

用 NOAA/AVHRR 数据估算上海地区水稻种植面积*

吴健平

杨星卫

(华东师范大学地理系, 上海 200062) (上海市气象科学研究所, 上海 200030)

提 要

应用模糊监督分类方法对 NOAA/AVHRR 数据进行象元分解, 并估算了上海市 10 个郊县 1992 年和 1993 年两年的水稻种植面积. 估算的结果与统计数据比较后表明, 全市总的估算精度(采用面积比精度)分别为 92.0% 和 95.4%.

关键词: NOAA/AVHRR 数据; 混合象元; 模糊监督分类; 面积比精度.

前 言

作物面积的估算是进行作物估产的一项不可缺少的工作, 它的精度直接影响到对一个地区作物总产量的估算精度. 目前, 利用高的空间分辨率的遥感图象(如 SPOT、TM 或 MSS)进行作物面积调查能达到比较高的精度, 这在方法上是可行的. 但高空间分辨率的遥感图象成像周期较长, 在作物的生长期中, 往往得不到完整的无云覆盖图象, 加上成本较高, 在实际工作中受到许多限制, 特别对大范围区域来说更是如此. 与之相比较, NOAA/AVHRR 数据尽管其空间分辨率较低(星下点为 $1.1 \times 1.1 \text{ km}^2$), 但其具有覆盖周期短、价格低、数据易获取等优点, 研究结果表明^[1,2], 混合象元分解的方法, 能克服 AVHRR 数据空间分辨率低的缺陷, 并能提取一定精度的作物面积信息.

本文利用 AVHRR 数据对上海地区水稻种植面积进行估算. 考虑到水稻生长期上海地区土地覆盖状况比较复杂, 为此, 本研究采用模糊监督分类方法对单时相的 AVHRR 数据进行分类, 提取水稻种植面积.

1 模糊监督分类

模糊监督分类方法在一些文章中已有介绍^{[3][4]}. 它首先选择样本象元以确定分类参数. 和传统监督分类不同, 选择的样本象元并不严格要求是均质的, 但要根据样本象元的地面实际情况, 确定每一样本象元对各种地类的隶属度. 根据样本象元计算各类型的模糊均值矢量和模糊协方差矩阵. 每一类的模糊均值矢量 V_i^* 为:

* 国家“八五”攻关课题资助项目. 1995-09-12 收到, 1995-11-15 收到再改稿.

$$V_j^* = \sum_{i=1}^n \mu_{ij} X_i / \sum_{i=1}^n \mu_{ij}$$

式中, n 为样本象元数; μ_{ij} 是第 i 个象元属于第 j 类的隶属度; X_i 为第 i 个样本象元的象元值矢量, $X_i = (X_{i1}, X_{i2}, \dots, X_{im})$, m 是象元光谱值矢量的维数. 模糊协方差矩阵 $\sum_j^*$ 为:

$$\sum_j^* = \sum_{i=1}^n \mu_{ij} (X_i - V_j^*) (X_i - V_j^*)^T / \sum_{i=1}^n \mu_{ij}$$

当 $\mu_{ij}=0$ 或 1 时, 上面两式就成为一般的均值矢量和协方差矩阵, T 为转置矩阵.

确定分类参数 V_j^* 和 $\sum_j^*$ 后, 接着对每一象元根据它们光谱值进行模糊监督分类, 即确定象元对各类地物的隶属度. 它可用最大似然分类算法确定, 这时, V_j^* 和 $\sum_j^*$ 代替通常的均值矢量和协方差矩阵, 象元属于某一类 j 的隶属度 μ_{ij} 定义如下:

$$\mu_{ij} = P_j^*(X_i) / \sum_{k=1}^c P_k^*(X_i)$$

这里 $\sum_{k=1}^c P_k^*(X_i)$ 为归一化系数, 其中

$$P_k^*(X_i) = \frac{1}{(2\pi)^{\frac{m}{2}} |\sum_k^*|^{\frac{1}{2}}} \cdot \exp\left[-\frac{1}{2} (X_i - V_k^*)^T \sum_k^{*-1} (X_i - V_k^*)\right]$$

c 是预先定义的分类数, $1 \leq j \leq c$, k 表示第 k 类, $1 \leq k \leq c$.

2 研究区域和 AVHRR 数据

我们选择了上海市所属的 10 个郊县(区)作为研究区域, 它们是宝山、上海、嘉定、川沙、南汇、奉贤、松江、金山、青浦和崇明. 10 个县(区)的总面积为 6129.1 km².

研究所用的 AVHRR 数据的成象日期分别为 1992 年 6 月 19 日和 1993 年 9 月 9 日. AVHRR 数据共有 5 个通道, 其中通道 3、4、5 为热红外波段. 由于热红外信号是反映物体的热辐射特性, 信号的大小除了和物体的类型有关外, 与影响物体热变化特性的其它因素如含水量、表面粗糙度等也有很大关系, 同一类型地物的热红外通道数据往往会有较大差异. 因此, 在分类中只选择通道 1 和通道 2 的数据. 但热红外图象由于地物的亮度差异较大, 图象较清晰, 在研究中用通道 3 的图象与通道 1 及通道 2 的图象进行彩色合成, 以便选取样本象元.

3 实验过程

本研究的实验过程如图 1 所示.

3.1 图象预处理

原始的 AVHRR 图象存在着几何误差和辐射误差. 在使用前, 通过气象卫星图象处理软件, 对图象进行了几何和辐射纠正, 为了使图象容易和地图资料配准以及便于与其它遥感图象之间的对照, 利用双线性内插法对象元进行内插, 使内插后的图象象元大小对应

于实地面积值(约 $250 \times 250 \text{m}^2$)。

在地理信息系统支持下,把一张上海市的行政界线图输到计算机中(包括行政界线和水系),在显示屏幕上与 AVHRR 图象套合。从显示的 AVHRR 图象上可以看到上海市上空局部地区有云,云的存在将影响分类精度,需要对云的影像进行处理。方法是在 AVHRR 通道 1 图象上,选取某一象元值作为阈值。由于在通道 1 上,云为亮色调,象元值大于其它地物的象元值,所以取最暗的云的象元值为阈值。打印出每个县的象元值直方图,判断是否有大于阈值的象元值存在,如有的话,则以相邻的非云象元值进行内插。

3.2 土地利用类型分类

尽管本研究只是估算水稻种植面积,但模糊监督分类需要知道研究区域内各类地物的模糊均值矢量和模糊协方差矩阵。只有这样,才能计算出每一象元中水稻的比例。

地物的类型一般按土地利用分类,一级土地利用分类共有 8 个类型,它们是农田、园地、有林地、牧草地、居民地、交通过地、水体和未利用土地。由于模糊监督分类是从地物光谱角度来进行分类的,一方面,有些类型的光谱特征相近,在分类中不易区分;另一方面,同一类型的地物内部也可能存在着较大的光谱差异。基于地物的光谱特征,实验的土地利用分类作了如下调整:农田中的旱地、菜地同园地、有林地具有相近的植被光谱特征,合并为一类,其中主要是菜地,因此,类型定为菜地;水体的光谱特征明显,单独作为一类;农田中的水田,由于成象时间上海地区的水稻栽插不久,既有植被的光谱信息,也有水体的光谱信息,从农田中分出来;居民地的光谱特征较复杂。大的居民地,如城镇,由于内部建筑物密集,水蒸汽不易散发,形成城镇的湿岛效应,具有类似水体的光谱特征;小的居民地,如农村中的居民住宅,光谱特征主要是这些住宅屋顶颜色的反映,和城镇居民地的光谱特征有较大的差异,因此,把这两种居民地区分开;交通过地的分布面积不大,大多数交通过地的光谱特征与农村居民地的光谱特征类似,把它们合并到农村居民地中;牧草地和未利用土地在研究区内分布很少,在分类中不予考虑。因此,实验的土地利用分类共有 5 个类型:菜地、水田、城镇居民地、农村居民地和水体。

3.3 选择样本象元和确定隶属度

选择的样本象元是否有代表性以及确定的样本象元隶属度是否符合实际情况对最后

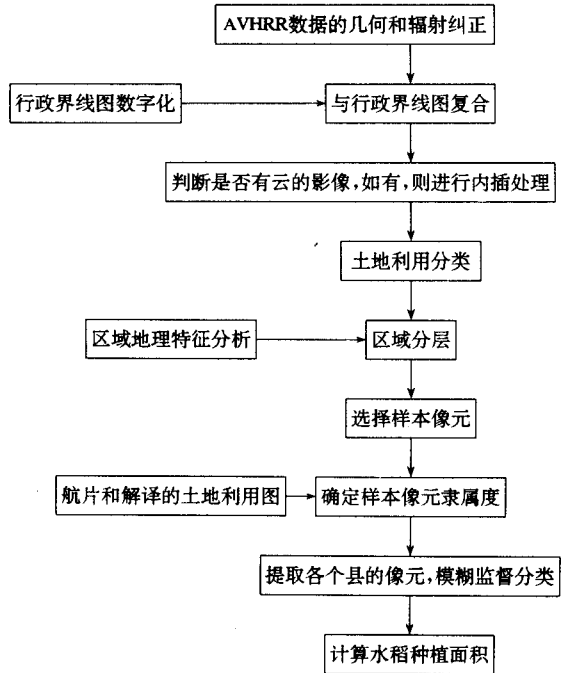


图 1 实验过程框图

Fig. 1 Flow chart of estimating paddy area from NOAA/AVHRR data

的分类精度有很大影响。由于各县的区域特征并不完全一致,区域特征的差异会造成土地利用类型的同物异谱、同谱异物的现象,因此,为了减少这种影响,在选择样本象元前,首先对区域进行分层。通过区域地理特征及图象分析,把 10 个县分成 3 个部分,即西部的松江、金山和青浦,东部的宝山、上海、嘉定、川沙、南汇和奉贤以及长江上的崇明。

区域进行分层后,每一部分分别选取样本象元,并通过对照航空像片(1989 年 11 月下旬成象,比例尺 $1:1 \times 10^4$)和根据航空像片解译的分县土地利用图来确定样本象元的隶属度。样本象元的选择考虑了下面几个问题:①样本象元应位于航片和气象卫星图象成象时间间隔中,土地利用类型没有或很少变化的地方。②样本象元的选择要反映出各个土地利用类型的光谱特征。③样本象元在航片上的位置比较容易确定,如城镇、城镇附近的菜地、河流湖泊等。④样本象元不包括内插出来的象元。

选择出的样本象元与航片或地图进行对照来确定隶属度,对某些样本象元,需要在航片或地图的相应位置上,确定一个一定大小的区域(约 $10 \times 10 \text{cm}^2$),计算各个土地利用类型图斑面积,得出它们的比例。

3.4 分类及面积计算

由于水稻种植面积估算是以县为单位,因此,在分类前首先要把各个县的象元提取出来。通过行政界线图与 AVHRR 图象的叠置,分别提取 10 个县的象元(不包括双线性内插出的象元)。

利用选择的样本象元及 AVHRR 通道 1 和通道 2 数据,对每个县内各个象元进行模糊监督分类。分类后分别得出 10 个县 5 类土地利用类型的象元数及百分比。以每县的总面积作为控制,算出每县的水稻种植面积。

4 精度分析

水稻种植面积的精度分析采用面积比的方法,计算公式为:

$$P = 1 - \frac{|S - \bar{S}|}{\bar{S}}$$

式中 P 为估算精度; S 为 AVHRR 数据计算面积; $\bar{S}$ 为实际的面积。

在本文中,我们选用了各个县当年上报的水稻播种面积作为实际的面积。通过分析,得出 1992 和 1993 年分县及整个地区水稻种植面积的估算精度(表 1)。

表 1 上海 10 个县(区)的水稻面积估算精度(面积单位: km^2)
Table 1 Classification accuracy of paddy area (area unit: km^2)

县(区)	宝山	上海	嘉定	川沙	南汇	奉贤	松江	金山	青浦	崇明	合计
1992 年											
计算面积	149.8	143.4	168.3	175.1	219.9	195.5	290.7	287.2	283.7	322.0	2235.6
上报面积	129.9	140.7	158.8	148.8	206.7	204.9	260.8	260.1	251.5	281.1	2070.2
精度(%)	84.7	98.1	90.6	82.3	93.6	95.4	88.5	89.6	87.2	85.5	92.0
1993 年											
计算面积	103.4	101.4	185.8	138.1	200.3	180.5	281.2	225.3	185.5	225.7	1827.2
上报面积	110.9	110.2	164.6	129.6	198.2	197.0	249.9	251.7	229.6	273.5	1915.2
精度(%)	93.2	92.0	87.1	93.4	98.9	89.1	87.5	89.5	80.8	82.5	95.4

5 小 结

通过研究,我们认为利用 NOAA/AVHRR 数据估算上海地区水稻种植面积在方法上是可行的,特别是对整个地区的水稻种植面积进行估算是能够达到比较高的精度的.在面积估算过程中,图象的质量和时相,以及样本隶属度的确定是决定估算精度的关键.因此,应尽量选择无云覆盖,成象时间为水稻种植前期的图象,同时,研究人员要对研究区域深入了解,这样才能保证估算结果有较高的精度.

参 考 文 献

- 1 徐希骝等. 混合象元的因子分析方法及其在大面积冬小麦种植面积估算中的应用. 科学通报, 1990, **35**(4): 317~320.
- 2 Quarmby N A, Townshend J R G, Settle J J and White K H. Linear mixture modelling applied to AVHRR data for crop area estimation. *International Journal of Remote Sensing*, **13**: 415~425.
- 3 Wang F. Fuzzy supervised classification of remote sensing images. *IEEE Trans. Geosci. Remote Sensing*, **28**: 194~201.
- 4 Wang F. Improving remote sensing image analysis through fuzzy information representation. *Photogrammetric Engineering and Remote Sensing*, **56**: 1163~1169.

ESTIMATING PADDY AREA IN SHANGHAI REGION USING NOAA/AVHRR DATA

Wu Jianping

(Geography Department, East China Normal University, Shanghai 200062)

Yang Xingwei

(Shanghai Meteorological Institute, Shanghai 200030)

Abstract

The data of NOAA/AVHRR, because of their some advantages, have been applied in many fields including estimation of crop area. This research applies fuzzy supervised classification to decomposing the mixed pixels of NOAA/AVHRR data and estimates the paddy area of 10 counties in Shanghai during 1992 and 1993. As compared with the statistic data in the same period the results show that the overall accuracies (ratios of estimated to statistic area) are 92.0% and 95.4%, respectively.

Key words: NOAA/AVHRR data; Mixed pixel; Fuzzy supervised classification; Area ratio accuracy.