого прямоугольника с заданной точностью p подсчитываются k и n, а затем вычисляется NFA

$$\mathrm{NFA}(r) = (NM)^{5/2}\gamma \cdot \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j}.$$

Прямоугольники с $\mathrm{NFA}(r) \le \varepsilon$ проверяются как обнаружения.

Плотность выравненных точек

Метод τ-угловой терпимости может дать неправильную интерпретацию изображения, особенно если две прямые грани образуют меньший угол, чем допустимая τ. В таких случаях в алгоритме LSD обнаруживают проблемные области и разделяют их на две меньшие области. При обнаружении проблемы плотность выровненных точек в прямоугольнике становится низкой, и если плотность меньше заданного порога, то прямоугольник отклоняется. Для решения проблемы используются два метода: уменьшение угловой терпимости и уменьшение радиуса региона. При этом плотность рассчитывается как отношение количества выровненных точек к площади прямоугольника, и если она меньше заданного порога, то прямоугольник отклоняется.

Улучшение прямоугольника

Процедура улучшения прямоугольника с помощью LSD состоит из следующих шагов:

1. попробуйте более высокую точность
2. попробуйте уменьшить ширину
3. попробуйте уменьшить одну сторону прямоугольника
4. попробуйте уменьшить другую сторону прямоугольника
5. попробуйте еще более высокую точность

Если найден значимый прямоугольник (NFA $\leq \varepsilon$), процедура улучшения остановится после шага, на котором он был найден.

Вычислительная сложность

Сначала выполняется субдискретизация Гаусса и вычисление градиента изображения, что также может быть выполнено за время, пропорциональное числу пикселей в изображении. Затем пиксели псевдо-упорядочиваются с помощью классификации в бины, что может быть сделано за линейное время. Время выполнения алгоритма нахождения региона-поддержки линейно зависит от числа посещенных пикселей, которое остается пропорциональным общему числу пикселей в изображении. Обработка разделяется на два типа задач: первый тип пропорционален общему числу пикселей во всех регионах, второй тип пропорционален числу регионов. В итоге LSD алгоритм имеет вычислительную сложность, пропорциональную числу пикселей в изображении.

Принцип Гельмгольца

Не должно быть обнаружений на изображении шума.

contrario approach

contrario approach (противоположный подход) – это метод, который заключается в определении объекта путем исключения фона, то есть выделения объекта как всего, что не является фоном.

Теорема 1

Формулировка

$$E_{H_0}\left[\sum_{r\in\mathcal{R}} \mathbb{1}_{NFA(r,I)<\varepsilon}\right] \leq \varepsilon$$

где E – оператор ожидания, 1 – индикаторная функция, R – множество рассматриваемых прямоугольников, I – случайное изображение на H_0.

Теорема утверждает, что среднее число ε-значащих прямоугольников в модели a contrario H_0 меньше ε. Таким образом, количество обнаружений на шуме контролируется ε, и его можно сделать сколь угодно малым. Другими словами, это показывает, что LSD удовлетворяет принципу Гельмгольца.

Теорема 1

Доказательство

Мы определяем $\hat{k}(r)$ как

$$\hat{k}(r) = \min\left\{\kappa \in \mathbb{N}, P_{H_0}[k(r,I) \geq \kappa] \leq \frac{\varepsilon}{(NM)^{5/2}\gamma}\right\}$$

Тогда, $\mathrm{NFA}(r,i) \leq \varepsilon$ эквивалентно $k(r,i) \geq \hat{k}(r)$. Теперь,

$$E_{H_0}\left[\sum_{r\in\mathcal{R}} \mathbb{1}_{\mathrm{NFA}(r,I)\leq\varepsilon}\right] = \sum_{r\in\mathcal{R}} P_{H_0}[\mathrm{NFA}(r,I) \leq \varepsilon] = \sum_{r\in\mathcal{R}} P_{H_0}[k(r,I) \geq \hat{k}(r)].$$

Но, по определению $\hat{k}(r)$ мы знаем

$$P_{H_0}[k(r,I) \geq \hat{k}(r)] \leq \frac{\varepsilon}{(NM)^{5/2}\gamma}$$

и используя $\#\mathcal{R} = (NM)^{5/2}\gamma$, мы получаем

$$E_{H_0}\left[\sum_{r\in\mathcal{R}} \mathbb{1}_{\mathrm{NFA}(r,I)\leq\varepsilon}\right] \leq \sum_{r\in\mathcal{R}} \frac{\varepsilon}{(NM)^{5/2}\gamma} = \varepsilon$$

Преимущества и недостатки алгоритма

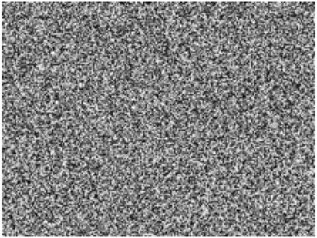
Преимущества:

- Высокая скорость
- Высокая точность
- Параметры настраиваются

Недостатки:

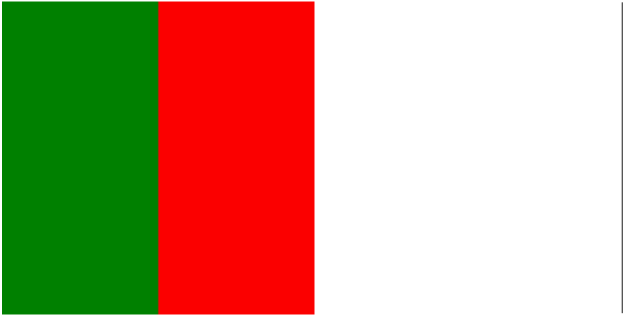
- Чувствительность к яркости
- Чувствительность к разрешению
- Неточность на некоторых типах линейных структур

Примеры работы алгоритма



Шум 220 x 165 пикселей

Примеры работы алгоритма

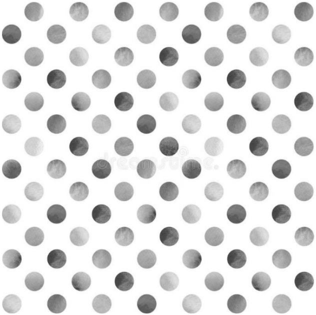


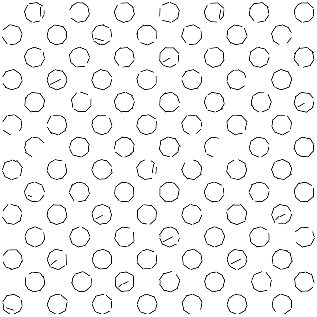
Цвета 518 x 517 пикселей

Примеры работы алгоритма

Небо 612 x 408 пикселей

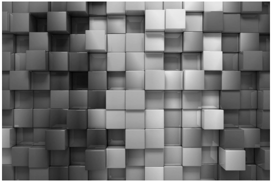
Примеры работы алгоритма

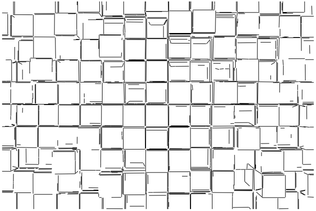




Круги 578 x 578 пикселей

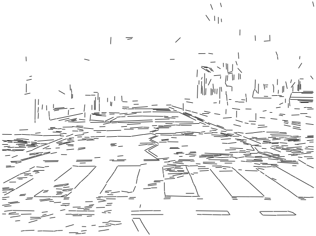
Примеры работы алгоритма





Квадраты 1620 x 1080 пикселей

Примеры работы алгоритма



Улица 935 x 701 пикселей

Спасибо за внимание!