

**Е.И. Поспелова**

## **ИСПОЛЬЗОВАНИЕ МЕТОДОВ ИНЖЕНЕРИИ ЗНАНИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ АТМОСФЕРНОЙ КОРРЕКЦИИ ДИСТАНЦИОННЫХ ИЗМЕРЕНИЙ**

Проведен обзор работ по применению методов инженерии знаний к обработке данных дистанционного зондирования. Дано описание экспертной системы, учитывающей влияние атмосферы на распространение теплового излучения в системе «атмосфера – подстилающая поверхность». Показано, как особенности предметной области повлияли на формирование основного принципа системы: необходимость участия пользователя в выборе релевантных данных и методов исследования. Указаны направления потенциального применения прототипов представленной системы для решения прикладных задач.

### **Введение**

Практическая значимость и актуальность работ по исследованию воздействия атмосферы на перенос излучения не вызывает сомнений. Эти исследования основываются на анализе зависимости радиационных характеристик от оптико-метеорологического состояния атмосферы и характеризуются следующими особенностями:

- большим объемом разнородной (спектроскопической, метеорологической и т.п.) информации, которую необходимо использовать практически на всех этапах исследования;
- сложностью описания физических механизмов трансформации излучения;
- отсутствием у исследователей единой точки зрения на природу некоторых атмосферных процессов.

Последние обстоятельства не позволяют формализовать знания специалистов об этих процессах, их интуитивные догадки и основанные на опыте предположения. В силу этого традиционные методы автоматизации, например разработка расчетных программ, программных систем и т.п., имеют ограниченную практическую эффективность. Зато известно, что аналогичные проблемы в ряде случаев удается разрешить путем использования достижений искусственного интеллекта (ИИ). Так, необходимые для решения любой информационно-насыщенной задачи процедуры систематизации, структурирования и сокращения релевантных данных (т.е. данных, относящихся к предметной области) относительно легко и безболезненно могут быть осуществлены с помощью эвристических алгоритмов. Если обработка данных требует учета большого количества независимых параметров, то роль того или иного фактора наиболее эффективно может быть определена самим исследователем как пользователем некой интеллектуальной системы. В последнее время стали появляться концепции средств обработки знаний, создаваться экспертные системы (ЭС), системы поддержки решений и т.п., убедительно доказывающие, что многие проблемы дистанционного зондирования (ДЗ) решаются с помощью систем с ИИ.

В данной статье приводится описание ЭС, предназначенной для изучения влияния атмосферы на перенос теплового излучения в системе «атмосфера – подстилающая поверхность». Но прежде чем обсуждать эту разработку, дадим краткий обзор исследований по инженерии знаний применительно к задачам ДЗ.

### **Экспертные системы для задач дистанционного зондирования**

В настоящее время активно ведутся работы по созданию различных интеллектуальных систем, которые могли бы использоваться для решения задач ДЗ. Среди этих работ можно выделить исследования, проводимые с целью определения места интеллектуальных средств в системах обработки спутниковой информации.

Рассматривая обусловленность ИИ в системах автоматизации космических работ, исследователи NASA делают вывод, что методы и принципы ИИ имеют много потенциальных применений для этой предметной области, т.к. она, по их мнению, характеризуется следующими чертами [1]:

- достаточно определенный объект исследований, цель изысканий вполне ясна и однозначна;
- наличие квалифицированных специалистов, готовых передать свои знания «электронному эксперту»;
- частая повторяемость и трудоемкость процесса исследования;
- непригодность или высокая стоимость традиционных программных решений.

В силу этого выделяются три основных направления использования ЭС, где такие системы обещают наибольшие преимущества: контроль бортовой аппаратуры; планирование и управление бортовыми ресурсами; облегчение работы исследователя в режиме человек-машина.

Предлагается строить ЭС с различными иерархическими уровнями: информационной управляющей системы, функциональных подсистем спутника и операционной управляющей системы. Модель принятия решений, разработанная в NASA для выполнения коммерческих космических операций, производит выбор и оценку оптимального использования ИИ для автоматизации конкретной космической станции [2]. В создаваемой ВНИИСХМ концепции построения интеллектуальных систем существенное внимание уделяется определению возможного использования ЭС в процессе обработки данных ДЗ. Их система ТОПИКА объединяет в себе подсистему поддержки разработки программ и интеллектуальные средства интерпретации данных ДЗ в интересах агрометеорологии [3].

Сегодня активно изучаются возможности интегрирования традиционных технологий обработки данных ДЗ и методов ИИ. Современный уровень развития регистрирующей аппаратуры определяет заметное усложнение структуры обрабатываемой информации, а увеличение знаний о предметной области и расширение круга решаемых задач ведет к росту числа правил и фактических данных, участвующих в процессе принятия решения. Эти обстоятельства делают актуальными попытки комбинировать технологии принятия решений, характерные для ЭС, и методы эффективной обработки больших объемов данных, относящиеся к области баз данных (БД).

ЭС, работающая на борту спутника, может быть непосредственно связана с системами получения информации. В этом случае поступающая информация архивируется в комплексной БД, уже содержащей и традиционные географические данные. Такая ЭС может осуществлять предварительную обработку прямо на борту (определение облачности, атмосферная коррекция, классификация изображений и т.п.), а использование дополнительно географической информации позволяет повысить точность и надежность результатов [4]. В этой же связи представляет интерес методология, позволяющая решить вопрос обработки больших массивов экспертных знаний, а не только классических данных, при помощи объединения технологии построения ЭС и системы БД. Проблемы и методы создания интерфейса между ЭС и БД обсуждаются в [5].

При проведении экспериментов по интегрированию географической информации о лесных массивах в провинции Британская Колумбия (Канада) с изображениями, получаемыми в системе анализа цифровых изображений спутников Landsat (LDIAS) и SPOT, возник ряд научных и технических проблем, для решения которых потребовалось внедрение символических описаний и экспертных оценок. Так появилась иерархическая ЭС, названная «аналитическим консультантом» (Analytical Advisor) [6, 7], которая управляет всеми устройствами и компьютерами системы LDIAS, а также всеми решаемыми в системе задачами. ЭС состоит из набора программ, каждая из которых моделирует работу отдельного эксперта-специалиста. Программы организуют иерархию с помощью коммутационных блоков общей памяти, используемых для связи между различными уровнями. Теперь органы управления этой провинции обеспечены своевременной и точной информацией о возобновляемых лесных ресурсах всего региона.

Для помощи в планировании заданий ИСЗ типа UARS разработана ЭС MOPA. Трудоемкий процесс составления планов ежедневных операций спутника состоит, главным образом, в определении режимов работы 10 основных приборов и их координации. Работа приборов имеет повторяющийся характер, ежедневно меняются лишь отдельные характеристики, что делает удобным формализацию знаний о работе приборов в виде фреймов. Действия приборов представляются в виде иерархической схемы, на верхних уровнях которой располагаются более общие виды действий, а на более нижних – элементарные детальные операции. Сначала выбирается нужный уровень, а затем на нем указываются непосредственные операции, при этом на каждом уровне осуществляется последовательное преобразование символического описания работы приборов в конкретные наборы команд [8].

Следует отметить, что все создаваемые ныне интеллектуальные системы для оказания помощи в решении задач ДЗ относятся к разряду так называемых специализированных систем. Последние характеризуются тем, что решают фиксированный набор задач, предопределенный при проектировании системы, и для своего использования требуют только наполнить базу знаний фактами и правилами, соответствующими выбранному приложению. Наряду со специализированными выделяются системы общего назначения, заведомо спроектированные для обработки знаний из достаточно широкого спектра предметных областей и в силу своей универсальности грешащие избыточностью (так называемые оболочки ЭС). В отличие от систем общего назначения специализированные системы обеспечивают максимально рациональное использование всех вычислительных ресурсов.

### **Экспертная система атмосферной коррекции данных дистанционного зондирования**

При разработке сложных интеллектуальных систем выработалась технология, основывающаяся на так называемом эволюционном методе [9]. Вкратце этот метод можно охарактеризовать следующим образом. Создается некоторая достаточно элементарная система-прототип, в которой реализуются основные свойства будущей ЭС. Работа прототипа всесторонне тестируется и на этом этапе выясняется, насколько эффективен выбранный способ представления знаний, уточняются требования к системе на предмет извлечения дополнительных знаний, усовершенствования механизма взаимодействия с пользователем и т.д.

Такой метод позволяет разработчику постоянно совершенствовать свое творение, получая на каждом этапе работающий программный продукт.

Теперь рассмотрим ЭС, разрабатываемую по принципам эволюционного подхода и предназначенную для изучения влияния атмосферы на перенос излучения в системе «атмосфера—подстилающая поверхность». В [10] описана первая версия системы, названная информационно-диалоговой (ИДС) и созданная для отработки принципов организации и использования информационного обеспечения, а также средств диалогового взаимодействия системы с пользователем. Функциональная наполненность ИДС была ограничена расчетом оптических характеристик атмосферы в ИК-окнах прозрачности. Для получения более детальной информации об алгоритмах и методиках, заложенных в ИДС, следует обратиться к [10, 11]. В настоящей статье представляется вторая версия системы, получившая название «экспертная система атмосферной коррекции данных дистанционного зондирования» (ЭСАК). Кратко описываются алгоритмы и функции системы, интересующие исследователя-физика, а основное внимание уделяется изложению принципов и примеров «интеллектуальности» ЭСАК.

Для осуществления атмосферной коррекции дистанционных измерений с помощью ЭСАК можно определять вертикальные профили коэффициентов ослабления, функции пропускания, оптическую толщину, интенсивность уходящего теплового излучения. Все эти характеристики могут быть получены с произвольным разрешением по спектру, в любом частотном диапазоне, при любых метеорологических ситуациях и трассах распространения излучения. Интенсивность излучения системы «атмосфера—подстилающая поверхность» определяется как сумма

$$I = I_s + I_a + I_r,$$

при этом каждое слагаемое — это интенсивность соответственно подстилающей поверхности, атмосферы и отраженная. Их значения в интервале  $\Delta\nu = \nu_2 - \nu_1$  определяются следующим образом:

$$I_s = \frac{1}{\Delta\nu} \int_{\nu_1}^{\nu_2} \delta_\nu B(\nu, T_0) \exp\left(-\int_0^L K(l, \nu) dl\right) d\nu;$$

$$I_a = \frac{1}{\Delta\nu} \int_{\nu_1}^{\nu_2} \int_0^L B(\nu, T(l)) K(l, \nu) \exp\left(-\int_0^l K_t(l', \nu) dl'\right) dl d\nu;$$

$$I_r = \frac{2}{\Delta\nu} \int_{\nu_1}^{\nu_2} \left\{ (1 - \delta_\nu) \exp\left(-\int_0^L K_t(l, \nu) dl\right) \int_0^L B(\nu, T(l)) K(l, \nu) \left[ \int_0^1 \exp\left(-\frac{1}{\mu} \int_0^L K(h', \nu) dh'\right) d\mu \right] dl \right\} d\nu,$$

где  $\delta_\nu$  — излучательная способность подстилающей поверхности;  $B(\nu, T)$  — функция Планка, определяемая по формуле:

$$B(\nu, T) = \frac{c_1 \nu^3}{\exp\left(\frac{c_2 \nu}{T}\right) - 1},$$

в которой  $c_1 = 0,1191 \cdot 10^{-4}$ ,  $c_2 = 1,43879$  — первая и вторая радиационные константы. Далее,  $L$  — длина неоднородной трассы;  $T$  — температура;  $\mu = \cos\theta$ ,  $\theta$  — угол визирования;  $K(l, \nu)$  — коэффициент поглощения;  $K_t(l, \nu)$  — полный коэффициент ослабления. Для получения функции пропускания используется формула

$$P_\nu = \frac{1}{\Delta\nu} \int_{\nu_1}^{\nu_2} \exp\left(-\int_0^L K(l, \nu) dl\right) d\nu.$$

Оптическая толщина  $K_{\Delta\nu}$  равна  $-\ln(P_\nu)$ . Полный коэффициент ослабления считается как сумма

$$K_t(l, \nu) = K(l, \nu) + K_s,$$

где  $K_s$  — коэффициент молекулярного и аэрозольного рассеяния. Для определения его молекулярной компоненты используются методики, приведенные в [12].

$$K(l, \nu) = K_a + K_m,$$

слагаемые определяют поглощение соответственно аэрозольными и молекулярными составляющими атмосферы. Поглощение и рассеяние на аэрозоле учитывается в виде вертикальных профилей значений соответствующих коэффициентов [13]. Поглощение молекулярными компонентами представляется в виде суммы селективного и континуального поглощения [14]:

$$K_M = K_{sel} + K_{cont}.$$

Для расчета коэффициентов селективного и континуального поглощения используются данные о тонкой структуре спектров поглощения атмосферными газами. Это количественные характеристики спектральных линий: положение центра, интенсивность, полуширина, энергия нижнего состояния и информация об изотопах; они организованы в каталог исходной спектроскопической информации. Учет континуального поглощения осуществляется по формулам обобщенного контура спектральной линии [15]. Во избежание в процессе работы ЭСАК больших временных затрат на вычисления по этим достаточно сложным формулам используется каталог континуального поглощения. Этот каталог содержит предварительно рассчитанные значения коэффициентов континуального поглощения водяным паром для случаев самоуширения и уширения азотом, а также коэффициенты интерполяционных кубических сплайнов. Данные структурированы по равномерной частотной и температурной сетке. Физические основы и целесообразность создания такой информационной базы приведены в [11], там же отмечается, что этот каталог является достаточно универсальным и может использоваться для решения любых задач, где требуется учет континуального поглощения водяным паром.

В качестве примера принятия экспертного решения рассмотрим, как ЭСАК при непосредственном участии пользователя выбирает алгоритм и релевантные данные для определения коэффициента селективного поглощения. В общем виде формулу для расчета коэффициента селективного поглощения можно записать так:

$$K_{sel} = \sum_{i=1}^N \sum_{j=1}^{n_i} f_j q_i,$$

где  $q_i$  — концентрация  $i$ -го газа,  $i = \overline{1, N}$ ;  $n_i$  — число спектральных линий  $i$ -го газа;  $f_j$  — контур спектральной линии.

В ЭСАК реализованы следующие методы определения формы контура спектральной линии: Лоренца, полного Лоренца, Фойгта, Наумова—Жевакина и обобщенный контур. В литературе (см., например, [12], [14]) приводятся соображения относительно правомерности использования той или иной формы контура в той или иной ситуации, причем разные исследователи придерживаются на этот счет различных мнений. Надо отметить, что именно такая неоднозначность знаний о предметной области и определила основной принцип функционирования ЭСАК: решение задачи под контролем исследователя. Поэтому, ведя диалог с системой, пользователь постоянно находится в курсе того, как решается его задача, какие используются исходные данные алгоритмы. ЭСАК предоставляет ему исчерпывающие сведения о возникшей в процессе работы ситуации, о состоянии исследуемого объекта, о наличии релевантной информации, о способах расчета и возможных модификациях хода исследования. Эти сведения двух уровней:

- оперативные, рассчитанные на квалифицированного специалиста и обученного общению с ЭВМ пользователя. Здесь кратко описывается ситуация и предлагается принять решение;
- более полные сведения, которые пользователь может получить после нажатия специальной клавиши в случае, если оперативной информации недостаточно и требуется более детальная проработка вопроса.

В довершение всего ЭСАК выдает рекомендацию относительно того, какой из возможных алгоритмов расчета предпочтительнее, а далее дело пользователя — принять рекомендацию системы или выбрать любой другой из предлагаемых ЭСАК в виде меню алгоритмов. На каждом этапе работы пользователь имеет возможность вернуться на предыдущий шаг и изменить принятое ранее решение.

Теперь рассмотрим вопрос о получении информации относительно вертикальных профилей температуры, давления и концентраций атмосферных газов. ЭСАК дает пользователю возможность самому выбрать источник такой информации: либо усредненные по сезонам и регионам метеомодели, либо реальные результаты измерений, проведенных на метеостанциях. Данные метеомоделей взяты из различных литературных источников, такого рода информация достаточно часто используется и хорошо известна специалистам. Что касается БД метеопараметров, наполненной данными с метеостанций, то на ее структуре и особенностях ее использования стоит остановиться подробнее. Deskрипторными ключами, т.е. ключами поиска по этой базе, являются географические координаты и время измерения. Пользователь сам делает запрос по указанным ключам, причем он может либо выбрать в режиме меню координаты одной из зарегистрированных в БД метеостанций, либо указать произвольную точку земного шара. В последнем случае из базы выбираются профили температуры, давления и концентраций газовых составляющих с ближайших к указанной точке метеостанций, а для пространственной коррекции этой информации используется метод оптимальной интерполяции [16]. Содер-

жающиеся в БД профили в силу возможных сбоев в работе радиозондов могут иметь пропущенные значения. Для их восстановления ЭСАК использует барометрическую формулу, либо методы, приведенные в [17] и [18]. Каждый из этих методов в зависимости от состояния исходного профиля и наличия корреляционной информации дает большую или меньшую точность. Поэтому пользователь так же, как и при определении формы контура спектральной линии, поддерживаемый рекомендациями ЭСАК, сам выбирает способ восстановления.

ЭСАК постоянно контролирует пользователя, указывая и объясняя его ошибки в случае их возникновения. При этом проводится проверка правильности нажатия функциональных клавиш и допустимости значений вводимых числовых параметров.

Концептуальные знания об алгоритмах, о возникших в процессе работы системы ситуациях и т.п. выражены в терминах предметной области на естественном языке. Они представляют собой вербальные семантические конструкции (т. е. смысловые текстовые выражения) со свободными переменными, значения которых заполняются конкретным содержанием в зависимости от результатов расчетов и анализа ситуации, а затем предлагаются пользователю либо в виде меню в альтернативные моменты, либо в виде подсказки в случае его ошибочных действий. Эти знания организованы в ситуационные совокупности, что позволяет легко модифицировать функциональную направленность ЭСАК. Можно, например, безболезненно отсекая отдельные ветви, тем самым получать средства решения задач для более узких приложений, а можно, используя ЭСАК в качестве ядра и включая в ее структуру новые алгоритмы, данные и знания, «наращивать мощность» системы для решения более глобальных задач. В скором времени на основе представленного здесь базового варианта ЭСАК планируется:

– создание системы экологического и природо-ресурсного регионального мониторинга лесных массивов на основе обработки аэрокосмической информации;

– проведение экологической экспертизы лесов-торельников Александровского района Томской области по данным аэрокосмических наблюдений.

Как уже отмечалось, представленная здесь версия ЭСАК является развитием ранее описанной информационно-диалоговой системы [10] и, в свою очередь будет совершенствоваться в направлении расширения функциональных возможностей и формализации новых знаний.

## Заключение

Несмотря на большое количество разработок в области инженерии знаний, использование ИИ при обработке данных дистанционного зондирования находится на первоначальной стадии развития. Формулируются концепции «интеллектуальной» интерпретации дистанционных измерений, выясняются основные закономерности, на которых должны основываться структуры баз данных и знаний, изучаются оптимальные режимы взаимодействия с пользователем, делаются попытки формализовать новые знания. Такое положение обусловлено прежде всего особенностями предметной области, ее неоднозначностью и информационной насыщенностью. При этом существует достаточное количество исследователей, так называемых необученных пользователей, которые плохо владеют навыками работы с вычислительной техникой, но в то же время испытывают потребность в таком инструменте, как компьютерная система, сложность общения с которой не превышала бы сложности общения с учителем или коллегой. Наличествуют и узкие специалисты, готовые применить свои знания в решении малоизвестных для них задач, получив при этом возможность познакомиться с новой предметной областью. Поэтому работы по «интеллектуализации» ДЗ, по созданию экспертных систем, снабженных дружественным интерфейсом и способных воспринимать новые знания, актуальны и перспективны.

1. Boarnet M., Culbert C., Savely T. R. //IEEE Int. Conf. Rob. and Autom. Raleigh, N. C., 1987. Proc. V. 1. P. 131–136.
2. Vernon Grose L. //Proc. Soc. Photo-Opt. Instrum. Eng. 1985. V. 580. P. 21–27.
3. Федоров В. В. //Тр. ВНИИ с.-х. метеорол. 1985. № 25. С. 70–83.
4. Ester J. E., Friedl M. A., Star J. L. //Proc. Soc. Photo-Opt. Instrum. Eng. 1988. V. 924. P. 297–304.
5. Reuter A. //Informationstechnik 1988. B. 29. № 3. S. 164–175.
6. Голдберг М., Гуденаф Д. Г., Алво М. //ТИИЭР, 1985. Т. 73. № 6. С. 119–130.
7. Goodenough D. G. //Photogramm. Eng. and Remot. Sens. 1988. V. 54. № 2, P. 167–176.
8. Schuetzle J. G., Zoch D. R. //Exp. Syst. Gov. Symp. McLesn, Va. 1986. P. 210–220.
9. Heng M. S. H. //Future Generat. Comput. Syst. 1987. V. 3. № 2. P. 103–109.
10. Гендрин А. Г., Новикова Н. В., Поспелова Е. И. Информационнодиалоговая система атмосферной коррекции данных дистанционного зондирования параметров подстилающей поверхности. Томск, 1989. 73 с. (Препринт /ИОА СО АН СССР, № 22).
11. Афонин С. В., Гендрин А. Г. //Информационно-программное обеспечение задач атмосферной оптики. Новосибирск: Наука, 1988. С. 38–64.
12. Зуев В. Е. Распространение видимых и инфракрасных волн в атмосфере. М.: Советское радио, 1970. 496 с.
13. Креков Г. М., Рахимов Р. Ф. Оптико-локационная модель континентального аэрозоля. Новосибирск: Наука, 1982. 200 с.
14. Зуев В. Е. Распространение лазерного излучения в атмосфере. М.: Радио и связь, 1981. 288 с.
15. Телегин Г. В., Фомин В. В. О возможности аналитической аппроксимации формы линии в колебательно-вращательных спектрах молекул. Томск, 1979. 52 с. (Препринт /ИОА СО АН СССР, № 26).

16. Рубинштейн К.Г., Шильев В.Б. //Метеорология и гидрология. 1987. № 9. С. 26–33.
17. Мак-Картни Э. Оптика атмосферы. М.: Мир, 1979. 421 с.
18. Рейтенбах Р.Г., Шерстюков Б.Г. Применение статист, методов в метеорол. //Тр. 5 Всесоюз. совещ. Казань, 1985. С. 11–113.

Институт оптики атмосферы СО АН СССР,  
Томск

Поступила в редакцию  
30 апреля 1991 г.

**E. I. Pospelova. Use of Knowledge Engineering Methods for Solving the Problem on Correcting Remote Measurements Data for the Atmospheric influence.**

An overview of the paper devoted to application of the knowledge engineering methods to processing of the remote sensing data is presented. An expert system allowing for influence of the atmosphere on the propagation of thermal source radiation in the «atmosphere – underlying surface» system is described. It is shown how the special features of the subject region affect the formation of the basic principle of the system, i.e. on the necessity of user's interference into the process of selection of relevant data and investigation methods. The directions of potential applications of this system prototypes to solve applied problems are shown.