维普资讯 http://www.cqvip.com

第4卷第4期 遥感学报 Vol. 4, No. 4
2000年11月 JOURNAL OF REMOTE SENSING Nov., 2000

文章编号：1007-4619（2000）04-0299-05

多重主成分分析及在地质构造信息提取中的应用

TP751
P618.130.8

朱小鸽
（石油勘探开发科学研究院 遥感地质所，北京 100083）

摘 要： 提出一种多重主成分分析方法，是对原始遥感数据经过图像变换或运算处理后，再有针对性地对专题信息进行二次乃至多次提取的图像处理方法，同时也是对多种不同类型，不同分辨率的遥感图像进行综合处理的尝试。应用于柴达木盆地西部山区提取地质构造信息获得显著效果。图像上新发现了一个鼻状圈闭及一组连接上、下两个断裂带的弧形纹理。

关键词： 遥感图像处理；多重主成分分析；地质构造信息
中图分类号： TP79/TE122 **文献标识码：** A

1 引 言

柴达木盆地位于青藏高原东北隅，是中国西部主要含油气盆地之一。试验区处于盆地西北缘，属于高原山地地貌。主体为一北西向展布的山脉，海拔3000—4000m；荒山秃岭此起彼伏，深沟陡坡纵横交错，地形相对高差可达300—400m，基本无植被覆盖，地表条件相当恶劣。由于无法进行连网地震勘探，目前的地质研究及勘探成果仅限于周边地区、对山区内部复杂山地地区的地质构造研究仍很不够。因此，选择此地区进行了遥感数据提取石油地质专题信息研究。研究方法重点放在遥感数据的多方法特殊处理上，从遥感地质的角度探讨油气勘探中急需认识的地质问题。

在大多数情况下，油气田的形成与地质构造密不可分，在有生油条件的盆地中，圈闭构造往往可以起到储集油气从而形成可开采油气田的作用。圈闭构造有许多种，最典型的如地层隆起形成的背斜圈闭（局部高），地层断裂形成的断块圈闭，不同岩性形成的岩性圈闭等。而断裂本身既可以形成圈闭储集油气，又可以成为破坏圈闭导致油气运移的重要因素。因此在对含油气盆地的石油地质勘探中，最重要的目标就是圈闭构造及断裂。圈闭与断裂是石油地质勘探中最重用的专题信息。

2 多重主成分分析

所谓多重，指对主成分分析的结果与其它处理结果及原始数据经过有针对性地选择后组合起来再次进行主成分分析。也可以是几种不同主成分分析结果的组合。

主成分分析的原理，着眼于变量之间的相互关系，是尽可能不丢失信息地用几个综合性指标汇集多个变量的测量值进行描述的方法。“在数学含义上，它是一种基于图像统计特征的多维正交线性变换”[1]，特点是数据压缩与去相关。而不是针对特征信息的处理方法。

张满郎先生在“金矿蚀变信息提取中的主成分分析”[2]一文中，介绍了由Crosta A.P.等人（1989）提出的“选择主成分分析（SPCA），Frazer S. J., Green A. A.（1987）提出的“直接主成分分析方法（DPCA）”，还提出主成分分析与植被掩膜及滤波，主成分逆变换几种不同的主成分分析的应用方法[2]。郭德方、叶和飞在《油气资源遥感》一书中提出了“特征主成分选择”[3]。以上几种不同主成分分析的应用技巧，目的都是提取地表不同的蚀变岩性信息进行找矿研究。

收稿日期：1999-03-22；修订日期：2000-01-17

作者简介：朱小鸽（1951— ），女，北京市人，高级工程师，1982年毕业于长春地质学院（现长春科技大学），获工程学士学位，从事遥感图像处理、遥感与GIS综合应用、石油遥感及遥感环境监测等方面工作，发表论文20多篇。

维普资讯 http://www.cqvip.com

本研究中所应用的多重主成分分析，主要是针对地质构造信息，特别是圈闭与断裂这两个石油地质专题信息，重点是隐伏构造信息。而隐伏构造信息可以通过地表地层含水量、水系、地形、岩性等差异特征表现出来，反映在遥感图像上，就是影纹结构、颜色色调的变化与异常。但这些信息往往被掩盖在地表物质光谱的强反射信息下，十分微弱，在原始图像中不易发现，在主成分分析中也往往处于较高分量中，信噪比低，信息量少，难以进行高质量的分析解释。

多重主成分分析方法与上述各种主成分分析方法的不同之处在于，首先应用各种手段，包括一般主成分分析、选择主成分分析(特征主成分选择)、波段间比值等，尽可能提取虽然微弱，但确实存在于遥感图像中的石油地质专题信息，再把那些对专题信息显示最好或较好的结果选择出来进行二次处理。其处理方法包括两种：一是进行不同的彩色组合或叠加以突出专题信息；第二种就是选取对专题信息提取最有利的结果和/或原始波段数据再次进行主成分分析，进行信息的二次提取与压缩。由于不同处理结果(如主成分分析与波段间比值的处理结果)之间以及与原始数据之间仍存在一定的相关性，又由于第二次主成分分析是在第一次图像处理的基础上有针对性地选择输入信息，是针对专题信息的数据压缩与信息分离，因此多重主成分分析是对微弱的专题信息的二次提取与增强。

3 多重主成分分析提取地质构造信息

3.1 主成分分析方法

封三图版Ⅰ图1为实验区山脉东部区域TM7，4，1波段组合，其北西西—南东东方向纹理十分清晰，并有一组北东向张性断裂。右上角为山前倾斜平原，其它信息较杂乱。对封三图版Ⅰ图1进行TM1—5，7六个波段主成分分析处理后，第三分量PCA(3)显示出一组清晰的向南西方向突出的弧形纹理，此组弧形纹理把上下两组北西西向纹理连接起来，并在北东向断裂处形成明显的转折。这一组信息在原始图像上十分微弱，基本看不出来。这一发现对于该地区的构造运动及断裂发育历史情况提供了新的十分重要的地质信息。

此外，通过主成分分析还发现在该区右上角存在一个鼻状局部隆起圈闭构造，需进一步的工作证实。

3.2 多重主成分分析方法

由于断裂和圈闭构造是石油地质勘探中特别重要的信息，尤其是圈闭构造，在已知的含油气盆地中，是石油勘探的首要目标。因此必须对其进行进一步的研究。

处理方案 通过对已知油气田构造研究后发现，地下构造在地表能以水系、地形、岩石类型、含水量、破碎情况等异常特征反映出来，遥感图像能够识别水系、地形、一些不同岩性分布等空间信息，特别是通过图像上影纹结构的变化、色调异常，能够追踪不同规模的线性单元或体系。总结这些线性体地质性质的研究后得出的结论是，线性体是不同类型、不同埋藏深度的断层位移的地表表示[4]。根据其在遥感影像上显示的清晰程度，与其所处的构造位置、相互关系以及与线性构造的空间分布关系，结合地质构造固有的地质属性进行综合分析，可揭示地壳的隐伏构造。因此关键是把有用信息提取出来。

由上述主成分分析的结果还不能清楚反映出该构造的内部特征及周边接触关系，因此采用多重主成分分析方案进行进一步特征信息提取。

实验区TM1—5，7六个波段间相关系数如表1。

表1 图1TM1—5，7波段间相关系数

Table 1 Co-relationship coefficient of TM1—5，7 of image 1

波段	TM1	TM2	TM3	TM4	TM5	TM7
TM1	1					
TM2	0.978	1				
TM3	0.9588	0.9883	1			
TM4	0.9515	0.9833	0.9954	1		
TM5	0.8711	0.9028	0.9177	0.9204	1	
TM7	0.8469	0.8931	0.9139	0.9193	0.9795	1

维普资讯 http://www.cqvip.com

由表1看出此地区TM数据波段间相关性很高。

对封三图版Ⅰ图1进行多重主成分分析处理，用PCA表示主成分分析，PCA( )括号内数字表示主成分分析结果的某个分量。

PCA1—主成分分析1　输入TM1--5,7；　输出PCA1（1—6）

PCA2—主成分分析2　输入TM1,4,5,7；　输出PCA2（1—4）

PCA3—主成分分析3　输入TM4,5,7；　输出PCA3（1—3）

PCA4—主成分分析4　输入PCA1（1,3,4），TM6,TM5/TM7；　输出PCA4（1—5）

其中，PCA1是对原始TM数据进行主成分分析处理，由于TM6热红外波段分辨率大大低于其它波段，因此暂时不用。特征向量矩阵如表2。

PCA2主要提取粘土矿物，碳酸盐岩矿物信息[3]。

PCA3则主要提取岩性信息。

对这3个原始数据的不同输入组合的主成分分析结果进行对比如表3。

表2　图1TM1—5,7特征向量矩阵

Table 2　Eigenvectors of covariance matrix of TM1—5,7 of image 1

输入 / 主分量	TM1	TM2	TM3	TM4	TM5	TM7	特征值/%
PCA1(1)	0.32631	0.24963	0.36629	0.36013	0.62670	0.41783	94.52
PCA1(2)	0.50392	0 30235	0.35928	0.32412	-0.52897	-0.37511	4.27
PCA1(3)	0.59230	0.00715	-0 32530	-0 41066	0.41490	-0.45002	0.77
PCA1(4)	-0.45188	-0.05111	0.26755	0.29292	0.39336	-0.69359	0.33
PCA1(5)	0.27338	-0.77852	-0.21247	0.52317	-0.01514	0.00966	0.07
PCA1(6)	-0.09914	0 48736	-0.71708	0.48693	0.01814	-0.03201	0.05

表3　不同主成分分析结果对比

Table 3　Comparison of different PCA results

PCA	PCA(1)	PCA(2)	PCA(3)	PCA(4)
PCA1	基本相同	鼻状构造、中部断裂突出，呈暗色调	分割鼻状构造的断裂带明显，呈暗色调	鼻状构造内部走向线明显、构造头部断裂呈亮色调，构造呈暗色调
PCA2		信息与上述基本相同	信息与上述基本相同，但断裂带呈亮色调	与上述基本相似，但构造线呈亮色调与上述相反
PCA3		鼻状构造、中部断裂呈亮色调与PCA1互补	与PCA1(4)相似，但构造头部信息呈暗色，信息相补	

对比发现第一分量均基本相同，代表总亮度，而其后几个分量所提取出的信息或是类似，或是相补，既其灰度值的高低（图像上亮暗）正好相反，而这些相补的信息有时正是含有特征地质构造的信息。由于这些特征信息通常是含在较强的背景信息之中，因此一般情况下其存在于主成分分析的较高分量中，信息较弱，信噪比大大降低。选择对目标信息提取最突出的分量进行彩色组合，由于图像上亮、暗色调的反差加大，比仅用普通的一种主成分分析结果，更能突出特征信息。

PCA4则是在以上处理分析的基础上，选出较好反映了鼻状地质构造信息的分量及其他处理结果，再次进行主成分分析，目的是信息压缩与专题信息再提取。选择依据为：

主成分分析的特点是第一分量包含的信息量最大，噪音降低，有利于细节特征的增强和分析，适用于线性特征的增强与提取[5]。而第三、四分量则主要包含着相关程度较高的波段之间的差别信息。TM5与TM7均为短波红外，适用于岩性分类，地质调查研究，但它们之间的相关性很强（在此地区高达0.9795），由表2的特征向量矩阵表明，第三、四分量主要反映它们的差异信息，这正是提取地质特征信息所需要的；其图像也证明了此两分量反映的构造特征信息较多。同理，选择TM5与TM7的比值也是

维普资讯 http://www.cqvip.com

为了提取地质、岩性差异信息。

由于地质构造信息往往在原始 PCA 分析的后几个分量，信噪比已经降低，而 TM6 波段分辨率虽较低，但其热红外光谱波段往往能够反映地下断裂的情况，因此参与二次主成分分析运算，基本是在相同的信息量基础上进行信息提取与压缩，它们之间的相关系数如表4。特征向量矩阵如表5。

表4说明，TM5/TM7 比值结果与主成分分析的第三、第四分量 PCA1(3,4)之间，原始波段 TM6 与主成分分析的第三分量 PCA1(3)之间，仍然存在一定的相关性，虽然与原始波段之间的相关性相比，其相关性已大大降低。

表4 图1 PCA1(1,3,4)，TM6，TM5/TM7 间相关系数

Table 4 Co-relationship coefficient of PCA1 (1,3,4,), TM6, TM5/TM7 of image 1

波段号	PCA1(1)	PCA1(3)	PCA1(4)	TM6	TM5/TM7
PCA1(1)	1				
PCA1(3)	-0.00283	1			
PCA1(4)	0.00134	0.00	1		
TM6	0.2586	0.4146	-0.0729	1	
TM5/TM7	-0.1274	0.7108	0.6049	0.197	1

表5 图1 PCA1(1,3,4)，TM6，TM5/TM7 特征向量矩阵

Table 5 Eigenvectors of covariance matrix of PCA1 (1,3,4), TM6, TM5/TM7 of image 1

主分量 \ 输入	PCA1(1)	PCA1(3)	PCA1(4)	TM6	TM5/TM7	特征值/%
PCA4(1)	-0.99935	0.00018	-0.00002	-0.03375	0.01270	96.41
PCA4(2)	-0.01688	0.51940	0.08138	0.68971	0.49761	2.04
PCA4(3)	-0.03032	-0.24781	-0.34671	0.66710	-0.61029	1.08
PCA4(4)	-0.00712	-0.65902	0.67179	0.27891	0.19117	0.43
PCA4(5)	-0.00673	0.48427	0.64951	-0.01881	-0.58585	0.04

表5的特征向量矩阵表明，多重主成分分析(这里是二次)的结果，第一分量 PCA4(1)包含了绝大部分的信息，特征与输入的原主成分分析的第一分量 PCA1(1)基本相同，信噪比高，信息量大。第二分量 PCA4(2)则负载了原主成分分析的第三分量 PCA1(3)、原始 TM6 波段与 TM5/TM7 的综合信息。第四分量 PCA4(4)对原主成分分析的第三、四分量 PCA1(3,4)是强负载，但符号相反，更突出了它们的差异，同时载有原始 TM6 波段的信息。此结果使得对专题信息有利的输入信息得到再次综合、压缩与分离。封三图版Ⅰ图2表明，鼻状构造及内部细节信息得到增强，达到了信息二次提取与增强的目的。

应用结果 对封三图版Ⅰ图1进行多重主成分分析处理的结果如封三图版Ⅰ图2，图中显示出一个明显的鼻状构造，并使特征信息更加突出，鼻状圈闭构造与断裂及周边的接触关系也更清晰。图像注记中用"P"代表"PCA"，其中：

封三图版Ⅰ图2A：P1(2,3,4)，显示出鼻状构造的轮廓①，北东向断裂②的切割关系，构造内部及与周边的接触关系不很清晰。

封三图版Ⅰ图2B：P1(3,2)P2(3)与封三图版Ⅰ图2D：P1(2)P2(3)P3(3)，有相似之处，从图像上可看出，鼻状构造与其周边的接触关系较清晰，构造南部的北西—南东向裂隙影像③突出。但前者弧形纹理④较清晰，但构造内部信息不够突出。而后者构造内部的信息较丰富，尤其是构造东南倾末端⑤形态更清晰，但弧形纹理及构造边界不甚清晰，尤其是弧形纹理的显示很微弱。

封三图版Ⅰ图2C：P1(4)P2(3)P3(3)，突出了构造边界①、弧形纹理④及其明显的转折端影像。构造内部信息也较丰富。

封三图版Ⅰ图2E：P4(1,4,2)与封三图版Ⅰ图2F：P4(1,4)P2(3)，有相似之处，构造内部信息清晰、丰富。断裂影纹②突出，有立体感。尤其是构造内部与构造走向基本垂直的小断裂⑥的发育情况反映的十分明显。鼻状构造的形态与周边的接触关系也更清楚，尤其是构造西北部与另一个构造的接触关系⑦，北部的沙漠覆盖以及图像东北角的环形(仅露

维普资讯 http://www.cqvip.com

出一角)影纹⑧比其它几个图像都更清晰。

初次发现这个鼻状构造后，经地质解译，已知它的东北部即是一个已知含油背斜构造。因此进一步的信息提取不但要求尽可能揭示构造内部的走向、断裂发育情况等详细的信息，更重要的是揭示它与周边地质体，尤其是北部已知含油构造的接触关系。封三图版Ⅰ图2就是应用多重主成分分析方法，针对石油地质勘探的要求，重点提取构造本身及其周边地质体、接触关系等地质信息的尝试。

对以上图像的地质解译表明，此鼻状构造与其东北部的已知含油背斜圈闭有关。由于已知含油背斜构造处于试验区山脉外缘的较平坦地区，构造形态、特征、岩层已被物探、钻井等资料证实。经过多重主成分分析处理后新发现的地质构造，有明显的标志层可以追踪，鼻状构造影像特征明显。经过地质解译剖析了油田最关心的它与已知含油构造之间的关系，认为两个构造有相似性，同处于一个背斜带中段，两者以一大型北西西向断层相接触。周边地区的研究表明，此背斜中有一连串的局部隆起构造，鼻状构造为其中的一个。这一发现与地质解译的结果，为此地区的石油勘探提供了重要信息。

4　结论与讨论

应用多重主成分分析的方法提取柴达木盆地西部山脉地区的地质构造信息取得了较好的效果，封三图版Ⅰ图2中的弧形转折与鼻状构造都是此次研究中发现的新构造，因此，多重主成分分析是特征地质信息提取中一个有效的手段。

由于数据源的限制，在此次研究的具体实现中，是把原始遥感数据(如TM、MSS等)与经过某些处理后的结果数据，如经过主成分分析处理后的某些分量，共同作为分析对象再次进行主成分分析。由于这种方法是在对原始数据已经过一次处理与选择的基础上，有针对性地对专题信息进行再提取处理，因此实际上是一种二次及至多次信息提取方法，其针对性、目的性更强。

本文中的多重主成分分析也可以作为对各种不同类型、不同分辨率的遥感数据及其它多元数据综合起来进行处理、分析的一种尝试。

致谢　本研究的地质解译由王世红、张友炎完成，在此表示感谢。

参考文献(References)

[1] Zhu Liangpu. Remote Sensing Geology [M]. Beijing: Gelogical Publishing House, 1994. [朱亮璞. 遥感地质[M]. 北京：地质出版社，1994.]

[2] Zhang Manlang. Principal Component Analysis in Extracting of Gold Mine Alteration Information [J]. *Remote Sensing Technology and Application*. 1996, 11(3). [张满郎　金矿蚀变信息提取中的主成分分析[J]. 遥感技术与应用，1996，11(3).]

[3] Guo Defang, Ye Hefei. Petroleum Resources Remote Sensing [M]. Hangzhou: Zhejiang University Press, 1995. [郭德方，叶和飞. 油气资源遥感[M]. 杭州：浙江大学出版社，1995.]

[4] G. I. Amurskii. Remote Sensing for Research of Structure in Oil-and gas-bearing Construction [M] Ding Shubai *et al*. Translation. Beijing: China Mine University Press, 1993. [G. I.阿莫尔斯基等. 丁树相等译. 含油气建造内构造研究的遥感方法[M]. 北京：中国矿业大学出版社，1993.]

[5] Feng Maosen. Digital Processing of Remote Sensing Image [M]. Beijing: China Geology University, 1988. [丰茂森. 遥感图像数字处理[M]. 北京：中国地质大学，1988.]

Extracting Geological Structure Information by Multi-Principal Component Analysis

ZHU Xiao-ge

(*Remote Sensing Geology Dept., of Research Institute of Petroleum Exploration & Development, Beijing* 100083, *China*)

Abstract: A new way of multi-principal component analysis is proposed. The image processing technology of information re-extraction focuses on the special objects based on the original remote sensing data being processed by some conventional methods. Meantime, it is a way to process comprehensively different remote sensing data with different spatial resolutions and spectral ranges. The application of the method has shown good result for extracting geological structure information in west Qaidam Basin of China.

Key words: remote sensing image processing; multi-principal component analysis; geological structure information