

TM 图像的光谱信息特征与最佳波段组合

戴昌达 雷莉萍

(中国科学院遥感卫星地面站)

1989年6月17日收稿

摘 要

本文分析了北自黑龙江省寒温带缓岗平原、南至广东省亚热带丘陵等9个不同景观类型样区的 TM 图像数据,查明 TM 图像的光谱信息具 3—4 维结构,其物理含义相当于“亮度”、“绿度”和“热度”、“湿度”。在 TM7 个光谱图像中,一般以第 5 波段包含的地物信息最丰富。3 个可见光波段(即第 1,2,3 波段)之间,两个中红外波段(即第 5,7 波段)之间相关性很高,表明这些波段的信息中有相当大的“重复性”或“冗余性”。第 4,6 波段则颇特殊,尤其是第 4 波段与各波段的相关性都很低,表明这个波段的信息有很大的独立性。计算 20 种组合的熵值的结果表明,由一个可见光波段、一个中红外波段及第 4 波段组合而成的彩色合成图像,一般具有最丰富的地物信息,其中又常以 4,5,3 或 4,5,1 波段的组合为最佳。

关键词 信息结构 信息量 最佳波段组合

一、引 言

美国于 1982 和 1984 年相继发射了陆地卫星 4 号和 5 号。星上安置的新一代传感器 TM (专题成像仪)与第一代传感器 MSS 相比,空间分辨率由 79 米提高到了 30 米;探测波谱通道由 4 个增加为 7 个,并调整了可见光与红外通道的光谱带范围;卫星姿控、定位及扫描成像方式等也相应做了改进。这些调整与改进对图像的信息特征、处理方法、应用效果等会产生什么重大影响?美国航宇局于陆地卫星 4 号发射前后,主持开展了 TM 图像数据质量分析计划(LIDQA),有关专家的分析研究结果表明:TM 图像的辐射测量和几何保真度与 MSS 相比都有很大提高。空间分辨率也与设计指标一致^[1]。其后,包括笔者在内的一些遥感应用研究者,在不同地区,结合不同应用目的,陆续开展了有关 TM 图像信息特征及最佳波段组合的研究^[2,3],这些研究者根据地物波谱测试数据,TM 图像的可判性分析,以及应用处理的实践,得出一些相似的看法,认为 TM 新增加的两个中红外波段(即 TM5,7),包含着极丰富的地物信息,尤其是 5 波段,信息量之丰富,往往是各可见光波段所远不及的;由中红外、近红外、可见光各 1 个波段(例如 TM4,5,3)组合而成的彩色图像,对于许多应用目的来说,往往比其它组合的合成片具有更佳的判读性。但由于以上研究结果一般都是作者根据一二个试区资料得出的,缺乏较全面、广泛和系统的分析论证,故并未引起有关方面的足够重视。至今世界上能接收处理 TM 图像的几个地面站,销售数量较大的合成图像产品仍沿用 TM4,3,2 赋红、绿、蓝色的组成方案。这种组合的光谱响应与原来 MSS7,5,4 赋红、绿、蓝的常规合成片一致。符合许多判读者的习惯,且由于 TM 的空间分辨率比 MSS 提高了一倍多,这三个波段的光谱带范围也做了适当调整,故这种图像包含的信息量与可判性明显优于 MSS。从分析植被长势、量算

大宗作物面积、探测水深、水质变化、区分水陆界线等应用目的来说,也不失为最佳组合之一。但是这种组合排斥了信息量很丰富的中红外波段,毕竟是个损失。

我国遥感卫星地面站建成投产以来,已接收处理出数千景 TM 图像,覆盖了中国一半多面积,包括各种景观类型地区。订购 TM 图像的用户越来越多。为了充分开发 TM 图像所包含的丰富信息,进一步提高 TM 图像的处理、分析水平与实际应用效果,笔者感到很有必要,也具备了条件对 TM 图像的信息特征及最佳波段组合等问题进行更广泛深入的分析研究,以便获得更精确、更具普遍意义的认识。本文是对我国北起黑龙江寒温带缓岗平原,南至广东省亚热带丘陵山地,包括9个不同景观类型样区的 TM 图像,进行分析研究的主要结果。

二、样区概述

样区的编号、位置及有关基本情况如下。

样区 1 黑龙江省大兴安岭北坡,漠河县,图号 121—23。样区范围 1024×1024 像元。1987年5月23日成像,属寒温带、缓岗地貌。在黑龙江沿岸冲积平原有些农地尚未播种。大面积覆盖以兴安落叶松为优势种的原始森林。1987年5月6日发生的特大森林火灾,使部分林地与居民点夷为焦土,成像时局部地方尚在燃烧。

样区 2 黑龙江省穆稜县,图号 115—29,样区范围 512×512 像元,1986年6月11日成像。属温带河谷低丘。河谷中多稻田,刚进入播种插秧期,小部分坡地森林保存尚好,大部分辟为农地,春播作物处于苗期。

样区 3 黑龙江省东宁县,图号 115—29,样区范围 256×256 像元,1986年6月11日成像。属温带山地,林木葱郁。局部地段开矿挖石,岩、土出露。

样区 4 北京市平谷县,图号 123—32,样区范围 512×512 像元,1987年9月26日成像,属暖温带山前平原。主要作物——玉米已基本收割完,正陆续播种冬麦。近年发展的小面积旱稻及县城村镇附近的菜地、果园仍是一片葱绿,但面积比重不大。

样区 5 河北省兴隆县,图号 123—32,样区范围 512×512 像元,1987年9月26日成像,属暖温带山地,林木、草、灌的覆盖度不很高,长势开始衰败。试区内有座山间水库。

样区 6 河北省南皮县。图号 122—34,样区范围 512×512 像元,1987年5月14日成像,属暖温带冲积大平原。约 $1/3$ 面积种植冬麦,正处齐穗期,果园、人工草地郁郁葱葱,春播作物或尚未出土,或处苗期。连年干旱,河流渠道大多干涸。

样区 7 山东省利津县,图号 121—34,样区范围 512×512 像元,1987年4月19日成像,属暖温带滨海平原,包括海域、滩涂及部分黄河老三角洲耕地,正处春播期,小面积冬麦已拔节,生机盎然。

样区 8 杭州市,图号 119—39,样区范围 512×512 像元,1985年2月12日成像,属北亚热带平原低丘。样区包括山水相映、景色如画的西湖风景区及濒湖伸展的杭州市与郊区¹⁾。

1) 样区 8 资料摘自浙江农业大学赵元洪、陈岚等及浙江农业大学王人潮、王深法等人报告。

样区 9 广东省高要县,图号 123—44,样区范围 512 × 512 像元,1987 年 9 月 10 日成像。属南亚热带丘陵山地。丘间谷地均辟为水田,双季晚稻已经封行。山坡上林、灌、草长势不一,局部地段采矿挖石,岩、土裸露。池塘、水库、村镇似天上繁星散布在农田与林草植被中。

三、TM 图像的光谱信息结构

TM 图像共有 7 个光谱响应通道,构成 7 维光谱空间,但正如 MSS 图像具有 4 维光谱空间却仅为二维光谱信息结构一样, TM 图像的光谱信息结构也不是 7 维。从前述样区的主成分分析结果(表 1)看,它基本上属于 3—4 维结构。有些样区,如南皮等,前面 3 个主分量的贡献累计值已达 98%,以后几个主分量的贡献都不足 1%,这类地区 TM 图

表 1 TM 光谱值主成分分析结果

Table 1 Principal Component Analysis of TM Multispectral Values

	主分量 序号	特征值	贡献累计 %	特 征 向 量						
				1	2	3	4	5	6	7
样 区 3	1	5.50	79.5	0.41	0.39	0.42	-0.26	0.36	0.37	0.41
	2	0.88	92.0	-0.05	0.27	-0.01	0.83	0.46	-0.13	0.04
	3	0.24	95.4	-0.14	-0.07	-0.22	0.22	-0.10	0.91	-0.16
	4	0.21	98.5	-0.33	-0.58	-0.18	-0.15	-0.54	0.04	0.46
	5	0.06	99.3	0.52	-0.48	0.12	0.39	-0.40	0.01	0.42
	6	0.03	99.7	-0.62	0.02	0.66	0.12	-0.29	0.05	0.21
	7	0.02	100.0	-0.20	0.46	-0.51	0.02	-0.33	-0.04	0.61
样 区 5	1	5.07	72.4	0.42	0.43	0.42	0.29	0.42	0.14	0.43
	2	0.93	85.8	-0.30	-0.09	-0.31	-0.31	-0.17	0.93	-0.11
	3	0.78	97.0	0.05	0.11	0.03	0.76	0.28	0.33	-0.01
	4	0.14	99.1	-0.51	-0.30	0.02	-0.30	-0.36	0.01	0.65
	5	0.03	99.6	0.27	-0.14	-0.21	-0.35	0.75	0.01	-0.42
	6	0.02	99.9	-0.58	0.29	0.59	-0.05	0.16	0.02	-0.04
	7	0.01	100.0	0.25	-0.76	0.57	0.14	0.05	0.01	-0.08
样 区 6	1	5.317	76.0	0.97	0.98	0.99	-0.38	0.98	0.60	0.98
	2	0.980	90.0	0.12	0.13	0.11	0.84	0.15	-0.44	0.09
	3	0.602	98.5	-0.06	-0.06	-0.07	0.38	-0.02	0.66	-0.05
	4	0.055	99.3	-0.17	-0.03	-0.02	-0.01	0.11	-0.01	0.11
	5	0.027	99.7	0.08	-0.09	-0.08	-0.00	0.06	-0.00	0.04
	6	0.013	99.9	0.01	-0.04	0.04	0.01	-0.01	0.00	0.07
	7	0.07	100	-0.00	-0.05	0.06	-0.00	0.03	0	-0.33
样 区 9	1	4.94	70.5	0.43	0.43	0.423	0.43	0.42	0.19	0.48
	2	0.94	83.9	-0.24	-0.09	-0.28	0.86	0.26	-0.24	-0.02
	3	0.83	96.0	-0.15	-0.17	-0.18	-0.13	0.04	0.94	-0.04
	4	0.23	99.2	-0.35	-0.45	-0.02	-0.31	0.48	-0.10	0.58
	5	0.04	99.7	-0.36	-0.26	0.84	0.16	-0.04	0.03	-0.27
	6	0.03	99.9	0.15	-0.28	0.08	0.26	-0.69	0.00	0.59
	7	-0.02	900.0	-0.68	0.65	-0.02	-0.04	-0.22	0.03	0.26

像的光谱信息,可认为是3维结构,有些样区,如高要、穆梭等,需前面4个主分量的贡献累计才达98%以上;这类地区TM图像的光谱信息可认为是4维或3维半结构。总之, TM 图像的光谱信息结构,尽管随地区不同、季相不同,会有些差异变化,但总体看均可认为系3—4维结构。

进一步分析表1的数据可以看到,各样区第1主分量7个波段的特征向量值,一般均为正值或均为负值,仅有个别样区个别波段的符号不一。这意味着该分量系各波段亮度值之和的函数,它反映了该样区各像元总辐射水准,即总亮度的高低,据此可简称其为“亮度”分量;第2个主分量7个波段特征向量值中,无例外地都以第4波段(植物叶绿素的陡坡效应波段)的最高,而第3,1波段(植物叶绿素的强吸收带)则很低,甚至为负值,表明这个分量在很大程度上取决于第4波段植被反射强度与1,3波段植被吸收强度之比,它反映了该样区各像元叶绿素含量水平,据此可简称其为“绿度”分量;第3主分量7个特征向量值中多以第6波段的最高,而3个可见光波段的特征向量值则不高,或多为负值,表明这个分量基本上反映了红外波段的热辐射水准,而红外辐射强度在很大程度上取决于温度高低,故此可简称其为“热度”分量;第4个主分量的特征向量值中,通常两个中红外波段为正高值,其它波段常为负值或小值,表明这个分量在颇大程度上取决于5,7波段亮度值的高低。而5,7波段对地物含水量,包括OH⁻多寡比较敏感,据此可简称其为“湿度”分量;其它3个分量的贡献都很小,而且其特征向量值,随地区不同,时相不同而摆动颇大,很难一般地分析其物理意义。因此,可以认为这些分量受噪声影响很大。

综上所述, TM 图像的光谱信息结构可以概括为“亮度”、“绿度”、“热度”和“湿度”四要素。其中“热度”、“湿度”两个要素密切相关;有时可以互相表征,合二为一,有时则各具某些特色,不能完全互相取代。看来这就是 TM 图像信息具3维或4维结构的实质所在。

四、TM 图像各波段信息量的比较

TM 图像各波段所包含的地物信息量,一般采用该波段图像覆盖的辐射量化级,即亮度值范围或称亮度差(=最大亮度值-最小亮度值)来衡量。表2列举了上述9个样区的 TM 光谱统计资料。从中可以看出,不同景观类型或不同时期的 TM 图像,其各波段的最大与最小亮度值,就绝对值及分布来说,可能波动颇大,但就其差值即覆盖范围来说,几乎无例外地都以第5波段最高,仅样区1由于成像时1987年5月6日的特大林火尚未彻底扑灭,有几处大火还在熊熊燃烧,火头温度高,其辐射峰值进入第7波段的光谱响应范围,从而达到最高辐射水准,获得饱和亮度值(255),而样区内的水体对第7波段全吸收,呈“0”辐射水准,亮度值为“0”,这样就使7波段的亮度值覆盖范围达256级极限,从而超过5波段跃居第一。在其它各样区,第7波段的亮度差都小于第5波段,甚至还可能小于第4波段(如样区3)或第1波段(如样区9)。第4波段的亮度差无例外地都大于3个可见光波段及热红外波段。却都小于第5波段及第7波段的大部分。由此可以有把握地说,在绝大多数情况下,第5波段图像包含的信息量最丰富,不同地物的影像反差最大,这样能识别区分的种类就可能最多。第7,4波段包含的信息量依次有所减少,识别区

表 2 TM 光谱值的统计结果
Table 2 Statistical Result of TM Multispectral Values

样区	波段号	最小值	最大值	均值	标准差	亮度值范围
1	1	75	189	104.7	15.78	114
	2	25	96	42.2	6.43	71
	3	22	109	39.8	7.12	87
	4	13	82	43.9	10.48	79
	5	2	186	63.9	23.20	184
	6	82	219	146.5	14.15	137
	7	0	255	31.7	13.93	255
2	1	63	141	83.9	9.34	78
	2	23	71	38	7.20	48
	3	18	93	40.5	12.6	75
	4	10	121	67.1	18.8	111
	5	0	194	91.7	37.0	194
	6	125	201	161.4	8.5	76
	7	1	113	44	23.0	112
3	1	57	115	66.6	5.02	58
	2	22	55	29.8	2.86	33
	3	15	64	23.73	5.61	49
	4	20	136	89.1	14.28	116
	5	12	150	78.8	13.58	138
	6	117	199	139.7	15.7	82
	7	6	88	25.2	10.62	82
4	1	46	132	66.7		86
	2	14	69	30.3		55
	3	10	84	35.0		74
	4	6	120	49.7		114
	5	3	147	71.3		134
	6	114	141	126.8		27
	7	0	98	36.2		98
5	1	41	119	57.4	7.57	78
	2	11	64	24.3	5.91	53
	3	8	81	25.1	10.25	73
	4	2	114	51.2	15.65	112
	5	0	133	61.7	23.54	133
	6	102	141	122.3	6.21	39
	7	0	91	25.6	13.75	91
6	1	85	174	110.9	10.92	89
	2	33	97	53.6	8.38	64
	3	30	119	63.6	14.47	89
	4	31	133	83.8	16.03	102
	5	25	188	114.9	23.8	163
	6	126	181	158.7	9.74	55
	7	8	115	66.3	20.05	107

续表 2

样区	波段号	最小值	最大值	均值	标准差	亮度值范围
7	1	83	180	108.8		97
	2	32	92	49.8		60
	3	27	109	51.6		82
	4	11	92	36.1		81
	5	3	173	49.0		170
	6					
	7	1	120	27.3		119
8	1	35	228	89.5	13.31	193
	2	17	125	39.5	8.08	108
	3	10	173	39.5	13.30	163
	4	1	183	67.1	28.35	182
	5	0	202	64.2	25.59	202
	6	150	192	164.7	6.94	42
	7	0	125	25.6	12.57	125
9	1	53	163	68.1	5.38	110
	2	17	79	29.5	4.68	62
	3	11	99	24.5	6.95	88
	4	8	144	29.3	14.70	136
	5	1	199	68.2	19.00	198
	6	124	147	131.6	3.53	23
	7	1	95	19.71	8.18	94

以第 2 及第 6 两个波段更甚。因此这两个波段的信息量相对最少。第 1,3 两个波段则介于中间,低于 5,7,4 波段,而高于 2,6 波段。

进一步分析表 2 数据还可看出,所有样区第 1 波段的亮度值都高出第 2,3 波段很多,甚至高出第 4 波段。但根据野外或样品室内测定的地物波谱特性资料,除水体以外,其它各类地物在第 1 波段光谱区间的反射率都普遍低于第 4 波段,大部分或一部分也低于第 3、第 2 波段,与卫星探测器获得的辐射数据很不一致,显然,这是由于大气对短波光的强烈散射,增加了投入卫星第 1 通道探测器上的辐射能量,从而大幅度提高了第 1 波段图像的亮度值。

表 2 数据还说明,第 6 波段也受到大气的强烈影响。这个波段的亮度均值,在寒温带至南亚热带的所有样区都很高。与其它各波段相比,无例外地均居首位。而且不同样区之间,相差不大(在样区 1 成像时,黑龙江尚未开冻,而局部林地熊熊烈火尚在燃烧,故 6 波段的亮度值变幅明显增大)。如果据此认为,在如此辽阔的地域,各类地物表面温度和热辐射系数差异都不大,显然是不可思议的,究其原因,应归咎于大气热辐射的干扰,增加了第 6 通道探测器接收到的辐射能量,从而提高了这个波段的亮度值,并在相当程度上压缩了地物热特性的差异变化。当然, TM 成像时间均在当地时间 9 点半至 10 点,这时地物温差不处在最大时刻,肯定也是使第 6 波段亮度值覆盖范围不宽的影响因素之一。

因此,第 2 与第 6 波段信息的充分开发利用,应更认真考虑大气辐射的干扰问题。

五、TM 图像各波段信息的相关性

TM 图像 7 个波段包含的地物信息量固然多寡不一,但基本上有一定的顺序。各波段之间信息的重叠与分异程度,也表现出明显的规律性,表 3 给出了上述试区 7 个波段亮度值的相关分析结果。可以看出,3 个可见光(即第 1,2,3)波段之间的相关性都很高,尤其是 1,2 波段及 2,3 波段之间,相关系数几乎都在 0.9 以上,仅样区 1 和 7 因植被信息比重太低,相关系数未达 0.9。但也明显高出与其它 4 个红外波段之间的相关系数,表明 3 个可见光通道取得的信息彼此重叠很多,有相当大的“一致性”,或称“冗余性”。

两个中红外(即第 5,7)波段之间的相关性也很高。样区 6,7 的相关系数高达 0.99;样区 1 最低,为 0.78,一般都在 0.9 以上。可见 5,7 波段的信息也有极大相似性。第 4,6 波段相当特殊,与其它各波段,包括 4,6 两波段之间的相关系数都不高,有的甚至为负相关,表明这两个波段的光谱信息有较大的独立性。

综合 TM 图像的单波段光谱信息量和各波段间光谱信息相关性分析结果,可以认为第 5 波段提供的光谱信息最丰富当无疑问;7 波段的信息量就单波段来说,常居第二,可是与第 5 波段的信息重叠过多,在许多情况下,几乎可被 5 波段取代,因此,总的来说其重要性肯定不及第 4 波段,甚至不及第 3 或第 1 波段,只是在探测森林火灾、岩矿蚀变带及

表 3 TM 光谱值相关性分析结果

Table 3 Correlation Analysis Result of TM Multispectral Values

样区	波段号	1	2	3	4	5	6	7
1	1	1.00						
	2	0.95	1.00					
	3	0.84	0.89	1.00				
	4	-0.32	0.01	0.23	1.00			
	5	-0.29	0.01	0.37	0.77	1.00		
	6	0.01	0.04	0.03	0.41	0.05	1.00	
	7	0.08	0.22	0.43	0.27	0.78	0.48	1.00
2	1	1.00						
	2	0.98	1.00					
	3	0.96	0.97	1.00				
	4	0.26	0.46	0.33	1.00			
	5	0.69	0.77	0.73	0.74	1.00		
	6	0.19	0.60	0.66	0.56	0.77	1.00	
	7	0.88	0.87	0.87	0.55	0.95	0.79	1.00
3	1	1.00						
	2	0.91	1.00					
	3	0.97	0.92	1.00				
	4	-0.61	-0.36	-0.60	1.00			
	5	0.77	0.84	0.82	-0.22	1.00		
	6	0.83	0.76	0.82	-0.58	0.68	1.00	
	7	0.91	0.84	0.94	-0.57	0.88	0.81	1.00

续表 3

样区	波段号	1	2	3	4	5	6	7
4	1	1.00						
	2	0.96	1.00					
	3	0.91	0.96	1.00				
	4	0.09	0.18	0.04	1.00			
	5	0.69	0.78	0.82	0.35	1.00		
	6	0.40	0.46	0.53	0.01	0.57	1.00	
	7	0.78	0.86	0.93	0.07	0.92	0.64	1.00
5	1	1.00						
	2	0.97	1.00					
	3	0.96	0.98	1.00				
	4	0.45	0.54	0.44	1.00			
	5	0.78	0.85	0.87	0.81	1.00		
	6	0.27	0.26	0.25	0.14	0.23	1.00	
	7	0.88	0.92	0.94	0.55	0.90	0.27	1.00
6	1	1.00						
	2	0.97	1.00					
	3	0.97	0.99	1.00				
	4	-0.29	-0.29	-0.31	1.00			
	5	0.96	0.98	0.97	-0.25	1.00		
	6	0.50	0.49	0.50	-0.35	0.51	1.00	
	7	0.96	0.97	0.98	-0.31	0.99	0.52	1.00
7	1	1.00						
	2	0.97	1.00					
	3	0.84	0.91	1.00				
	4	0.73	0.77	0.94	1.00			
	5	0.22	0.76	0.90	0.97	1.00		
	6							
	7	0.77	0.80	0.92	0.96	0.99		1.00
8	1	1.00						
	2	0.96	1.00					
	3	0.97	0.96	1.00				
	4	-0.24	-0.14	-0.23	1.00			
	5	0.11	0.13	0.14	0.78	1.00		
	6	0.48	0.31	0.45	0.13	0.30	1.00	
	7	0.55	0.49	0.57	0.35	0.81	0.65	1.00
9	1	1.00						
	2	0.95	1.00					
	3	0.96	0.94	1.00				
	4	0.31	0.44	0.25	1.00			
	5	0.78	0.83	0.79	0.66	1.00		
	6	0.35	0.31	0.32	0.15	0.37	1.00	
	7	0.88	0.86	0.89	0.44	0.94	0.37	1.00

土壤粘土矿物类型等方面有特殊的作用。

六、TM 彩色合成图像最佳波段组合的选择

由于遥感数据自动识别分类技术尚处于试验研究阶段,精度常达不到生产实用要求,因此,目前以至今后相当长时间内,处理出信息丰富的彩色合成图像仍是遥感图像处理的主要任务。目前投入使用的图像数字处理系统与光机合成仪都采取三色合成原理,即在三个通道上安置三个波段图像,然后分别赋以红、绿、蓝或其补色黄、品、青,叠合在一起而得彩色图像。TM 共有 7 个波段,考虑到第 6 波段分辨力比其它各波段低 4 倍,且受大气干扰严重,故这个波段一般不参与合成。这样,从 6 个波段中任选 3 个进行组合,可产生 20 组合成方案。评价各种组合优劣的方法虽然不少^[4,5],但相比之下,从衡量各种波段组合的信息量出发,引用 Sheffield^[6]提出的计算 N 维数据熵值原理,进行优选最佳组合的方法,却是一种既全面、效果又较好且较简便的方法。其算式为

$$S = \sum_{i=1}^M P_i(x) \ln P_i(x) \quad (1)$$

式中: S ——熵值, $P_i(x)$ ——变量概率密度函数。在正态分布条件下:

$$P_i(x) = 1/K_i \exp[-(x - \bar{x})^T M_i^{-1} (x - \bar{x})/2] \quad (2)$$

式中: $K_i = (2\pi)^{N/2} |M_i|^{1/2}$; M_i ——样区协方差矩阵; x ——图像变量,即像元亮度值; \bar{x} ——图像均值变量,即像元平均亮度值; N ——波段数; M ——样区像元总数。

遥感图像变量近似正态分布,故可使用(2)式。将(2)代入(1)式得:

$$S = \ln(K_i) + \frac{1}{2} \sum_{i=1}^M x^T \cdot M_i^{-1} x \cdot P_i(x) \quad (3)$$

对于无偏估计,由(3)式得

$$S = N/2 + \ln(K_i) = N/2 + N/2 \ln(2\pi) + 1/2 \ln |M_i| \quad (4)$$

由(4)式可看出,熵值 S 随变量协方差矩阵行列式值 $|M_i|$ 的变化而变化。因此只需计算三个波段组合的协方差矩阵行列式,其数值大小就可表征该组合的信息量多寡。

依据上述方法计算了五个样区、由 3 个波段可能组合成的 20 个组合的协方差矩阵行列式值并列于表 4。由表 4 可以看出, TM3,4,5 及 1,4,5 这两种波段组合的信息量最大,不是名列第一,就是名列第二。其中 3,4,5 组合在 3 个样区名列第一; 1,4,5 组合在两个样区名列第一; TM1,2,3 波段(即自然彩色)组合的信息量很少,无例外地名列倒数第一; TM2,3,4 (即常规)组合的信息量也偏低,列于第 13 至 17 位不等;其它如 4,5,7; 3,4,7; 1,4,7; 2,4,5 等组合居中上,都名列前 7 名之内。这进一步表明, TM 图像彩色合成的最佳波段组合,必须有中红外(5 或 7 波段)及近红外(4 波段)参与,才能充分显示其所含的丰富光谱信息。

最佳波段组合选出后,要想得到最佳彩色合成图像,还必须考虑赋色问题。对人眼,最敏感的颜色是绿色,其次是红、蓝色,因此,应将绿色赋予方差最大的波段。按此原则,采取 4,5,3 波段分别赋红、绿、蓝色合成的图像,色彩反差明显,层次丰富,而且各类地物的色彩显示规律与常规合成片相近,符合过去常规片的目视判读习惯。如把 4,5 两波段

的赋色对调一下,即 5,4,3 分别赋红、绿、蓝色,则获得近似自然彩色的合成图像,适合于非遥感应用专业人员使用。

表 4 TM 各种波段组合的信息量比较

Table 4 Information Quantity Comparison of Different Combinations of TM Bands

序号	样区 2		样区 4		样区 5		样区 6		样区 9	
	波段组合	行列式值	波段组合	行列式值	波段组合	行列式值	波段组合	行列式值	波段组合	行列式值
1	3,4,5	12 808 098	1,4,5	628 866	3,4,5	775 922	1,4,5	191 956	3,4,5	565 241
2	1,4,5	7 601 620	3,4,5	618 014	1,4,5	711 610	3,4,5	191 481	1,4,5	446 057
3	4,5,7	6 154 186	4,5,7	389 588	4,5,7	438 165	4,5,7	174 472	2,4,5	312 818
4	3,4,7	4 468 085	1,4,7	327 338	1,4,7	354 395	1,4,7	138 300	4,5,7	207 559
5	2,4,5	4 378 767	2,4,5	285 835	2,4,5	316 282	3,4,7	115 625	3,4,7	101 642
6	3,5,7	1 940 928	3,4,7	225 578	3,4,7	231 431	2,4,5	63 412	1,4,7	74 910
7	1,4,7	1 679 713	1,5,7	129 517	1,5,7	179 213	1,5,7	59 374	2,4,7	65 884
8	2,4,7	1 634 883	2,4,7	126 674	2,4,7	137 649	3,5,7	50 705	3,5,7	55 678
9	2,5,7	811 797	3,5,7	86 126	3,5,7	108 756	2,4,7	46 492	1,5,7	19 714
10	1,3,5	331 504	1,3,4	67 130	1,3,4	91 227	1,3,4	44 749	1,3,4	19 610
11	2,3,5	259 218	1,3,5	54 571	1,3,5	81 305	1,3,5	38 641	2,5,7	18 085
12	1,3,4	168 172	2,5,7	53 356	2,5,7	71 548	1,3,7	22 794	2,3,5	15 233
13	2,3,4	91 098	2,3,5	16 576	2,3,5	27 037	2,5,7	180 705	1,3,5	15 094
14	1,2,5	71 600	2,3,4	14 928	2,3,4	24 583	1,2,4	15 645	2,3,4	14 958
15	1,5,7	68 702	1,3,7	12 060	1,2,4	151 199	1,2,5	12 631	1,2,4	2 525
16	2,3,7	57 328	1,2,5	10 431	1,2,5	13 521	1,2,7	11 246	2,3,7	1 613
17	1,3,7	57 299	1,2,4	9 683	1,3,7	6 492	2,3,4	8 006	1,3,7	1 376
18	1,2,7	19 480	1,2,7	4 300	2,3,7	2 829	2,3,5	6 465	1,2,5	815
19	1,2,4	12 106	2,3,7	3 743	1,2,7	2 445	2,3,7	4 091	1,2,7	45
20	1,2,3	1 292	1,2,3	656	1,2,3	533	1,2,3	1 517	1,2,3	21

七、结 语

(1) TM 图像系包括“亮度”、“绿度”、“热度”和“湿度”等信息要素的 3—4 维信息结构。

(2) TM7 个波段的信息量多寡不一,总的看以 5 波段最丰富;7,4,3,1 波段其次;2,6 波段最差。各波段间信息有重叠,尤以 1,2,3 波段间及 5,7 波段间相关性很高,彼此存在过多冗余信息,在相当大程度上可互相取代。4,6 波段则比较特殊,信息有较大独立性。但 6 波段与 1 波段分别受大气辐射和散射影响严重,需特别予以注意。

(3) TM 彩色合成图像的最佳组合必须包括中红外(5 或 7)波段和近红外(4)波段,总的看,以 3,4,5 波段的组合最佳。1,4,5 波段组合的信息量,有时可能略超过 3,4,5 组合,但要注意大气散射对 1 波段亮度值的影响。目前通用的常规合成方案(即 2,3,4 波段组合),未能充分反映 TM 的丰富光谱信息。除某些应用目的已能满足之外,多数情况下,应以 3,4,5 或 1,4,5 波段等组合取代之。

参 考 文 献

- [1] Williams D. L., et al., A statistical evaluation of the advantages of Landsat Thematic Mapper data in comparison to multispectral scanner data, *IEEE Trans. Geoscience and Remote sensing*, GE-22(3), 1984.
- [2] Towsherd Y. R. G. et al., Preliminary analysis of Landsat-4 Thematic Mapper products, *International Journal of Remote Sensing*, Vol. 4, No. 4, pp. 817—823, 1983.
- [3] 戴昌达等, TM 数据的信息特征, 遥感信息, (2) 1987.
- [4] Chavez P. S. et al., Statistical method for selecting Landsat MSS ratios. *Journal Applied photographic Engineering*, Vol. 8, No. 1, pp. 23—30, Feb. 1982.
- [5] Pat S. Chavez, Jr., Digital Processing techniques for image mapping with Landsat TM and SPOT simulator data, *Eighteenth International Symposium on Remote sensing of Environment*, pp. 101—106, Paris, France, October 1—5, 1984.
- [6] Sheffield C., Selecting band combinations from multispectral data, *Photogrammetric Engineering & Remote Sensing*, Vol. 51, No. 6, pp. 681—687, 1985.

The Information Characteristics of Thematic Mapper Data and the Optimal Band Combination

Dai Changda Lei Liping

(Remote Sensing Satellite Ground Station, Chinese Academy of Science)

Abstract

In this article, TM image data covering nine sample areas of various landscape from the northern frigid temperate zone undulating plain to the southern subtropics hilly land one analysed and compared in many ways. The results indicate that the spectral information of TM image is essentially three to four dimensional structure which physical meaning are match with Brightness, Greenness, Temperature and Humidity. In the seven TM bands, band 5 have the most ground information. There is the strong correlation between the three visible bands (band 1, 2, 3) and between the middle-infrared bands (band 5, 7), which indicate that there is very much "repeatability" or "redundacy" information in these bands. Between band 6 especially band 4 and the other bands is there very low correlation, which is shown that the information of the two bands is with very large independence. Computing twenty kind of combinations indicate that the color images by one of the visible bands, one of the middle-infrared bands and band 4 is generally with the most ground information, in which the combination of band 4, 5 and 3 or band 4, 5 and 1 is the optimal combination.

Key words information content information structure optimal band combination