УДК 621.391: 519.67

А.И. Новиков

АЛГОРИТМЫ ОБНАРУЖЕНИЯ ГРАНИЦ ПОЛЕЗНЫХ СИГНАЛОВ

На примере задачи обнаружения и выделения границ однородных по интенсивности изображения (спектра) областей показаны способы конструирования операторов (масок, шаблонов) с произвольными, но фиксированными размерами и с заранее заданными свойствами.

Введение. В различных областях прикладных исследований возникает необходимость обнаружения полезных сигналов и определения границ областей этих сигналов. Такие задачи приходится решать в спектрометрии [1], в радиолокации [2], в биологии и в медицине. Известно большое число алгоритмов, применяемых для решения названных задач: калмановская фильтрация, различные варианты градиентных методов, операторы с различными схемами формирования весовой функции.

В настоящей статье обсуждаются два очень простых, но, вместе с тем, достаточно эффективных алгоритма решения названных задач. Первый из них относится к блоку градиентных методов и основан на следующей идее. Отрезок прямой конечной длины, построенный по измеренным значениям, например методом наименьших квадратов, в области фона будет иметь близкое к постоянному значение углового коэффициента b $(y = a + bx)$. Для многих спектров и изображений $\hat{b} \approx 0$. При переходе из области фона в область полезного сигнала в скользящем режиме значение коэффициента b начнет возрастать. При соответствующем подборе длины T отрезка прямой $y = a + bx, x \in [t,\, t+T]$ точкам t_0 начала и t_1 – окончания полезного сигнала будут отвечать максимальные значения оценок $\hat{b}$ коэффициента b. Увеличение длины T отрезка скольжения увеличивает одновременно сглаживающий эффект, т.е. повышает надежность оценки $\hat{b}$. Решение о принадлежности некоторой точки на прямой (для плоского спектра) или на плоскости (для изображения) границе области полезного сигнала принимается на основании сравнения значений коэффициента b с пороговыми значениями.

Второй алгоритм основан на применении симметричных операторов специальной структуры, трансформирующих исходные измерения в квазистационарную последовательность с выраженными локальными минимумами в граничных точках.

1. Градиентный метод. Для обработки изображений

$$z = f(x, y),\ (x, y) \in D,$$

где f – функция яркости, рассматриваемый алгоритм можно применять в схеме построчной (или столбцовой) обработки. В этом случае можно принять следующую модель измерения на дискретной сетке

$$z_t = v_t + \sum_{k=1}^{r} u_t^{(k)} + \zeta_t, \quad t = \overline{1, n}, \tag{1}$$

где z_t – измеренное значение сигнала;

v_t – значение фоновой составляющей;

$u_t^{(k)}$ – значение к - й полезной составляющей в точке t;

ζ_t – случайная составляющая.

При этом

$$u_t^{(k)} = \begin{cases} u_t^{(k)} > 0, & при\ t \in \left[t_0^{(k)},\, t_1^{(k)}\right] = T^{(k)}; \\ 0, & при\ t \notin T^{(k)}. \end{cases}$$

Задача обнаружения заключается в оценивании значений $t_0^{(k)}$ – начала и $t_1^{(k)}$ – окончания k - го полезного сигнала.

Будем считать, что в области фоновой составляющей, т.е. $t \notin \bigcup_{k=1}^{r} T^{(k)}$, измеряемый сигнал хорошо описывается линейной моделью на каждом отрезке $[t_i, t_i + m - 1]$ длины m

$$z_t = a_i + b_i(t - t_i + 1) + \zeta_t \tag{2}$$

т.е. $v_t = a_i + b_i(t - t_i + 1), \quad t \in [t_i; t_i + m - 1]$.

Осуществим замену

$$\tau = t - t_i + 1, \quad t \in [t_i; t_i + m - 1].$$

Тогда

$$z_{\tau + t_i - 1} = a_i + b_i \tau + \zeta_{\tau + t_i - 1}. \tag{3}$$

Оценку интересующего нас коэффициента $\hat{b}_i$ в точке t_i найдем с помощью стандартного МНК из условия минимума функции

$$F(a;b)=\sum_{\tau=1}^{m}\left(z_{\tau+t_i-1}-\left(a_i+b_i\tau\right)\right)^2 .$$

Получим

$$\hat{b}_i=\frac{12}{m\left(m^2-1\right)}\left(\sum_{\tau=1}^{m}\tau\, z_{\tau+t_i-1}-\frac{m+1}{2}\sum_{\tau=1}^{m}z_{\tau+t_i-1}\right). \quad (4)$$

В частности, для четных значений m $(m=2k)$ и нечетных $(m=2k+1)$ из (4) получаем:

Если $m=2k$, то

$$\hat{\mathrm{b}}_{\mathrm{i}}=\frac{3}{\mathrm{k}(2\mathrm{k}-1)(2\mathrm{k}+1)}\left(-(2\mathrm{k}-1),-(2\mathrm{k}-3),\ldots,\right.$$

$$\left.-3,-1,1,3,\ldots,(2\mathrm{k}-1)\right)\begin{pmatrix} \mathrm{z}_{\mathrm{t_i}} \\ \mathrm{z}_{\mathrm{t_i}+1} \\ \vdots \\ \mathrm{z}_{\mathrm{t_i}+2\mathrm{k}-1} \end{pmatrix}. \quad (5)$$

Если же $m=2k+1$, то

$$\hat{b}_i=\frac{3}{k(k+1)(2k+1)}\times$$

$$\times\left(-k,-k+1,\ldots,-1,0,1,k-1,k\right)\begin{pmatrix} z_{t_i} \\ z_{t_i+1} \\ \vdots \\ z_{t_i+2k} \end{pmatrix}. \quad (6)$$

Например, при $m=7$ $(k=3)$ из (6) имеем

$$\hat{b}_i=\frac{1}{28}\left(-3,-2,-1,0,1,2,3\right)\begin{pmatrix} z_{t_i} \\ z_{t_i+1} \\ \vdots \\ z_{t_i+6} \end{pmatrix},$$

А при $m=8$ $(k=4)$ из (5:)

$$\hat{b}_i=\frac{1}{84}\left(-7,-5,-3,-1,1,3,5,7\right)\begin{pmatrix} z_{t_{i_i}} \\ z_{t_i+i} \\ \vdots \\ z_{t_i+7} \end{pmatrix}.$$

Градиентные методы обработки изображений $f(x,y)$ основаны на получении оценок градиента $\nabla f=\left(\frac{\partial f}{\partial x};\frac{\partial f}{\partial y}\right)$ функции яркости. Рассмотренный выше алгоритм дает оценку частной производной $\frac{\partial f}{\partial x}$ при построчной обработке и $\frac{\partial f}{\partial y}$ – при столбцовой обработке. Действительно, если построчный срез изображения в некоторой окрестности $U\left(\bar{x}_{0,}\delta\right)$ точки $\left(x_0,y_0\right)$ имеет вид

$$f\left(x,y_0\right)=a+bx+\varepsilon_x, \quad x\in U\left(x_0,\delta\right),$$

и найдены оценки $\hat{a}$, $\hat{b}$, т.е.

$$\hat{f}\left(x,y_0\right)=\hat{a}+\hat{b}x,$$

то
$$\frac{\partial \hat{f}\left(x,y_0\right)}{\partial x}=\frac{\partial\left(\hat{a}+\hat{b}x\right)}{\partial x}=\hat{b}.$$

То есть оценки коэффициента $\hat{b}$ (5), (6) являются оценками либо частной производной $f_x\left(x,y_0\right)$, либо $f_y\left(x_0,y\right)$.

Обнаружение границ перепада яркости в рамках рассматриваемого градиентного метода осуществляется по следующей схеме.

1. В каждой точке $M\left(x_i,y_j\right)\in D$ вычисляются оценки частных производных $\hat{f}_x(M),\hat{f}_y(M)$ и модуля вектора градиента

$$\hat{h}(M)=\left(\left(\hat{f}_x(M)\right)^2+\left(\hat{f}_y(M)\right)^2\right)^{1/2}.$$

2. Если $M\in D_1$ - области фоновых значений функции яркости $\left(h(M)\le\Delta\right)$, то вычисляются оценки математического ожидания $\hat{m}_h$ и среднеквадратического отклонения $\hat{\sigma}_h$.

3. В каждой точке $M\in D$ текущая оценка модуля градиента $\hat{h}(M)$ сравнивается с пороговым значением Δ: $\Delta=\hat{m}_h+\alpha\hat{\sigma}_h$, где α – коэффициент, значение которого задается заранее или может корректироваться адаптивно. Такое сравнение необходимо для исключения срабатывания алгоритма обнаружения граничных точек на ложных локальных максимумах функции $h(x,y)$.

4. Если $\hat{h}\left(M_0\right)>\Delta$ и

$$M_0=\arg\max\hat{h}(M), \quad M\in U\left(M_0,\delta\right),$$

то точка M_0 идентифицируется как граничная точка.

Пример применения описанного алгоритма к обнаружению границ перепада яркости на радиолокационном изображении приведен в [3] в сравнении с известными операторами Превитта и Собеля [2], относящимися к классу градиентных методов. Операторы Превитта и Собеля практически не подавляют шум и формируют «размытые» границы областей. Предлагаемый оператор существенно уменьшает шум и формирует «тонкую» границу ∂D_k областей с однородной яркостью.

2. Метод симметричных операторов. Оценки (5), (6) коэффициента b (частных производных $f_{x,}\ f_y$) могут быть записаны в виде оператора с конечной памятью

$$\mathrm{A}z=\sum_{j=1}^{m}\alpha_j z_{t+j-1}=\hat{z}_t\ ,$$

где α_j – весовые коэффициенты оператора A.

В частности, при $m=2k$ из формулы (5) имеем

$$\alpha_j=\frac{3}{k(2k-1)(2k+1)}(2(j-k)-1),\quad j=\overline{1,2k}\,,$$

а при $m=2k+1$ из формулы (6) следует, что

$$\alpha_j=\frac{3}{k(k+1)(2k+1)}(j-k-1),\quad j=\overline{1,2k+1}\,.$$

Известно, что весовые коэффициенты оператора A можно подбирать так, что он будет обладать некоторым заданным свойством: аннулировать низкочастотную составляющую, выделять периодические колебания, обеспечивать заданный уровень снижения случайных колебаний.

Одними из наиболее интересных и с теоретической точки зрения, и с позиций практического применения являются операторы с симметричными весовыми коэффициентами:

$$A=\sum_{j=-k}^{k}\alpha_j P^j\ , \tag{7}$$

где P^j – оператор сдвига: $P^j z_i=z_{i+j}$;

$$\alpha_{-j}=\alpha_j\quad \forall j=\overline{1,k}\,;$$

в зависимости от целей применения весовые коэффициенты симметричного оператора A удовлетворяют одному из условий нормировки:

$$\sum_{j=-k}^{k}\alpha_j=1,\qquad \text{либо}\qquad \sum_{j=-k}^{k}\alpha_j=0\,.$$

Если коэффициенты α_j получены из условия несмещенного оценивания многочленов степени r, $r\le\ell$, где ℓ - заданное число, т.е.

$$A_{2k+1}^{\ell}\{P_r(t)\}_{t=t_0-k}^{t_0+k}=P_r(t_0)\quad \forall r\le\ell\,, \tag{8}$$

то выполняются условия:

$$\sum_{j=-k}^{k}\alpha_j=1\ \text{ и }\ \sum_{j=-k}^{k}\alpha_j^2=\alpha_0\ . \tag{9}$$

Вектор $\overline{\alpha}$ весовых коэффициентов α_j оператора A_{2k+1}^{ℓ} со свойствами (8), (9) является, как известно [4], элементами первой строки матрицы $G=\left(T^T T\right)^{-1}T^T$, где

$$T=\begin{pmatrix} 1 & -k & k^2 & \cdots & (-1)^n k^n \\ 1 & -k+1 & (k-1)^2 & \cdots & (-k+1)^n \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & -1 & 1 & \cdots & (-1)^n \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & k & k^2 & \cdots & k^n \end{pmatrix}.$$

Центральный коэффициент α_0 в составе вектора

$\overline{\alpha}=(\alpha_{-k},\alpha_{-k+1},\ldots,\alpha_{-1},\alpha_0,\alpha_1,\ldots,\alpha_{k-1},\alpha_k)$ характеризует степень подавления шума оператором A_{2k+1}^{ℓ}. Так, если

$$D[\zeta]=\sigma^2,\quad \eta=A_{2k+1}^{\ell}\{\zeta\},$$

то $\quad \mathrm{D}[\eta]=\alpha_0\sigma^2=\alpha_0\mathrm{D}[\xi]$.

Иначе: $\quad \sigma_\eta=\sqrt{\alpha_0}\,\sigma_\zeta$.

Назовем коэффициент $\gamma_l=\dfrac{1}{\sqrt{\alpha_0}}$ – коэффициентом сглаживания оператора A_{2k+1}^{ℓ}, $\ell\ge1$. Для оператора A_{2k+1}^{ℓ} он равен $\gamma_1=\sqrt{2\mathrm{k}+1}$ и является неулучшаемым среди всех операторов с заданной памятью (окном $m=2k+1$ скольжения). Однако оператор A_{2k+1}^{ℓ} дает регулярное смещение на многочленах $P_r(t)$ степени $r\ge2$. Устранить регулярное смещение можно за счет выбора операторов A_{2k+1}^{ℓ} с бóльшими значениями показателя ℓ - максимальной степени многочлена, оцениваемого этим оператором без смещения. Но это, в свою очередь, приведет к увеличению значения центрального коэффициента $\alpha_0^{(\ell)}$ и соответственно к снижению сглаживающего эффекта от действия оператора A_{2k+1}^{ℓ}, $\ell\ge2$.

Так, для оператора A_{2k+1}^{3} коэффициент $\alpha_0^{(3)}$ равен

$$\alpha_0^{(3)}=\frac{3(3k^2+3k-1)}{(4k^2-1)(2k+3)},$$

а коэффициент сглаживания

$$\gamma_3=\sqrt{\frac{(4k^2-1)(2k+3)}{3(3k^2+3k-1)}}\,.$$

В асимптотике (при $k\to\infty$) имеем

$$\frac{\gamma_3}{\gamma_1}=\sqrt{\frac{(2k-1)(2k+3)}{3(3k^2+3k-1)}}\ \xrightarrow[k\to\infty]{}\ \frac{2}{3},$$

т.е. сглаживающий эффект оператора A^3_{2k+1} составляет не более 67 % от сглаживающего эффекта оптимального в этом смысле оператора A^{ℓ}_{2k+1} с весовыми коэффициентами

$$\alpha_j = \frac{1}{2k+1}, \quad j = \overline{-k_1 k}\,.$$

Таким образом, обеспечение несмещенного оценивания оператором A^{ℓ}_{2k+1} многочленов максимальной степени и одновременно максимальное подавление шума – требования противоречивые.

Выбор симметричных операторов для решения задачи обнаружения и выделения границы полезных сигналов обусловлен еще одним свойством этих операторов. Если A_{2k+1} – *произвольный* симметричный оператор, I – тождественный оператор $\left(I\{z_t\}_{t=t_0-k}^{t_0+k} = z_{t_0}\right)$, то оператор $B = I - A$ будет аннулировать многочлены первой степени $\left(\mathrm{B}\{a_0 + a_1 t\}_{t=t_0-k}^{t_0+k} = 0 \quad \forall t_0 = k+1, k+2, \ldots\right)$, а на многочленах второй степени будет давать постоянное смещение Δ:

$$\Delta = a_2 \sum_{j=-k}^{k} \alpha_j j^2 = 2a_2 \sum_{j=1}^{k} \alpha_j j^2\,,$$

т.е.

$$B\left\{a_0 + a_1 t + a_2 t^2\right\}_{t=t_0-k}^{t_0+k} = \Delta \quad \forall t_0 = k+1, k+2, \ldots.$$

Учет перечисленных свойств симметричных операторов позволил сконструировать оператор B с вектором $\overline{\beta}$ весовых коэффициентов

$$\overline{\beta} = \left(\underbrace{\frac{-1}{2s}, \ldots, \frac{-1}{2s}}_{s}, \underbrace{0, \ldots, 0}_{p},\right.$$

$$\left.\underbrace{\frac{1}{2r+1}, \frac{1}{2r+1}, \ldots, \frac{1}{2r+1}}_{2r+1}, \underbrace{0, \ldots, 0}_{p}, \underbrace{\frac{-1}{2s}, \ldots, \frac{-1}{2s}}_{s}\right),$$

где $2(s+p) + 2r + 1 = m$ – память (окно скольжения) оператора B.

Оператор B с такой весовой функцией аннулирует низкочастотную составляющую, а полезную $u_t^{(k)}$ в составе модели (1) трансформирует так, что границам $t_0^{(k)}$ и $t_1^{(k)}$ k-го полезного сигнала отвечают локальные минимумы в квазистационарной последовательности $y_t = B\{z_t\}$.

Решение об объявлении некоторой точки M_0 точкой границы ∂D_k некоторой k-й области осуществляется на основании сравнения значения y_t с пороговым значением Δ. Пример применения оператора B к обработке реального радиолокационного изображения с подробным комментарием приведен в [5].

Заключение. Рассмотренные выше способы конструирования операторов с заданными свойствами для построчной или столбцовой обработки изображений легко могут быть модифицированы для обработки изображений одновременно в строках и в столбцах. Например, оператор

$$A^3_{5\times 5} = \frac{1}{70}\begin{pmatrix} 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 \\ -3 & 12 & 34 & 12 & -3 \\ 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 \end{pmatrix}$$

будет обеспечивать сглаживание измерений и при этом будет давать несмещенные (без искажений) оценки измерений, задаваемых алгебраическими функциями до третьей степени включительно.

$$\mathrm{P}_3(x, y) = a_1 x^3 + a_2 x^2 y + a_3 x y^2 + a_4 y^3 +$$
$$+ b_1 x^2 + b_2 x y + b_3 y^2 + c_1 x + c_2 y + c_3\,.$$

Оператор $B^3_{5\times 5} = I_{5\times 5} - A^3_{5\times 5}$ ($I_{5\times 5}$ – тождественный оператор)

$$B^3_{5\times 5} = \frac{1}{70}\begin{pmatrix} 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & -12 & 0 & 0 \\ 3 & -12 & 36 & -12 & 3 \\ 0 & 0 & -12 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \end{pmatrix}$$

можно использовать для оценки уровня помех.

Действительно, если измерения $z(x, y)$ задаются моделью

$$z(x, y) = P_r(x, y) + \zeta(x, y), \quad r \le 3\,,$$

где $P_r(x, y)$ – многочлен степени r $(r \le 3)$ относительно двух переменных $x, y; \zeta(x, y)$ – случайная составляющая, то

$$B^3_{5\times 5}\{z(x, y)\} = B^3_{5\times 5}\{\zeta(x, y)\} = \eta(x, y)\,.$$

Далее, можно найти оценку среднеквадратической ошибки σ_η трансформированной случайной составляющей $\eta(x, y)$ и пересчитать ее в оценку $\hat{\sigma}_\zeta$ среднеквадратической ошибки случайной составляющей ζ по формуле

$$\hat{\sigma}_\zeta = \frac{1}{\sqrt{\beta^2_{0,0} + \sum\limits_{i=1}^{k} \beta^2_{0,i}}}\,\hat{\sigma}_\eta\,,$$

где $\beta_{0,0}, \beta_{0,1}, \ldots, \beta_{0,k}$ – весовые коэффициенты оператора $B^m_{(2k+1)\times(2k+1)}$. Для оператора $B^3_{5\times 5}$:

$$\beta_{0,0} = \frac{18}{35}; \;\; \beta_{0,1} = -\frac{12}{35}; \;\; \beta_{0,2} = \frac{3}{35} \;\text{ и }\; \hat{\sigma}_\zeta \approx 1{,}603\,\hat{\sigma}_\eta\,.$$

Конструирование шаблонов (масок) любого размера и с заранее заданными свойствами как для построчной (столбцовой), так и для плоскостной обработки изображений, можно осуществлять целенаправленно, опираясь на геометрические свойства производных различных порядков и свойства оценок метода наименьших квадратов.

Библиографический список

1. *Новиков А.И.* Использование линейных операторов в анализе оже–спектров //Математические методы в задачах управления и обработки данных: Межвуз. сб. – Рязань; РРТИ, 1986. – С. 91 – 96.

2. *Бакут П.А., Колмогоров Г.С.* Сегментация изображений: методы выделения границ областей // Зарубежная радиоэлектроника. 1987. №10. С. 25 – 47.

3. *Новиков А.И., Конкин Ю.В., Федорович Я.А.* Применение градиентных методов в задачах обработки радиолокационной информации // Математические методы в научных исследованиях. Рязань. РГРТА, 2006. С. 55 – 63.

4. *Кендалл М. Дж., Стьюарт А.* Многомерный статистический анализ и временные ряды. М.: Наука, 1986.

5. *Новиков А.И., Конкин Ю.В., Архипов С.А.* Применение операторов с симметричной весовой функцией в задачах обработки радиолокационной информации // Математические методы в научных исследованиях. Рязань. РГРТА, 2006. С. 46 – 55.