

天气图图形数据库系统的研究

俞善贤

王耀生

(浙江省气象科学研究所, 杭州 310021) (北京气象学院)

吴尧祥

潘小凡 毛燕军 卢家麟

(浙江省气候应用所)

(浙江省气象台)

提 要

提出了天气图图形数据库的设计思想和具体实现的技术路线,并介绍了已建成系统的主要功能。

关键词: 天气图图形数据库 数据结构 相似检索 表变换法

引 言

现有的资料库中存放的气象数据都是探测数据,它包括仪器直接探测与遥感数据。人们从已有的气象资料库中并不能直接查询、检索有关槽、锋、高低压中心等气象信息,尽管这些信息对气象工作者来说是十分重要的,它们与天气的变化密切相关。

由于计算机存贮的数据与知识之间存在明显的差异,这就产生两个突出的问题:(1)对这类信息的加工处理,始终只能采用手工方式,不能很好地利用计算机等现代化工具,不论在精度、效益以及客观化方面都受到很大限制,因而严重阻碍了气象预报业务的现代化;(2)目前每天的气象数据大致有几百兆字节,计算机的速度和存贮量与数据量之间的矛盾日益尖锐,在我国由于经济条件的限制,计算机资源和通讯条件改善均受到较大的限制。因而问题愈显突出。

1 设计思想

天气图图形数据库是一个面向省级气象台的、通用型实时预报应用的业务系统。从省

级台拥有的计算机资源出发,立足于实时预报的需要,实现计算机对天气图资料的分析处理。

数据库建设要达到三个目的:(1)实现图形、图像数据的微机存贮管理;(2)缩短计算机应用与气象知识的距离,提高业务预报客观化与智能化程度;(3)为气象智能预报系统奠定基础。

由于在微机上实现,因此设计中必须解决好信息与微机容量与速度的矛盾。

因为是通用型,所以系统设计中必须考虑各方面的共同要求,同时还要兼顾各地区的不同特点,系统必须具有相当的灵活性。

实时预报有紧迫的时间要求,因而高效率是系统设计的中心环节。

2 系统的总体结构与功能

目前建成的系统为达到设计中提出的要求,在软件设计上,系统以 MS-C7.0、SDK 作为开发语言,在 Windows 3.1 操作系统中运行,具有完全 Windows 应用程序风格,用户界面友好,操作方便、直观、灵活、安全,提

* 参加此项工作的还有,广西壮族自治区气象台、黑龙江省气象台及山东省气象台的杨望月、于振东、曾奋等。

供一种不同于 DOS 系统命令的操作手段,利用“对话”、“图标”、“菜单”等图形画面和符号来实现操作;充分利用了 Windows 与设备无关的图形图像操作,使开发的系统能适用各种计算机,便于移植,充分利用 Windows 的内存管理技术,为进一步开发打下了良好基础。

系统的总体结构如图 1。

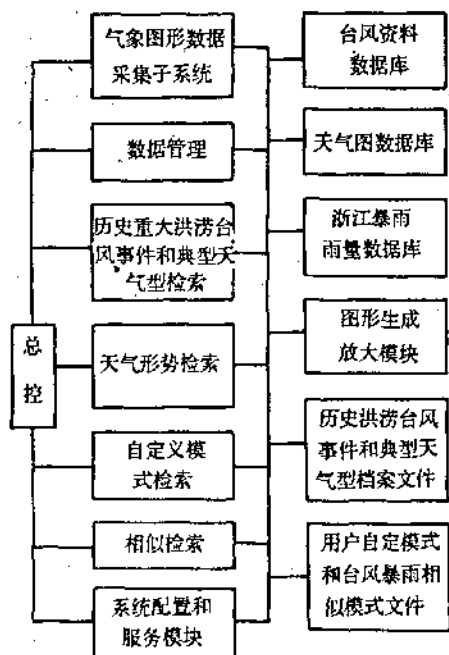


图 1 系统总体结构框图

系统框图中各框的功能如下:

(1)气象图形数据采集子系统:利用数字化仪采集图形数据。对数字化仪采集到直角坐标系变换成经纬度坐标,然后转换成相应图形数据,实现图形数据的半自动化,大大提高图形数据资料录入速度和精度,为输入大量的天气图、雨量图提供十分方便、快速的工具。

(2)数据管理模块:具有全屏幕编辑功能的数据查询、增删、修改及批数据文件输入及磁盘数据压缩功能。

(3)历史重大台风暴雨灾害事件和典型天气形势检索:对历史上产生重大灾害的台

风暴雨二个例和典型天气类型建立一个档案文件,利用 Windows 的列表框查阅该文件,随时均可显示天气形势图、台风路径图、雨量图和典型天气类型图。

(4)天气形势检索:用户可通过菜单和对话框,对历史上任何一天任何一层的天气图进行检索,如果有台风还将显示这一天以前的台风路径图,可以一屏幕显示一张大图,也可一屏显示四张小图。使用热键(或称功能键)可以方便地实现层次间的切换,大小图切换,或前一天或后一天的天气图检索。从打入检索命令到图形显示仅需 2—3 秒钟时间。

(5)用户自定义检索:用户可通过对话框来定义需要检索的内容,定义的图景(天气系统成员),可以前后几天,上下各层,并叠加到同一屏幕显示。例如可以一屏显示一天各层的槽线,也可以一屏显示连续几天的槽线。可利用该功能研究某一或多个天气系统成员的时间和空间变化。

(6)相似检索:首先确定相似的模式,用逻辑表达式来表达,然后确定天气模式的各个成员的相似条件。系统首先对逻辑表达式进行语法检查,然后把逻辑表达式展开成一张搜索表,根据前一成员的相似结果,按搜索表确定下一步搜索的内容,这样尽可能提高检索速度。实际试验表明:在 386 微机上,在 900 张天气图中搜出相似天气图约需 2—3 分钟。

(7)系统配置和服务模块:提供改变显示地图范围的功能,即改变兰勃脱投影的参数。另外显示系统使用的调色板。

2.1 气象图形数据采集子系统

系统结构如图 2。

(1)接口模块:实现数字化仪和计算机之间的通讯。分别编制了 Super, Kurta, Summa 3 种型号的数字仪接口模块。

(2)坐标转换:把直角坐标转换成经纬度坐标,用理论公式往往有一定的误差,还需采用统计订正方法加以修正。

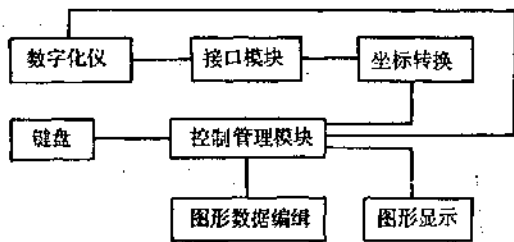


图2 气象图形数据采集子系统

(3)控制管理:对数字化仪,键盘输入数字,命令之间进行控制。

(4)图形数据编辑:具有全屏编辑功能,对点、线进行删除、修改、插入等编辑,系统随时提示需要键盘还是数字化仪操作,便于使用。

(5)图形显示:对输入数据随时可以生成图形,用来核对输入数据的正确性和合理性,直观明了。

(6)适用的地图及最大误差:

- a. 历史天气图(兰勃脱投影) $\leq 1^\circ$ (经、纬度);
- b. 浙江省区域图(雨量图) $\leq 0.03^\circ$ (经、纬度);
- c. 海浪图(莫卡托投影) $\leq 0.4^\circ$ (经、纬度)。

(7)适用的基本数据类型有:

- a. 等值线型
- b. 特征线型
- c. 中心点型
- d. 矢量型。

3 关键技术及解决途径

3.1 数据结构及文件组织

天气图数据是一类直接与气象知识相联系的数据,数据结构具有三个“不定”的特征,即“不定记录项、不定记录个数、不定记录长度”。对这样的数据结构要实现数据管理,随机存取等操作,一般的关系数据库(dBASE)就很难胜任。虽然PROLOG(人工智能)语言支持这类嵌套表结构数据的处理,但考虑到该语言其它方面的局限性,我们还是采用C

语言来实现这类数据的管理。

一张天气图数据的结构形式描述如下:

<图>::=日期<线>

<线>::=<中心类><槽><特征线><线>

<中心线>::=<中心名><中心字段>

<中心名>::=冷中心 暖中心 高中心 低中心

<中心字段>::=经度、纬度、强度 经度、纬度、强度<中心字段>

<槽>::=槽标识符<位置>

<特征线类>::=<特征线名>特征值<位置>

<特征线名>::=等高线 等温线

<位置>::=<点><点><位置>

<点>::=经度、纬度

有了上述定义后,计算机就能对每一个数据识别处理。多张天气图就构成一个无序的流水文件。为了实现随机检索,在数据库内模式(物理模式)中,为处理变长记录,把“记录位置表”和“记录长度指示器”这两个方法结合起来使用,充分发挥各自的优点,用索引文件记录每张天气图开始位置,用记录长度指示器记录每一个记录的长度。用上述方法管理数据的另一个优点是:这样设计的数据库,数据库的外模式(用户模式)和内模式独立,这样当数据物理组织发生变化时,对外模式不必修改,对内模式发生变化时,内模式也不必变化。

索引文件的格式是一个顺序文件,这样可按顺序文件方式进行搜索,我们采用对分法。

3.2 相似检索

3.2.1 识别天气系统成员相似方法

在计算机上进行天气系统相似检索时,首先要将预报员识别某一天气系统成员的相似知识进行描述,然后确定一个能够在计算机上处理的方法,同时要尽可能同预报员识

别系统成员的结果相吻合。

天气系统成员相似识别主要同层次、时间、位置、位置容许偏差、强度和强度偏差这些内容有关。其中位置和位置容许偏差可演化成各种不同形式的相似识别指标。

(1)中心类位置表示方法：用中心位置和东、西容许偏差经度和南、北容许偏差纬度这些条件来表示，图形上就是一个矩形区域。

例如，定义涡的相似识别条件是：中心位置，偏差范围为东西各3个经度，南边1个纬度，北边2个纬度(图3)。

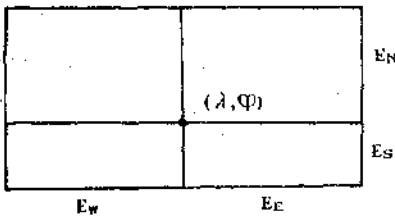


图3 定义涡的矩形区域示意图

只要使 $E_e = E_w = 3$ 、 $E_s = 1$ 、 $E_n = 2$ ，凡落入该区域内的涡则称为相似。

中心类天气系统成员有：热带气旋、低涡、北方高压、日本海低压、阻塞高压等。

(2)特征线类位置表示方法：

例如，定义切变线的相似识别条件为：用东端与 120°E 相交点的纬度及西端与 110°E 相交点的纬度，其偏差范围在西端为南北2个纬度，东端为1个纬度。即：西端点 110°E 上， $E_n = E_s = 2^\circ$ ；东端点 120°E 上， $E_n = E_s = 1^\circ$ 。

又如：槽北段(与 35°N 交点)经度与南段(与 30°N 交点)经度表示，偏差范围为北段东西各3个经度，南段东西各2个经度。即北段 35°N 上， $E_e = E_w = 3^\circ$ ；南段 30°N 上， $E_e = E_w = 2^\circ$ 。凡符合条件的则为相似。

这类主要天气系统成员有：切变线、槽、横槽、倒槽、东北冷槽等。

(3)等值线型位置表示法：在计算机处理

时只要使用多个矩形，类似于以上两类的表示方法。

用上述方法来模拟预报员识别天气系统成员的过程，能够在计算中统一处理，并能灵活转换成各种形式，满足各种需要，实际试验表明，这样的处理结果，同预报识别的结果相吻合。

3.2.2 系统成员相似识别的算法

相似识别最终需要判别某一图形元素(特征线、槽线)是否在所考虑的矩形区域内，这就是计算机图形学中的图形裁剪问题。裁剪处理常常涉及大量的图形元素，故裁剪原理虽并不复杂，但对处理的速度有十分高的要求。相似检索中除了读盘速度，大部分的计算时间是用在裁剪处理上，如果算法不佳，就会出现不能容忍的等待。比较几种算法后，我们采用梁友栋-Barky 方法。

设考虑的区域为正规矩形 $X_{\min} \leq X \leq X_{\max}$ ， $Y_{\min} \leq Y \leq Y_{\max}$ (见图4)。由于平行于 X 和 Y 的线段的线段裁剪十分简单，因此只考虑剪线段 AB 既不平行于 X 轴，也不平行于 Y 轴， A, B 的坐标分别为

$$A(x_A, y_A), B(x_B, y_B)$$

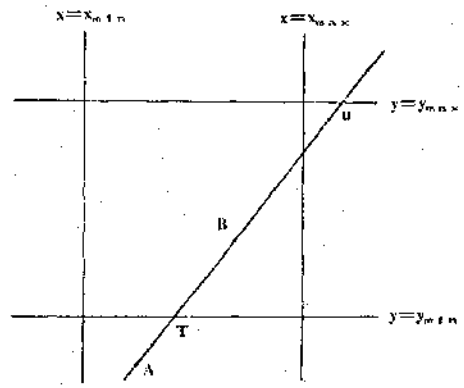


图4

由于 $x_A \neq x_B$ $y_A \neq y_B$ 得直线方程为

$$x = \frac{x_B - x_A}{y_B - y_A}(y - y_A) + x_A$$

求出:

$$x_T = \frac{x_B - x_A}{y_B - y_A}(y_{\min} - y_A) + x_A$$

$$x_u = \frac{x_B - x_A}{y_B - y_A}(y_{\max} - y_A) + x_A$$

可以证明线段 AB 有可见部分的充分必要条件为

$$\max\{\max[x_{\min}, \min(x_A, x_B)], \min(x_T, x_u)\} \\ \leq \min\{\min[x_{\max}, \max(x_A, x_B)], \\ \max(x_T, x_u)\}$$

令

$$L = \max[x_{\min}, \min(x_A, x_B)]$$

$$R = \min[x_{\max}, \max(x_A, x_B)]$$

则充分必要条件又可以表示为:

$$\begin{cases} L \leq R \\ L \leq \max(x_T, x_u) \\ \min(x_T, x_u) \leq R \end{cases}$$

若 $k = \frac{y_B - y_A}{x_B - x_A} > 0$ 时, 则 $\max(x_T, x_u) = x_u$, $\min(x_T, x_u) = x_T$ 因此 $k > 0$ 时, 不等式组又可写为:

$$\begin{cases} L \leq R \\ L \leq x_u \\ x_T \leq R \end{cases}$$

因此, 我们可以先测试 $L \leq R$ 是否成立, 如果不成立, 则根本不存在可见线段, 对 $k > 0$ 先计算 x_u , 再测试 $L \leq x_u$ 是否成立, 若成立再计算 x_T , 测试 $x_T \leq R$ 是否成立。总之, 只有在需要时才计算交点, 从而减少计算量而提高了裁剪处理的效率。

3.2.3 天气形势和天气过程的相似检索

根据天气系统成员的相似识别的结果, 可以用某一逻辑表达式组合成识别相似天气形势图和天气过程图。在具体实现时两个主要问题: (1) 如何解决任意逻辑表达形式的相似检索, 以满足各类天气形势相似检索的要求; (2) 尽可能提高检索的速度。上述的第一

问题也就是计算机如何来理解任一表达式的含义? 具体又怎么做? 对此我们采用了菊池敏典提出的“表变换法”。表变换法将一个逻辑式列成一个展开表。最基本的展开表形式见附表。表中“一致指向”, “不一致指向”规定检索的顺序, 即某一检索词符合条件时, 下一步做什么; 不符合条件时, 又做什么。一致指向, 不一致指向栏的内容只可能是三种取值: 命中、不命中、下一次检索词的地址。

表变换过程分为“前处理”、“后处理”两部分, 每一部分填好表的一半。

“前处理”主要工作如下: 顺序扫描逻辑表达式, ①遇检索词送入展开表; ②遇“(”, level(级位工作单元)=level+1; 遇“)”level=level-1。level 初值为 0。③遇“×”把下一检索词的地址置入当前检索词的“一致指向”栏; level 值送入当前检索词的“级位”栏。④遇“+”把下一检索词的地址置入当前检索词的“不一致指向”栏, level 值送入当前检索词的“级位”栏。⑤当逻辑表达式扫描结束后, 把“命中”置入最后一个检索词的“一致指向”栏, 把“不命中”置入最后一个检索词的“不一致指向”栏。

“后处理”主要工作如下:

后处理根据“级位”值从最后一个检索词开始, 反填表的空白处, 一直处理到第一个检索词为止。

在反填过程中, 如果遇到“不一致指向”栏为空, 则填入某级位小于或等于空白对应检索词级位且位于当前检索词之后的最邻近的检索地址, 不一致指向栏的内容。如果遇到“一致指向”栏为空, 则与当前检索词级位值大于该检索词之后的检索词级位值时, 把最邻近的检索词地址的一致指向栏的内容填到当前检索词的“一致指向”栏, 但当两者级位值相等时, 要继续往下扫描, 直到后继检索词级位值小于当前检索词的级位, 或级位相等。

但当后继检索词就是表达式的最后一个检索词,这时,把后继检索词的一致指向栏的内容填入当前检索词的一致指向栏。

例如:对于逻辑表达式

$$Q = A * (B + C) + D * ((E * F) + G)$$

表展开过程见附表

附表 表展开过程

检索词	序号	一致指向	不一致指向	级位
A	1	(1)2	(4)4	0
B	2	(13)命中	(2)3	1
C	3	(12)命中	(3)4	0
D	4	(4)5	(11)不命中	0
E	5	(5)6	(10)7	2
F	6	(9)命中	(6)7	1
G	7	(7)命中	(8)不命中	0

通过展开处理,用展开表代替原来的逻辑式,计算机可以顺着展开表进行检索,一旦得到了确定性的答案后,可立即停止,而不一定要求表达式中所有条件都一一处理。总之,最大限度地节约了计算机处理的时间。

具体检索过程中分三步进行:①对逻辑表达式进行语法检查;②表变换过程;③各天气系统相似识别过程。

4 结论

4.1 建立天气图图形数据库系统是气象部

门的一项首创性的工作,没有成功的经验成果可以借鉴,一切都必须自己探索。实际试验表明,对这一系统的总体设计思想是切实可行的,系统的结构是合理的,运行是高效率的。实现了大量的天气图、雨量图等图形资料在微机上的存贮分析处理,改变长期以来只能靠人工对这些资料进行检索、分析的状况,有利于提高气象业务的自动化和客观化。

4.2 气象图形数据采集系统的建立,不仅提高了数据采集的速度和精度,而且也为气象图形资料的传输提供了新的方法,具有信息量少传输速度快,出图及显示速度快的特点。

4.3 系统的总体设计思想和部分成果可直接适用于海洋、水文、地质、地理等领域的大量图形资料的计算机存贮、分析和处理。由于存贮的数据与专业知识相结合,大大压缩了图形资料的信息量,使微机中存贮处理大量的图形资料成为可能。

4.4 系统的应用研究方面还是处于初步阶段,系统还具有很大的应用开发潜力,随着应用的深入,还需有一个进一步完善提高的过程。

4.5 系统处理的只是静态的、二维图形数据,如何处理三维的动态图象、图形数据,则是今后气象关系数据库要开发解决的。

(参考文献略)

A study on the System of Synoptic Chart Database

Yu Shanxian

(Zhejiang Meteorological Bureau, Hangzhou 310021)

Wang Yaosheng

(Beijing Meteorological College)

Wu Yaoxiang Pan Xiaofan Mao Yanjun Lu Jialin

(Zhejiang Meteorological Bureau)

Abstract

The designs and some concrete practices of the synoptic chart database are introduced. The primary functions of the system constructed are given.

Key Words: synoptic chart database data structure analog-retrieval list transform