

А.А. Мицель

ВЫБОР ОПТИМАЛЬНЫХ СПЕКТРАЛЬНЫХ КАНАЛОВ ДЛЯ РЕШЕНИЯ ЗАДАЧ АБСОРБЦИОННОГО ГАЗОАНАЛИЗА И ЛОКАЦИИ

Часть 1. Общий алгоритм и критерий поиска

Предложен общий алгоритм выбора оптимальных спектральных каналов для решения задач локации и абсорбционного газоанализа. Алгоритм основан на вычислении информационного расстояния в пространстве состояний газового объекта. Рассмотрены условия, при которых задача вычисления информационного расстояния решается аналитически и получены четыре алгоритма для практического использования.

Введение

Информация об информативных спектральных каналах является чрезвычайно важной при проектировании оптических устройств для работы в атмосфере. Для анализа спектров поглощения атмосферных и примесных газов в широком диапазоне длин волн и выбора оптимальных спектральных каналов необходимы эффективные математические алгоритмы. В задачах локации выбор каналов осуществляется из требования переноса энергии излучения по атмосферным трассам с минимальными потерями. Эти участки спектра соответствуют «окнам прозрачности». В задачах газоанализа, основанного на спектроскопии поглощения, участки выбирают в полосах поглощения исследуемых газов для обеспечения максимальной чувствительности и селекции по газам.

В работе [1] предложен алгоритм выбора информативных спектральных каналов применительно к задаче газоанализа с помощью ОАД с лазерным возбуждением. Этот алгоритм основан на анализе среднего риска принятия одной из двух альтернативных статистических гипотез H_1 либо H_2 (например, H_1 — анализируемый газ отсутствует в смеси, H_2 — присутствует).

В данной статье излагается новый алгоритм выбора оптимальных спектральных каналов для решения задач локации и абсорбционного газоанализа.

Бейсовский критерий различимости состояний

В задачах абсорбционного газоанализа связь измеряемых сигналов с анализируемым параметром в общем виде может быть представлена выражением

$$y = y_0 \eta f(Kx + \beta), \quad (1)$$

где y , y_0 — измеренные энергетические характеристики излучения на выходе и входе в исследуемую среду; η — чувствительность измерительной системы; $f(\cdot)$ — некоторая функциональная (в общем случае нелинейная) зависимость от концентрации исследуемого газа x и коэффициента β поглощения других газов и аэрозоля, присутствующих в воздухе; K — коэффициент поглощения на единицу концентрации газа.

В общем случае под y и y_0 будем понимать n -мерные случайные векторы с компонентами $y_i = y(\lambda_i)$, $y_{0i} = y_0(\lambda_i)$, $i = 1, \dots, n$; x — m -мерный вектор концентраций анализируемых газов смеси с компонентами x_j , где x_j — концентрация j -го газа, $j = 1, \dots, m$, $m \leq n$; β — n -мерный вектор поглощения посторонних газов; $K(n \times m)$ — матрица коэффициентов поглощения.

Пусть имеется два класса состояний Y_1 и Y_2 , к одному из которых может принадлежать y . Классы состояний Y_1 и Y_2 задаются условными плотностями вероятностей $P(y/Y_1)$ и $P(y/Y_2)$. Принадлежность сигнала y к классу исходов Y_r связана в свою очередь с принадлежностью вектора концентрации газов x к классу состояний x_r , $r = 1, 2$. Пусть q и $p = 1 - q$ априорные вероятности того, что неизвестный параметр y принадлежит классам Y_1 и Y_2 . Кроме того, будем считать известной плотность вероятности $P(y_0)$ мешающего параметра y_0 .

Для проверки гипотез о принадлежности y к классам Y_1 либо Y_2 воспользуемся бейсовским решающим правилом, минимизирующим ошибку решения [2]. В дальнейшем нас будет интересовать только ошибка решения ϵ , равная

$$\epsilon = q\epsilon_1 + p\epsilon_2,$$

где

$$\varepsilon_2 = \int_{-\infty}^{\mu} p(h/Y_2)dh, \quad \varepsilon_1 = \int_{\mu}^{\infty} p(h/Y_1)dh;$$

$$\mu = \ln(q/p), \quad h = \ln[\tilde{p}(y/Y_2)/\tilde{p}(y/Y_1)].$$

Здесь $p(h/Y_r)$ – плотность вероятности случайной величины h : $\tilde{p}(y/Y_r)$ – усредненная по $p(y_0)$ плотность случайного вектора y .

Зададим исходные распределения $p(y/Y_r)$ и $p(y_0)$ нормальными с параметрами (My_r, V_y) и (M_0, V_0) , где My_r – вектор среднего значения в классе Y_r , $r = 1, 2$; $V_y = \text{diag}\{\sigma_y^2(1), \dots, \sigma_y^2(n)\}$ – диагональная матрица ковариации вектора y , одинаковая для обоих классов Y_1 и Y_2 ; M_0 – вектор среднего значения y_0 ; $V_0 = \text{diag}\{\sigma_0^2(1), \dots, \sigma_0^2(n)\}$ – матрица ковариации вектора y_0 .

Учитывая, что погрешности измерений ξ_y и ξ_0 независимы и σ_0 не зависит от y_0 , получим нормальное распределение $\tilde{p}(y/Y_r)$ с параметрами (M_z, V_r) , где

$$M_r(i) = M_0(i) \eta_i f_i(Kx_r + \beta), \quad r = 1, 2, \quad (2)$$

$$V_r = \text{diag}\{\sigma_r^2(1) \dots \sigma_r^2(n)\}, \quad r = 1, 2. \quad (3)$$

Здесь x_r – вектор концентраций газов в классе x_r , $r = 1, 2$; $M_0(i)$, η_i , f_i – i -е компоненты векторов M_0 , η , f .

Значения дисперсий $\sigma_r^2(i)$ рассчитываются по формуле

$$\sigma_r^2(i) = \sigma_y^2(i) + \sigma_0^2(i) \eta_i^2 f_i^2(Kx_r + \beta). \quad (4)$$

Для вычисления ошибки классификации ε необходимо знать распределение вероятностей логарифма отношения правдоподобия $p(h/Y_r)$. Введем следующее допущение: в обеих гипотезах дисперсии $\sigma_r^2(i)$, $r = 1, 2$ принимаются одинаковыми и независимыми от значения h , т.е.

$$\sigma_1^2(i) = \sigma_2^2(i) = \sigma^2(i), \quad i = 1, \dots, n.$$

Тогда мы имеем $V_1 = V_2 = V$ и распределение $p(h/Y_r)$ будет нормальным с параметрами (η_r, Σ) , $r = 1, 2$ $\eta_1 = -\eta_2 = -E$; $\Sigma = 2E$, а выражение для ε примет вид при $q = p = 0,5$

$$\varepsilon = 0,5 \left[1 - \Phi\left(\frac{\sqrt{E}}{2}\right) \right], \quad (5)$$

где

$$E = \frac{1}{2} (M_2 - M_1)^T V^{-1} (M_2 - M_1); \quad (6)$$

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

Входящие в (6) векторы M_1, M_2 задаются формулой (2); индекс «т» означает транспонирование; элементы диагональной матрицы V определяются формулой (4), в которой из соображений разумности следует взять в качестве векторов x_1 либо x_2 тот из них, для которого элементы $\sigma^2(i)$ будут максимальными

$$\sigma^2(i) = \max \{ \sigma^2(a_i, x_1), \sigma^2(a_i, x_2) \} \quad (7)$$

или

$$V(x) = \max \{ V(x_1), V(x_2) \}, \quad (7a)$$

где a_i – i -я строка матрицы коэффициентов поглощения K . Взяв из двух значений $\sigma^2(i)$ ее максимальное, мы получим завышенное значение ε , что позволит в свою очередь повысить надежность различения сигналов $y(x_1)$ и $y(x_2)$.

Связь ошибки классификации ε с информационным расстоянием в пространстве x

Рассмотрим величину $d(Y_i, Y_j) = 2E(M_i, M_j)$, которая представляет собой меру различимости парных состояний Y_i и Y_j в пространстве измерений (см. (6));

$$d(Y_i, Y_j) = d(i, j) = (M_i - M_j)^T V^{-1} (M_i - M_j), \quad (8)$$

и удовлетворяет условиям а) $d(i, j) \geq 0$, $d(i, j) = 0$ при $i = j$; б) $d(i, j) = d(j, i)$;
в) $\{d(i, j)\}^{0.5} \leq \{d(i, k)\}_{0.5} + \{d(k, j)\}^{0.5}$.

Первые два условия очевидны, третье условие (неравенство треугольника для парных состояний (i, j) , (i, k) , (k, j) легко доказывается с использованием неравенства Коши [3]. Величина d , удовлетворяющая условиям «а», «б», «в» представляет собой квадрат расстояния в пространстве исходов между классами Y_i и Y_j .

Пусть класс X_1 — газовый объект, характеризуемый состоянием x , а X_2 — характеризуемый состоянием $x + dx$, где dx — дифференциал вектора x . Тогда в пространстве исходов вектор $(M_2 - M_1)$ можно представить в виде

$$M_2 - M_1 = F(K(x + dx) + \beta) - F(Kx + \beta) = Rdx, \quad (9)$$

где F — n -мерный вектор с компонентами

$$F(i) = M_0(i) \eta_i f \left(\sum_{j=1}^m K_{ij} x(j) + \beta(i) \right);$$

R — $(n \times m)$ -мерная матрица с компонентами

$$R_{ij} = \frac{dF(i)}{dx(j)}, \quad i = 1, \dots, n; \quad j = 1, \dots, m.$$

Подставим (9) в (8) и в результате получим

$$(dL)^2 = dx^T R^T V^{-1} R dx = dx^T G dx, \quad (10)$$

где $G = R^T V^{-1} R$.

Выражение (10) представляет собой дифференциальную квадратичную форму от переменных x_1, \dots, x_m , являющихся проекциями вектора x , и определяет квадрат расстояния между состояниями x и $x + dx$ в пространстве X . Элементы матрицы G_{ij} задают метрический тензор в римановом пространстве [4]. Конечное расстояние между точками, задаваемыми радиус-векторами x_1 и x_2 , определяется как наименьшая длина параметрической кривой $x(t)$, $t_1 \leq t \leq t_2$

$$L = \min_{x \in X} \int \sqrt{z^T G z} dt. \quad (11)$$

Минимум берется по всем непрерывным дифференцируемым кривым $x(t)$, концы которых закреплены, т. е. $x(t_1) = x_1$, $x(t_2) = x_2$ [4]. Компоненты вектора z заданы в виде $z(j) = dx(j)/dt$. Элементы матрицы G зависят от t через проекции вектора x , т. е.

$$G_{ij} = G_{ij} [x_1(t), \dots, x_m(t)].$$

Отметим, что впервые риманову информационную метрику для оценок предельных возможностей спектрофотометрических методов было предложено использовать В.П. Козловым [5, 6].

Из формул (6)–(8) и (10) следует соотношение, связывающее ошибку классификации ε с информационным расстоянием L в пространстве состояний X

$$\varepsilon(i, j) \geq \frac{1}{2} \left[1 - \Phi \left(\frac{L(i, j)}{2\sqrt{2}} \right) \right]. \quad (12)$$

или

$$\varepsilon \geq \frac{1}{2} \left[1 - \Phi \left(\frac{L}{2\sqrt{2}} \right) \right]. \quad (12)$$

Вычисление информационного расстояния L

Рассмотрим способ вычисления L . Минимизация (11) эквивалентна решению системы дифференциальных уравнений (уравнений Эйлера) [7]

$$\frac{\partial^2 \Theta}{\partial z_j^2} z_j' + \frac{\partial^2 \Theta}{\partial x_j \partial z_j} z_j + \frac{\partial^2 \Theta}{\partial t \partial z_j} - \frac{\partial \Theta}{\partial x_j} = 0, \quad (13)$$

где $\Theta = \sqrt{z^T G z} = \Theta[x_1(t), \dots, x_m(t); z_1(t), \dots, z_m(t)]$.

Аналитическое решение системы (13) удается получить лишь в случае линейной связи вектора y с вектором x , т. е. $y = Kx + \beta$. Для этого случая элементы матрицы $G_{ij} = \{R^T V^{-1} R\}_{ij}$ не зависят от $x_j(t)$. Поэтому второе и четвертое слагаемые в (13) обращаются в нуль и мы получим следующую систему дифференциальных уравнений:

$$\left[G_{ij} \Theta^2 - \left(\sum_{l=1}^m G_{il} z_l \right) \right] z_j' + \sum_{l=1}^m G_{il} z_l' \Theta^2 - \sum_{l=1}^m G_{il} z_l \sum_{j,k} z_j' G_{ik} z_k = 0.$$

Решение этой системы имеет вид

$$z_j = \text{const} = a_j, \quad x_j = a_j t + b_j, \quad j = 1, \dots, m,$$

а константы интегрирования a_j, b_j определяются из условий

$$x_j(t_1) = a_j t_1 + b_j = \rho_1^j; \quad x_j(t_2) = a_j t_2 + b_j = \rho_2^j; \quad t_1 = 0, \quad t_2 = 1.$$

В результате получим

$$x_j(t) = (\rho_2^j - \rho_1^j) t + \rho_1^j; \quad z_j(t) = (\rho_2^j - \rho_1^j);$$

$$L = \int_0^1 \sqrt{z^T G z} dt = \sqrt{z^T G z}, \quad (14)$$

где ρ_1^j, ρ_2^j — границы диапазона изменения концентрации j -го газа. Сравнивая (14) с (6), мы видим, что в случае линейной связи вектора y с концентрациями газов $x(y = Kx + \beta)$ информационное расстояние $L = \sqrt{2E}$.

В случае нелинейной связи y с вектором x аналитическое решение системы (13) получить не удастся. Однако можно указать три практически важных частных примера, когда вариационная задача (11) имеет решение для нелинейной модели.

Первый пример связан с однопараметрическим состоянием ($m = 1$). Выражение для L при $m = 1$ приобретает вид

$$L = \int_{\rho_1}^{\rho_2} \sqrt{R^T V^{-1} R} dx, \quad (15)$$

т.е. величина L может быть вычислена обычным интегрированием функции одной переменной.

Второй пример связан с одноволновым методом измерения y ($n = 1$). В этом случае элементы матрицы G_{ij} распадаются на произведение компонент вектора Q , т.е. $G_{ij} = Q_i Q_j$, и выражение (11) примет вид

$$L = \sum_{j=1}^m \int_{\rho_1^j}^{\rho_2^j} Q_j(x_1, \dots, x_m) dx_j. \quad (16)$$

И наконец, последний пример аналитического решения вариационной задачи (11) связан с ситуацией, когда $n = 1, m = 1$. Для этого примера из (11) немедленно следует результат

$$L = \frac{M_0 \eta}{\sigma} \int_{\rho_1}^{\rho_2} \left| \frac{\partial f}{\partial x} \right| dx, \quad (17)$$

где функция f задана формулой (1).

Критерии поиска оптимальных спектральных каналов

Информационная длина L зависит от характеристик измерительной системы ($\sigma_y, \sigma_0, M_0, \eta$) и матрицы коэффициентов поглощения K , которая, в свою очередь, зависит от длин волн. Рассмотрим для простоты случай, когда $n = 1, m = 1$. Величина L определяется формулой (17) и является функцией длины волны λ . С изменением λ значение L будет также меняться. Чем больше L , тем меньше ошибка классификации ϵ газового объекта, а значит, более надежно может быть обнаружен исследуемый газ данным прибором. И напротив, чем меньше L , тем больше ϵ , а следовательно, менее перспективен данный спектральный участок для измерения газа. Если для всех атмосферных газов, имеющих полосы поглощения в заданном спектральном участке, величина L окажется малой (желательно $L \ll 1$), то, очевидно, этот участок спектра перспективен для задач локации.

Таким образом, задача поиска информативных участков сводится к вычислению длины L . Критерий поиска можно представить для задач газоанализа в виде

$$L(\lambda) \rightarrow \max_{\lambda}, \text{ либо } L(\lambda) \geq L_0; \quad (18)$$

для задач распространения (задач локации) в виде

$$L(\lambda) \rightarrow \max_{\lambda}, \text{ либо } L(\lambda) \ll 1; \quad (19)$$

В (18) L_0 – пороговое значение информационного расстояния, которое однозначно связано с пороговой величиной ϵ_0 ошибки классификации. На практике обычно задают $\epsilon_0 = 0,05$, что соответствует $L_0 = 3,3$.

В [6] предложено величину L интерпретировать как число попарно различных состояний газового объекта в интервале $[x_1, x_2]$, при этом в качестве порогового значения авторы [6] рекомендуют брать $L_0 = 1$. При $L < L_0$ состояния x_1 и x_2 не различимы данным прибором, при $L > L_0$ число попарно различных состояний равно $L+1$. Не умаляя значимости предложенной авторами [6] интерпретации L , следует, однако, заметить, что в ней присутствует субъективный фактор, связанный с выбором величины $L_0 = 1$. При $L_0 = 1$ значение ошибки $\epsilon = 0,3$, что составляет достаточно большую величину. На практике, как указывалось выше, величина ϵ задается обычно на уровне 0,05 и меньше, что соответствует $L = 3,3$ и больше.

Заключение

Предложенный критерий выбора оптимальных спектральных каналов является упрощением рассматриваемого ранее нами подхода [1]. В отличие от [1] в этой статье использовались простые гипотезы в обоих классах состояний X_1 и X_2 информационного параметра x . Кроме того, в этих гипотезах дисперсии были приняты равными и не зависящими от сигналов, что позволило связать ошибку решения ϵ с одним параметром E (в работе [1] вычислялся средний риск, который выражается через два параметра g_1 и g_2). Такой подход дает возможность строить эффективные алгоритмы поиска для более сложных физических моделей описания сигналов, в частности, для многомерных и нелинейных моделей при любом спектральном разрешении.

Введение дифференциальной формы информационного расстояния $(dL)^2$ (см. (10)) позволило перейти от дискретных состояний x , к непрерывным. Это привело к появлению параметра L , имеющего более глубокий смысл меры различимости состояний в задаче классификации объектов. Величину L можно рассматривать как число попарно различных состояний в заданном диапазоне концентраций $[r_1, r_2]$. Впервые информационное расстояние в задаче классификации было введено в [5, 6] для определения предельных возможностей спектрофотометрических методов измерения газов. Для практического использования авторами [6] был предложен алгоритм вида (17) ($n = 1, m = 1$). В данной статье рассмотрены условия, при которых вариационная задача (II) решается аналитически и получены четыре общих алгоритма (14)–(17) для практического использования. На основе этих алгоритмов во второй части работы получены формулы для конкретных спектральных приборов, работающих по принципу поглощения.

В заключение автор выражает благодарность Г.Н. Глазову и С.Д. Творогову за полезное обсуждение работы на начальном ее этапе и высказанные критические замечания.

Приложение

Введем некоторые определения.

Газовый объект – газовая смесь из определенного набора газов, имеющих спектры поглощения в оптическом диапазоне.

Состояние газового объекта (в пространстве информационного параметра X) – нахождение газовой смеси при некоторых фиксированных значениях концентраций газов, задаваемых вектором x .

Вектор измерений (вектор наблюдений) y — совокупность измеренных энергетических характеристик газового объекта (например, ОА-сигналы на различных длинах волн, генерируемых ОА-ячейкой с газовой смесью при прохождении через нее оптического излучения).

Состояние в пространстве измерений Y (в пространстве исходов) задается вектором измерений y , однозначно связанным с вектором x . Вектор y представляет собой некоторое (в общем случае нелинейное) преобразование вектора состояний газового объекта x .

Класс состояний в пространстве X — совокупность случайных состояний газового объекта, задаваемых вектором x .

Класс состояний в пространстве Y — совокупность случайных состояний, задаваемых вектором y . Определяется как состояниями газового объекта, так и состояниями прибора.

Рассмотрим вкратце задачу классификации и ее связь с задачей выбора информативных длин волн.

Будем считать вектор измерений y случайным с условной плотностью вероятности $p(y/Y_r)$, $r = 1, \dots, l$. Плотность вероятности $p(y/Y_r)$ зависит от принадлежности вектора y к определенному классу Y_r . Если условная плотность вероятности известна для каждого класса Y_r , то задача классификации газовых объектов сводится к задаче статистической проверки гипотез.

Рассмотрим случай двух классов ($r = 1, 2$). Необходимо проверить гипотезу H_2 о принадлежности вектора y к классу Y_2 против альтернативы $H_1: y \in Y_1$. Байесовское решающее правило, минимизирующее ошибку решения, имеет вид [2]

$$p(y/Y_2)p(Y_2) \geq p(y/Y_1)p(Y_1) \rightarrow y \in \begin{cases} Y_2 \\ Y_1 \end{cases} \quad (\text{П1})$$

или

$$l(y) = \frac{p(y/Y_2)}{p(y/Y_1)} \geq \frac{p(Y_1)}{p(Y_2)} \rightarrow y \in \begin{cases} Y_2 \\ Y_1 \end{cases}, \quad (\text{П2})$$

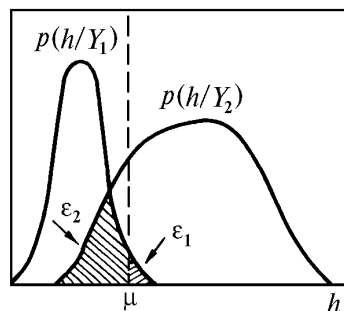
где $p(Y_1)$, $p(Y_2)$ — априорные вероятности принадлежности вектора y к классам Y_1 и Y_2 соответственно. По знаку неравенства принимается решение о принадлежности объекта, характеризуемого вектором y , к классу Y_1 либо Y_2 .

Величина $l(y)$ называется отношением правдоподобия, а $P(Y_1)/P(Y_2)$ — порогом для данного решающего правила. Во многих случаях бывает удобнее использовать вместо $l(y)$ логарифм отношения правдоподобия $h(y)$

$$h(y) = \ln(l(y)) \geq \mu = \ln(p(Y_1)/p(Y_2)) \rightarrow y \in \begin{cases} Y_2 \\ Y_1 \end{cases}. \quad (\text{П3})$$

Решающее правило (П3) не гарантирует безошибочной классификации. Вероятность того, что мы ошибочно отнесем измеренный вектор y к данному классу Y_r , называют вероятностью ошибки решения, или просто ошибкой решения. На рисунке схематично показано распределение $p(h/Y_r)$. Затрихованные площади под кривыми распределений представляют собой ошибки первого и второго рода ε_1 и ε_2 , вычисляемые по формулам

$$\varepsilon_2 = \int_{-\infty}^{\mu} p(h/Y_2) dh, \quad \varepsilon_1 = \int_{\mu}^{\infty} p(h/Y_1) dh. \quad (\text{П4})$$



Распределение логарифма отношения правдоподобия

Общая ошибка определяется как средневзвешенная сумма этих ошибок

$$\varepsilon = p(Y_1) \varepsilon_1 + p(Y_2) \varepsilon_2. \quad (\text{П5})$$

Из (П4) и (П5) нетрудно заметить, что величина ϵ не зависит от измеряемого вектора y и поэтому может быть рассчитана до проведения эксперимента. Так как $p(h/Y_r)$ зависит от характеристик измерительной системы, коэффициентов поглощения и концентраций газов, то ϵ может рассматриваться как обобщенный параметр потенциальных возможностей прибора для анализа газовой смеси. Существенным признаком для рассматриваемой в данной работе задачи является зависимость ϵ от коэффициентов поглощения, которые в свою очередь зависят от длины волны. Это обстоятельство используется в дальнейшем при построении алгоритма выбора оптимальных спектральных каналов.

1. Катаев М. Ю., Мицель А. Д., Тарасова С. Р. // Оптика атмосферы. 1990. Т. 3. С. 832–841.
2. Фукунага К. Введение в статистическую теорию распознавания образов. М.: Наука. 1979. 367 с.
3. Беллман Р. Введение в теорию матриц. М.: Наука, 1969. 367 с.
4. Рашевский П. К. Риманова геометрия и тензорный анализ. М.: Наука, 1967. 568 с.
5. Козлов В. П. // Докл. АН СССР. 1966. Т. 166. № 4. С. 779–783.
6. Гальцев А. П., Козлов В. П., Ривин М. М. // Проблемы физики атмосферы. Л.: ЛГУ, 1967. № 5. С. 129–145.
7. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. М.: Наука, 1974. 783 с.

Институт оптики атмосферы СО РАН,
Томск

Поступила в редакцию
31 марта 1992 г.

A. A. Mitsel. Selection of Optical Spectral Channels for Solving the Problems of Absorption Gas Analysis and Remote Sensing. Part I. General Algorithm and Criteria of Searching.

A generalized algorithm of searching for optimal spectral channels to be used for solving problems of remote sensing and absorption gas analysis is proposed. The algorithm is based on calculating information length in the space of a gaseous object states. Conditions under which the problem on calculation of the information length can be solved analytically are considered and four algorithms that could be of practical interest are proposed.