

Санкт-Петербургский политехнический университет Петра Великого
Институт прикладной математики и механики
Высшая школа прикладной математики и вычислительной физики

Научно-исследовательская работа

"Выделение границ объектов на биологических изображениях с
помощью вейвлетного преобразования"

Работу
выполнила:
студентка I курса
магистратуры
Добрецова Е.В.
Группа:
5040102/40101
Преподаватель:
Козлов К.Н.

Санкт-Петербург
2025

1. Введение

Обнаружение краев изображений является одной из фундаментальных задач компьютерного зрения, обеспечивающая упрощенное представление сцены для таких задач, как обнаружение цели, сегментация и 3D реконструкция. Но эффективность обнаружения краёв падает из-за влияния фонового шума, в связи с чем появилась необходимость разработки нового, устойчивого к шуму метода. Одним из таких методов стал алгоритм обнаружения краёв на основе локально-оптимального сплайн-вейвлета (LOSW-ED алгоритм), основанный на возникшем в конце 1990-х годов вейвлетном анализе. Алгоритм LOSW-ED сочетает в себе модульные максимумы сплайн-вейвлетов, морфологическую фильтрацию и анализ структурной неопределенности (Structural Uncertainty-Aware Modulus Maxima) - дополнительный фильтр, оценивающий уровень неопределённости и адаптивно усиливающий или подавляющий края с использованием морфологических операций.

2. Методология

2.1. Локально оптимальный сплайн-вейвлет (LOSW)

2.1.1. Алгоритм локально-оптимального сплайна

LOSW строится на основе локально оптимального сплайна $L(t)$, который определяется как линейная комбинация сдвинутых кубических В-сплайнов:

$$L(t) = \left(\frac{9}{5}\delta_0 - \frac{19}{15}\delta_1 + \frac{8}{15}\delta_{3/2} - \frac{1}{15}\delta_2 \right) * \beta_3(t),$$

где $\beta_3(t)$ – кубический В-сплайн, а δ_i – оператор сдвига.

Функция $\beta_3(t)$ определяется как:

$$\beta_3(t) = \sum_{i=0}^4 \frac{(-1)^i}{4!} \binom{4}{i} \left(t + \frac{n}{2} - i\right)^{n-1} \cdot \mu\left(t + \frac{n}{2} - i\right),$$

где $\mu(t)$ – ступенчатая функция:

$$\mu(t) = \begin{cases} 0, & t < 0, \\ 1, & t \geq 0. \end{cases};$$

а $\delta_i(t)$ - оператор сдвига:

$$\delta_i * \beta_3(t) = \frac{(\beta_3(t-i) + \beta_3(t+i))}{2}$$

Эквивалентное выражение уравнения для $L(t)$:

$$L(t) = \frac{9}{5}\beta_3(t) - \frac{19}{30}\beta_3(t+1) + \frac{4}{15}\beta_3\left(t + \frac{3}{2}\right) - \frac{1}{30}\beta_3(t+2) - \frac{19}{30}\beta_3(t-1) + \frac{4}{15}\beta_3\left(t - \frac{3}{2}\right) - \frac{1}{30}\beta_3(t-2),$$

$L(t)$ может унаследовать все полезные свойства $\beta_3(t)$, включая аналитичность, центральную симметрию, локальную поддержку и гладкость высокого порядка.

Фурье-образ функции $L(t)$:

$$\hat{L}(\omega) = \left(\frac{9}{5} - \frac{19}{15} \cos \omega + \frac{8}{15} \cos \frac{3\omega}{2} - \frac{1}{15} \cos 2\omega \right) \left(\frac{\sin \frac{\omega}{2}}{\frac{\omega}{2}} \right)^4.$$

2.1.2. Построение локально-оптимального сплайн-вейвлета (LOSW)

В-сплайн $\beta_m(t)$ является масштабируемой функцией соответствующего анализа с несколькими разрешениями. $L(t)$ как масштабируемая функция может построить новый LOSW $\psi(t)$. Обозначим за $L^*(t)$ двойственную масштабирующую функцию $L(t)$ и за $\psi^*(t)$ двойственный вейвлет $\psi(t)$.

Двухмасштабные уравнения:

$$\begin{cases} L(t) = \sum_{n \in \mathbb{Z}} p_n L(2t - n), \\ \psi(t) = \sum_{n \in \mathbb{Z}} q_n L(2t - n), \end{cases} \quad \begin{cases} L^*(t) = \sum_{n \in \mathbb{Z}} h_n L^*(2t - n), \\ \psi^*(t) = \sum_{n \in \mathbb{Z}} g_n L^*(2t - n). \end{cases}$$

Репрезентация в частотной области:

$$\begin{cases} \hat{L}(\omega) = P(e^{-i\frac{\omega}{2}}) \hat{L}(\frac{\omega}{2}), \\ \hat{\psi}(\omega) = Q(e^{-i\frac{\omega}{2}}) \hat{L}(\frac{\omega}{2}), \end{cases} \quad \begin{cases} \hat{L}^*(\omega) = H(e^{-i\frac{\omega}{2}}) \hat{L}^*(\frac{\omega}{2}), \\ \hat{\psi}^*(\omega) = G(e^{-i\frac{\omega}{2}}) \hat{L}^*(\frac{\omega}{2}). \end{cases}$$

Определение двухмасштабных символов:

$$\begin{cases} P(z) = \frac{1}{2} \sum_{n \in \mathbb{Z}} p_n z^n, \\ Q(z) = \frac{1}{2} \sum_{n \in \mathbb{Z}} q_n z^n, \end{cases} \quad \begin{cases} H(z) = \frac{1}{2} \sum_{n \in \mathbb{Z}} h_n z^n, \\ G(z) = \frac{1}{2} \sum_{n \in \mathbb{Z}} g_n z^n, \end{cases}$$

где $z = e^{-i\frac{\omega}{2}}$.

$$M(z) = \begin{bmatrix} P(z) & P(-z) \\ Q(z) & Q(-z) \end{bmatrix}$$

Вычислим определитель этой матрицы:

$$\Delta_{P,Q}(z) := \det M_{P,Q}(z),$$

$\Delta_{P,Q}(z) \neq 0, (|z| = 1)$:

$$H(z) = \frac{Q(-z)}{\Delta_{P,Q}(z)}, \quad G(z) = \frac{-P(-z)}{\Delta_{P,Q}(z)}.$$

Низкочастотный фильтр $\psi(t)$ может быть получен следующим образом:

$$P(\omega) = \frac{\hat{L}(2\omega)}{\hat{L}(\omega)} = \frac{27 - 19 \cos \omega + 8 \cos \frac{3\omega}{2} - \cos 2\omega}{27 - 19 \cos \frac{\omega}{2} + 8 \cos \frac{3\omega}{4} - \cos 4\frac{\omega}{2}}.$$

Можем найти подходящее значение $Q(z)$ в частотном диапазоне, удовлетворяющее условию $\Delta_{P,Q}(z) \neq 0$, следующим образом:

$$Q(\omega) = e^{-i\omega} \sin^2 \frac{\omega}{2}.$$

Обратным преобразованием Фурье уравнений для $P(z)$ и $Q(z)$ получаем следующее:

$$p_k = \frac{1}{2\pi} \int_0^{2\pi} P(\omega) e^{ik\omega} d\omega,$$

$$q_k = \frac{1}{2\pi} \int_0^{2\pi} Q(\omega) e^{ik\omega} d\omega,$$

аналогично

$$h_k = \frac{1}{2\pi} \int_0^{2\pi} H(\omega) e^{ik\omega} d\omega,$$

$$g_k = \frac{1}{2\pi} \int_0^{2\pi} G(\omega) e^{ik\omega} d\omega.$$

2.2. Обзор метода LOSW-ED

Поступившее на вход изображение раскладывается с использованием LOSW для получения её низкочастотной компоненты cA и высокочастотных компонент cH , cV , cD . Применяется MSANM (Multi-structure Anti-Noise Morphology, мультиструктурная антишумовая морфология) к cA для получения низкочастотной карты признаков E_d со структурными подробностями, а затем SUAMM для определения более тонкого высокочастотного края M_f с богатой семантической структурой, включая cH' , cV' , cD' . Затем используется M_f для выполнения LOSW-реконструкции для получения E_r . Наконец, получаем окончательный вывод F_r с помощью выполнения морфологической реконструкции по E_r и E_d .

Диаграмму LOSW-ED алгоритма можно видеть на рисунке 2.1.

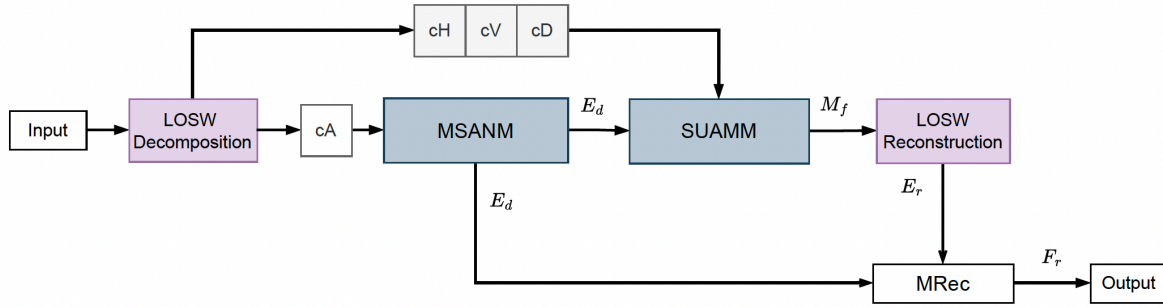


Рисунок 2.1. Диаграмма LOSW-ED алгоритма

2.3. Мультиструктурная антишумовая морфология (MSANM)

В этом разделе используются три различных структурных элемента λ_1 , λ_2 и λ_3 , ориентированных в разных направлениях для учета особенностей границ объектов. $g(x)$ - поступающая на вход чёрно-белое изображение, $\lambda(x)$ - структурные элементы. С учётом равномерности интенсивности отклика в разных направлениях краев вводим их следующим образом:

$$\lambda_1 = \mu \begin{bmatrix} 0.5 & 1 & 0.5 \\ 1 & 2 & 1 \\ 0.5 & 1 & 0.5 \end{bmatrix}, \quad \lambda_2 = \mu \begin{bmatrix} 0 & 0.5 & 0 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.5 & 0 \end{bmatrix}, \quad \lambda_3 = \mu \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

Здесь вес μ усиливает яркость области, значение $\mu = 2$. Выражение для упрощённого мультиструктурного антишумового оператора:

$$E_d = (O_{\lambda_3} (E_{\lambda_2} (D_{\lambda_1} (g)))) - (E_{\lambda_3} (E_{\lambda_2} (D_{\lambda_1} (g)))),$$

где: $D_{\lambda}(x)$ - дилатация (расширение), $E_{\lambda}(x)$ - эрозия (сужение), $O_{\lambda}(x)$ - открытие.

E_d обозначает изображение края, полученное с помощью оператора, который подавляет шум края, сглаживает границы более крупных объектов и сохраняет важную структурную информацию.

2.4. Метод максимумов модуля с учётом структурной неопределённости (SUAMM)

Отличие данного метода от стандартного метода максимумов модуля (ММ) состоит в том, что он учитывает не только локальные различия соседних пикселей, но и всю структуру изображения в целом.

1. Для каждого пикселя вычисляются горизонтальные и вертикальные составляющие градиента:

$$C_x = \frac{\partial C_1(x, y)}{\partial x}, \quad C_y = \frac{\partial C_1(x, y)}{\partial y}.$$

2. Модуль градиента:

$$M_u(x, y) = \sqrt{|C_x|^2 + |C_y|^2}.$$

3. Угол градиента:

$$A_u = \arctan\left(\frac{C_y}{C_x}\right).$$

4. Направления градиента, определяющие, в каком направлении находятся локальные максимумы модуля:

- Горизонтальная компонента

$$\Theta_{cH} = \left\{ \theta \mid \theta \in [0, \frac{\pi}{8}) \cup [\frac{15\pi}{8}, 2\pi) \cup [\frac{7\pi}{8}, \frac{9\pi}{8}) \right\}$$

- Вертикальная компонента

$$\Theta_{cV} = \left\{ \theta \mid \theta \in [\frac{3\pi}{8}, \frac{5\pi}{8}) \cup [\frac{11\pi}{8}, \frac{13\pi}{8}) \right\}$$

- Диагональная компонента

$$\Theta_{cD} = \left\{ \theta \mid \theta \in [\frac{\pi}{8}, \frac{3\pi}{8}) \cup [\frac{9\pi}{8}, \frac{11\pi}{8}) \cup [\frac{5\pi}{8}, \frac{7\pi}{8}) \cup [\frac{13\pi}{8}, \frac{15\pi}{8}) \right\}$$

5. Адаптивный порог:

$$T_\delta = \frac{\delta_{Dmax} + \delta_{Dmin}}{2}.$$

Таким образом, получили выражение, позволяющее определить, является ли точка краевой:

$$\{C'_H, C'_V, C'_D\} = \begin{cases} C_H, C_V, C_D, & \text{если } \theta_{m,n} \in \Theta \text{ и } \delta_{m,n} > T_\delta, \\ 0, & \text{иначе.} \end{cases}$$

Финальная морфологическая реконструкция:

$$F_r = \text{imreconstruct}(E(Ed), E_r),$$

где E_d – карта границ из MSANM, а E_r – восстановленные границы после SUAMM и вейвлетной обработки.