

场地类型的图象识别

蔡红卫

(国家地震局兰州地震研究所)

马苏引

(甘肃省建筑勘察设计院)

摘 要

本文采用改进的连续亨明方法对兰州市城关区及张掖地区进行场地类型的图象识别。结果表明,将图象识别方法用于场地类型的划分是可行的,它不仅改进了以往经验的、宏观的工作方法,而且在方法上实现了数值和计算机化,在增进识别能力的同时,提高了识别的准确度和速度。利用图象识别方法可以综合分析地质构造、场地土、地形地貌等多种因素或特征,并对特征进行选择,定量地对场地类型进行识别分类。

关键词: 图象识别; 连续亨明方法; 场地类型

一、引言

在历次强震的震害调查中,人们已清楚地认识到,地震破坏不仅仅与地震的震级和震中距离有关,而且在某一小区范围内,也与当地的场地条件有着密切的关系。在认识到场地条件的影响后,便开始在工程建筑中对它进行考虑,目前,对场地条件的考虑还仅限于地震小区划和具体的工程场地的选择中***。而场地类型在工程地震中又是一个至关重要的问题。一个地区的震害取决于两方面的因素—地基失效和地震动强度***。目前,有关地震动强度研究,大多是采用一系列的分析方法,所得结果大多也是一系列的定量参数,如PGA、PGV、反应谱等等,而对地基失效的研究,仍基本上局限于宏观的、定性描述的方法。

对地基失效的判定决定场地的选择,在场地对抗震是否有利这一问题上;按抗震规范的要求和实际需要习惯上将场地划分为三级:有利场地、不利场地和危险场地,以作为对场地的分类评价。但场地的性质是多种多样的,可以通过很多方面的特征去识别它们。而各类的分界线并不十分明确,不仅如此,对于同一因素,在场地评价中所起的作用也难于一致。总的看来,目前的场地评价还是以经验的、宏观的方法为主,评价中难免带有较强的主观意识。

* 胡聿贤,地震工程学,第十一章,1983.

** 孙崇绍、陈丙午,兰州市建筑场地抗震区划,1981.

*** 胡聿贤,地震工程学,第四章,1983.

对于这一问题, 国内外都曾有人进行过研究, 国家地震局地质研究所的王立功同志首先应用易损性分析方法识别一些地区的场地, 进行地震小区划(如徐州)* 以及建筑物的震害预测(如乌鲁木齐)**, 当时采用的方法是根据易损性因子分区打分, 然后进行综合评定, 其中也包括了识别和聚类分析的过程, 并取得了相应的进展。

基于以上认识, 本文主要采用王碧泉等提出的改进的连续亨明方法^[1], 对兰州市城关区以及张掖地区的张掖、临泽、高台三县进行场地类型的图象识别, 旨在探讨用图象识别方法来识别场地类型的可行性及具体实施步骤。

二、图象识别方法简介

亨明方法的核心就是求解亨明核和计算亨明距离。设全部样品集合分为危险点集合 D 类和安点集合 N 类。两集合所含样品个数分别为 N_1 和 N_2 , 则总样品个数 $N = N_1 + N_2$, 每个样品 X_i 为一个 m 维向量, 由 m 个连续特征描述:

$$X_i = (X_{i1}, X_{i2}, \dots, X_{ij}, \dots, X_{im})^T. \quad (1)$$

其中 $i = 1, 2, \dots, N; j = 1, 2, \dots, m$ 。

具体计算步骤包括:

(1) 将连续特征数据标准化

$$a_j = \max[X_{ij}] - \min[X_{ij}]. \quad (2)$$

其中 X_{ij} 为第 i 个样品中的第 j 个特征。令

$$X'_{ij} = (X_{ij} - \min[X_{ij}]) / a_j, \quad (3)$$

作为各个样品的第 j 个新特征。

(2) 计算亨明核

分别计算 D 类和 N 类每一特征的均值:

$$\begin{cases} \bar{X}_j(D) = (\sum_{X \in D} X_{ij}) / N_1, \\ \bar{X}_j(N) = (\sum_{X \in N} X_{ij}) / N_2 \end{cases} \quad (4)$$

和每类样品的各特征的方差平方和:

$$\begin{cases} S^2(D) = \sum_{j=1}^m S_j^2(D) = \sum_{j=1}^m \left[\sum_{X \in D} (X_{ij} - \bar{X}_j(D))^2 \right] / (N_1 - 1), \\ S^2(N) = \sum_{j=1}^m S_j^2(N) = \sum_{j=1}^m \left[\sum_{X \in N} (X_{ij} - \bar{X}_j(N))^2 \right] / (N_2 - 1) \end{cases} \quad (5)$$

若 $S^2(D) \leq S^2(N)$, 则计算 D 类样品的亨明核, 否则, 计算 N 类样品的亨明核。

$$Z = (Z_1, Z_2, \dots, Z_j, \dots, Z_m)^T. \quad (6)$$

* 徐州市城乡建设委员会, 徐州市城市抗震防灾规划, 1984.

** 新疆地震局, 乌鲁木齐地震烈度小区划及震害预测报告, 1984.

$$\text{其中 } Z_j = \begin{cases} \bar{X}_j(D), & \text{当 } S^2(D) \leq S^2(N) \\ \bar{X}_j(N), & \text{当 } S^2(D) > S^2(N) \end{cases}$$

(3) 计算亨明距离 $d_i(X_i, Z)$

$$d_i(X_i, Z) = \left[\sum_{j=1}^m W_j (X_{ij} - Z_j)^2 \right]^{\frac{1}{2}} \quad (7)$$

式中 W_j 为权系数。

$$W_j = \frac{|\bar{X}_j(D) - \bar{X}_j(N)|}{[\max|\bar{X}_j(D) - \bar{X}_j(N)| + 0.5]}$$

(4) 分类判别

根据实际需要, 我们对场地进行了三级划分, 选取了两个阈值 T_1 、 T_2 , 且 $T_1 < T_2$ 。

对于 $S^2(D) \leq S^2(N)$ 的情况:

$$\begin{cases} X_i \in D, & d_i(X_i, Z) \leq T_1 \\ X_i \in N_1, & \text{当 } T_1 \leq d_i(X_i, Z) < T_2 \\ X_i \in N_2, & d_i(X_i, Z) \geq T_2 \end{cases} \quad (8)$$

对于 $S^2(D) > S^2(N)$ 的情况

$$\begin{cases} X_i \in N_2, & d_i(X_i, Z) \leq T_1 \\ X_i \in N_1, & \text{当 } T_1 < d_i(X_i, Z) < T_2 \\ X_i \in D, & d_i(X_i, Z) \geq T_2 \end{cases} \quad (9)$$

其中 D 为危险场地* (不利场地 (II)); N_1 为不利场地 (I); N_2 为安全场地。

三、改进的连续亨明方法在场地图象识别中的应用

为了研究图象识别方法在识别、分类场地中的应用, 我们选取了资料丰富且对场地已进行过详细研究的兰州市城关区和资料不太丰富、研究程度相对较差的张掖地区作试验区域, 进行场地类型的图象识别。

1. 兰州市城关区场地类型的识别

(1) 研究区的场地条件

兰州市位于青藏高原东北边缘黄河谷地之中, 市区沿黄河两岸分布, 大河谷地和黄土覆盖这两个因素决定了该市区的场地条件的基本特点。我们的研究范围西起雷坛河口, 东到桑园峡口, 南北均到南北两山, 包括整个的城关行政区, 连同黄河河床总面积共约 70 平方公里。地震危险性分析结果表明, 兰州市 50 年基准期内 0.1 超越概率为 VIII 度, 故识别以 VIII 度为标准。

研究区位于兰州东盆地之内, 区内主要构造有北北西向的雷坛河断层, 此外, 西盆地北西西向金城关断层和孔家崖断层东端均伸入了本区。由于区内的断层长度都不很大, 按照断层长度—震级统计关系, 区内的这些断层最大可能发生 5.5 级地震。这种等级的地震在地表不至引起开裂, 所以, 在图象识别时主要考虑断层带自身的影响。见图 1。

* 这里所说的“危险”是指图象识别中的 D 类, 但是否属于真正的“危险”场地还应作具体分析。

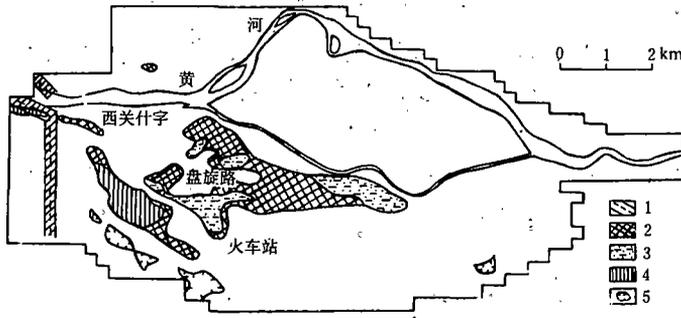


图1 饱和砂土、淤泥质土分布图

1.断层；2.厚度大于0.5米的饱水粉砂、细砂及轻亚粘土分布区；3.厚度大于1米的饱水粉砂、细砂及可液化轻亚粘土分布区；4.厚度大于0.5米的饱水淤泥质土分布区；5.崩塌、错落

Fig.1 Distribution of saturated sandy soil and silt

区内可能引起地基失效的场地土主要是饱和砂土、淤泥质土、人工填土和疏松、可能产生震陷的大孔土(图1)。饱和砂土和淤泥质土的分布范围有限,不连片且短距离之内厚度和粒度成份变化迅速。人工填土的分布范围也不广,厚度很少超过5米,一般以素填土为主,仅个别点有较厚的杂填土。

表1 场地多指标图象识别标志

场地条件	特征内容(识别标志)	类型
构造因素	是否有 M>6 级地震发震断层通过	二值
场地土因素	基盘面的埋深	连续
	人工填土	连续分档
	地下水埋深	
	黄土厚度	连续
	是否有淤泥质土分布	二值
	是否有可液化层(饱和砂土)分布	
地形地貌	地形坡度	连续分档
	是否邻近有可能发生小规模崩塌的斜坡	二值
	是否有小规模滑坡、崩塌发育	
	是否有大规模滑坡、崩塌发育	
	是否邻近危险的斜坡	
	山坡上有没有地下水出露或浸润	

为一个样品,样品总数 N=1110。

研究区地处黄河中上游山间盆地之中,南北两侧紧邻黄河的高阶地及剥蚀丘陵区。黄河两岸断续发育有五级阶地,各阶地之间的结构、土质都有一定的差别,且高差较大。其中区内南侧皋兰山前的山坡,是兰州市区最高陡的斜坡,由于山坡高陡、坡面向阴,山坡土体湿度较大,从火车站至五泉山一带,山坡下方断续分布有错落体。

(2) 样品与特征

在考虑了研究区的场地特征后,识别是在 1:25000 的图上进行的,采用两组平行线(间距为 0.25km),将研究区划分为 1110 个小区,其中一组平行线近于北北西向,与本区地貌界线 and 构造线的总体走向一致,另一组则与之相垂直。每一小区作

在样品类别的划分上，我们选定以场地土作为划分依据，另外考虑到区内三种可能受震失效的场地土的分布及其影响范围，在学习过程中，选择饱和砂土和淤泥质土，以及人工填土厚度大于 5 米作为划分标准，若某一样品中含有这三种土的任何一种，则将此样品命名为危险类 D 类样品，否则为安全类 N 类样品。

在研究了区内地质构造、场地土、地形地貌等自然条件以及它们的相互关系之后，为了进一步分析地震时场地的性能，确定场地的稳定性以及其本身可能产生的地基失效，共选取了十三个特征对研究区进行场地类型的图象识别，见表 1。

(3) 识别结果

用改进的连续亨明方法对兰州市城关区的场地进行图象识别时，由于 $S^2(D) > S^2(N)$ ，需要计算 N 类样品的亨明核。对于阈值 T_1 、 T_2 的选取是采用试算的方法，反复检验，选择使分类结果的误识率 ϵ 达到最小的阈值。误识率 ϵ 的计算公式为：

$$\epsilon = (P(\omega_1)\tau_1 + P(\omega_2)\tau_2) / N$$

式中 $P(\omega_1)$ 、 $P(\omega_2)$ 分别为 D 类和 N 类样品的先验概率，N 为样品总数， τ_1 、 τ_2 分别为 D 类和 N 类样品被错误分类的样品数。

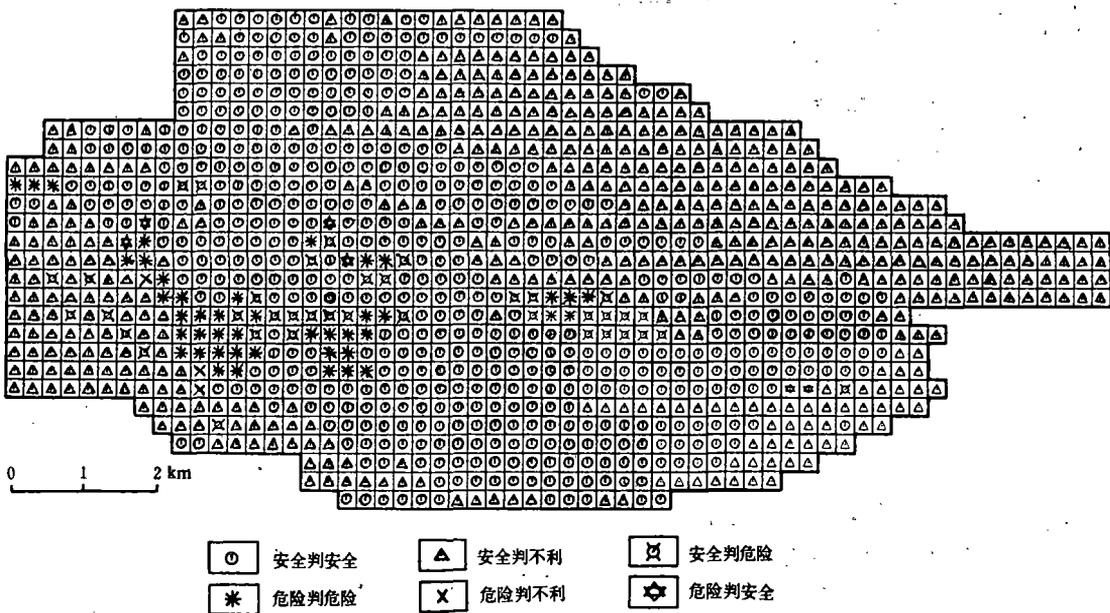


图 2 兰州市城关区场地图象识别结果

Fig.2 Recognizing result of site type in Chengguan-district of Lanzhou City

从分区结果图 2 可以看出，兰州市城关区图象识别所判定的对建筑物抗震“危险”区，主要分布在天水路以西，沿民主东路一线分布，以及盘旋路兰州饭店附近部分地区；对建筑物抗震不利区主要分布在大雁滩、小雁滩、骆驼滩一带，以及白银路以西，华林坪以东

表 2 兰州市城关区各类地层标准贯入试验资料

工程名称	试验地点	受试地层	校正后 $N_{63.5}$ (击)			N_{cr}	判别情况
			$N_{63.5}$ (max)	$N_{63.5}$ (min)	$N_{63.5}$		
市城建开发公司六号楼	西关什字北	第三系强风化砂岩	120	45	88		不液化
省气象局办公楼	五里铺桥东	第三系强风化砂岩	89	66	77	26	不液化
市质量中心综合楼	中央广场黄家园	第三系强风化砂岩	91	42	71	13	不液化
省新厦公司办公楼	庆阳路·静宁路口	第三系强风化砂岩	99	52	75	17	不液化
飞天商场	盘旋路广场	第三系强风化砂岩	84	52	68	10	不液化
兰山商场扩建工程	兰州火车站	第三系强风化砂岩	162	46	104		不液化
省物资局办公楼	南城巷	第三系强风化砂岩	187	120	150		不液化
靖远矿务局驻兰办事处	定西路东段	第三系强风化砂岩	122	71	90	18	不液化
亚欧商厦	西关什字	第三系强风化砂岩	97	77	85	14~17	不液化
兰澳花园大厦	永昌路北段	第三系强风化砂岩	97	54	76	14~17	不液化
省职工财经学院 I 期工程	东岗镇 小河边	黄土状轻亚粘土	5	2	3		液化
省教育学院 I 期工程	雁滩	饱和砂层	3	2	2.6	9.3	液化

地带。除此以外，兰州市城关区绝大部分地区均是对建筑物抗震有利区。

在这里值得注意的是，所谓的“危险”实际上是较厚的不良场地土的分布区域，如饱水粉细砂、淤泥质土等，这些场地土在遭受Ⅷ度地震袭击时，可能产生地基失效，影响建筑物的安全，但可以采取工程措施予以防治或消除。按抗震规范的标准（地震时可能发生滑坡、崩塌、地陷、泥石流等及发震断裂带上可能发生地表位错的部位为危险场地）。尽管在图象识别中，我们使用了“D类”（危险）这一专业术语，然而就其本质而言，应判为“不利”场地。所以，根据识别的结果，将兰州市城关区的场地划分为三类：（1）、有利场地；（2）、不利场地（Ⅱ）（判别时为危险场地）；（3）、不利场地（Ⅰ）（判别时为不利场地），两个不利场地在程度上有一定的差别。

表 2 为兰州市城关区一些地段各类地层标准贯入试验结果，有助于对以上判别的检验。

表 3 张掖地区场地识别标志

场地条件	特征内容 (识别标志)	类型
场地土因素	是否有软弱土层或饱水细砂层分布	二值
	地下水埋深	连续分档
构造因素	是否有 $M < 5.5$ 级地震发震断层通过	二值
	是否有 $M < 6.5$ 级地震发震断层通过	
	是否有 $M < 6.0$ 级地震发震断层通过	

大型建筑的工程勘察中，基础多采用大开挖，对较薄土层不作原位测试，仅对强风化砂岩作标准贯入。可以看到：兰州市城关区绝大部分地区上部第三系砂岩虽然风化极强且表层饱水，仍不致地基失效，与抗震规范有关液化的初判条件相符。而在土层较厚的东岗、雁滩等地段，则存在液化的可能，为抗震不利区。

2. 张掖地区场地类型的图象识别

张掖地区是在大范围内进行场地识别的试验区，工作范围包括张掖、临泽、高台及肃

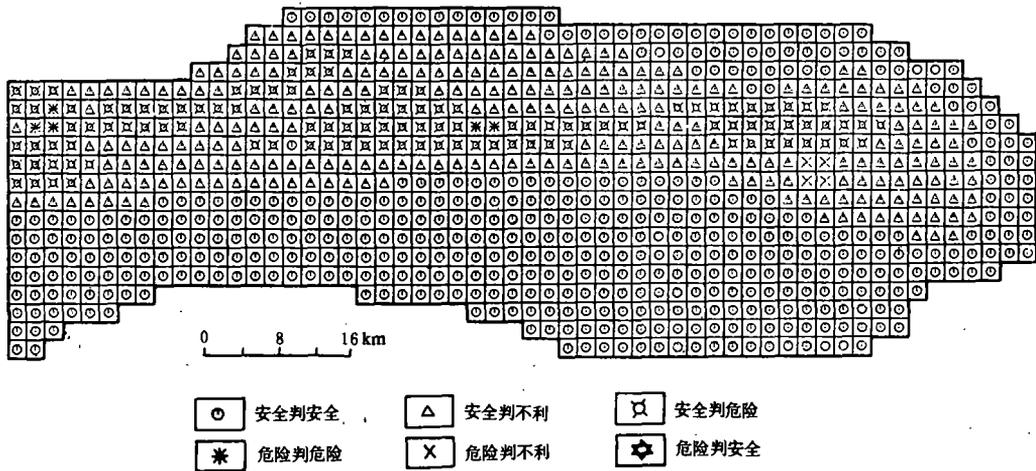


图 3 张掖地区图象识别结果

Fig. 3 Recognizing result of site type in Zhangye area, Gansu Province

南县的一部分，总面积约 3400 平方公里。

总的原理与识别途径与兰州市场地识别相仿，这里不再叙述。

对大范围的场地识别, 只选择区域性的控制特征, 划分不宜过细, 因此识别工作是在 1:200000 图上进行的, 以 2 公里为间距, 用两组平等线将识别区划分为 848 个小区, 即 $N=848$ 。

根据当地的条件, 在控制区域性场地条件的指标中, 选取了五个特征作为分类判别的标志, 见表 3。

从分区结果 (图 3) 可以看出, “危险”区主要分布在张掖市城区及火车站以北的部分地区, 临泽县城以西新华农场的大片地区, 黑河河漫滩附近; 不利区主要分布在张掖东南部山丹河附近, 临泽县城经北黑河河漫滩附近及高台县以南高台农场一带; 有利区主要分布在研究区南部大片地区。

四、结论及识别结果的讨论

通过对兰州市城关区及张掖地区的场地类型的图象识别, 可以看出:

(1) 采用改进的连续亨明方法对场地进行识别还是一个尝试, 从识别结果来看, 这种方法比宏观的、经验的方法有很大的优越性。首先, 对于事先分解出的样品特征, 不论在任何地点, 任何情况下都能“一视同仁”地衡量它所起的作用。而宏观经验或半经验方法则受到判断者的人为因素影响较大。其次, 当判断的因子数很大时, 图象识别方法能根据事先确定的每个因子的权重, 判断各因子所起的作用。

(2) 从识别结果来看, 选定合理的阈值, 能将识别结果划分得更细微一些, 而各档次之间的差别又能较好的反映场地条件的差异, 因而能使小区划工作做得更详细。

(3) 图象识别方法使宏观小区划工作便于程序化、计算机化、改进了以往经验的、宏观的工作方法, 在提高准确度、增进识别能力的同时, 提高了识别速度。

(4) 图象识别方法能够比较各特征在识别过程中所起的作用。根据所计算的每个特征因子 D 类和 N 类的均值, 比较两类均值的差, 差值越大, 该特征在识别过程中所起的作用也就越大, 比如在对兰州市城关区及张掖地区各识别因子进行标准化后, 依次可以看出, 软弱土层及可液化土层的分布 (或饱水粉细砂层的分布) 在识别过程中起着很大的作用。

(5) 对图象识别结果要进行具体的分析和认识, 即对所分类的等级作出相应的场地条件的说明, 因为识别出的同一类型可能包含有不同性质的问题, 通过场地分析, 可以将识别结果进一步具体化。

由于我们的工作还带有一定的探索试验的性质, 从兰州市城关区及张掖地区的场地识别结果总体来看, 还是比较理想的, 但也存在着某些问题有待进一步解决。就兰州市城关区的场地而言, 首先, 五泉山附近的不稳定边坡体没有识别出来, 从场地条件分析来看, 这些不稳定边坡体应为“危险”场地 (或不利 (II) 区), 而识别结果却为不利场地即不利 (I) 区。造成这种结果的原因可能是由于特征的提取存在着某些问题, 从表 3 中可以看出, 关于边坡的几个识别特征, 在整个识别过程中, 所起的作用很小。其次是黄土层的识别, 在一般情况下, 黄土的厚度越大越危险, 而识别结果却是黄土层薄的地方反而“危险”, 这主要是由于黄土层分布主要在兰州市的南北两山, 识别区内黄土层的厚度差别很小, 黄土层薄的地方“潜水埋深”变浅, 后一因素所起的作用超过前者。这一识别结果与实

际情况是符合的。当黄土层厚度相差不大时,它不是决定场地类型的主要因素,而当厚度差别增大时,图象识别方法完全可以充分估计在场地分类中所起作用。

亨明方法只是图象识别中的基本方法之一,在场地类型的识别当中,也取得了比较满意的结果,但是否将其它方法用于场地类型的划分其结果会更好,还有待于我们的进一步研究和改进,总之,将图象识别用于场地类型的划分是可行的。

本文得到兰州地震研究所孙崇绍研究员的指导,在此表示谢意。

参考文献:

- [1]王碧泉、陈祖荫,模式识别理论、方法和应用,地震出版社,1989.

PATTERN RECOGNITION OF SITE TYPE

Cai Hongwei

(Seismological Research Institute of Lanzhou, SSB, China)

Ma Suyin

(Institute of Architectural Reconnaissance and Design of Gansu)

Abstract

By using the improved consecutive Hamming method, the site types in the Chengguan district of Lanzhou City and Zhangye area have been classified. The results show that it is feasible to classify site type using the method of pattern recognition. It not only improves the past empirical and macroscopic methods, but also makes it numeralization and computerization. The present method also increases the accuracy and velocity of recognition while improving the recognition ability. It is possible to comprehensively analyze and pick up the various factors such as geology, topography, geomorphology, site soil features, etc, and to identify the site types quantitatively.

Key Words: Pattern Recognition; Consecutive Hamming Method; Site type