

В.М.Голубев

СИСТЕМЫ ГЕОМЕТРИЧЕСКИХ ПОСТРОЕНИЙ.
ПОСТАНОВКИ ЗАДАЧ И МЕТОДЫ ИХ РЕШЕНИЯ

Под программными средствами геометрических построений (ПСГП) в настоящей работе понимаются программные средства, предназначенные для моделирования плоских геометрических объектов, составленных из примитивов типа точка, прямая, окружность, дуги кривых второго порядка и т.п.

Существуют две основные области применения ПСГП:

- в САПР машиностроительного типа для подготовки и выпуска чертежно-конструкторской документации;

- в 3Д-системах геометрического моделирования (для построения границ плоских областей с целью последующей генерации базовых пространственных форм типа тел смещения, тел вращения и т.п.).

В первой части работы проводится обзор современных отечественных и зарубежных ПСГП. Следует сразу же отметить, что в рамках настоящего обзора не ставилась цель детального описания и сравнения существующих ПСГП. Основное назначение статьи - анализ ПСГП с единой точки зрения, позволяющей дать строгие постановки задач ПСГП и методов их решения. Для достижения этой цели в качестве основополагающего понятия было выбрано понятие отношения между примитивами, явно или неявно использующееся в любом ПСГП.

Во второй части работы описывается новый подход к реализации средств геометрических построений, разработанный в

рамках интерактивной системы ИНСМ. Характерные особенности подхода: функциональная полнота системы геометрических построений, простой и удобный интерфейс с оператором, малый объем программного кода.

I. Обзор систем геометрических построений

В настоящее время используется два основных подхода к реализации ПСПП [1]: операторный (ПСПП первого типа) и непрограммный, в основе которого лежит понятие отношения (ПСПП второго типа).

В основе ПСПП операторного типа лежит понятие оператора построения примитивов. В зависимости от реализации ПСПП это может быть оператор некоторого языка, вызов подпрограммы из библиотеки, ввод команды и т.п. Каждый такой оператор позволяет добавить к модели объекта новый примитив, причем в качестве входной информации обычно используются параметры ранее построенных примитивов и/или некоторый набор числовых значений. Такого рода ПСПП реализованы во многих отечественных системах машинной графики и геометрического моделирования: ИНКАНЭЛ, ФАП-КФ, ГРАФИТ, ГРАФОР, РЕДГРАФ, СФОРГИ, BRAVO, CATIA, LABCAD, AUTOCAD и др. Независимо от методов реализации таких систем (библиотеки программ, пакеты программ, оригинальные языки, интерактивные системы и т.д.) ПСПП в них фактически являются средствами, позволяющими смоделировать обычный процесс последовательного вычерчивания контурных и вспомогательных линий чертежа за кульманом.

ПСПП второго типа предоставляют возможность описания сразу всех примитивов объекта путем ввода некоторой системы отношений, связывающих параметры примитивов. Основная задача ПСПП такого рода заключается в предоставлении средств вычисления параметров примитивов путем решения системы отношений. Таким образом, ПСПП второго типа являются средствами непрограммного (декларативного) описания объектов моделирования. В настоящее время интерес к таким ПСПП растет [1-7]. В качестве отношений, как правило, используются размеры и геометрические условия, регламентированные стандартами ЕСКД для

чертежей машиностроительного типа [8]: расстояния между примитивами, углы, радиусы; условия касания, параллельности симметрии и т.д.

I.I. Примитивы и отношения

Пусть T – конечное множество типов примитивов, и пусть A_t ($t \in T$) обозначает множество всех примитивов типа t . Предположим, что каждое множество примитивов A_t параметризовано, т.е. каждому конкретному примитиву поставлена в соответствие точка в некотором вещественном евклидовом пространстве. Не теряя общности, будем считать, что для каждого $t \in T$ $A \subset \mathbb{R}^{|t|}$, где $|t|$ обозначает размерность пространства параметров. Для простоты будем пока предполагать параметры независимыми, т.е. считать, что каждое множество примитивов A_t является областью в пространстве $\mathbb{R}^{|t|}$.

Введем обозначения:

$$A_* = \bigcup_{t \in T} A_t, \quad T_* = \bigcup_{n \geq 1} T^n, \quad Z_* = \bigcup_{k > 0, n > 0} \{1, 2, \dots, n\}^k.$$

(Здесь и ниже верхние индексы, если не оговорено противное, используются для обозначения декартовых степеней множеств.) Элементы множества T_* будем называть мультитипами, а элементы множества Z_* – мультииндексами. Далее, для произвольных мультитипа $\alpha = (t_1, t_2, \dots, t_n)$ и мультиинекса $\theta = (m_1, m_2, \dots, m_k)$ положим

$$A_\alpha = A_{t_1} \times A_{t_2} \times \dots \times A_{t_n}, \quad |\alpha| = \sum_1^n |t_i|,$$

а через θ' и θ'' будем обозначать отображения $\theta' : T^n \rightarrow T^k$ и $\theta'' : A_\alpha \rightarrow A_{\theta'(\alpha)}$, такие, что

$$\theta' : \alpha = (t_1, t_2, \dots, t_n) \in T^n \rightarrow$$

$$\rightarrow \theta'(\alpha) = (t_{m_1}, t_{m_2}, \dots, t_{m_k}) \in T^k,$$

$$\theta^n : a = (a_1, a_2, \dots, a_n) \in A_\alpha \rightarrow$$

$$\rightarrow \theta^n(a) = (a_{m_1}, a_{m_2}, \dots, a_{m_k}) \in A_{\theta^n(\alpha)}.$$

Определение 1. Элементы множества A_α ($\alpha \in T^*$) будем называть составными геометрическими объектами типа α (или более коротко – объектами типа α).

Определение 2. Гладкую числовую функцию $f : R^n \rightarrow R$ будем называть А-функцией, если в ее области определения есть подобласть, в которой градиент функции отличен от нуля.

Определение 3. $(n+1)$ -местное отношение $u \subset A_\alpha \times R$ будем называть n -местным отношением из класса размеров, определенным на мультитипе $\alpha \in T^n$ (или более коротко – размером), если существует А-функция $f : A_\alpha \times R \subset R^{|\alpha|+1} \rightarrow R$, такая, что

$$(a, d) \in u \Leftrightarrow f(a, d) = 0 \quad (a \in A_\alpha, d \in R).$$

Определение 4. n -местное отношение $u \in A_\alpha$ будем называть n -местным отношением из класса условий, определенным на мультитипе $\alpha \in T^n$ (или более коротко – условием), если существует А-функция $f : A_\alpha \times R \subset R^{|\alpha|+1} \rightarrow R$, такая, что

$$a \in u \Leftrightarrow f(a) = 0 \quad (a \in A_\alpha).$$

Определение 5. n -местное отношение $u \subset A_\alpha$ будем называть n -местным отношением из класса ограничений, определенным на мультитипе $\alpha \in T^n$ (или более коротко – ограничением), если существует А-функция $f : A_\alpha \subset R^{|\alpha|} \rightarrow R$, такая, что

$$a \in u \Leftrightarrow f(a) > 0 \quad (a \in A_\alpha).$$

Определение 6. Множество $\{m, k, \alpha, \{(u_i, \theta_i) : i = 1, 2, \dots, m+k\}\}$ будем называть системой из $m+k$ отношений, заданной на мультитипе α , если:

а) для любого $1 \leq i \leq m$ $u_i \subset A_{\theta_i^1(\alpha)} \times R$ – отношение из класса размеров;

б) для любого $m + 1 \leq i \leq m + k$ $u_i \subset A_{\theta_i^1(\alpha)}$ – отношение из класса условий:

Определение 7. Будем говорить, что составной объект a типа α и вектор числовых значений размеров $d = (d_1, d_2, \dots, d_m) \in R^m$ удовлетворяют системе отношений $\{m, k, \alpha, \{u_i, \theta_i\}\}$, если:

$$(\theta_i^1(a), d_i) \in u_i \subset A_{\theta_i^1(\alpha)} \times R, (i = 1, 2, \dots, m);$$

$$\theta_i^2(a) \in u_i \subset A_{\theta_i^2(\alpha)}, (i = m + 1, m + 2, \dots, m + k).$$

Пусть $\alpha = (\alpha_1, \alpha_2)$ – некоторый мультитип, составленный из двух мультитипов: α_1 и α_2 (α_2 может быть пустым).

Определение 8. Систему отношений $\{m, k, \alpha, \{(u_i, \theta_i)\}\}$, заданную на мультитипе (α_1, α_2) , будем называть разрешимой относительно мультитипа α_1 , если существуют область $Q \subset A_{\alpha_2} \times R^m$, область $P \subset A_{\alpha_1}$ и гладкое взаимно-однозначное соответствие $F : Q \rightarrow P$ (которое будет называться оператором построения объектов типа α), такое, что для любых $(a, d) \in Q$ объект $(F(a), a)$ и вектор d удовлетворяют системе отношений.

Замечание I. По теореме о неявной функции для разрешимости системы отношений необходимо выполнение следующих условий:

а) число отношений в системе (т.е. число уравнений) должно совпадать с общим числом параметров у примитивов из объекта типа α_1 , т.е. $m + k = |\alpha_1|$;

б) каждому уравнению в системе должна соответствовать непрерывно дифференцируемая функция, в области определения которой имеется по крайней мере одна точка с ненулевым относительно α_1 градиентом (для А-функций f_i , соответствующих отношениям u_i , это требование выполняется, если только мультииндекс θ_i содержит хотя бы одну компоненту из множества $\{1, 2, \dots, |\alpha_1|\}$).

З а м е ч а н и е 2 (теорема о неявной функции). Если существуют объекты $a_1 \in A_{\alpha_2}$, $a_2 \in A_{\alpha_2}$ и вектор $d \in R^m$, такие, что объект (a_1, a_2) и вектор d удовлетворяют системе отношений, и, кроме того, якобиан $\det [\partial f_i / \partial p_j] \neq 0$ в точке (a_1, a_2, d) , то система отношений разрешима. (Здесь f_i – А-функция, соответствующая отношению u_i , $p \in A_{\alpha_1}$.)

З а м е ч а н и е 3. Если система отношений разрешима, то, вообще говоря, могут существовать сразу несколько операторов построения, и поэтому в зависимости от точки $(a, d) \in A_{\alpha_2} \times R^m$ могут существовать один, несколько либо ни одного объекта $x \in A_{\alpha_1}$, такого, что (x, a) и d удовлетворяют системе отношений.

I.2. Постановка задач и методы их решения

В этом пункте даются строгие математические постановки задач синтеза и анализа составных геометрических объектов, описываются методы их решения в ПСПП первого и второго типов.

З а д а ч а I (синтез составного объекта). По заданному объекту $a \in A_{\alpha_2}$ и вектору числовых значений размеров $d \in R^m$ найти объект $x \in A_{\alpha_1}$, такой, что объект (x, a) и вектор d удовлетворяют заданной системе отношений $\{m, k, (\alpha_1, \alpha_2), \{(u_i, \theta_i)\}\}$, разрешимой относительно мульттипа α_1 .

Решение задачи I предоставляет в распоряжение пользователя возможность накопления библиотеки типовых фрагментов чертежа, причем каждый фрагмент является семейством составных геометрических объектов, параметризованных набором числовых значений размеров и/или параметрами ранее построенных примитивов. В общей постановке эта задача в ПСПП пока не решается, но она имеет несколько частных постановок, играющих важную роль в ПСПП первого и второго типов.

З а д а ч а Ia (построение примитива). По заданному составному объекту $a \in A_{\alpha}$ типа α и вектору числовых значений $d \in R^m$ найти примитив $x \in A_t$ типа t , такой, что объект (x, a) и вектор d удовлетворяют системе отношений $\{m, k, ((t), \alpha), \{(u_i, \theta_i)\}\}$, разрешимой относительно примитивов типа t .

(Ниже такого рода разрешимые системы будут называться системами отношений первого типа.)

Подобного рода задачи решаются ПСГП операторного типа, причем каждое такое ПСГП предназначено для решения только некоторого фиксированного набора систем отношений первого типа (этот набор определяется на стадии разработки ПСГП, исходя из требований ко входному языку). С помощью методов аналитической геометрии для каждой системы отношений, как правило, нетрудно получить ее явное аналитическое решение, т.е. выразить параметры искомого примитива через функции от параметров ранее построенных примитивов (из объекта a) и числовых значений размеров.

Задача Iб (синтез составного объекта по системе отношений второго типа). По заданному вектору числовых значений размеров $d \in R^m$ найти объект $x \in A_\alpha$, такой, что вектор d и объект a удовлетворяют системе отношений $\{m, k, \alpha, \{(u_i, \theta_i)\}\}$, разрешимой относительно мультитипа α (ниже такого рода разрешимые системы будут называться системами отношений второго типа).

Задача Iб является основной задачей ПСГП первого и второго типов. Поскольку число отношений в системе может варьироваться, то в общем случае ее аналитическое решение найти невозможно. Поэтому для решения таких систем в ПСГП второго типа используются итерационные методы, например метод Ньютона в [2].

Для решения задачи Iб с помощью ПСГП операторного типа пользователь вынужден преобразовывать исходную систему отношений в некоторую эквивалентную ей, решение которой сводится к последовательному решению ряда систем первого типа. Кроме того, ему, как правило, приходится добавлять в исходную систему отношений целый ряд вспомогательных примитивов и отношений [1]. В некоторых ПСГП такое эквивалентное преобразование делается автоматически, но при этом на вид исходной системы накладываются очень сильные ограничения. Например, в [4, 5] рассматриваются системы, которые путем перестановки в них отношений можно привести к системе с блочно-трехугольной формой якобиана, причем каждый диагональный блок должен

содержать все параметры одного из примитивов. Естественно, что после такого преобразования задача Iб непосредственно сводится к решению последовательности систем первого типа.

Аналогичный прием используется для решения так называемой задачи натягивания контура [9] (построения группы элементов в [10], построения полигона в [11] и т.д.). Эта задача заключается в построении непрерывной линии, составленной из отрезков прямой и дуг окружностей заданного радиуса, сопрягающих последовательность ранее построенных примитивов. И хотя постановка этой задачи самая общая (см. задачу I), очевидно, что ее решение сводится к последовательному решению систем первого типа, причем без перестановки отношений.

Перейдем теперь к обсуждению постановок задач на разрешимость систем отношений.

Что касается систем отношений первого типа, то их разрешимость вытекает из существования аналитического решения, и поэтому проблема разрешимости для таких систем решается еще на стадии разработки ПСПП (при этом в состав ПСПП первого типа, естественно, включаются только разрешимые системы).

В системах же второго типа, а также в общей постановке задачи I конкретный вид и число отношений в системе, а также числовые значения размеров варьируются, и поэтому может оказаться необходимым (это зависит от используемого метода решения системы) сначала проверить разрешимость системы. Кроме того, проверка разрешимости системы может иметь самостоятельный интерес (без поиска ее решения), например, с целью получения параметрического описания типовых фрагментов (см. выше).

Задача 2 (разрешимость системы отношений). Доказать (или опровергнуть) разрешимость относительно мульттипа α_1 системы отношений $\{\pi, k, (\alpha_1, \alpha_2), \{(u_i, \theta_i)\}\}$, т.е. доказать существование объектов $a_1 \in A_{\alpha_1}$, $a_2 \in A_{\alpha_2}$ и вектора $d \in R^m$, таких, что (a_1, a_2) и d удовлетворяют системе отношений и $\det [\partial f / \partial p_j] \neq 0$ в точке (a_1, a_2, d) , где f_i — А-функция, соответствующая отношению u_i , $p \in A_\alpha$.

Как уже отмечалось в п. I.I, разрешимая система отношений, вообще говоря, может иметь несколько решений, одно решение

или вообще ни одного. В связи с этим возникает проблема выбора нужного варианта решения.

В ПСГП операторного типа для систем отношений первого типа эта проблема решается следующим образом. На стадии разработки ПСГП проводится анализ системы отношений, в результате которого находятся все операторы построения примитивов. Затем пространство искомых примитивов A_t (см. постановку задачи Ia) разбивается на конечное число подобластей, каждая из которых соответствует какому-либо одному из операторов построения. Выбор нужного варианта построения искомого примитива осуществляется пользователем путем выбора одной из этих подобластей. Фактически этот выбор эквивалентен добавлению к системе отношений первого типа одного или нескольких отношений из класса ограничений (соответствующие а-функции будут неравенствами), связывающих искомый примитив с некоторыми из ранее построенных примитивов.

Для задачи Iб (и тем более задачи I) выбор нужного варианта решения усложняется. В настоящее время используется только один метод: в диалоговом режиме работы оператор вводит приближенный эскиз решения, который затем используется в качестве начального приближения для метода Ньютона. Заметим, что такой метод годится и для решения систем отношений первого типа, и именно такой метод используется в интерактивной системе моделирования (ИНСМ) (см. п. 2 данной работы). При этом пока остаются открытыми как вопрос о сходимости метода вообще (теория гарантирует сходимость только для случаев "достаточно" близкого расположения начального приближения к решению системы [12]), так и о сходимости его именно к нужному решению.

Следующая группа задач связана с анализом уже построенного составного геометрического объекта.

Задача 3. (метрический анализ объекта). По заданному объекту $a \in A_\alpha$ найти числовое значение размера $d \in R$ такое, чтобы объект a и величина d удовлетворяли заданной системе из единственного отношения (размера) $\{1, 0, \alpha, \{(u, \theta)\}\}$, т.е. чтобы

$$(\theta''(a), d) \in u \subset A_{\theta''(\alpha)}.$$

Задача 4 (позиционный анализ объекта). Для заданного объекта $a \in A_\alpha$, отношения u из класса ограничений и мультииндекса θ проверить выполнение соотношения

$$\theta''(a) \in u \subset A_{\theta'}(\alpha).$$

2. Геометрические построения в системе ИНСМ

Интерактивная система моделирования (ИНСМ) предназначена для моделирования плоских геометрических объектов и чертежей машиностроительного типа на ЭВМ класса IBM PC и СМ-1420. Интерактивная версия системы ИНСМ состоит из трех подсистем: 1) подсистемы геометрических построений, 2) подсистемы 2Д-моделирования плоских областей с многосвязной границей (аффинные преобразования, теоретико-множественные операции, эквидистанты и т.д.) и 3) подсистемы оформления чертежа в соответствии с требованиями ЕСКД (простановка различного рода размеров, обозначений и надписей). В основе последних двух лежат соответствующие подсистемы пакетов программ ГРАФИТ и СПЕЙС [II-14], для которых был разработан диалоговый интерфейс.

2.1. Решение систем отношений первого типа

В рамках подсистемы геометрических построений системы ИНСМ развивается предложенный в [I, I5] новый подход к реализации ПСГП, объединяющий в себе черты ПСГП первого и второго типов. С одной стороны, построение примитивов ведется последовательно, что характерно для ПСГП операторного типа, с другой - в основе предлагаемого подхода лежит понятие отношения между примитивами.

В настоящее время в системе ИНСМ реализовано около 60 одно- и двухместных отношений, предназначенных для построения точек, отрезков прямых и окружностей (см. Приложение). Характерная особенность всех отношений: соответствующие им А-функции являются полиномами от параметров примитивов, причем порядок полиномов относительно параметров искомого примитива не выше второго.

Для построения нового примитива оператор вводит (в режиме диалога) одно или несколько отношений, связывающих параметры нового примитива с параметрами некоторых из ранее построенных примитивов и, возможно, с некоторым набором числовых значений. После того как число введенных отношений сравняется с числом параметров нового примитива, ввод отношений прекращается.

Если среди уравнений, полученных в результате ввода отношений, есть хотя бы одно нелинейное относительно параметров нового примитива, то может существовать сразу несколько вариантов решения. Для выбора нужного варианта оператору достаточно указать на экране его приблизительное расположение, которое затем будет использовано системой в качестве начального приближения для итерационного процесса.

Решение введенной оператором системы отношений осуществляется с помощью итерационного метода Ньютона [12].

Запишем систему отношений первого типа $\{m, k, ((t), \alpha), \{(u_i, \theta_i)\}\}$, $m + k = |t|$ в виде системы, вообще говоря, нелинейных уравнений:

$$f_i(p, q, d) = 0, \quad (i = 1, 2, \dots, |t|), \quad (1)$$

где t – тип нового примитива; f_i – A-функция, соответствующая отношению u_i ; $p \in R^{|t|}$ – вектор неизвестных параметров нового примитива; $q \in R^{|\alpha|}$ – вектор параметров ранее построенных примитивов; $d \in R^m$ – вектор числовых значений размеров (может отсутствовать). Тогда n -е приближение $p^{(n)}$ определяется путем решения следующей системы линейных алгебраических уравнений:

$$f_i(p^{(n-1)}, q, d) + (\text{grad}_i(p^{(n-1)}, q, d), p^{(n)} - p^{(n-1)}) = 0,$$

$$i = 1, 2, \dots, |t|. \quad (2)$$

Из (2) видно, что для решения этой системы достаточно для каждого отношения u_i вычислить значение функции f_i и ее частных производных. Для вычисления этих значений в системе ИСМ

имеется специальная подпрограмма, на вход которой поступают: тип нового примитива; тип ранее построенного примитива и массив его параметров; тип отношения, словесное значение размера, а также полученный на предыдущей итерации массив параметров нового примитива. Таким образом, каждому отношению в системе ИНСМ соответствует блок всего лишь из 5-6 операторов (для вычисления значения функции и ее частных производных).

Если в процессе выполнения итераций возникает система линейных уравнений с вырожденной матрицей, то оператор получит сообщение о невозможности решения введенной им системы либо ввиду ее вырожденности (функциональной зависимости А-функций), либо вследствие неудачного ввода начального приближения.

Если для некоторого n модуль невязки $\sum_i |f_i(p^{(n)}, q, d)|$ меньше некоторого ϵ , то итерационный процесс считается успешно завершенным, вычисленные параметры нового примитива сохраняются в модели объекта, а изображение примитива выводится на экран дисплея.

Если же число итераций i или число случаев роста модуля невязки превышают некоторые пороговые значения (в текущей версии системы - 20 и 5 соответственно), то оператор получает сообщение о том, что система отношений решения не имеет.

Вообще говоря, метод Ньютона гарантирует сходимость итераций только в случае достаточной близости начального приближения к решению (I). Многочисленные эксперименты показали, что если система уравнений (I) разрешима и имеет решение, то итерационный процесс сходится при любом начальном приближении, и, кроме того, если существует более чем один вариант построения, то процесс сходится обязательно к ближайшему решению.

Замечание 4. Для некоторых типов примитивов удобно использовать зависимые параметры (в системе ИНСМ это реализовано для прямых). При построении любого примитива такого типа к системе (I) добавляется соответствующее число функциональных зависимостей на его параметры, т.е. каждая такая зависимость рассматривается как некоторое одноместное отношение.

З а м е ч а н и е 5. Для ускорения работы оператора в систему ИНСМ введено несколько "составных" отношений, каждому из которых соответствует два уравнения.

2.2. Сравнение двух подходов

В настоящем пункте проводится сравнение реализованного в рамках системы ИНСМ подхода с используемым в ПСПП второго типа подходом.

Хотя понятие отношения в ПСПП операторного типа в явном виде обычно не фигурирует (за исключением языков ИНКАНЭЛ [10], ОГРА [16], а также входного языка [17]), тем не менее практически каждый из реализованных в них операторов построения фактически является программой, предназначеннной для решения некоторой конкретной системы отношения первого типа. В то же время в системе ИНСМ, с одной стороны, вообще нет операторов построения, но с другой – представляется возможность решения любой разрешимой системы первого типа (составленной из реализованных в ИНСМ отношений). Этот факт позволяет говорить о ИНСМ как о функционально полной системе в области геометрических построений.

Несложные расчеты показывают, что реализованные в системе ИНСМ отношения предоставляют в распоряжение пользователя 44, 51 и 766 различных способов построения точек, прямых и окружностей соответственно. (При подсчете числа способов учитывались только разрешимые системы отношений, а все системы, различающиеся только порядком отношений, рассматривались как одна и та же система.) Для сравнения: аналогичные данные для ПСПП ФАП-КФ [9] составляют 24, 16 и 35 способов соответственно, а для ПСПП ГРАФИТ [11] – 24, 7 и 17 способов (здесь при подсчете числа способов построения не учитывались аффинные преобразования, так как в них используются трехместные отношения).

В отличие от системы ИНСМ практически для всех ПСПП операторного типа характерна их функциональная неполнота. Например, если выписать все отношения, используемые в операторах построения окружностей в ПСПП ФАП-КФ или ПСПП ГРАФИТ, а затем подсчитать, сколько должно было бы быть в этих пакетах способов

бов построения окружностей, соответствующих всем возможным разрешимым комбинациям этих отношений, то получится, что реально доступное число способов построения составляет менее 10%.

Другое важное преимущество предложенного подхода — простота и удобство интерфейса для пользователя. Действительно, построение каждого примитива сводится к выбору из нескольких меню и, возможно, к указанию на несколько ранее построенных примитивов и вводу числовых значений размеров. При этом порядок ввода отношений может быть любым. Отметим также простоту принятого в системе ИНСМ метода выбора варианта построения (достаточно просто указать на экране дисплея приблизительное расположение нужного варианта).

В то же время функциональная неполнота, малоинформационные имена операторов построения, фиксированный порядок operandов, сложные и громоздкие методы выбора варианта построения — все эти характерные особенности несомненно затрудняют как изучение, так и взаимодействие с ПСП операторного типа.

Один из основных принципов, положенный в реализацию системы ИНСМ, — принцип независимости обработки отношений (как на уровне ввода системы отношений, так и на уровне ее решения). В результате программный код системы геометрических построений занимает сравнительно малый объем оперативной памяти. Он содержит 13 меню, программу решения системы уравнений методом Ньютона, по 5–6 операторов на каждое отношение (для вычисления значений соответствующей функции и ее частных производных) и средства указания и отображения для каждого типа примитивов.

Независимость обработки отношений позволяет предельно упростить ввод в систему ИНСМ новых типов отношений. Для этого достаточно ввести в одно из меню дополнительную альтернативу и добавить блок из 5–6 операторов для вычисления значений а-функции и ее частных производных. Кроме того, добавление только одного типа отношения, как правило, дает сразу несколько новых способов построения примитива.

Заключение

Основным недостатком текущей версии системы ИНСМ является практически полное отсутствие отношений, связанных с отрезками и дугами окружностей (если не принимать во внимание операцию "усечения" прямых и окружностей по их точкам пересечения с другими примитивами). Другой важный недостаток — отсутствие трехместных отношений. Поэтому в следующей версии системы ИНСМ планируется существенное расширение функциональных возможностей системы путем введения одно-, двух- и трехместных отношений, связывающих отрезки прямых и дуги кривых второго порядка.

Другое важное направление развития системы ИНСМ связано с расширением системы геометрических построений в область пространства R^3 . В числе основных примитивов планируются точка, прямая, плоскость, система координат (репер в R^3), шар, цилиндр и некоторые другие. Основные отношения: условия инцидентности, касания, симметрии, параллельности; расстояния, углы и т.п. На основе вышеизложенного подхода предполагается разработка мощных и удобных средств размещения трехмерных объектов в пространстве.

Л и т е р а т у р а

1. Голубев В.М. Формальное представление геометрии деталей на чертежах// Машинная графика и ее применение. — Новосибирск, 1987. — С. 70-97.
2. Light R., Gossard D. Modification of geometric models through variational geometry // Comput. Aided Des.— 1982.- Vol. 14, N 4. — P. 209-214.
3. Брунс О.П. Система размерного кодирования плоских изображений// Автоматизация технической подготовки производства оснастки. — Рига, 1981. — С. II-17.
4. Параметризация чертежей по размерам, основанная на задании образца. Постановка задачи/ А.И.Горлин, В.Н.Коваленко, Е.В.Хухлаев. — М., 1987. — 22 с. — (Препринт/ АН СССР. ИПМ им. М.В.Келдыша; 66).
5. Параметризация чертежей по размерам, основанная на задании образца. Построение экземпляра методом поэлементного рас-

- чата. - М., 1987. - 20 с. - (Препринт / АН СССР. ИПМ им. М.В.Келдыша; 184).
6. Система геометрического моделирования "Гром"/ Е.В.Биряльцев, А.М.Гусенков, И.Р.Насыров, А.А.Савельев// Машинная графика "89": Тез. докл. Всесоюз. конф. - Новосибирск, 1989. - С. II3.
 7. Голубев В.М. Моделирование геометрии деталей с помощью чертежей// Тез. докл. Всесоюз. конф. по проблемам машинной графики, Протвино, сент. 1987. - Серпухов, 1987. - С. 173.
 8. ГОСТ 2.305-68. Единая система конструкторской документации. Общие правила выполнения чертежей. Издание официальное. - М., 1971.
 9. Горелик А.Г. Пакет программ машинной графики для ЭВМ. - М.: Машиностроение, 1986.
 10. Стародетко Е.А. Элементы вычислительной геометрии. - Минск: Наука и техника, 1986.
 11. Сиротин В.Г. Язык ГРАФИТ для описания, редактирования, хранения и визуализации моделей двумерных геометрических объектов и чертежей// Машинная графика и ее приложения. - Новосибирск, 1983. - С. 63-91.
 12. Орtega Дж., Рейнbold B. Итерационные методы решения нелинейных систем уравнений со многими неизвестными/ Пер. с англ. - М.: Мир, 1975.
 13. Упольников С.А. Алгоритм построения распадающейся эквидистанты// Машинная графика "89": Тез. докл. Всесоюз. конф. - Новосибирск, 1989. - С. 76.
 14. Голубев В.М., Упольников С.А. ПП СПЕЙС на СМ ЭВМ// Там же. - С. I09.
 15. Голубев В.М. Геометрические построения в системе ИНСМ// Там же. - С. 75.
 16. Зозулевич Д.М. Машинная графика в автоматизированном проектировании. - М.: Машиностроение, 1976.
 17. Ксенофонтова Л.Н., Рябов Е.К. Интерактивный синтез топологической модели изображения чертежа// Машинная графика "89": Тез. докл. Всесоюз. конф. - Новосибирск, 1989. - С. 84.

Приложение

Список отношений, реализованных в системе ИНСМ

Одноместные отношения на точки: x -координата, y -координата.

Одноместные отношения на прямые: угол наклона к оси Ox , угол наклона к оси Oy , вертикальность, горизонтальность, x -координата точки пересечения с осью Ox , y -координата точки пересечения с осью Oy .

Одноместные отношения на окружности: x -координата центра, y -координата центра, радиус, диаметр.

Двухместные отношения, связывающие две точки: расстояние, разность x -координат, разность y -координат, равенство x -координат, равенство y -координат.

Двухместные отношения, связывающие точки и прямые: точка лежит на прямой, расстояние.

Двухместные отношения, связывающие точки и окружности: точка лежит на окружности, точка - центр окружности (два уравнения), расстояние от точки до центра, разность x -координат точки и центра, разность y -координат, равенство x -координат, равенство y -координат.

Двухместные отношения, связывающие две прямые: расстояние между параллельными прямыми (два уравнения), угол, параллельность, перпендикулярность.

Двухместные отношения, связывающие прямые и окружности: касание, центр лежит на прямой, расстояние от центра до прямой.

Двухместные отношения, связывающие две окружности: внешнее касание, внутреннее касание, концентричность (два уравнения), расстояние между центрами, разность x -координат центров, разность y -координат, равенство x -координат, равенство y -координат, равенство радиусов, разность радиусов, разность диаметров, центр одной окружности лежит на другой.

Н.С.Щупта

СИСТЕМА КОНСТРУКТИВНОГО ГЕОМЕТРИЧЕСКОГО МОДЕЛИРОВАНИЯ
ТВЁРДЫХ ТЕЛ

I. Введение

В настоящей статье описывается система конструктивного геометрического моделирования твёрдых тел, являющаяся частью системы геометрического моделирования для САПР машиностроительного профиля, разрабатываемой ВЦ СО АН СССР и НГУ.

При разработке первой версии системы были использованы:

- опыт разработки системы СМОГ-85 [1-4],
- спецификации ряда отечественных и зарубежных систем геометрического моделирования (БПМО АСК, BRAVO! и CATIA[5-8]);
- методы и алгоритмы, изложенные в литературе [9-15].

2. Задачи системы геометрического моделирования

Основными задачами, решаемыми системой геометрического моделирования для САПР, являются:

- визуальный анализ моделей;
- расчёт метрических и интегральных характеристик (численный анализ) моделей;
- построение видов.

Визуальный анализ необходим для контроля правильности модели на различных этапах проектирования. При визуальном контроле используются два способа построения изображений: векторный и растровый.

Векторная визуализация – изображение модели твёрдого тела в виде совокупности линий на плоскости. Возможны два режима векторной визуализации: с удалением и без удаления невидимых линий.

Растровая визуализация – изображение модели твёрдого тела в виде совокупности закрашенных областей на плоскости.

К численному анализу относятся операции расчёта характеристик моделей твёрдых тел, которые можно разбить на следующие группы:

- определение принадлежности точки пространства моделируемому телу;
- вычисление длин, площадей, объёмов, расстояний и габаритов для моделируемых тел;
- расчёт центров, осевых и центробежных моментов инерции моделируемых тел в предположении их однородности и единичной плотности.

Вид – это специальная каркасная трёхмерная модель объекта, предназначенная для построения чертежей. Вид строится на основании конкретного преобразования проецирования и содержит линии (составные кривые) следующих типов:

- видимые действительные линии (границы кривые и линии пересечения);
- невидимые действительные линии;
- видимые линии очерка;
- невидимые линии очерка;
- помеченные контуры (границы сечений, вырывов, разрезов и т.п.).

3. Конструктивная геометрическая модель твёрдого тела

Основное требование к конструктивной модели – обеспечить высокую точность выполнения операций над моделями твёрдых тел, что достигается путём сохранения точной информации о проектируемом изделии (геометрические, физические и структурные особенности).

Конструктивная геометрическая модель (КГМ) представляет собой:

- некоторое множество примитивных объектов (примитивов);

- конструктив (формула построения) – бинарное дерево построения конструктивной модели из примитивов с использованием теоретико-множественных операций;

- конечное множество специальных параметров – атрибутов, связанных с примитивами, конструктивами и/или объектами и отражающих особенности объекта или модели.

Работа с конструктивной моделью (в отличие, например, от работы с полиэдральной моделью) не зависит от точности представления модели. Примитивными объектами конструктивной модели служат базовые геометрические формы с соответствующими атрибутами. Базовые геометрические формы (БГФ) – точечные, гранично-заданные, регуляризованные, ограниченные подмножества трёхмерного евклидова пространства. Задание границ БГФ осуществляется перечислением ограничивающих его элементарных поверхностей.

Чаще всего предусматривается следующий набор базовых геометрических форм:

- прямоугольный параллелепипед,
- прямой эллиптический конус,
- прямой эллиптический цилиндр,
- шар,
- тор.

Над примитивами определены операции позиционирования в пространстве (поворот, сдвиг и масштабирование), задаваемые матрицей аффинного преобразования системы координат примитива (ПСК).

К атрибутам примитива относятся:

- положение примитива в пространстве;
- точность аппроксимации примитива для построения граничной модели.

Возможно появление дополнительных атрибутов, например, допуск значений размеров (точность задания), цвет, габариты и т.п.

Конструктивное представление описывает сложный объект как композицию, полученную из примитивных с использованием операций построения.

В качестве операций построения в бинарном дереве построения конструктивной модели используются теоретико-множественные операции: объединение, пересечение и вычитание.

К атрибутам сложного объекта, описываемого конструктивной геометрической моделью (КГМ), относятся: положение в пространстве; ряд дополнительных атрибутов, отражающих его особенности (цвет, плотность, материал и т.д.) или используемых для повышения эффективности алгоритмов работы системы (габариты).

Построение конструктивной модели включает:

- задание примитивов;
- размещение заданных примитивов в пространстве;
- задание теоретико-множественной операции над примитивами или уже построенными конструктивами и получение из них нового конструктива.

4. Алгоритмы работы с КГМ

Основным алгоритмом работы с конструктивной моделью является алгоритм "оценивания" лучом.

Суть метода "оценивания" лучом [14] состоит в "протыкании" конструктивной модели объекта семейством лучей, т.е. нахождении точек пересечения лучей с моделью. "Протыкание" лучом заключается в:

- определении точек пересечения луча со всеми примитивами модели;
- формировании на луче семейства отрезков, лежащих внутри примитивов модели;
- выполнении одномерных теоретико-множественных операций над отрезками луча в соответствии с формулой конструктивной модели (получение "следа" луча).

Для решения конкретных задач осуществляется анализ "следов" семейства лучей.

Временные характеристики метода "оценивания" лучом линейно зависят от числа примитивов и числа лучей.

4.1. Определение множества лучей

Определение множества лучей в конструктивном процессоре базируется на понятии "экран".

При определении экрана будем считать, что все модели формируются в некотором "кубе", расположенном в трёхмерном евклидовом пространстве, в котором определена глобальная система координат (ГСК).

Экран - плоскость, расположенная в трехмерном евклидовом пространстве. С экранной плоскостью связана экранная система координат (ЭСК) так, что плоскость $Z = 0$ ЭСК совпадает с экранной плоскостью и модель находится в положительном по оси Z полупространстве. Проекция пространства моделирования на экранную плоскость задает границы экрана.

Весь экран разбит на площадки - пиксели. Число пикселей экрана по осям X и Y задает размер раstra. Общее число лучей равно числу пикселей.

Возможны два способа проецирования: центральное и параллельное. При параллельном проецировании каждый луч начинается в центре соответствующего пикселя экрана и все лучи параллельны одному направлению (оси Z ЭСК). При центральном проецировании все лучи начинаются в одной точке - центре проецирования и каждый луч проходит через центр соответствующего пикселя. Лучи задаются параметрически.

В случае параллельного проецирования уравнение луча LI выглядит следующим образом:

$$LI(t) = PI + t * D,$$

где $PI = (x_I, y_I, 0)$ - координаты точки начала луча (центр первого пикселя); $D = (0, 0, 1)$ - вектор направления лучей.

В случае центрального проецирования уравнение луча LI имеет вид

$$LI(t) = P + t * DI,$$

где $P = (0, 0, -z_0)$ - координаты точки начала лучей (точки проецирования); $DI = (x_I, y_I, z_0)$ - вектор направления луча, определяемый точкой проецирования и центром первого пикселя.

4.2. Пересечение луча с моделью

Пересечение луча с конструктивной моделью производится путем обхода бинарного дерева построения конструктивной модели в концевом порядке. При обходе дерева выполняются следующие действия:

- вычисление точек пересечения луча с листьями дерева (примитивами) - нахождение элементарных следов луча;
- выполнение операции "комбинирования" уже полученных следов луча в соответствии с теоретико-множественными операциями из бинарного дерева - построение следа луча.

Аффинные преобразования параметрически заданного луча не меняют его параметризации. Это позволяет (при решении задачи нахождения точек пересечения луча с примитивом) заменить преобразование поверхностей примитивов в ГСК преобразованием луча в систему координат поверхности (ЛСК). Таким образом, задача пересечения луча с примитивом сводится к:

- преобразованию луча в систему координат примитива;
- подстановке параметрического уравнения луча в каноническое уравнение поверхности и решению полученного уравнения относительно параметра луча t ;
- проверке (если это необходимо) соответствия точек пересечения ограничениям на каноническое уравнение поверхности;
- классификации точек луча относительно примитива - внутри/вне.

След луча задает множество отрезков луча, содержащихся целиком внутри примитива, и содержит информацию о границах, пересекаемых лучом.

Операция "комбинирования" следов определяет участки луча, принадлежащие модели, и граничные поверхности. Она включает:

- выполнение теоретико-множественной операции над отрезками, задаваемыми точками следов луча;
- копирование указателей граничных поверхностей;
- классификацию точек результирующего следа.

Пример пересечения луча с моделью приведен на рис. I.

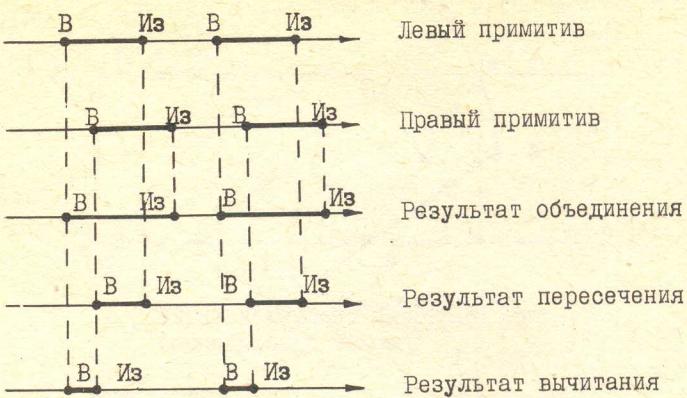


Рис. I. Пересечение луча с моделью

5. Решение задач геометрического моделирования

5.1. Построение видов

Следы лучей используются для нахождения линий пересечения граничных поверхностей, граничных и очерковых линий.

Для пояснения процесса нахождения точек линий пересечения, граничных или очерковых линий, рассмотрим следы двух соседних лучей. Пусть лучи L_1 и L_2 пересекают поверхности I и J . Возможные случаи пересечения приведены на рис. 2-7, где использованы следующие обозначения: T_1, T_2 - точки следа луча L_1 ; S_1, S_2 - точки следа луча L_2 ; I, J - номера (указатели) поверхностей; R - точка линии пересечения, граничной или очерковой линии; P - приближенная точка, полученная системой.

Все получаемые линии - точечно-заданные. Информация о типе поверхностей, порождающих линии, может позволить определить элементарные кривые, составляющие полученную линию. Например, при пересечении плоскости и сферы по точечно-заданной линии может быть восстановлена дуга окружности.

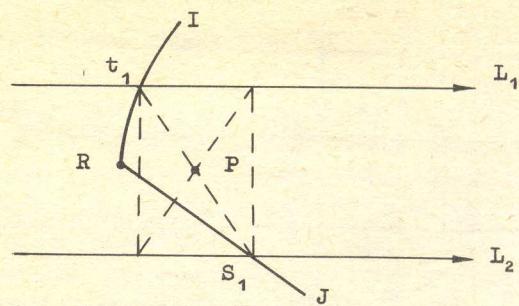


Рис. 2. Точка граничной линии
(плохое приближение)

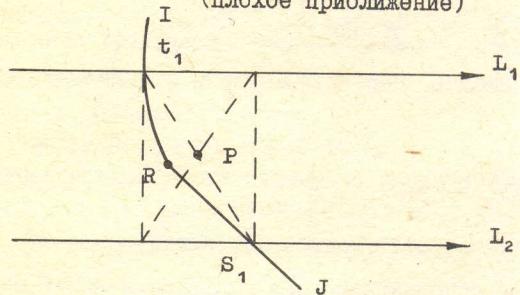


Рис. 3. Точка граничной линии
(хорошее приближение)

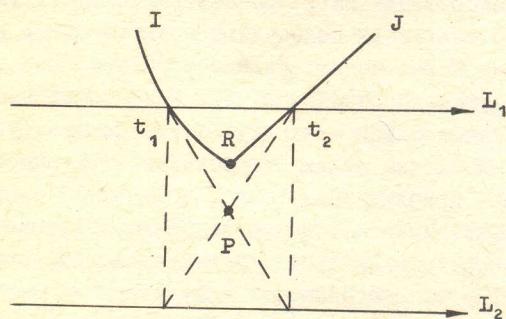


Рис. 4. Точка граничной и очерковой линий

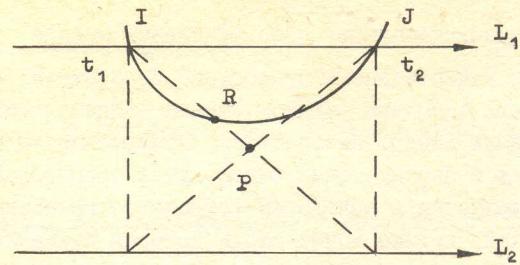


Рис. 5. Точка очерковой линии

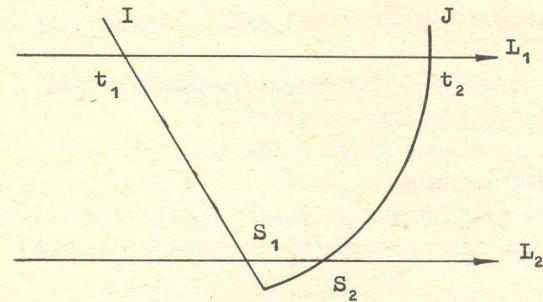


Рис. 6. Между лучами нет граничной или очерковой линии

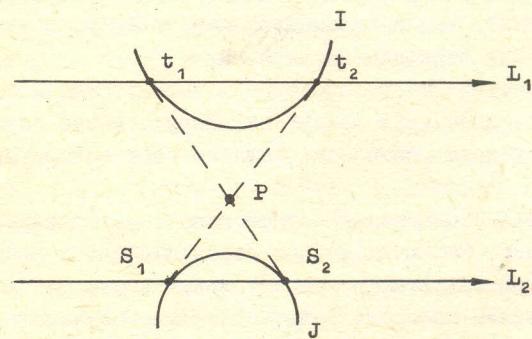


Рис. 7. Лишняя точка граничной линии

5.2. Визуальный анализ конструктивной модели

При получении качественного векторного изображения по конструктивной модели основным режимом вывода является вывод с удалением невидимых линий. Изображение формируется на основе анализа ближайших к экрану точек следа луча. Формирование изображения заключается в выводе на устройство границы пикселей соседних лучей, первые точки следа которых принадлежат разным поверхностям. Для вывода без удаления невидимых линий используются все точки следа каждого луча.

Качественное растровое изображение конструктивной модели формируется методом "трассировки" луча [14]. При этом задаются:

- интенсивность и расположение источников света,
- интенсивность рассеянного света,
- параметры отражения света объектом,
- направление взгляда.

На основании этих данных по модели Фонга [12, 15] для видимых точек следа каждого луча рассчитывается яркость.

5.3. Численный анализ конструктивной модели

Для определения принадлежности точки к объекту, модель объекта "протыкается" лучом, проходящим через точку, и анализируется след луча. Точка, принадлежащая объекту, попадает в один из внутренних отрезков следа луча.

Задавая множество лучей, проходящих через точку и сферу, вокруг неё (т.е. проецируя объект на сферу), можно определить расстояние между двумя объектами, однако этот алгоритм слишком медленный.

Пусть в методе "оценивания" лучом используется параллельное проецирование. Заключим каждый отрезок каждого луча, содержащийся внутри объекта, в прямоугольный параллелепипед (стержень), сечение которого равно пикслю. Полученный набор стержней будем использовать в качестве приближенного представления объекта для дальнейших расчетов.

Объём, центр и моменты инерции тела определяются как объем, центр и моменты инерции совокупности стержней. Знание моментов позволяет определить тензор инерции.

6. Замечания о реализации

К настоящему времени реализована первая очередь системы конструктивного геометрического моделирования, включающая:

- подсистему ввода и визуализации геометрической модели на ПЭВМ типа РС;

- алгоритмы лучевого оценивания конструктивной модели.

На основе лучевых алгоритмов реализованы программы качественной растровой визуализации КГМ на ПЭВМ типа РС с графическим адаптером EGA и ЭВМ типа СМ с дисплеем ГАММА-4.2. Примеры растровых изображений приведены на рис. 8-10.

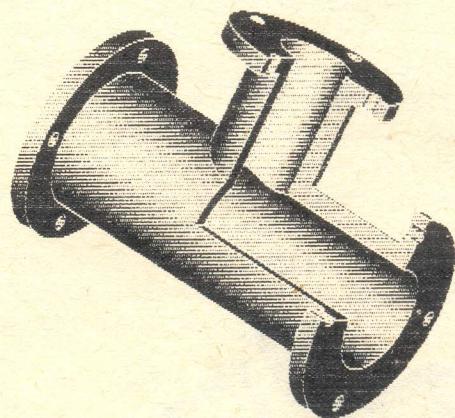


Рис. 8. Тройник

Все алгоритмы реализованы на языке Фортран.

Подсистема формирования КГМ позволяет формировать модель в режиме графического диалога. Для быстрой визуализации формируемой модели используется векторная визуализация примитивов без удаления невидимых линий. Активные примитивы выделяются цветом. Кроме того цвет изображения примитива зависит от способа (операции) его включения в дерево (объединение, пересечение или вычитание).

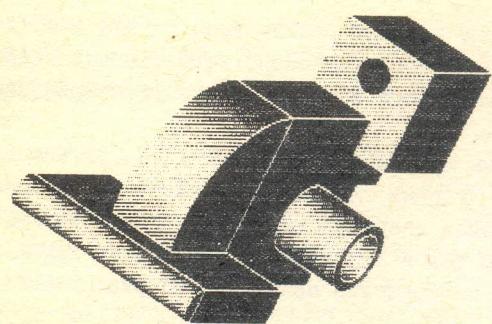


Рис. 9. Станина

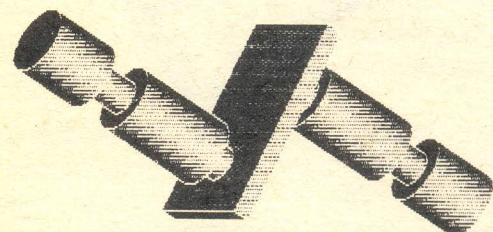


Рис. 10. Пластина

Готовая модель может быть через файл передана программе качественной визуализации.

7. Дальнейшее развитие

При развитии системы предполагается:

- реализовать программы численного анализа модели;
- разработать и реализовать алгоритмы восстановления линий пересечения и очерковых линий;
- разработать на основе "оценивания" лучом алгоритмы перевода конструктивной геометрической модели в полиэдralное представление.

Подробнее поясним суть последних алгоритмов. Множество точек пересечения лучей с КГМ объекта задает на его поверхности сетку, по которой может быть построен набор треугольных граней, аппроксимирующий поверхность объекта. Информация о типе точек (точка входа или выхода луча) позволяет ориентировать грани в пространстве и получить модель, а при полиэдralном - нет.

Множество точек пересечения лучей с КГМ объекта задает на его поверхности сетку, по которой может быть построен набор треугольных граней, аппроксимирующий поверхность объекта. Информация о типе точек (точка входа или выхода луча) позволяет ориентировать грани в пространстве и получить полиэдralную модель объекта [16].

При этом возникает задача минимизации числа граней аппроксимирующего многогранника и множества лучей для получения многогранного представления. При визуализации или численном анализе КГМ лучи задаются равномерно. При аппроксимации КГМ полиэдром или при получении видов можно в областях экрана, в которых проекция объекта не имеет мелких особенностей (аппроксимация) или не содержит контурных или граничных линий (построение видов), увеличивать расстояния между лучами, а при приближении к границе или особенности - увеличивать плотность лучей.

Решение этой задачи может проводиться в интерактивном режиме. После получения "грубого" приближения пользователю демонстрируется полученное изображение и используемая сетка лучей. Затем пользователь задает необходимое стяжение или разрежение сетки.

Л и т е р а т у р а

1. Дебелов В.А., Мацокин А.М. Система машинной графики СМОГ-85//Упр. системы и машины.- 1986.- Т. 4.- С. 107-122.
2. Мацокин А.М., Упольников С.А. ИПП SPACE - средство формирования и обработки моделей трехмерных объектов// Программные средства машинной графики.- Новосибирск, 1984.- С. 77-90.
3. Упольников С.А. Реализация алгоритмов обработки полиэдров// Там же.- С. 91-103.
4. Мацокин А.М. Упольников С.А. Построение плоских образов многогранных поверхностей// Машинная графика и ее приложения.- Новосибирск, 1987.- С. 37-50.
5. Система "BRAVO" фирмы "Априкон"/ Есида М.; ВЦП.- № И-30182.- М.- 10с.- Пер. ст. из журн.: Кикай нео когу.- 1983.- Vol. 27, № 6.- Р. 15-18.
6. Руководство пользователю системой моделирования физических тел. П/ВЦП.- № Н-20636.- М., 06.04.87.- 171 с.- Пер. ст. из журн.: Applicon Schlumberger. - 1985.
7. Справочное руководство к системе "Солидс моделинг - П"/ ВЦП.- № Н-20637.- М., 09.04.87.- 76 с.- Пер. ст. из журн. Applicon Schlumberger. - 1985.
8. CATIA. Основы пространственного конструирования: Руководство для пользователя. Т. I./ ВЦП.- № М-14564.- М., 10.07.86.- 329 с.- Пер. материала фирмы: IBM.- 1985.- 385 р. США.
9. Requicha A.A.G. Representation for rigid solids: Theory, methods and systems //Computing Surveys. - 1980.- Vol. 15, N 4. - P. 437-464.

- I0. AYS Y14.26M - 1981 . Digital representation for communication of product definition data. - ANSY, 1981.
- II. Фоли Дж., Вэн-Дэм А. Основы интерактивной машинной графики.- М.: Мир, 1985.- Т. I.
- I2. То же.- Т. 2.
- I3. Гардан И., Лука М. Машинная графика и автоматизация проектирования.- М.: Мир, 1987.
- I4. Roth S.D. Ray casting for modelling solids //Computer Graphics and Image Processing.- 1982 .- Vol. 18, N 2.- P. 109-144.
- I5. Bui-Tuong Phong. Illumination for computer generated pictures //Communications of the ASM.- 1975. - Vol. 18. - N 6. - P. 311-317.
- I6. Упольников С.А. Реализация алгоритмов обработки полизадров// Программные средства машинной графики.- Новосибирск, 1984.- С. 91-103.

В.Г.Сиротин, Е.Е.Витяев

ПРОЕКТ ЭКСПЕРТНОЙ СИСТЕМЫ ЭЛЕКТРОННОЙ НАВИГАЦИИ

Проблематика электронной навигации средств наземного и водного транспорта охватывает широкий спектр задач, простирающийся от задачи оптимального управления отдельным транспортным средством до задачи глобального управления транспортными потоками в пределах определенного района.

В настоящее время в ряде стран анонсированы либо находятся в стадии реализации проекты комплексов электронной навигации наземных и водных транспортных средств [1-4].

Представляемый проект имеет своей целью формирование основных принципов построения систем электронной навигации и конкретную программную реализацию ряда наиболее важных подсистем, а также инструментальных средств, необходимых для такого рода разработок.

В общем случае в системе электронной навигации можно выделить следующие основные подсистемы:

- 1) получения информации о транспортных ситуациях;
- 2) анализа (распознавания);
- 3) мониторинга;
- 4) прогноза;
- 5) принятия решений.

В рамках описываемого проекта внимание концентрируется на подсистемах 2-4). При этом за чертой рассмотрения остаются специфические для каждой конкретной системы электронной навигации.

гации проблемы сбора исходной информации (зависящие от применяемых технических средств) и проблемы принятия решений (зависящие от назначения системы в целом). Выделенные задачи (анализ-распознавание, мониторинг, прогноз) относятся к классу задач, для которых, по-видимому, единственным способом практического решения является создание соответствующей экспертной системы [5]. Именно данный подход и намечено реализовать в описываемом проекте.

I. Основные определения

I.I. Понятие транспортной обстановки

Говоря о транспортной обстановке, будем различать ее стационарную и динамическую части. Под стационарной частью имеем в виду рельеф рассматриваемого участка местности, расположющиеся на нем естественные и искусственные объекты, объективно влияющие на движение транспортных средств и на качество работы устройств получения информации о них. Под динамической частью будем понимать движущиеся в пределах участка транспортные средства, которые ниже называем динамическими элементами (ДЭ).

I.I.I. Стационарная часть транспортной обстановки. Пусть $\Omega = [a, b] \times [c, d] \subset \mathbb{R}^2$ – прямоугольник на плоскости. Над Ω задана функция $z = f(x, y)$. Множество точек $(x, y, z=f(x, y))$, $(x, y) \in \Omega$ назовем опорной поверхностью.

Графом возможных перемещений будем называть планарный граф $G = (V(G), E(G))$ с вершинами $v \in V(G)$ и ребрами $e \in E(G)$, $e = (v_1, v_2)$, $v_i \in V(G)$. Одна пара вершин может быть соединена более чем с одним ребром. С каждой вершиной v связана точка $(x_v, y_v) \in \Omega$, а с каждым ребром e – гладкая кривая, целиком принадлежащая Ω , т.е. можно говорить о наложении графа возможных перемещений G на опорную поверхность. Каждое ребро графа характеризует числовой атрибут, интерпретируемый как его качество.

I.I.2. Динамическая часть транспортной обстановки. В любой момент времени t на поверхности участка Ω находится N_t динамических элементов u_k , $k = 1, N_t$. Терминалные узлы графа

возможных перемещений могут поглощать существующие и порождать новые динамические элементы.

Состояние каждого динамического элемента описывается кортежем $u_k = \langle Q_k, M_k \rangle$. Q_k будем называть объективным компонентом состояния ДЭ, а M_k - внутренним (поведенческим).

$$Q_k = (x_k, y_k, z_k, v_k, a_k),$$

где (x_k, y_k, z_k) - координаты динамического элемента, v_k - его скорость, a_k - ускорение. (Поскольку движение ДЭ может проходить только вдоль ребер графа возможных перемещений, далее вместо векторных представлений v_k и a_k будем рассматривать абсолютные значения проекций этих векторов на касательные к ребру графа возможных премещений в точке расположения данного ДЭ).

$$M_k = (c_k, m_k, r_k, s_k),$$

где c_k - идентификатор маршрута данного динамического элемента; m_k - идентификатор стратегии выбора (смены) маршрута элементом; r_k - идентификатор стратегии выбора ускорения элементом; s_k - идентификатор априорной принадлежности элемента некоторой группе динамических элементов.

Под маршрутом далее будем подразумевать упорядоченный набор ориентированных ребер графа возможных перемещений. В этом наборе конец предыдущего ребра должен обязательно совпадать с началом следующего. Кроме того, в наборе не должно быть одинаковых ребер и циклов. Понятия стратегии и априорной принадлежности элемента группы описаны ниже.

I.2. Поведенческий компонент

Как уже отмечалось, состояние каждого динамического элемента может быть зафиксировано и описано в терминах (параметрах) его объективного компонента $Q_k = (x_k, y_k, z_k, v_k, a_k)$. Изменение этих параметров будем называть эволюцией ДЭ. Каждый ДЭ эволюционирует в зависимости от определенного

набора правил и в соответствии с закономерностями, которые определяются параметрами его поведенческого компонента. Будем предполагать, что эволюция всех динамических элементов протекает следующим образом:

1. Находясь в неконечных точках какого-либо ребра графа возможных перемещений, любой ДЭ может в определенных пределах через равные промежутки времени скачкообразно менять свое ускорение a_k . В результате этого в последующем по законам кинематики будут меняться значения v_k и x_k, y_k, z_k . В частности, ДЭ может остановиться либо изменить направление движения на противоположное.

2. Попав в вершину графа возможных перемещений, ДЭ в соответствии со своим поведенческим компонентом может выбрать для дальнейшего движения то или иное его ребро, примыкающее к данной вершине, либо прекратить движение. Эволюция первого типа определяется стратегией выбора скорости, а второго — стратегией выбора маршрута.

I.3. Понятие стратегии движения

Как уже указывалось выше, любой ДЭ, находясь на каком-либо ребре графа возможных перемещений в достаточной удаленности от его концов, может скачкообразно менять значение своего ускорения через равные промежутки времени. Эти изменения производятся в соответствии с правилами, которые задаются системой вероятностных продукций. Вероятностными продукциями называют тройки вида

$$\{P_i(x_1, x_2, \dots, x_{n_i}), E_i, F_i(y_1, y_2, \dots, y_{m_i})\}. \quad (1)$$

Предикат $P_i(x_1, x_2, \dots, x_{n_i})$ будем называть определителем ситуации, а x_1, x_2, \dots, x_{n_i} — ее параметрами. Например, значение некоторого такого предиката можно считать истиной, если расстояние от данного ДЭ до ближайшего к нему другого ДЭ на данном ребре, движущегося в том же направлении и находящегося впереди него, больше, чем заданное.

Второй элемент, входящий в каждую продукцию, - это функция-генератор случайного числа из множества {0,1} с матожиданием выпадания единицы M_i .

С помощью функции $F_i(y_1, y_2, \dots, y_{m_i})$ вычисляется новое значение a_k .

В целом система продукции действует на основании следующего метаправила:

Шаг 1. $i = 1$.

Шаг 2. Применить предикат $P_i(x_1, x_2, \dots, x_{n_i})$, если $P_i(x_1, x_2, \dots, x_{n_i}) = 1$, то на шаг 3, иначе исключить продукцию i из дальнейшего рассмотрения (считать неприменимой к данному ДЭ в данной ситуации) и на шаг 5.

Шаг 3. Вычислить с помощью E_i значение случайной величины ξ и, если $\xi = 1$, то на шаг 4, иначе на шаг 5.

Шаг 4. Вычислить с помощью функции

$$F_i(y_1, y_2, \dots, y_{m_i}) \quad (2)$$

новое значение ускорения для данного ДЭ и на ВЫХОД.

Шаг 5. Имеются ли в системе продукции, которые ранее не были помечены как неприменимы? Если да, то i присвоить номер следующей (циклически) из них, и - на шаг 2, если таких продукции нет, то на ВЫХОД с информацией об ошибке.

Формально стратегия выбора маршрута, так же как стратегия выбора ускорения (скорости), представляет собой систему вероятностных продукции (1), действующую на основании метаправила (2). При выборе маршрута могут использоваться глобальные данные. Например, при сравнении маршрутов может быть выбран тот, который оптимальнее других по критерию минимизации времени достижения определенной вершины на графике возможных перемещений.

I.4. Постановки задач

В дальнейшем относительно способов получения информации о ДЭ будем предполагать следующее. Информация о находящихся на Ω ДЭ поступает порциями (кадрами) через определенные ин-

тервалы времени в виде неупорядоченных n -ок чисел $\{w_j\}$, $w_j = (w_j^1, w_j^2, \dots, w_j^n)$. С известной вероятностью можно утверждать, что наличие в кадре n -ки означает наличие в определенном участке Ω одного либо нескольких близкорасположенных ДЭ. Другими словами, существует некоторое неоднозначное отображение G множества ДЭ в множество, которое далее будем называть множеством отметок. Отображение G известно приблизительно и носит вероятностный характер.

I.4.1. Различие и распознавание групп ДЭ и отдельных ДЭ. На неформальном уровне эта задача может быть определена как задача отыскания разбиения множества отметок на группы, которое будет в некотором смысле эквивалентно разбиению на группы множества располагающихся на Ω динамических элементов. Дадим более формальное определение. На основании критериев, учитывающих близость координат и скоростей, все располагающиеся на Ω динамические элементы можно разбить на несколько групп. (При этом, возможно, некоторые группы будут содержать единственный элемент.) Каждая группа будет характеризоваться кортежем

$$T_1 = (m_1, H_1, B_1, v_1),$$

где m_1 – количество ДЭ в данной группе; H_1 – описание геометрии области, на которой располагаются ДЭ данной группы (например, координаты центра и размеры габаритного прямоугольника для точек расположения ДЭ на Ω); v_1 – средняя скорость ДЭ данной группы; B_1 – параметры, характеризующие форму данной группы.

Задача различия состоит в разбиении множества отметок на группы (возможно, с отфильтровыванием части отметок) таким образом, чтобы затем, на этапе распознавания по параметрам отметок, включенных в каждую группу, оценить параметры для ДЭ, образы которых в виде отметок вошли в данную группу.

I.4.2. Мониторинг. После распознавания и анализа некоторой последовательности кадров по каждому из них будет получен кортеж

$$\langle n_k, \{t_1^k\} \rangle,$$

где n_k – количество групп отметок на k -м кадре измерений;
 t_1^k – характеристический кортеж 1-й группы отметок k -го кадра измерений.

Далее в тексте каждый такой кортеж будем называть кластером, подразумевая группу ДЭ, образы которых в виде отметок позволили вычислить значения $\{t_1^k\}$. Задача мониторинга состоит в контролировании параметров каждого кластера в отдельности и всей совокупности кластеров в целом в течение определенного периода времени (на определенной последовательности кадров). Для этого необходимо прежде всего соотносить кластеры с $k-1$ -го и k -го кадров, определяя, в частности, возникновение новых, исчезновение, расщепление и слияние старых кластеров.

В случае, если с определенной вероятностью установлено соответствие между некоторым кластером на k -м кадре и кластером на $k-1$ -м кадре, можно говорить о произошедшей за период между измерениями эволюции кластера.

Между кластерами на последовательностях кадров могут возникать "динамические отношения", например: сближение, удаление, следование одного за другим и т.п.

I.4.3. Прогноз. Задачи прогноза делятся на три большие группы. В первую группу входят запросы типа: "Где будут находиться ДЭ, составляющие данный кластер через интервал времени δt ?".

Во вторую группу входят запросы типа: "Окажутся ли ДЭ данного кластера в некоторой подобласти Ω в интервал времени $[t_0, t_1]?$ "

Третья группа содержит запросы типа: "Сложится ли в интервал времени $[t_0, t_1]$ ситуация, определяемая предикатом $P_N(a_1, a_2, \dots, a_{m_N})?$ ". (Как можно видеть, запросы этого типа являются обобщением запросов второго типа.)

2. Концепция экспертной системы

Эффективность и качество системы в целом существенно зависят от возможностей подсистемы прогнозирования. В связи с этим рассмотрим основные предположения, лежащие в основе создания этой подсистемы.

Фундаментальным предположением, лежащим в основе используемого подхода к прогнозированию, является предположение о том, что закономерности, определяющие эволюцию группы динамических элементов, могут быть аппроксимированы системой вероятностных продуктов. При этом для эволюции разных групп могут быть использованы различные системы продукции.

На практике данный принцип предполагается реализовать следующим образом.

В работе системы будем различать этап накопления знаний и последующие сеансы непосредственной работы.

Этап накопления знаний достаточно продолжителен. На этом этапе эксперты и инженер знаний формируют универсальный (т.е. рассчитанный на максимально возможное число ситуаций) набор (архив) элементарных продуктов. (Элементарная продукция отличается от вероятностной отсутствием функции вычисления случайного числа.) На этом же этапе продукции связываются в ассоциации. Каждая ассоциация характеризует определенный класс групп динамических элементов. Члены класса считаются близкими друг другу по параметрам своего поведенческого компонента.

Во время сеанса непосредственной работы подсистема распознавания стратегий пытается на основании мониторинговых данных подобрать для каждой группы ДЭ одну из подготовленных ранее ассоциаций продуктов либо создать новую ассоциацию. По мере увеличения объема фактических данных об эволюции конкретной группы ДЭ уточняется порядок расположения продуктов и устанавливаются вероятностные коэффициенты (матожидания), в результате чего ассоциация элементарных продуктов превращается в систему вероятностных. При этом, возможно, на определенном этапе придется отказаться от ранее подобранный системы продуктов, если она войдет в противоречие с мониторинговыми данными.

Прогнозирование как таковое заключается в построении на основании мониторинговых данных и сформированной системы вероятностных продукций последовательности возможных состояний данной группы ДЭ (прямой ход) либо проверке гипотез с помощью этой же системы продукций (обратный ход).

3. Архитектура системы и назначение основных подсистем

Архитектура описываемой экспертной системы представлена на рис. I.

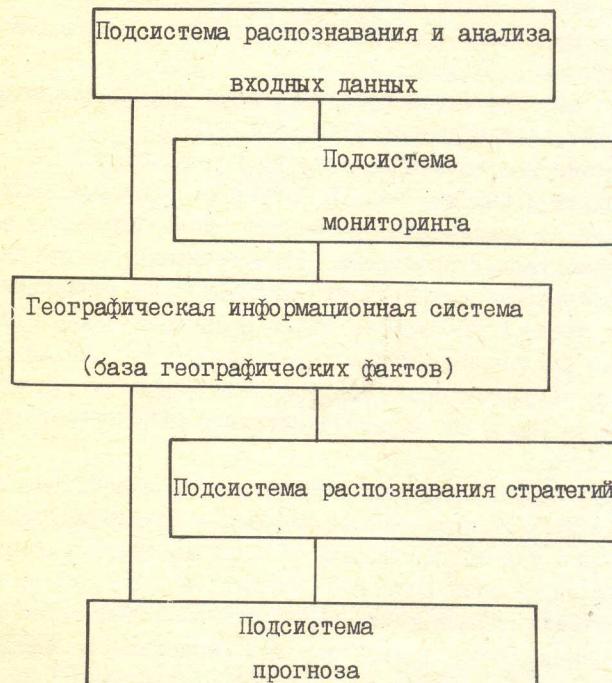


Рис. I. Архитектура ядра экспертной системы
электронной навигации

Подсистемы 1) и 2) работают в режиме конвейера и должны успевать обработать всю поступившую информацию за интервал времени между поступлениями очередных кадров. Подсистема прогноза работает по запросам, генерируемым подсистемой принятия решений. Подсистема распознавания стратегий работает в своем собственном темпе.

Ниже кратко описывается назначение основных подсистем.

3.1. Подсистема анализа и распознавания входных данных

Подсистема анализа и распознавания входных данных выполняет следующие основные функции: фильтрацию, восстановление, кластеризацию, классификацию.

Фильтрация исходных данных производится с целью устранения шумов и помех. Поскольку применяемые в системе алгоритмы фильтрации не всегда дают для оставшихся данных верные значения параметров, в ряде случаев их можно восстановить по географическим данным. В этом заключается задача блока "Восстановление".

Задача кластеризации состоит в разбиении множества отмечок на группы (кластеры) таким образом, чтобы получившееся разбиение максимально соответствовало группированию ДЭ.

Задача классификации состоит в отнесении кластера к некоторому типу объектов и вычислении параметров формы и положения.

3.2. Подсистема мониторинга

Задачи, которые должна решать подсистема мониторинга, перечислены в п. 1.4.2. Их можно разбить на два основных класса.

Задачи первого класса заключаются в протоколировании и обобщении данных, поступающих от системы распознавания и анализа. При этом должны быть построены структуры данных, приспособленные к извлечению интегрально-ретроспективных сведений, а также информации о ситуации в конкретный момент времени. Например, подсистеме прогноза в одних случаях может потребоваться восстановить историю эволюции некоторого кластера, а в других – получить информацию о расположении всех кластеров на последнем кадре измерения.

Задачи второго класса заключаются в проверке входных данных на выполнение определенных критериев (предикатов) и построения на их основе новых сущностей. Например, после анализа ряда кадров могут быть обнаружены два кластера, один из которых следует за другим. Тем самым должен быть построен новый объект - "цепь связанных кластеров", который в дальнейшем сам является объектом мониторинга.

В определенных случаях возможна обратная связь подсистемы мониторинга к подсистеме распознавания и анализа с целью уточнения и/или корректировки результатов работы последней. Это может оказаться необходимым, если результаты распознавания и анализа k -го кадра значительно отличаются от аналогичных результатов $k-1$ -го кадра. Попытка исправить ситуацию может быть проведена или с помощью новых правил и алгоритмов, или/и с помощью оператора.

Подсистема мониторинга генерирует и обновляет структуры данных: протокол эволюции обстановки; перечень объектов мониторинга; карты географии эволюции кластеров и сущностей.

Протокол эволюции обстановки и перечень объектов мониторинга подсистема поддерживает с помощью собственных файловых средств. Географические карты поддерживаются с помощью географической информационной системы.

Протокол эволюции обстановки представляет собой индексно-последовательный файл с записями переменной длины. Каждая запись соответствует мониторинговым данным для соответствующего кадра измерений. Запись делится на подзаписи, каждая из которых описывает либо отдельный кластер, либо некоторую сущность (отношение). В описаниях кластеров используются координаты центров кластеров в данный момент времени, являющиеся, фактически, ссылками на карту эволюции кластеров. При необходимости при описании географических компонентов сущностей некоторых типов из соответствующей подзаписи могут быть сделаны ссылки на отвечающую этой сущности карту.

3.3. Географическая информационная система

На концептуальном уровне географическую информационную систему [6, 7] можно представить как средство поддержки "географической матрицы". Для стационарного случая эта матрица трехмерна. Два ее измерения отвечают плоской системе координат (топоснове), а третье – списку характеристик. На каждом "слое" матрицы, получаемом за счет фиксации значения на дискретной оси характеристик, располагается элементарная тематическая карта, которая содержит данные строго одного географического и геометрического типа. Переход к хранению нестационарных объектов возможен либо за счет добавления к атрибутам объектов, соответствующих разным моментам времени, еще одного атрибута – "момент создания" (например, так можно кодировать развивающиеся во времени траектории), либо за счет увеличения карт в базе.

Поддержка матрицы заключается во внесении на тот или иной ее слой необходимых данных, считывании информации и групповых операциях (теоретико-множественные операции, визуализация и т.п.).

Предполагается, что любые географические объекты, подлежащие хранению и обработке с помощью географической информационной системы, могут быть геометрически интерпретированы как объекты одного из следующих типов: точечных, линейных, площадных либо однозначных функций двух переменных. Это позволяет ограничиться достаточно небольшим набором универсальных операций хранения и обработки.

3.4. Подсистема распознавания стратегий

Как отмечалось в п. I.3, стратегии делятся на два класса – выбор маршрута и выбор ускорения.

Суть распознавания стратегий:

- а) вычисление и динамическое обновление коэффициентов доверия для каждой из стратегий;
- б) переупорядочивание продукции внутри системы и назначение каждой продукции своего матожидания.

В случае стратегии выбора ускорения процесс распознавания сводится к:

- 1) вычислению по законам кинематики реальных изменений ускорения (скорости);
- 2) определению продукции, с помощью которых подобное изменение могло бы быть произведено;
- 3) статистическому анализу полученных результатов.

Для распознавания стратегии выбора маршрутов выполняются операции, аналогичные шагам 2 и 3 описанной выше процедуры.

4. Средства машинной графики

в экспертной системе электронной навигации

В описываемой экспертной системе в качестве основного способа взаимодействия с операторами и инженерами знаний предполагается использовать графический диалог. Такое решение объясняется двумя основными причинами. Во-первых, это требование работы в реальном времени, а во-вторых, - возможность геометрически наглядно интерпретировать большинство возникающих в процессе работы ситуаций. Требование работы в реальном времени подразумевает не только использование аппаратно-программных средств и прикладных программ, обеспечивающих соответствующую производительность системы, но и достаточно быструю ответную реакцию оператора. Это, в свою очередь, означает, что информация должна представляться в виде, удобном для восприятия человеком, а управление системой производится в рамках объектно-ориентированного подхода. Для операторов, работающих с системой, должна быть создана среда понятий и образов, адекватно описывающих предметную область. При этом трудности возникают как в случае понятий, имеющих аналоги среди окружающих естественных или искусственных объектов, так и с чисто абстрактными понятиями. Например, при изображении рельефа и объектов на нем зачастую не требуется добиваться максимальной достоверности изображения. Однако нередко требуется акцентировать внимание оператора на каком-то участке или объекте. Подобный подход порождает весьма специфические требования к используемым в системе алгоритмам

машинной графики. Абстрактные объекты, такие, например, как тенденции, связи и т.д., требуют большой изобретательности при проектировании для них графических образов.

Одной из центральных в этой связи становится проблема построения генераторов новых сущностей и адекватных им графических образов.

Специфические для экспертных систем проблемы возникают из-за необходимости объяснять принимаемые системой решения с целью контроля их оператором либо поиска причин принятия неверных решений инженером знаний. В этом случае система должна показать, "как ее научили решать данную задачу", т.е. проиллюстрировать использование знаний из своей базы знаний.

Это можно сделать, если использовать архитектуру экспертных подсистем, представленную на рис. 2. С "внешним миром" такая система общается, получая от него запросы в виде директив и в таком же виде выдавая обратно ответы на них. При этом директивы, генерируемые какими-либо подсистемами по инициативе оператора, никак не отличаются от директив, генерируемых автоматически. Транслятор директив производит их синтаксический и семантический анализ и обращается к экспертной оболочке с целью порождения необходимой цепочки обращений к собственно прикладным модулям. Например, запрос на оценку проходимости некоторого участка местности для автосредств определенного типа может потребовать десятков обращений к модулям географической информационной системы (см. [7]), при этом последовательность вызова модулей и возникающие в процессе отработки запроса промежуточные сущности ("холмистость", "наличие дорожной сети" и т.п.) зависят от географической ситуации. При необходимости логическая цепочка вызова модулей и генерированные промежуточные сущности с помощью ретранслятора директив и "обратных" средств интерактивной графической программы (см. рис. 2) могут быть продемонстрированы оператору или инженеру знаний.

В настоящее время начата реализация ряда основных блоков описываемой подсистемы [7-9]. В целом разработку системы на уровне демонстрационного прототипа планируется завершить в 1992 г.

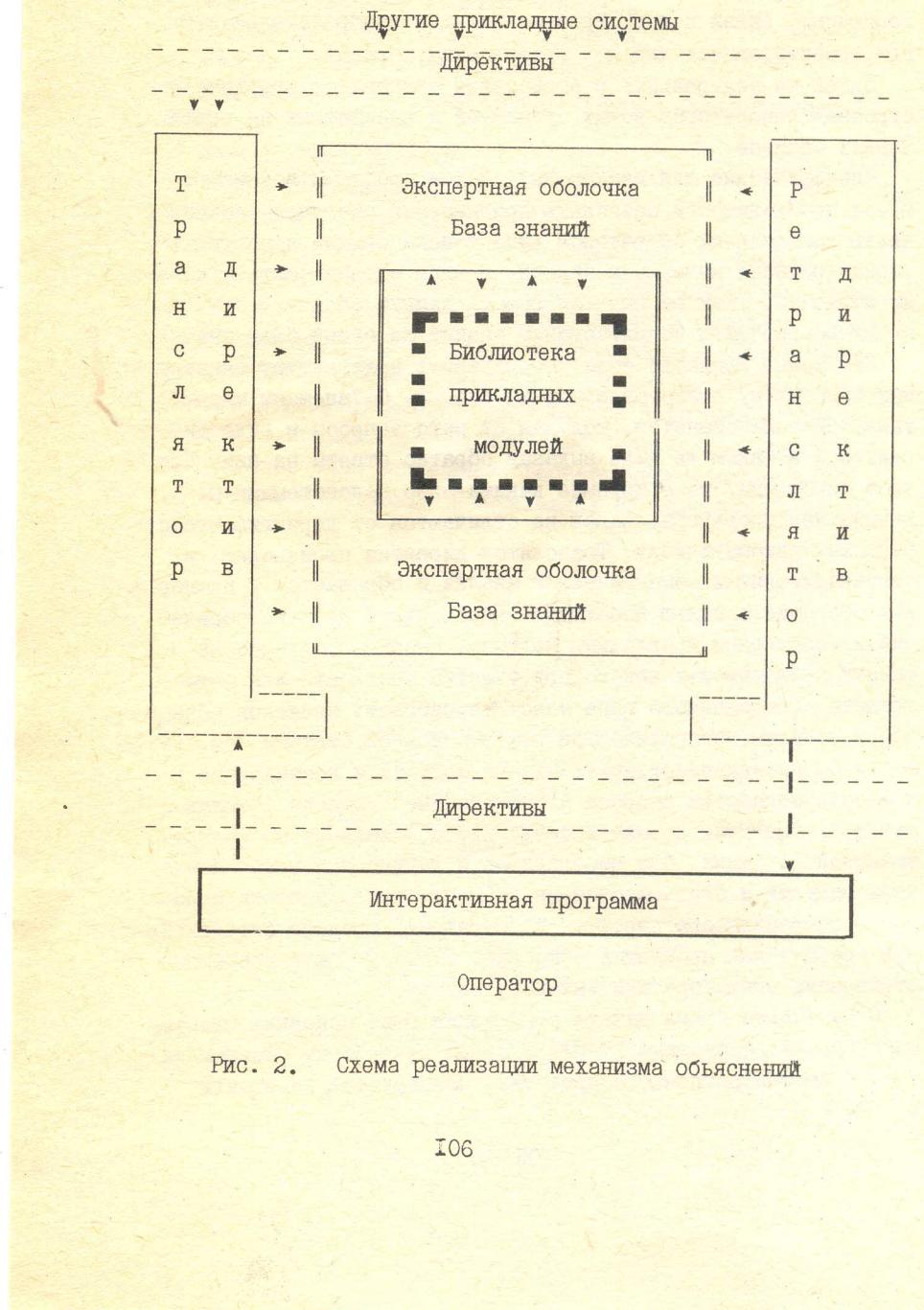


Рис. 2. Схема реализации механизма объяснений

Л и т е р а т у р а

1. Lachance S. Electronic ice navigation in the Laurentian region//Charting a course for electronic navigation display in Canada.: - Proc. Seminar. - Toronto, 1987. - Р 8-13.
2. Nichol D.G. a o. Some image processing aspects of a military geographic information system//Austral. Comput. J. - 1987. - Vol. 19, N 3. - P. 154-160.
3. Kobatake Hidefumi, Inone Yoshiaki, Namai Tatsuro, Hamba Nobuhito. Measurement of two dimensional movement of traffic by image processing//ICASSP '87: Proc. Int. Conf. Acoust. Speech and Signal Process., Dallas, Tex, Apr. 6-9, 1987. - N.Y., 1987. - Vol. 1. - P. 614-617.
4. Богемский В.А. Концепция построения базы данных, отображающей динамику предметной области, для систем управления железнодорожным транспортом// Упр. системы и машины. - 1988, № 2. - С. 77-82.
5. Уотерман Д. Руководство по экспертным системам. - М.: Мир, 1989.
6. Сиротин В.Г. Географическая информационная система АЛЬБА// Тез. докл. Междунар. симпозиума по разработке и использованию персональных ЭВМ "INFO-89".-Минск, 1989.- С.720-725.
7. Сиротин В.Г., Коновалов А.В. Библиотека программ для представления и обработки графических данных методом квадро-деревьев. - Наст. сб. - С. I08-I2I.
8. Торшин В.И. Интерактивная графическая система для подготовки транспортных обстановок. - Наст. сб. - С.
9. Упольников С.А. Визуализация рельефа местности. - Наст. сб. - С. I35-I53.