

# 卫星云图的计算机分析识别

王耀生

王燕<sup>1)</sup>

(北京气象学院 100081)

(北京大学)

冯晓娟<sup>2)</sup>

师春香

(卫星气象中心)

## 提 要

提出了采用数学形态学与句法模式识别结合,将云系结构与天气系统联系起来的卫星云图计算机分析识别方法。克服了传统的统计识别方法不能很好地反映云系整体性结构的缺陷。对一次台风过程不同阶段云系发展演变特征,实现了计算机分析识别试验。结果显示,在台风发展的不同阶段,云系的结构、形状和组成是不同的。

关键词: 卫星云图 台风云系 数学形态学 句法模式识别

## 引 言

卫星云图在气象中的作用,特别是在灾害性天气预报中的重要性,已成为气象工作者的共识。在我国早期的云图应用中,一直采用的是经验性的人工定性判断,近年来卫星云图的计算机加工处理获得比较广泛的应用,但就方法而言,大都还停留在一般的简单处理,如分屏、漫游、动画、增强和不同光谱通道的简易叠加合成等。如何使云图应用趋向客观定量化处理,是当前提高灾害性天气预报水平的一个重要环节。

卫星云图作为图象的一种,它既有一般图象的共性,又有某些特殊性。一般图象处理的目标物往往是动的不变、变的不动,而云却是既动又变,有的还变动很快。云图的另一个特征是云没有清晰的轮廓边界。这些都给进一步计算机加工处理增加了难度。

## 1 方法简述

对图象的分析识别主要采用两种方法:

统计法和结构法(也叫句法方法)。统计法的主要特点是对图象模式的统计特征提取,和对这些特征量值集合的识别,而结构法则是选取能反映图象结构特征的基元,并将这些基元按一定的规则来进行组织、排列,以期实现识别。如果在某些图象的分析识别时它的结构信息并不重要,只是简单的分类,那么统计法是比较合适的,反之则采用结构法更恰当一些。然而在实际问题中,往往是处于这两种极端之间,也就是说将这两种方法结合运用是适宜的。

自傅京孙<sup>[4]</sup>在70年代提出句法模式识别(syntactic pattern recognition)以来,尽管已建立了一套理论和严密的推理法则,但只在一些有限范围应用。究其原因,主要症结在于如何对图象进行分割,提取具有结构信息的基元。假如不能或没有提取出合理的基元,结构识别就失去了基础。然而,由于图象中的景物往往是很复杂的,很难得到符合要求的基元,这是结构法难于在实际中应用的根源所在。

1) 现在美国纽约州立大学。

2) 现在在英国。

云的计算机分析识别在气象上并不是一个新的命题,已有几十年的历史,但过去所用的方法几乎都采用统计法,即使在有了卫星云图的今天,情况并没明显的改变。以80年代后期L·Garand<sup>3)</sup>的工作为例,他将卫星云图分割成 $128 \times 128$ 的网格区域,称之为单元,经过统计,选取了用作云分类的13个特征,它们是总云量、低云量、中云量、高云量、云顶高度、单元有云的数、单元的云反照率、有云的数目、多层次指数、背景的连通性、云的连通性,以及用能谱表示的纹理因子和波长在20—40km的强频谱量等。采用两步分类识别,先用前11个特征进行分类识别,再用后两个特征进行第二步识别。据介绍对高、中、低云及对流云的识别准确率可达90%以上。李俊<sup>4)</sup>选用光谱辐射强度和纹理结构两类特征,采取树分类法,先用阈值进行预分,再用动态聚类,然后再用阈值或动态法进行再分,这样逐级细分,可以把低云、中云、卷云、积云等分开。这种以地面观测的云分类为基础建立的云识别方法,不能很好地反映卫星云图上云的整体性结构特点,存在着内在的局限性。(1)气象工作者从云图上识别云,首先考查它与天气系统的联系,如低槽云系、台风云系等,考察它们的形状、范围、结构特点等,如台风云系则看它的螺旋云带情况,密蔽云墙及是否形成眼区等。(2)云图上一个完整的云系或云团被零碎地分割到不同的单元中,这样即使统计识别完成了,人们也不能建立起对云的整体认识。正是由于上述原因,统计识别结果仅能用于气候分析,难于在实际预报业务中应用。

## 2 数学形态学(Mathematical Morphology)

### 原理简介

数学形态学是近十几年才发展起来的一门新兴学科,是一种用于数字图象分析和识别的新理论和新方法。又由于它的算法与计算机并行体系结构密切关联,因而有着广阔的前景。

一幅图象(譬如云图)包含着很多信息,然而由于人们的目的、要求和观点的不同,往往需要抽取不同的信息,因此在对图象进行信息提取时,必须寻找与所要解决的特定问题有关的特征。为此必须先确定图象的结构,逐一试探图象各部分之间的关系,并进行检验。为了达到这一目的,数学形态学设计了一种收集信息的“探针”,称为“结构元素”,在图象中不断移动结构元素,便可考察图象各部分之间关系。

数学形态学设计了一整套很独特的几何变换,它的最巧妙之处在于这套变换仅是由两种基本变换(腐蚀和膨胀)不同的组合方式所构成。而且将这种几何变换转变为对应的代数运算,这就为计算机应用奠定了基础。

综上所述,可以把数学形态学的基本原理和主要特点归结为:(1)提出了结构元素的新概念。(2)设计了两种基本变换(腐蚀、膨胀),在此基础上构筑了整套几何变换。(3)将几何变换转变成对应的代数运算。

### 2.1 二值数学形态学

所谓二值数学形态学,就是将数学形态学用于处理二值图象,而二值图象就是图象中只有0、1二值,不存在其他数值的图象。

#### 2.1.1 结构元素

所谓结构元素是人们为了了解一个复杂图象,收集相应信息而设计的一种结构单纯、形状规则的图象单元。由于结构元素作为图象几何形状的探针,在变换运算中处于主导地位,因此在抽取图象特征时,选取恰当的结构元素就显得十分重要。然而如何选取合适的结构元素又往往与一定的先验知识有关,因此也有人称这为先验知识法。

在数字图象情况下,虽然结构元素的选择受到一定的原则上的限制(如拓扑同构要求),还是有相当大的灵活性,目前一般在计

3) L·G的博士论文。

4) 李俊的硕士论文。

算机中已预先设定一些结构元素作为选取的模板。

### 2.1.2 基本运算——腐蚀变换和膨胀变换

腐蚀变换 (Erosion) 和膨胀变换 (Dilation) 是两种最基本的运算, 它们有着很直观的几何形象。如图 1 所示,  $X$  是被研究的原始图象, 小圆  $B$  就是结构元素。令小圆  $B$  沿着  $X$  的内边移动一周, 圆心轨迹所围的阴影部分就叫做  $X$  被  $B$  腐蚀的结果 (图 1a)。由图 1a 可知,  $X$  被  $B$  腐蚀的结果保留了原图象  $X$  的基本形状, 却去除了两端凸出不平滑的一部分。如果小圆  $B$  沿图 1b 中  $X$  的外边移动一周, 圆心轨迹所围的阴影区就是  $X$  被  $B$  膨胀的结果。同样地膨胀也保留了原图的基本形状, 只是填平了一些原图两端凹陷不平的部分。

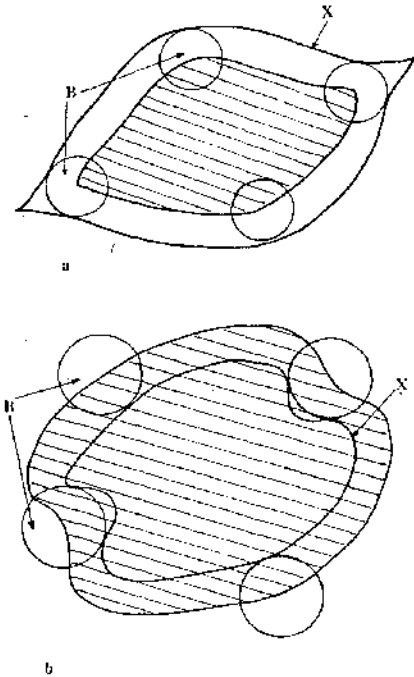


图 1 数学形态学基本运算示意

a. 腐蚀变换; b. 膨胀变换

图中阴影部分表示原图  $X$  被结构元素  $B$  腐蚀(膨胀)的结果

由图 1 可以形象地看到, 所谓腐蚀变换就是将原图按小圆半径缩小了一圈, 而膨胀变换则是扩大了一圈, 它们好象是被小圆“腐蚀”、“膨胀”了似的。由图还可以知道, 膨胀与腐蚀之间并不是一种逆变换的关系。

### 2.1.3 常用运算

由腐蚀与膨胀这两种基本运算可以组成 5 种最常用的运算。

2.1.3.1 开、闭运算。所谓开运算就是对图象先腐蚀, 后膨胀。而闭运算则是对图象先膨胀, 后腐蚀。它们的基本作用是对图象进行平滑, 开运算可以去掉图象中的一些孤立子域和毛刺。闭运算则可以填满一些小洞和将两个邻近的目标连接起来。

2.1.3.2 击中(中的)运算。为了研究图象  $X$  的结构, 可以利用结构元素  $B$ , 用它来判定哪些成份应属于图象, 哪些不属于“的”, 这就是中的运算名称的由来。

设  $X$  是被研究图象, 结构元素  $B$  是由两个不相交的子结构元素  $B_1, B_2$  之和组成。所谓  $X$  被  $B$  击中的结果就是  $X$  被  $B_1$  腐蚀的结果与  $X$  被  $B_2$  的反射集  $\tilde{B}_2$  膨胀结果之差。

2.1.3.3 细化、加厚运算。所谓细化就是将图象  $X$  减去被结构元素  $B$  击中的结果。而加厚则是  $X$  加上被  $B$  击中的结果。

人们在图象分析时, 经常希望得到某一图象只有一个象素宽的中轴(或叫骨架), 而且在抽取中轴时又要保持图象的拓扑等价性, 细化算法就是为了实现这一目的而设计的。加厚则是细化的对偶算法。

## 2.2 讨论

通过上面的简介可以看到, 尽管数学形态学具有严谨的数学背景, 然而它的基本观念却是如此简单直观, 特别巧妙的是它把极其复杂多变的图象处理问题归结为非常简单的两种变换及其组合。

数学形态学是一种非线性变换, 它与传

统的 FFT 变换、统计处理及卷积等相比,由于它是根据图象景物的几何特点,通过选取结构元素作变换的,因此它能较好地提取反映图象结构方面的一些特征。

在数学界对数学形态学也还存在争议,其焦点在于数学形态学是建立在连续空间上的,在它转向离散空间应用时,存在不严格之处。但从实用角度考察,它的巨大应用价值则是无可否认的。

本文没有涉及灰值与连续空间问题,因为这超出本文应用讨论的范畴,有兴趣者可查阅文献。从实用考虑,灰值问题也可转化为多二值来处理。

### 3 句法(结构)模式识别概要<sup>5)</sup>

#### 3.1 句法模式识别系统

图 2 给出了句法模式识别系统的框图。它包括四个基本部分:即预处理、分割、基元(以及关系)识别和句法分析。预处理的目的是为了得到一份可用的图象模式数据。分割则是要将图象中的目标物与背景分开,并进一步将目标物分割成基元。基元识别就是以一组样本基元(或子模式)为标准来辨认(或识别)基元和关系,并按指定的句法运算用一组基元来表述每个模式。这组基元所组成的模式是否属于给定的句法描述的模式类,则将由“句法分析程序”或“剖析程序”来完成。在进行句法分析或剖析时,通常采用“树”的形式,产生出模式的一个完整的句法描述,这表明模式在句法上是正确的,否则为不正确,进而确认模式的是否被拒识。

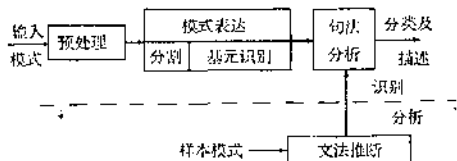


图 2 句法模式识别系统框图(引自[4])

#### 3.2 预处理

预处理包括数据压缩、过滤和增强等内容。数据压缩可采用傅氏变换(FFT)、整量化处理及编码压缩等方法。这里要指出的是数据压缩可区分为无损压缩与有损压缩两类。所谓无损压缩是数据可完全复原,无信息丢失的压缩,一般编码压缩属于这类;有损压缩是信息有所丢失,数据不能完全复原的,FFT 在离散情况下的应用一般为有损压缩。整量化处理则视所用的方法而定。

过滤和增强的目的虽都是为了改善图象的品质,然而这两者却是相反相成的。过滤就是要过滤(至少抑制)噪声,为此一般采用平滑处理或带通滤波等方法来实现。显然,采取这种方式处理的图象必然变得模糊或有畸变。增强是为了使图象的某些部分和特征变得更为清晰,一般采用微分算子来进行,最常用的是梯度算子、拉普拉斯算子和卷积运算及其组合。

#### 3.3 分割

图象分割是图象处理中很重要但又是容易处理好的一个问题,不是把背景中的某些景物也当作目标物给分割出来,就是在分割时,目标物产生了某种程度的失真。由于图象情况的复杂性,目前尚没有一个一般的分割方法,比较常用的是阈值分割和边缘跟踪分割。

阈值法的核心是阈值选取,这可以采用统计(频率法、众数法)选取,也可以通过人机交互、经验确定。

对于一些边缘特征已知或边缘分明的图象采用边缘跟踪是比较好的,如果边缘不很清楚或在边缘特征信息不多的情况,直接用边缘跟踪就可能遇到困难,在这时将阈值分割与边缘跟踪结合处理可能效果会好一些。

5) 本节主要根据文献[4,5]改写。

### 3.4 基元识别

基元作为句法模式识别中描述和构筑模式结构的基本单元,必须具有结构简单、含义明确两个特点,正如构成句子基元的单词,它一定有着确定的词类与词义。

正由于基元的结构简单,包含的结构信息相应地比较少,因此一般情况,采用模板匹配或对基元特征的多重阈值运算,就可进行基元识别,而且由于这种识别只是为了实现简单的分类,因而传统的统计方法也往往可以应用,如非参数决策分类(线性判别、最小距离分类、最近邻域分类及多项式判别等)和贝叶斯分类等。

### 3.5 句法分析

句法分析是句法模式识别中的核心和关键,因为即使基元识别全部实现了,如果不能完成句法分析,证明模式的句法描述是正确的,最终还是不能达到模式识别的目的。

由于形式语言中上下文无关文法的相对成熟性和理论上的严密性,以及计算机实现的语言条件限制,目前真正实用的是上下文无关语言的句法分析。

分析程序是一个自上而下的剖析算法,由一个产生算法和一个回溯算法组成,它能系统地描绘出与给定的上下文无关文法语言相关联的产生树或导出树。这种算法可对链形模式进行分析,并对被分析的链的每一种有效的产生式都将被找出并记录下来。一个可用的剖析程序流程图示于图3。需要指出的是,即使在一次剖析成功完成之后,剖析程序还必须继续工作,因为一般说来,事先并不知道文法是否含混,所以需要产生出所有剖析式。

### 3.6 讨论

句法识别作为模式识别的两个分支之一,对推进模式识别工作起重要作用。因为在实际的图象中,结构特征往往具有重要意义,

不识别和描述图象结构上的差异性,往往就缺乏实用的价值。而且在对结构的处理中人的智慧和技巧也更有用武之地,特别在今天,计算机的人机交互功能日益加强,这一特色将更显突出。

基元是句法模式识别的前提,如果不能从图象中提取出能反映模式结构的基元,那句法识别就无法进行,然而恰恰在这一基本上,过去一直没有很好地解决。本文提出采用数学形态学与句法模式识别结合的做法,正是为解决这一“瓶颈”所作的尝试。

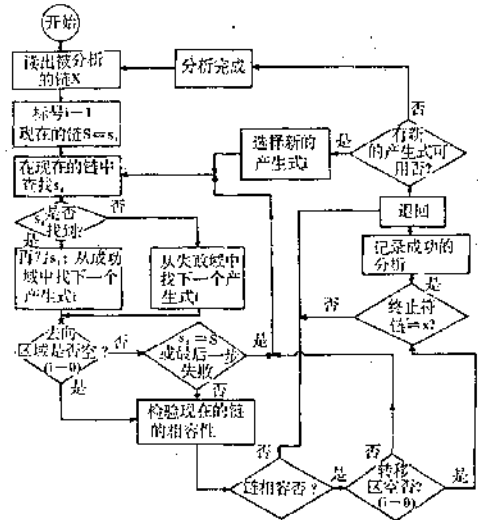


图3 上下文无关程序化文法的剖析程序(引自[4])

由于条件的限制,目前的句法识别局限于上下文无关情况,这对于静态识别问题还不明显,如果要进行复杂的动态识别(如云的演变识别),那么上下文有关的处理可能就是必要的了。

## 4 云图计算机分析识别

### 4.1 早期工作回顾

从1987年起开始了卫星云图计算机分析识别试验。当时的想法是希望改变在云图

分析使用时的主观定性判断的做法,使之逐步趋于客观定量,同时也为气象智能数据库建设及智能预报系统的实施创造条件。

国内外多年计算机云图识别的实践表明,采用单纯的统计识别很难达到上述目的。因此决定以结构识别为主,辅以统计处理的工作方案,几年来的实际试验结果表明,这一总的设计思路是正确可行的,已取得良好的进展,达到了预期的目标。

然而在实施过程中也遇到困难与矛盾,这主要表现在以下方面:

(1)云的状况是非常复杂的,很难找到一个恰当的统计方案能适用于各种云的情况,无论是分割还是平滑,都会发生类似的问题。

(2)在云的分析识别中,它的结构,特别是它的形态结构是很重要的特征条件,而在形态结构中骨架(也叫中轴)又是最突出的特征之一,正是在这样一些关键性特征抽取时,传统的方法不能保证拓扑同构这样一个重要条件的满足。

(3)缺乏一套规范化的操作,对每张图,每个系统都要经多次试验,多种探索,这就难于在实际业务中应用。

## 4.2 卫星云图分析识别框图

图4给出了卫星云图分析识别的基本框架。整个过程可区分为分析(主要由数学形态学完成)和识别(由结构法实现)两部分。分析由滤波、分割、光滑及基元提取四步组成。识别包括组句和句子分析两项内容。

### 4.2.1 云图分析

4.2.1.1 M 滤波 由数学形态学的开、闭运算组成的 M 滤波是一种非线性滤波算法,它与传统的线性滤波不同,在滤掉脉冲噪声时,可以避免同时滤去图象中有用的高频成分。这是因为 M 滤波是根据物景与噪音之间形态或拓扑结构上的差别来去伪存真的,它可以既去掉孤立的高频噪声,而又不影响

物景边界的清晰度。这是线性滤波器所不具备的特点。

4.2.1.2 分割 把要分析的云系从云图中抽取出来实现分割,这是云图分析中关键的一步。由于一个云系中往往包含有多种形态的云,而云系在发展的不同阶段,它所含云的情况又很不相同,这些表现为云图上云系有很广的灰度值变化范围,因此采用统计阈值是很难做好分割工作(这就是前述第一个难点)。

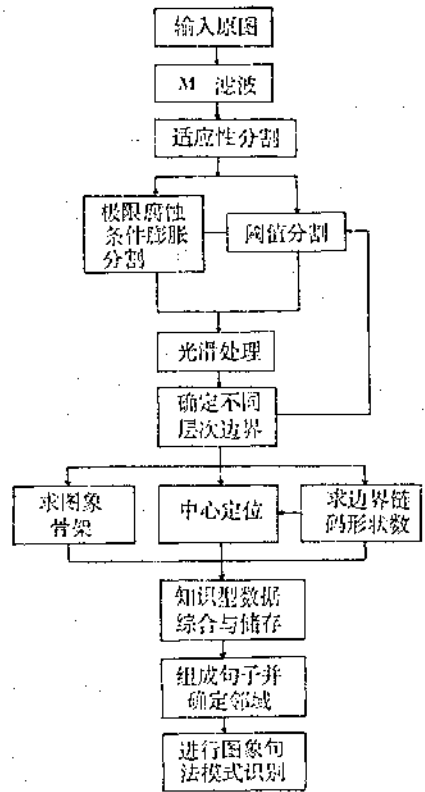


图4 卫星云图分析、识别流程

本文提出用数学形态学的适应性分割、阈值补充分割及极限腐蚀和条件膨胀分割的三步分割算法,可以较好地实现云系的分割。

4.2.1.3 光滑 通过上述分割得到的云系图象还可能边缘参差不齐,内部有噪声点等问题,需要进一步的处理。在光滑处理

中所用的方法就是腐蚀、膨胀、开、闭 4 种算法,根据分割云系出现的实际情况,灵活运用上述算法,达到光滑的目的。

这里要指出的是在光滑处理时,结构元素(一般采用正多边形)的半径不要大,以不改变云系的拓扑结构为宜。

**4.2.1.4 边缘抽取** 经过以上一系列处理后就得到一幅已经分割且边缘比较光滑的云系的二值图象(改变阈值对原图再重复进行上述运算,就可以将原来的灰色图象转变为多二值图象)。由于是边缘比较光滑的二值云系图象,因此采用一般的边缘跟踪就可实现边缘抽取。

**4.2.1.5 求骨架(中轴线)** 图象的骨架是描述图象的几何形状和拓扑性质的主要特征之一。对图象求骨架的过程通常叫做“细化”过程,细化的实施过程也就是对前述云系图象边缘进行层层剥落的过程,直至只剩下一个像素宽的一条线为止,这条线就是所求取的骨架线。在结构元素选取时,要注意细化过程中保持连通性质不变。

如果由于在光滑处理时不够完善,那么在细化求得的骨架中有可能出现残留的“毛刺”点,还需要作去“毛刺”的工作。

## 4.2.2 云图的识别

采用数学形态学的云图分析,完成了句法识别中的基元抽取工作,这里本文选取的图象基元是(1)云系的多二值边缘,(2)骨架及(3)闭合中心点。

在获取了上述三部分图象基元后,求它们的八方位链码和形状数。由于链码与所选的端点有关,因此采用形状数作为基元识别与句法分析的基础。

将边缘、骨架及中心点的形状数按组排列组成不同长度的数字链的句子,然后用前述的句子分析方法来进行判识。这里要强调的是,这种识别方法本质上是一种模式匹配,

由于种种原因可能导致图象基元抽取时出现差异,这种差异有可能使链码和形状数有一些不同。如果采取严格的数对数的匹配,就会有误判或不应有的匹配失败。为了避免或减少这种情况的发生,就要设定一定的偏差范围,或确定不同基元的识别重要性,改匹配为树状搜索。

## 5 一次台风过程云系的分析定位试验

选取台风云系作为试验的目标物,是由于台风是一种重大灾害天气,云图在台风的监视和预报中发挥了重要的作用,在海上时这种作用更为突出。也由于台风云系的特征比较明显,人们对它进行过不少研究,积累了不少经验,这一点对我们初次试验,缺乏实践认识者来说是尤其需要的。

螺旋云带、密蔽云墙和台风眼是构成台风云系的 3 个基本特征。实现对这 3 个特征的提取与描述是试验的基本要求。我们选取 1989.10.16—20 日的一次台风过程,这是一次比较典型的过程,它具有台风发展的 4 个阶段,每个阶段分别选取了两张云图,它们是:

发生期:1989.10.16 05:31—06:32;  
发展期:1989.10.16 11:31—17:31;  
成熟期:1989.10.18 20:32—23:31;  
衰亡期:1989.10.20 05:31—11:31。

从图 5a—d 给出了发展期(1989.10.16.11:31)从原图经 M 滤波、分割、光滑后的边缘提取及在边缘提取基础上的求骨架的全部处理过程。

将图 5b 与 5a 比较可以看到,在台风云系中已出现了一块空洞区,从进一步的处理可以看到这空洞区就是一个闭合涡旋的形成,这在经过 M 滤波处理的 5b 图上看到比原图 5a 要清晰得多了。

对照图 5c、5a 可见经分割处理,只保留了台风云系的主体部分,几乎完全去掉了与



图5 云图台风云系分析定位

a. 原图; b. M 滤波; c. 分割; d. 特征提取

分析关系不大的各种云,这一点在采用其他分割算法时是很难做到的。

由图 5d 的结果显示,运用数学形态学处理后求得的边缘与骨架的确能较好地勾画出台风云系的形状和拓扑结构。

图 6 给出台风发展不同阶段云系骨架特

征。它清楚地显示,这次台风过程在发生阶段的特点是由多条螺旋云带向闭合涡旋形成演变;而发展阶段的基本特点是闭合涡旋形成并发展加强的时期;成熟阶段则是涡旋相对扩大,内部形成眼区(图中没有表述)的阶段;到了衰亡阶段则眼区和闭合涡旋均已消失。这表明在台风发展的不同阶段,构成台风云系三个基本特征的结构、形状和组成是不同的。

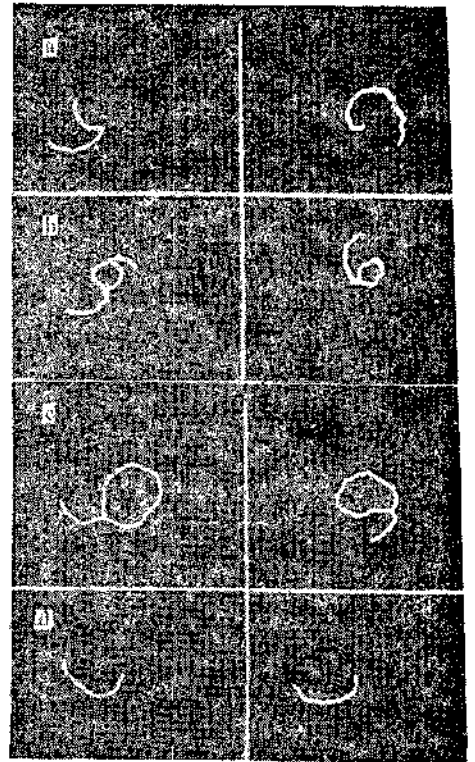


图6 台风发展不同阶段的云系骨架特征

a. 发生阶段; b. 发展阶段; c. 成熟阶段; d. 衰退阶段

台风定位是目前气象业务中容易引起争议的一个问题,这在飞机定位终止之后更为突出。本文在上述工作的基础上,进行了台风的计算机定位试验。

利用云图进行台风定位是基于以下 3 种不同的处理原则。

(1) 在台风的的发生与衰亡阶段,由于没有



闭合涡旋,根据螺旋云带的汇(辐)合点来定位。

(2)在台风的发展阶段,由闭合涡旋的几何中心定位。

(3)在台风的成熟阶段,由眼区内的最高温度点定位,在最高温度出现多个点时,取平均位置定位。

这次试验为客观、准确地台风定位作了一次有意义的尝试。

**致谢** 在本工作进行过程中,江吉喜同志帮助选取台风个例,董克勤同志提供了有关的知识和文献,中科院合肥智能所提供了良好的设备环境,韦穗同志提供了数学形态学的大量资料和他们的研究报告,并帮助指导了部分研究生论文工作,孔兵参与具体算法的讨论,黄李本同志、张进同志在硬件保证、图片照相等方面给予许多帮助。在此表示衷心的感谢。

### 参考文献

1 王耀生.智能预测系统设计及应用.北京:气象出版社.

1990年2月.

2 王耀生.利用卫星多通道红外遥感资料进行云的识别试验.北京气象.1985年第4期.

3 Wang Yaosheng. Intelligent Prediction System Proceedings of the 2nd International Conference on East Asia and Western Pacific Meteorology and Climate P. 561-574. 1993年.

4 傅京孙.模式识别及其应用.北京:科学出版社,1990年10月.

5 R. C. Gonzalez 和 M. G. Thomason. 句法模式识别.北京:清华大学出版社,1984年. 洪群等译.

6 唐常青等编著.数学形态学方法及其应用.北京:科学出版社,1990年8月.

7 J. Seera. From Mathematical Morphology to AIICPR 1986.

8 徐建华编著.图象处理与分析.北京:科学出版社,1992年.

## Analysis and Recognition of Satellite Image with Computer

Wang Yaosheng

Wang Yan

(Beijing Meteorological College, 100081)

(Peking University)

Feng Xiaojuan

Shi Chunxiang

(Satellite Meteorology Centre)

### Abstract

A technique of satellite image recognition with computer, mathematical morphology combined with syntax pattern recognition, and cloud system structure connected with synoptic system, is introduced. With the technique, a test of recognition of cloud system of a typhoon process is made. The result shows that the structures, shapes and composition are various at different evolution stages of a typhoon. It is also shown that the new method has an advantage to reveal the overall structure features of a cloud system that is superior to the traditional statistical recognition method.

**Key Words:** satellite image typhoon cloud system mathematical morphology syntax pattern