

太原理工大学

博士学位论文

跳汰机人工神经网络控制技术研究

姓名：符东旭

申请学位级别：博士

专业：机械电子工程

指导教师：熊诗波

20070401

跳汰机人工神经网络控制技术研究

摘要

跳汰选矿是物理选矿的一种重要方法。我国目前入洗原煤的 50%左右是用跳汰机分选的。

跳汰机作为跳汰选煤的关键设备，其控制包括分层和排料两个方面，控制效果的好坏直接决定了分选效率的高低。由于传统的控制方法是基于精确的系统数学模型的控制，适于解决线性、时不变等相对简单的问题，而对于跳汰机这样复杂的非线性、大滞后系统，建立精确的数学模型极其困难，因此传统的控制方法很难达到良好的控制效果。本文的主要研究工作包括：

在详细分析了影响跳汰机分选效果的因素后，指出：尽管分层是排料的前提，但跳汰是一个连续的过程，分层和排料是同时进行的，排料的好坏直接影响着分层效果，分层和排料的共同作用决定了跳汰机分选效率的高低，跳汰机的控制必须将二者结合考虑，而目前的控制系统仍然是简单的、粗放的控制，问题的关键是没有解决跳汰床层分层状态的在线检测。

分析了跳汰床层 γ 射线分层状态检测系统的构成及实现方法，有效解决了分层状态的在线检测问题，为跳汰机有效控制系统的建立确定了基本的、可靠的平台。通过采集分层状态典型样本数据，使用人工神经网络和支持向量机学习算法，建立了跳汰机分层、排料一体模型。

提出了把跳汰机分选效果（终极目标—矸石带煤）作为目标函数，即

把性能函数直接作为目标函数，而把分层和排料的众多过程变量作为输入，利用人工神经网络的分布并行处理、非线性映射、自适应学习和鲁棒容错等特性，通过实验确定了适合跳汰机特性的网络拓扑结构、隐含层数目、隐含层节点数，并用遗传算法对神经网络的权值和阈值进行了优化。

研究了基于统计学习理论的支持向量机算法在跳汰机研石段的模式识别问题，发现专门针对有限样本情况的支持向量机比人工神经网络具有更大的优越性，表现在测试误差进一步变小。对支持向量机的核函数进行了线性函数、多项式函数、径向基函数和S型函数的对比，并利用最小二乘支持向量机进行了仿真，针对此项研究，发现了径向基函数作为核函数时训练误差和测试误差均较小。

由于预测控制注重的是模型功能，而不是结构形式，过程的描述可以通过简单的实验获得，不需要深入了解过程的内部机理。因此，预测控制算法改变了现代控制理论对模型结构较严格的要求，更着眼于根据功能要求，按最方便途径建立多样性模型。这一点对于跳汰机这种理论严重落后于实践的系统尤为适用。

从模型算法控制、动态矩阵控制和广义预测控制这三种预测控制算法和基于神经网络的预测控制分析入手，提出了利用事先离线训练好的神经网络预测器，通过预测下一时刻的研石带煤量实现跳汰机风阀周期预测控制的方法，并利用BP神经网络构建了研石带煤预测器以及优化器，利用已取得的样本数据进行了仿真，结果表明，研石带煤有了明显的降低。

关键词：跳汰机，人工神经网络，研石带煤，分层，排料，支持向量机

RESEARCH OF ARTIFICIAL NEURO NETWORK CONTROL TECHNOLOGY ON JIG

ABSTRACT

Ore jigging is one important method in physical preparation. About 50% of raw coal is processed in this method in China. As the key equipment on coal preparation, jig control includes the control of stratification and discharging. The separation efficiency of jig is related directly with its control results. As the conventional control is based on the precise mathematical model and used to solve problems of linear, invariant time system, it is difficult to get the satisfied effect on the complicated, nonlinear and great lag system like jig, for it is hard to find its precise mathematical model. The main work did in this paper includes:

After analyzed the influencing factors of jig separation efficiency, it is pointed out that, although stratification is the precondition of discharging, but the jigging process is continuous, the stratifying and discharging is happening at the same time, i.e. the discharging good or bad is affecting the stratifying results, so it's the combined action of stratification and discharging deciding the efficiency of jig, jig control must be considered with this combination. However the control system we are using now is simple and extensive for the online measurement of jig bed status is not solved yet.

The γ ray Stratification Status Measuring System is analyzed in details and it shows that the jig bed online measuring problem has been solved and it provided a basic virtual platform for the construction of jig control system. Through taking typical samples from site, the integration model of stratification and discharging is reached by using ANN and SVM.

It is pointed out that the Efficiency of Jig should be the final goal function (ultimate objective is the Coal Lost In Waste), i.e., the performance function is the goal function, and many variables in process make up of the inputs of the ANN. As the ANN has the advantages of distributed parallel treating, nonlinear mapping, self-adapting learning and robust, the topology structure, the number of hidden layer and the neuron number of hidden layer are determined. The weights and threshold is optimized through inherit arithmetic.

The Support Vector Machine (SVM) arithmetic based on Statistical Learning Theory is studied on pattern recognition of waste segment of jig and it is found that the SVM based on limit samples shows the better results compared with ANN in test error. After compared the SVM kernel with linear, polynomial, RBF and S functions, using simulation of least-square SVM, it is found that the training error and the test error is smaller when the RBF as the kernel.

As the predict control pay more attention to the model function rather than structure form, the description of process can be got through simple test and it is unnecessary to realize the inside mechanism of the process. So the predict

control algorithm changed the strict request of model structure from the modern control theory. To build multiformity model in most convenient route is established according to the function desire. This method is so suitable to the jig process, which its theory got great behind the practice.

Proceeding with model algorithm, dynamic matrix control, generalized predict control and the predict control based on NN analysis, the method is offered of how to use the off-line pretrained ANN to predict the Coal Lost In Waste to achieve the predict control of air valve cycle of jig. The predictor of Coal Lost In Waste and optimizer is constructed using BP Neural Network. The results of simulation using sample data shows that the Coal Lost In Waste is lower.

KEY WORDS: jig, artificial neuro network, coal lost in waste, stratification, Discharging, support vector machine

声 明

本人郑重声明：所呈交的学位论文，是本人在指导教师的指导下，独立进行研究所取得的成果。除文中已经注明引用的内容外，本论文不包含其他个人或集体已经发表或撰写过的科研成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律责任由本人承担。

论文作者签名： 符东旭 日期： 2008.1.24

关于学位论文使用权的说明

本人完全了解太原理工大学有关保管、使用学位论文的规定，其中包括：①学校有权保管、并向有关部门送交学位论文的原件与复印件；②学校可以采用影印、缩印或其它复制手段复制并保存学位论文；③学校可允许学位论文被查阅或借阅；④学校可以学术交流为目的，复制赠送和交换学位论文；⑤学校可以公布学位论文的全部或部分内容（保密学位论文在解密后遵守此规定）。

签 名： 符东旭 日期： 2008.1.24

导师签名： 王训林 日期： 2008.1.24

第一章 绪论

1.1 课题研究的背景与意义

1.1.1 煤炭在我国能源结构中的主体地位将是长期的

我国是世界产煤大国，煤炭是我国的主要能源，在一次能源生产和消费结构中，煤炭的比重一直都在 70% 左右^[1-2]。而且在未来相当长的时期内，以煤炭作为主要能源的战略地位不会改变。充分发挥煤炭资源优势、立足煤炭、立足国内，是我国能源工业发展的现实和必然选择。

我国煤炭资源具有很大优势：

(1) 分布广泛，总量丰富，为世界所瞩目。我国煤炭资源分布面积约 60 多万平方公里。根据第三次煤炭资源预测与评价，全国煤炭资源总量为 5.57 万亿吨，保有储量 10032.6 亿吨，其中已利用储量 3469.09 亿吨，尚未利用储量 6563.5 亿吨。

(2) 种类齐全，煤质优良。煤类从褐煤到无烟煤均有分布，低变质烟煤占 33%，其次为中变质烟煤、贫煤、无烟煤和褐煤。我国煤炭以特低硫、低硫煤为主，占 56%；低中硫、中硫煤占 33%。灰分中等，低中灰、中灰煤占查明资源量的 76%，特低灰、低灰煤也比较丰富，占查明资源量的 22%。发热量高，中高热值煤占 92%，中低热值煤很少。我国北方的中西部地区分布着丰富的优质动力用煤，其中特低灰煤占 36%，集中分布在陕北和内蒙古东胜地区，是全球少有的优质动力用煤。

(3) 具有其他能源无可比拟的优势，煤炭约占我国化石能源的 95%，储量的 90%。

(4) 煤系共伴生矿产资源丰富。煤系地层中具有煤层气、锗、铀、高岭土等多种矿产。我国煤炭地质总局组织的《全国煤层气评价》表明，我国 2000 米以浅煤层气资源量达 15 万亿立方米，内蒙古胜利煤田查明锗资源量 1600 吨，占全国查明资源量的近 1/3。新疆、内蒙古等省区煤田中含有丰富的铀矿。

然而煤炭资源与人口、区域经济和环境矛盾十分突出。我国人均占有煤炭储量仅为世界人均占有煤炭储量平均水平一半左右。煤炭资源分布与区域经济发展和消费需求极不适应。我国是世界第一产煤大国，也是煤炭消费的大国。

我国煤炭资源虽然比较丰富，但分布极不平衡。西部资源比较丰富，但地区发展

比较滞后，运输也不便利；东部资源较少，但市场需求大，运输条件比较好；北部资源集中，南部分散且较少，而需求较大；因此形成了西煤东运，北煤南调的格局，煤炭运量占铁路货运总量的 40%左右。煤炭行业已经成为国民经济高速发展的重要基础。

煤炭的基础性地位还体现在，与其他能源相比具有明显的成本优势。与所有石化能源相比，煤炭是最便宜的一种能源，同等的发热量，用煤的成本只相当于用油的 30%，天然气的 40%。专家预测到 2010 年，使用石油、天然气的成本将是煤炭的 8 倍，甚至更高。依靠大量使用石油、天然气来解决我国能源问题的设想，是不实际的。

我国经济发展的阶段性、区域经济的不平衡性和资源禀赋的特殊性，决定了煤炭是我国的主要能源。在可以预见的未来几十年内，煤炭仍将是我国的主要能源和重要的战略物资，具有不可替代性，煤炭工业在国民经济中的基础地位，将是长期的和稳固的。在最近完成的《我国可持续能源发展战略》研究报告中，20 多位中科院和工程院院士一致认为，到 2010 年煤炭在一次性能源生产和消费中占 60%左右；到 2050 年，煤炭所占比例不会低于 50%^[47, 48]。

同时，随着煤炭科学技术的发展，煤炭进一步转化成为新型洁净能源，减少开发利用过程中所造成的污染正在成为全社会关注的焦点。最近十多年来洁净煤技术产生重大突破，目前我国已经研究出脱硫、脱氮、除尘以及煤炭气化、液化、水煤浆等技术。这些技术的发展和推广，将使煤炭的应用基本上能够达到环保要求，从而树立起全新的形象。

此外，煤炭的应用领域不断拓宽，正在逐步摆脱单纯的一次性能源形象。在煤化工方面，目前已经可以从煤炭中提炼出硅、焦油、苯等上百种化工产品。煤炭液化和“煤变油”技术的开发，使煤炭改变了其原始的固体形态，使用领域相当广阔：

到 2020 年，中国规划投资逾 1 万亿元大力发展煤化工产业，这有助于解决中国日益紧张的能源需求问题。《煤化工产业中长期发展规划征求意见稿》显示，2010 年、2015 年和 2020 年，煤制油规划年产分别为 150 万吨、1000 万吨和 3000 万吨（2015 年和 2020 年，煤制油占成品油的比例分别为 4%和 10%）；掺烧于汽油的二甲醚的规划是，在上述三个时间点上年产分别为 500 万吨、1200 万吨和 2000 万吨。煤烯烃的规划是，在上述时间点上年产分别为 140 万吨、500 万吨和 800 万吨，占烯烃总量比例分别为 3%、9%和 11%。煤制甲醇的规划增长速度也很大，在上述时间点上年产分别达 1600 万吨、3800 万吨和 6600 万吨，到 2020 年，煤制甲醇占甲醇总量的 94%。二

甲醚和煤烯烃由甲醇转化而来，前者转化比例为 1.5:1，后者转化比例为 2.92: 1，两者合计在 2010 年、2015 年、2020 年共约用甲醇 1170 万吨、3300 万吨、和 5400 万吨。

中国计划在全国打造七大煤化工产业区，分别是黄河中下游、蒙东、黑东、苏鲁豫皖、中原、云贵和新疆。按规划初稿，黄河中下游、新疆、蒙东将形成大规模甲醇、二甲醚、煤制油生产基地。

综上，煤炭是我国最安全、最经济、最可靠的能源，我国煤炭资源总量远远超过石油和天然气资源，随着高新技术的推广应用，煤炭生产成本正在并将继续降低，洁净煤技术已取得重大突破，这些都将使煤炭成为更加廉价、洁净、可靠的能源，而所有煤炭利用技术的基础都来源于它的第一道工序：选煤。

1.1.2 我国选煤大发展是必然趋势

我国是世界上除南非和印度以外少数的一直以煤为主要能源的国家。目前我国煤炭消费存在的主要问题，首先是煤炭利用技术落后，效率非常低。其次是能源利用强度仍然大大高于发达国家与世界平均水平。我国煤炭的大量生产和消费，无论是对大气污染、酸雨等区域性环境，还是对气候的变化等全球环境来说都有很大影响，我国的“节能减排”形势将变得越来越严峻。因此，洁净煤技术的提出和推广逐渐成为煤炭行业的重要课题。洁净煤技术不仅能够有效地减少有害气体和烟尘的排放，而且还能提高煤炭的利用效率，节约资源。

洁净煤技术涉及面很广，贯穿煤炭生产、运输、加工、利用的全过程。我国洁净煤技术的重点不同于发达的工业化国家，在美国、欧洲各国、日本、澳大利亚等国，煤炭基本上都经过洗选加工后利用，而且主要用于发电，因此，他们的洁净煤技术重点是发展先进的燃烧器，实现洁净燃烧和控制污染。我国的情况则不同，全国原煤入选率仅为 30% 左右，70% 的原煤未经洗选就用于燃烧和其他利用，而且用于发电的部分仅占 30%，其余则用于国民经济各部门和人民生活，因此，目前阶段，洁净煤技术仍着重于燃烧前的洁净生产，即对原煤进行洗选加工、提高煤炭质量。

选煤技术是洁净煤技术中最成熟、投资最少、成本最低的技术，也是洁净煤技术的第一个环节、第一道工序。机械选煤可以除去煤中的杂质，包括灰分和 50~70% 的黄铁矿，减少燃煤的烟尘和 SO₂，建选煤厂的投资仅为电厂脱硫装置的 1/10。根据我

国的煤炭实践，选煤厂脱除 1t 的 SO₂成本约为 500~600 元，而电厂脱除 1t 的 SO₂成本约为 1400~1600 元。

纵观我国选煤的发展历程，选煤在中国起步较晚。中华人民共和国 1949 年成立时，全国只有 10 个选煤厂，年入洗原煤能力只有几百万吨。新中国成立后，选煤发展加快，大致分为三个阶段^[3]：

第一阶段为起步阶段（1949—1980），也称“洗煤保钢”阶段，在此期间建设的一大批选煤厂采用了波兰、苏联和中国自己制造的选煤设备，原煤入洗量由 510 万吨/年增加到 1142 万吨/年，原煤入洗比例由 7.67% 提高到 18.4%。

第二阶段为成长阶段（1981—2000），也称“洗煤节能”阶段，这一阶段，中国能源供应极度紧张，国家非常重视节约能源以及提高燃烧效率，把选煤作为提高煤炭质量、节约能源的重要手段来抓。1980 年，邓小平提出“开发煤炭，首先应当做出必须做的，是要提高洗煤比重”。1982 年，国务院颁发了“发展洗煤，节约能源”的第五号指令，对促进选煤的发展起到了重要作用。在此期间，我国先后从美国引进了平朔安太堡、兖州兴隆庄选煤厂，从波兰引进了钱家营选煤厂，从苏联引进了山西晋阳选煤厂等，这些选煤厂的工艺和设备大大提高了我国选煤的技术水平。全国原煤入洗能力也从 1.1 亿吨/年增加至 2000 年的 5.0 亿吨/年，原煤实际入洗量从 1981 年的 1.2 亿吨增加至 2000 年的 4.0 亿吨。

第三阶段为快速发展阶段，从 2001 年起，中国煤炭生产和消费大量增加，引起的大气污染和环境问题受到各方的关注和重视，许多省市出台了燃用煤炭严格的质量标准，促使煤炭企业重视煤炭质量，促进了动力煤的洗选。与此同时，中国钢铁工业的大发展也使炼焦煤的洗选大幅度增长。这一阶段，选煤作为洁净煤技术的基础和前提进入了快速发展阶段，入洗原煤能力从 2000 年的 5.0 亿吨/年提高到了 2005 年的 8.3 亿吨/年，原煤入洗量由 4.0 亿吨提高到了 7.03 亿吨，原煤入洗率达到了 32%。

尽管在“十五”期间煤炭洗选加工在我国得到了较快的发展，但是和世界发达国家相比，差距还是比较明显的。目前，全世界原煤平均入洗比在 50% 左右，国外一些发达国家都明显高于这一比例。虽然美国煤炭入洗比例只有 55%，但据统计，剩余部分煤质好，无需洗选即可符合用户要求。另外，这些国家的精煤产品质量高，炼焦精煤灰分一般 <7%，硫分 <1%，水分 <10%。分析其原因，除有的国家原煤易选外，主要原因是工艺系统合理，设备性能可靠，尤其是近年来强化了以“脱硫降灰”为宗旨

的细粒和极细粒级煤分选设备。另外一个原因是，在一些国家，重介选工艺逐渐升到主导地位。80年代国外所用的选煤方法中基本以跳汰选为主，到了90年代，一些国家重介选所占比重已由原来的第二位上升到第一位。如美国，在80年代，重介选占31%，跳汰选占49%，到1996年，重介选升到45%，跳汰选下降到35%。澳大利亚、南非跳汰选使用量很少，甚至不用。但德国由于煤易选，且巴达克跳汰机技术成熟，跳汰选一直是主要选煤方法。

可以看出，我国原煤入洗比例相当低，并且质量较差，但从国家发展规划上看出，“十一五”期间，煤炭入洗率要由目前的32%提高到50%，原煤入洗量提高到2010年的13亿吨。这样的发展速度在世界上是空前的。

我国的选煤技术在“十一五”期间还要取得明显的进步，研制一批具有自主知识产权的先进选煤技术和设备，大大提高选煤技术和装备水平，在重介质选煤方法发展迅速的今天，常用的跳汰选煤方法也必须发挥出应有的作用。

1.1.3 跳汰选煤研究意义重大

跳汰选煤是在垂直升降的变速介质流中按矿粒密度进行分选的过程^[5]。是否采用跳汰方法的关键是看原煤的可选性。而原煤的可选性即是原煤在洗选过程中获得一定质量产品的可能性和难易程度。它是一种工艺技术评价指标，主要评价在某种选煤过程中能否获得既定质量的产品。影响原煤可选性难易的因素主要是产品的质量要求、选煤方法和工艺技术。而可选性难易又决定采用哪种选煤方法。原则上，中等可选、易选的和极易选的3种原煤采用跳汰选煤方法。难选煤是用跳汰选还是用重介选，应通过技术经济比较后来确定，对极难选煤，原则上必须采用重介法，以求高水平、高质量和高效益。近年来重介选煤发展很快，尤其对难选煤具有较高的分选效率，但其建厂投资和生产成本要比跳汰高，生产管理和设备维护也比跳汰工艺复杂，两种方法各有优势，要根据具体条件确定。但跳汰选在现阶段乃至今后相当长的时间内，仍将是常用的选煤方法之一。

目前，我国各种选煤方法的比例为：跳汰45%，重介39.5%，浮选9.5%，其它6%^[3]，跳汰选仍占据主导地位。近几年来，我国重介选煤呈上升趋势，这主要是因为设备磨损问题、介质回收问题等都得到了解决，但对于易选煤来讲，跳汰选仍显示出工艺流程简单、处理量大、加工费用低的优势，随着我国煤炭洗选的大力发展，原煤入洗比

例将逐步向发达国家看齐，这是历史发展的必然趋势^[4]。因此，跳汰选在我国未来若干年内，随着重介选的比重上升，仍将占有较大的比例。

然而，和重介选煤方法相比，跳汰选煤最大的弱点就是其分选效率低，表现在“矸石带煤”上一般均在 5-12% 左右，与重介选煤方法的 2% 相差甚大，而在跳汰机制造水平相差无几的基础上，进一步完善检测手段以及提高智能控制水平进而深入研究物料分层机理、了解和掌握影响物料分选的相关因素成为重点和难点。那么，从全国范围看，即使把跳汰机的分选效率普遍提高 1%，也将产生巨大的社会效益和经济效益。

1.2 跳汰选煤研究的历史与现状

跳汰选的应用，最初见于德国矿冶学家阿格里科拉（Georgius Agricola）于 1550 年撰写的论述 16 世纪欧洲开采、冶金技术的巨著“矿冶全书”（De Re Metallica Libri XII），中译本定名为“坤舆格致”。书中介绍了用“动筛跳汰法”洗选矿石的情况，当时采用脉动式跳汰筛板浸入固定的水槽。到 19 世纪初，跳汰技术得到了进一步的发展，活塞跳汰法代替了动筛跳汰法。1891 年鲍姆申请空气脉动跳汰法专利，1930 年，该工艺与自动调控装置相结合才使其受到重视。到 1960 年将排料技术与透筛技术相结合，研制出了连续排料的跳汰机^[5]。

1.2.1 跳汰理论滞后于实践

1851 年法国的倍尔罗利特发表“对矿物机械加工的研究、确立现用操作制度和可能操作制度的理论试验”，它从个别颗粒的运动动力学角度探讨了跳汰分选原理。从这篇论文算起，跳汰理论的研究已经有 140 多年的历史^[93]，但有些问题仍未真正解决，尽管一些假说从各个不同侧面已经触及到问题的核心部分，但是始终未能揭开跳汰机理的实质。究其原因是，流体力学是一门非常复杂的科学，流体对固体的作用就更为复杂，很难甚至不可能用精确的数学表达式描述，只能通过大量实验，用经验方程描述；至今为止的各种假说^[5, 93]包括：

（1）古典理论：

研究颗粒在水中的运动特性，但却忽略了加速期内颗粒阻力牵引确定这一重要命题，事实上，实际跳汰过程中颗粒很少能达到它们的沉降末速。

（2）重介理论：

认为颗粒分层来自于对颗粒的重介作用，但没有注意到重介分选很少依靠颗粒与流体的相对作用。

(3) 位能理论：

由 Mayer 提出，认为跳汰床层在跳汰过程中寻求位能最低状态，只要能使床层松散、使颗粒能重新排列，就会得到好的分层效果，但没有考虑颗粒的大小在跳汰过程中的影响，也未提出如何达到颗粒松散的操作制度。

(4) 统计理论：

基于处理大批量、散状原煤的情况，利用统计理论来描述分层和分离现象，比较具有现实性，也易于接受，但至今对分层性能与操作制度在对应性研究上注意得不够。

近年来，国内外还有一些新的理论出现，如陈迹提出的等压强同层位分选理论，认为颗粒在跳汰床层中主要受压强作用而不是按密度分层等。

上述理论使人们对跳汰过程的认识逐渐加深，但相对于跳汰机型的不断更新、相对于控制系统的不断进步来说，跳汰理论已经远远落后于生产实践，不能有效指导实际的跳汰生产，也解释不了宽级别或不分级入选的现象。

但在长期的生产实践中人们已经对跳汰分层的现象取得了一些共识，正是这些共识使跳汰技术得到了发展。这些共识可以归纳为以下几点：

跳汰分层基本上是按密度分层，同时也受粒度的影响。分层后，相同粒度密度大的物料分布层位相对靠下，相同密度粒度大的物料分布层位也相对靠下。

跳汰过程离不开床层，床层是矿粒按密度分层的介质。

跳汰过程是周期性脉动过程，需保证一定的松散度范围，松散度上限过大或过小，对矿物按密度分层都不利。

人们经过长期实践得到这些共识的同时，跳汰机也从“活塞式机械脉动”发展为“空气脉动”，进而从“周期性运动机械风阀”发展成为“数控气动风阀”，以至出现了以德国“巴达克”为代表的跳汰机时代。

1.2.2 跳汰机过程控制的现状及技术难点

跳汰机过程控制主要是两个方面：一是重产物排出过程的控制，需要在保证床层分层稳定的前提下，达到一定的分离精度。二是床层分层过程的控制，应使选取的风水操作制度能满足物料按密度进行分层，尽量减少不同密度层间的污染。

多年来，重产物排出过程的控制一直是先由人工设定好重产物床层厚度期望值，用浮标传感器获取重产物层的厚度，与设定值相比较后，通过偏差对排料闸板或排料轮进行控制，一般采用简单的逻辑比较法或 PID 算法。这种控制的效果较差，具体表现为排料闸板或动作迟缓或大起大落，床层或超厚或排空，波动很大，严重时发生压斗子事故。目前，在邢台矿业集团的邢台矿和东庞矿选煤厂采用了由太原理工大学机械电子工程研究所开发的模糊排料系统，该系统运用模糊控制技术，克服了上述床层不稳定的状况，已取得良好的控制效果。

床层的分层过程控制的现状是：跳汰周期和频率等参数均通过数控箱的拨码盘由人工设定，操作工人完全依赖于手工探查和自己的经验来做调整。所谓手工探查即工人用探杆（木棍）插入动态床层中来感觉重产物层的厚度和床层的松散程度。显然，这样的控制很难实现合理的风水制度，以使分层效果达到最佳状态。这就导致了灰分控制不精确、数量效率低和大量的精煤流失。

事实上，床层按密度分层是一个非常复杂的液固两相流相互作用的过程，受到很多因素的影响（如风阀工作制度、脉动水流的速度、加速度、风压、顶水量等），加上检测环节不完善，所以控制的难度很大。主要技术难点集中在以下几个方面：

（1）要实现对床层分层状态的自动控制，首先要检测或估计出床层状态的优劣，但适合跳汰机分层状况检测的传感器的发展很不成熟。检测手段的落后，制约了跳汰机自动控制的发展。

（2）跳汰分选是一个非常复杂的过程，相关因素很多，且互相关联，互相影响，无法建立精确的数学模型，传统控制方法难以实现精确控制。

（3）尽管国内外对跳汰分选的作用机理进行了多方面的研究，取得了很多有益的结果，但对何种风水操作制度对床层的松散状况和分层效果更为有利，还没有取得共识。

1.3 自动控制技术的发展及对跳汰控制的影响

这里从古典控制理论和现代控制理论二个方面讨论自动控制技术的发展和可能应用于跳汰机控制的若干方面。

古典控制理论是以传递函数为基础，研究单输入单输出自动控制系统的分析和设计问题。这些理论研究较早，现在已经比较成熟。并且在工程实践中得到了广泛的应

用。

现代控制理论是二十世纪 60 年代在古典控制论的基础上，随着科学技术的发展和工程实践的需要而迅速发展起来的。其内容主要以状态空间法为基础，研究多输入多输出、变参数、非线性、高精度、高效能等控制系统的分析与设计问题。最优控制、最佳滤波、系统辨识、自适应控制等理论都是这一领域研究的主要课题。特别是近年来由于电子计算机技术和现代应用数学研究的迅速发展，使现代控制理论又在研究庞大的系统工程的大系统理论和模仿人类智能活动的智能控制、生物控制、模糊控制等方面有了重大进展^[8]。

工业生产过程控制技术是一种运用控制理论、仪器仪表、计算机和其它信息技术，对工业生产过程实现检测、控制、优化、调度、管理和决策，达到增加产量、提高质量、降低消耗、确保安全等目的的综合性技术，主要包括工业自动化软件、硬件和系统三大部分。工业控制自动化技术作为 20 世纪现代制造领域中最重要的技术之一，主要解决生产效率与一致性问题，对企业生产过程有明显的提升作用。

1.3.1 跳汰控制的发展

跳汰控制同过程控制的发展一样，大致经历了以下三个阶段：

第一阶段是 70 年代以前的简单控制系统。采用 PID 控制规律的单输入单输出反馈控制回路成为当时过程控制的核心，其基础理论是经典控制理论。单回路 PID 控制无法处理复杂工业过程中多变量间的耦合，而且控制系统的给定值由人工事先设定，不能根据生产过程的变化和相应的过程信息及时进行调整。这严重限制了连续工业过程的发展。

而对于跳汰控制技术而言，目前绝大多数的跳汰机仍然停留在这一阶段，即只是对复杂的跳汰过程中的排料部分进行简单的 PID 控制，从理论上和实践上都反映出了这种方式的局限性：重产物层厚度即使在跳汰机的操作变量稳定不变的情况下也是随着入洗原料煤的性质而波动的，人工给定值的确定不能及时反映这一变化，因而成为有限控制。在世界范围内比较著名的德国巴达克跳汰机也仅仅局限在这一层次。如图 1—1 所示。

第二阶段是 70 年代至 80 年代的先进控制系统。在这一时期，计算机技术的迅速发展，分布式工业控制计算机系统（DCS）的出现与成熟为在生产过程中实施先进控

制创造了技术基础。这一时期的控制理论在深度和广度上有了许多进展，像鲁棒控制、非线性控制、预测控制在理论上都有重大突破。但是，控制理论的完美性与实际过程控制的实用性的鸿沟依然存在，控制理论的发展仍不能满足工程实际的需要。这就迫使人们去研究生产过程的特点与难点，以缩小理论与应用的距离。

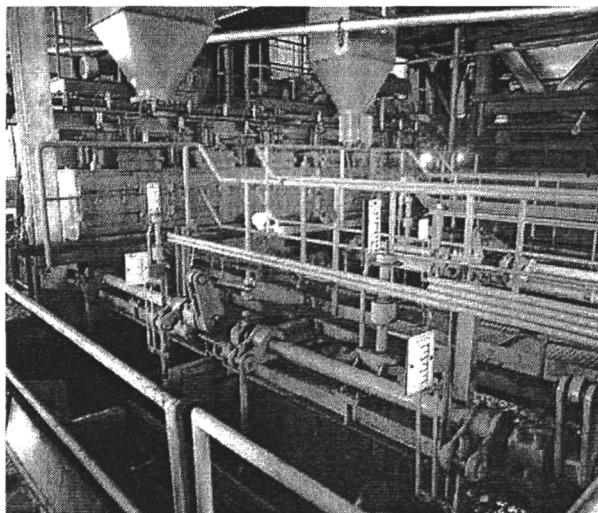


图 1—1 德国 $45m^2$ 巴达克跳汰机

figure 1—1 $45m^2$ Batac jig from Germany

跳汰控制在此时期开始在多参数检测方面有所发展，例如澳大利亚的 Apic 跳汰机可检测脉动速度、空气室液位、床层密度的核测量等，但对入料粒度、密度和脉动波形的影响不能分辨；日本的跳汰机能够测量床层脉动的波形、排料量，根据这些数据，可以控制高低压风的进风量和频率，设计出了能够满足不同煤质的可变波形跳汰机，使选煤效率得到提高，但仍然没有把最优分层状态和分选效率作为目标函数，仍然停留在对局部变量的控制上，因而没有得到推广应用。

第三阶段，二十世纪 80 年代后期特别是 90 年代以来，随着多学科如控制论、信息论、系统论和人工智能等学科的渗透与交叉，同时信号处理、数据库、计算机技术的迅猛发展，为实现高水平的工业过程自动控制提供了强有力的工具。这一阶段的特点是人工智能技术开始在连续过程控制中应用并取得成果。

此时的跳汰控制取得了一些成果，但主要是针对局部变量、个别问题的研究，从跳汰机整体控制研究方面，有代表性的当属由抚顺老虎台矿选煤厂与抚顺市三元工业测控技术研究所联合完成的“TKX—1 型跳汰机自动控制系统的研究”项目，该项目于 1999 年 10 月 6 日在老虎台矿通过了辽宁省科委组织的鉴定，鉴定认为该系统的特点是：紧紧抓住床层检测、给煤、排料三个重要环节（风、水为手动调节），实现“通

过控制跳汰机二段重物料层厚度的稳定来间接稳定精煤灰分”的目标^[95]。然而，该系统仍然属于通过浮标传感器实现跳汰机给、排料的粗放控制。

1.3.2 智能控制技术在跳汰控制中的应用现状

智能控制代表了自动控制的最新发展阶段^[6, 7]，也是应用计算机模拟人类智能实现人类脑力劳动和体力劳动自动化的一个重要领域。智能控制象征着自动化的未来，是自动控制科学发展道路上的又一次飞跃。

智能控制是人工智能和自动控制的重要部分和研究领域，人工智能的发展促使自动控制向智能控制方向发展。20世纪60年代中期，自动控制与人工智能开始交接，著名的美籍华人科学家傅京琼教授首先把人工智能的启发式推理规则应用于学习系统；然后又于1971年论述了人工智能与自动控制的交接关系，至此，智能控制开始发展应用。近几十年来，随着人工智能技术的快速发展，对智能控制的研究出现了一股新潮。各种智能决策系统、专家系统、学习控制系统、模糊控制系统和神经网络系统等已被应用于各类工业过程控制。

模糊控制是智能控制的一个活跃研究领域，扎德于1965年发表了他的著名的论文“模糊集合”，开辟了模糊控制的新领域。此后，在模糊控制的理论探索和实际应用两方面，都进行了大量的研究，并取得可喜的成果。模糊控制已应用于工业生产系统的众多方面。

神经网络又是智能控制的一个重要研究领域，早在1943年麦卡洛克和皮茨就开发了脑模型，其最初动机在于模仿生物的神经系统。随着超大规模集成电路、光电子和计算机技术的发展，人工神经网络已引起更广泛的注意。近年来，基于神经元控制的理论和机理已获得进一步的研究和应用。由于神经网络控制器具有学习能力、记忆能力、概括能力、并行处理能力、容错能力和适应于超大规模集成电路制造等重要特性，有许多基于神经网络的控制器被设计出来。这类控制器称为神经网络控制器。神经网络控制器具有并行处理、执行速度快、鲁棒性好、自适应性能强和适合应用等优点，因而具有广泛应用前景。以神经网络控制器为基础而构成的神经网络控制系统已在非线性和分布式控制系统，以及学习系统中得到不少成功的应用。

随着智能控制技术的成熟，智能控制作为一门独立学科已正式建立，而且自1987年以来，一些国际组织定期或不定期地举行各类有关智能控制国际学术会议，在一定

程度上反映了智能控制发展的势头。近十多年来，国内对智能控制的研究也比较活跃，相关学术组织不断出现，学术会议频繁召开，尤其是在现实生产中得到了应用，在跳汰控制中的应用条件逐渐成熟。

目前，人工智能的研究是与具体应用相结合进行的。基本上有如下方面：专家系统、机器学习、模式识别、理解自然语言、机器人学、智能决策支持系统以及人工神经网络。

一、专家系统的应用

专家系统是依靠人类专家已有的知识建立起来的知识系统，目前专家系统是人工智能研究中开展较早、最活跃、成效最多的领域，广泛应用于医疗诊断、地质勘探、石油化工、军事、文化教育等各方面。它是在特定的领域内具有相应的知识和经验的程序系统，它应用人工智能技术、模拟人类专家解决问题时的思维过程，来求解领域内的各种问题，达到或接近专家的水平。

专家系统在跳汰上的应用比较典型的是 1992 年底由原煤炭部组织的在东庞矿 309 跳汰机上的实验系统。

该系统通过在线检测跳汰机的床层厚度、排料闸板开度、斗子提升机电机电流、风压、循环水流量、给煤机转速、总精煤灰分、入洗原煤灰分等参数，将操作人员的经验包括风阀控制经验汇入知识库，在各参数高低限范围内，使跳汰机自动处于较稳定的工作状态。

该专家系统的内核试图通过床层水压传感器在线检测每个脉动周期内水压变化，通过计算有限颗粒群的运动轨迹预测出床层内的分层状况，但毕竟仅仅一个水压的信息，不足以反映其他变量对床层的影响；另外，床层水压的变化不仅仅是由床层物料的变化引起的，风阀的开闭速度以及给水压力、供风压力的变化对床层水压的影响很难计算出来。但这一思路仍为我们考虑跳汰床层状态检测提供了很好的借鉴。

二、人工神经网络的应用

跳汰过程的复杂性需要人工智能技术的支持，而近年来迅速发展起来的应用模糊数学模式、人工神经网络模式的方法逐渐取代传统的用统计模式和结构模式的识别方法，特别是神经网络方法在模式识别中取得了较大进展。人工神经网络是在研究人脑的奥秘中得到启发，试图用大量的处理单元（人工神经元、处理元件、电子元件等）模仿人脑神经系统工程结构和工作机理。

在人工神经网络中，信息的处理是由神经元之间的相互作用来实现的，知识与信息的存储表现为网络元件互连间分布式的物理联系，网络的学习和识别取决于和神经元连接权值的动态演化过程。

多年来，人工神经网络的研究取得了较大的进展，成为具有一种独特风格的信息处理学科。当然目前的研究还只是一些简单的人工神经网络模型。要建立起一套完整的理论和技术系统，需要作出更多努力和探讨。然而人工神经网络已经成为人工智能中极其重要的一个研究领域。

人工神经网络在跳汰上的应用近几年有不少学者进行了不同程度的研究。

朱金波等人在实际分选资料的基础上，利用人工神经网络学习操作因素和分选指标的映射关系，用计算机经 250000 次训练学习后，所映射的关系达很高的精度；预测了一定操作条件下分选指标，并确定了最佳操作参数，实验验证了用人工神经网络所得的结果。其中给出一定的分选指标确定操作因素和在一定的操作因素下会得出怎样的结果一直是选矿工作者研究的重要课题之一^[10]，但实验未给出测试误差。

钱立全，杜长龙提出了一种新的松散度表达形式，并在此基础上设计了床层松散度灰色预测控制器及神经网络控制系统。仿真结果表明该系统有可能实现对床层松散度的有效控制^[11]。

贾廷贵，刘剑针对跳汰选煤过程受多种非线性因素的制约、不能稳定精煤灰分这一现象，利用模糊优选 BP 神经网络从非线性角度对跳汰选煤及其影响因素进行了理论分析，并使用老虎台选煤厂现场数据对智能网络进行训练、学习和预测精煤快灰的稳定性^[12]。

太原理工大学机械电子研究所利用 BP 回归神经网络对跳汰风阀控制进行了实验室设计和试验，控制结果显示出人工神经网络可以满足现场对风阀的控制要求。

三、支持向量机

人工神经网络这种方法利用已知样本建立非线性模型，克服了传统参数估计方法的困难。尽管神经网络在理论上还有很多根本性的问题没有解决，但在实际中得到了较为广泛的应用。然而，这种方法缺乏一种统一的数学理论，在理论上缺乏实质性进展，与神经网络相反，统计学习理论（Statistical Learning Theory 或 SLT）是建立在坚实的数学基础之上的^[81]，具有较完整的理论体系，所以统计学习理论开始受到越来越广泛的重视。

Vladimir N.Vapnik 等人早在 20 世纪 60 年代就开始研究小样本情况下的机器学习问题。当时这方面的研究尚不十分完善，且数学上难于理解和接受，直到 90 年代以前还没有能够提出将其理论付诸实现的方法。加之当时正处于其他学习方法飞速发展的时期，因此这方面的研究一直没有得到足够的重视。90 年代中期，小样本情况下的机器学习理论研究逐渐成熟起来，形成较完善的理论体系——统计学习理论。而同时，神经网络等新兴的机器学习方法的研究则遇到了很多困难，在这种情况下，试图从本质上研究机器学习问题的统计学习理论逐步得到重视。

统计学习理论是建立在一套较坚实的理论基础之上的，为解决有限样本学习问题提供了一个统一的框架。它能将很多现有方法纳入其中，有望帮助解决许多原来难以解决的问题（比如神经网络结构选择问题、局部极小点问题等）；同时，在这一理论基础上发展了一种新的通用学习方法—支持向量机（Support Vector Machine 或SVM），已初步表现出很多优于已有方法的性能。

支持向量机是根据 Vapnik 提出的结构风险最小化原则来提高学习机泛化能力的方法，即由有限训练样本取得的决策规则对独立的测试集仍能够取得小误差的一种方法。支持向量机算法是一个凸二次优化问题，能够保证找到的极值解就是全局最优解。这些特点表明支持向量机是一种优秀的学习算法。支持向量机本质上是一种非线性数据处理工具，在模式识别、函数逼近、函数拟合、回归估计、数据挖掘、遥感图像分析和非线性系统控制中均有很好的应用。

鉴于所研究的跳汰系统问题取得大量样本的困难，本文在小样本情况下利用SVM 对跳汰分层进行分析探索。

四、基于模型的预测控制

1. 现代控制理论的不足

尽管以状态空间法为基础的现代控制理论解决了反馈控制理论所不能解决的复杂问题，尤其在航空航天领域取得了令人瞩目的成就，同时对自动控制技术的发展起到了积极的推动作用。但随着科学技术的迅速发展，对大型、复杂和不确定性系统例如跳汰机实行自动控制的要求不断提高，使得现代控制理论的局限性日益明显，出现这些问题的主要原因是：

(1) 现代控制理论过分依赖于被控对象的精确数学模型，而在现实工业过程中，往往很难建立精确的数学模型，即使一些被控对象能够建立起数学模型，但因其结构十

分复杂，难于设计和实现有效的控制。

(2) 由于生产环境的改变和外部扰动的影响，实际工业过程经常伴随着非线性、时变性和不确定性。在这种情况下，按理想模型设计的所谓最优控制系统只不过是数学意义上的最优罢了，而对实际工业过程而言失去了最优性，更有甚者会导致控制品质严重下降而无法正常生产。

(3) 在现代化复杂工业过程中，为了取得良好的经济效益和优良的调节品质，往往要求在多变量、多目标和有约束的情况下设计相应的控制系统，而以状态空间法为基础的最优控制难以满足这一要求。

这些来自工业实际的因素和要求，使得现代控制理论无法在工业过程中得到有效地应用。其结果是在工业过程领域，应用现代控制理论设计的过程控制器的控制效果，往往还不如按经典控制理论设计的 PID 调节器好，跳汰系统的控制就是一个非常好的例子。因此，迫切需要一种能解决实际问题的理论出现。

基于现代控制理论的不足，人们开始探索寻求新的方法来解决遇到的问题。20世纪 70 年代，美国和法国几家公司从管理学得到启示，融合滚动优化思想和反馈校正的策略提出了基于模型的预测控制技术，一经问世，就引起学术界的重视，不少研究人员开始研究发展这个理论，并且在石油、电力和航空等领域取得十分成功的应用，随后又出现了各种其它相近的算法。虽然这些算法的表达形式和控制方案各不相同，但基本思想非常类似，即以预测模型为基础，采用在线滚动优化性能指标和反馈校正的策略，来克服受控对象建模误差和不确定因素的影响，有效弥补了现代控制理论对复杂系统所无法避免的不足之处。

2. 预测控制的特点

预测控制发展至今主要算法有：Mehra(1979) 的模型算法控制(MAC)^[43, 44]、Cutter(1980) 的动态矩阵控制(DMC)、Clarke(1987) 的广义预测控制(GPC)等算法^[45]。虽然有不同的表示形式，但归纳起来，它的任何算法形式不外乎包括：①预测模型，②参考轨迹，③在线校正，④目标函数作为性能指标，⑤在线滚动优化等五个方面并且具有以下几个特征：

(一) 预测模型的多样性

在预测控制中，注重的是模型功能，而不是结构形式，过程的描述可以通过简单的实验获得，不需要深入了解过程的内部机理。因此，预测控制算法改变了现代控制

理论对模型结构较严格的要求，更着眼于根据功能要求，按最方便途径建立多样性模型。这一点对于诸如跳汰机这种理论严重落后于实践的系统尤为适用。

（二）滚动优化的时变性

预测控制采用在有限时域内的滚动优化策略，即在线反复进行优化计算这种时变性，虽然在每一时刻只能得到全局的次优解，然而却能使模型失配、畸变、干扰等引起的不确定性及时得到补偿，从而得到较好的动态控制性能。在跳汰过程中，这种时变性主要表现在大批量、散状的入洗原煤性质快速多变上。

（三）在线校正的鲁棒性

在预测控制中，把系统输出的动态预估问题分为预测模型的输出预测和基于偏差的预测校正两部分，由于预测模型只是对对象动态特性的粗略描述，而实际系统中通常存在非线性、时变性、模型失配与随机干扰等因素，因此，预测模型不可能与实际对象完全相符，预测模型的输出与实际系统输出间必然存在偏差，采用这种偏差进行在线校正，使系统构成具有负反馈环节的系统，则能广泛用在各种控制系统中，从而提高了预测控制系统的鲁棒性。对跳汰过程控制而言，即表现在精煤灰分的稳定性以及错配物含量的最小方面。

3. 预测控制的研究动向

预测控制早期研究的成果主要体现在理论分析、算法的改进与推广以及算法的比较分析几方面。理论分析包括解释预测控制的算法机理；推导预测控制系统的闭环传递函数；分析参数对其稳定性、鲁棒性的影响及其定量分析等。近几年来，预测控制的研究已经突破早期研究的框架，摆脱了单调的算法研究模式，从而开始了与极点配置、自适应控制、鲁棒控制、精确线性化、解耦控制和非线性控制相结合的一类先进预测控制策略的研究。随着智能控制技术的发展，预测控制也向着智能预测控制方向发展，并将人工智能、大系统递阶原理等引入预测控制，构成多层智能预测控制模式，由此，进一步增强了预测控制处理复杂对象、复杂任务的能力，拓展了应用领域。智能预测控制弥补了单纯预测控制算法在性能上精度不高、仅适用于线性系统、缺乏自学习、自组织、鲁棒性不强的缺陷，达到了日益提高的控制性能要求。这类控制主要有模糊预测控制和神经网络预测控制。模糊预测控制最早由日本学者安信等提出，并成功地应用于地铁的列车运行控制上。近几年来我国学者在这方面也取得了不少研究成果，如将模糊辨识与广义预测控制相结合进行在线修正的模糊预测控制算法。随着

神经网络发展成熟，预测控制与神经网络的结合日益受到人们的青睐，逐步广泛应用于各种控制系统中。

以跳汰床层密度在线探测为基础，基于跳汰床层动态辨识的预测控制研究目前仍是空白，从致力于测试误差最小化、进一步提高跳汰系统洗选效率出发，进行预测控制在跳汰过程控制中的应用研究是非常必要的。

1.3.3 跳汰机智能控制研究的必要性和可行性

从前述分析看出，煤炭是我国的基础能源和重要原料，在国民经济中占有重要的战略地位。未来几十年内，煤炭依然是我国的主要能源，以煤炭为主的能源结构将难以改变。然而，随着我国经济的高速增长，能源趋紧的信号时有出现，加大入选原煤比例、提高选煤效率成为可持续发展的关键所在。而选煤的基础工作则在于如何在现有工艺流程和生产设备的条件下，发展在线信息获取技术并利用过程信息，以经济效益为目标，实时在线确定生产过程参数，使整个生产过程实现自动运行并处于尽可能优化的状态，切实提高跳汰选煤的分选效果成为目前跳汰选煤过程自动控制所要解决的主要问题。

传统控制理论的思想是建立在精确数学模型基础上的，然而对像跳汰机这样的在实际应用中一般依靠经验进行操作的非线性、时变性、不确定性和不完全性的系统，迄今为止无法获得精确的数学模型。一般来讲，这类生产过程的自动化控制存在相当的难度，原因是没有过程模型，传统的优化控制策略就缺乏基础，许多工艺过程参数无法优化获取，只能靠操作人员根据经验设定，难免造成生产过程的不良波动和产品指标的恶化，这一点在跳汰选煤行业尤为突出。

在跳汰选煤过程中，由于分层机理复杂，影响因素众多，有很多变量同时在起作用，而且互相耦合，互相影响；实际生产过程中有关物料分层的问题无法用现有的跳汰理论来圆满解释；分层状态难以通过合适的检测手段来得到。同时对运行状态和过程信息了解甚少，环境干扰不确定等原因，生产工艺操作参数（如跳汰频率、跳汰周期）与系统所需达到的目标（如分层状态）之间难以用传统的机理模型描述。所以基于精确数学模型的经典控制方法对跳汰机分层过程无能为力^[23, 25]。

但是人的智能对复杂现象的分析和判断能力是很强的，对于许多复杂的难以用自动控制系统实现的生产过程，在有经验的操作人员的操作下可以得到较满意的控制效果。

跳汰生产过程中的风阀操作便是如此，大多数情况下，具有丰富跳汰生产经验的操作人员根据自己的经验能够使跳汰机达到较好的分选效果。这充分说明跳汰过程需要将人的知识和智能行为融入控制系统。而现代计算机控制技术的发展也为其实现提供了有效的工具。计算机在处理逻辑运算，模糊信息，模式识别，知识与经验的积累方面，完全可以取代人的操作。智能控制系统正是利用计算机实现的具有人工智能的控制系统。因此，利用智能控制技术对跳汰过程进行控制是值得探索的，用智能控制方法去解决一直困扰跳汰选煤自动化的问题有望得到稳定高效的控制效果。

未来几十年，煤炭仍将是我国的主要能源和重要的战略物资，具有不可替代性，煤炭工业在国民经济中的基础地位，将是长期的和稳固的。煤炭进一步转化成为新型洁净能源，减少开发使用过程中所造成的污染正在成为全社会关注的焦点。

选煤技术作为洁净煤技术的第一个环节，是洁净煤技术中最成熟、投资最少、成本最低的技术，在我国原煤入洗比例要由目前的32%提高到50%以上、原煤入洗量提高到2010年的13亿吨甚至更多的过程中，作为主要选煤方法的跳汰选煤技术，在理论滞后、控制方法原始的大环境下，利用智能控制技术改善跳汰分选效果条件成熟、意义重大。

1.4 本文结构及技术路线

全文共分五章，结构如图 1—2 所示。从与发达国家的差距及国民经济发展等角度看，中国的煤炭洗选将要得到很快的发展，跳汰仍占据很大的比重；而鉴于跳汰系统的复杂性，跳汰机分选效率低下的问题一直没有得到根本的解决。在分析了影响跳汰分选效果诸因素、控制现状以及智能控制技术的发展后，本文立足于研创的 γ 射线床层检测技术以及在此基础上的工业应用系统，以减少矸石带煤为直接控制目标的思路，在现场试验的基础上，采用 ANN 以及 SVM 等方法对跳汰过程进行仿真，并以此为基础进行基于神经网络的预测控制研究。

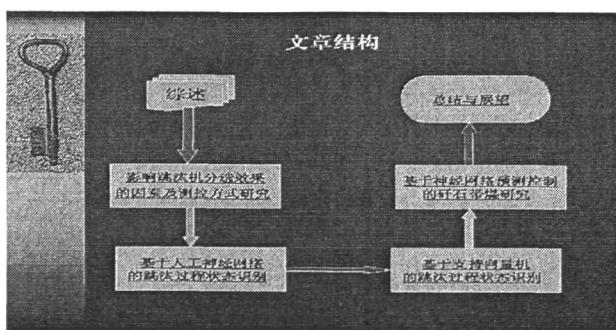


图 1—2 文章结构 figure 1—2 structure of the thesis

第二章 影响跳汰机分选效果的因素及测控方式研究

众多的矿粒混合物，在垂直升降的变速介质流中，按密度差异进行分层和分离的过程，即所谓跳汰选矿过程。跳汰选矿过程中，所使用的介质可以是水，也可以是空气。以水作为分选介质时，称为水介质跳汰或水力跳汰；若以空气为分选介质，则为风力跳汰。

当今，国内外选煤或选矿的工业生产中，水介质跳汰的应用最为广泛，相比之下，风力跳汰应用很少。从基本原理来看，风力跳汰与水力跳汰相似，突出的区别是由于空气易作单向运动，物料分选只能依靠在间断的上冲气流作用下来完成。

水力跳汰机类型很多，根据设备结构和水流运动方式不同，大致可分为五种：(1)活塞跳汰机；(2)隔膜跳汰机；(3)空气脉动跳汰机；(4)水力脉动跳汰机；(5)动筛跳汰机。

目前，国内外广泛采用的是空气脉动跳汰机。空气脉动跳汰机中的水流垂直交变运动是借助压缩空气进行的。按跳汰机空气室的位置的不同，可分为两类：筛侧空气室跳汰机和筛下空气室跳汰机。筛侧空气室跳汰机因空气室设在一侧，如增大跳汰面积将造成沿跳汰室沿宽度各点受力不均影响分选效果。而筛下空气室跳汰机的空气室安设在跳汰机机体内筛板的下方，这样可以克服筛侧跳汰机的不足，在占地面积相同的情况下，处理量提高，分选精度好。目前绝大多数的选煤厂采用的都是筛下空气室跳汰机，大型跳汰机更是如此。筛侧空气室跳汰机一般只用于少数跳汰室宽度不超过2.5m的小型跳汰机上。本文研究以国内广泛采用的筛下空气室跳汰机为主。

按照国标规定，跳汰机的分选效果主要用数量效率或错配物含量来表示。

数量效率是指：一定量的某种原煤经跳汰机分选，在生产出合格精煤的前提下，实际精煤产率与理论精煤产率之比的百分数。这里，一定量的原煤指的是跳汰机在运行一段时间后所处理的原煤数量，表明了对跳汰机的分选效果的考核只能在分选完成后进行，即为选后考核，到目前为止，还没有在线考核的规定，同时也是煤炭洗选行业的特殊性质决定的，即大批量、散状、物理性质多变的物料，在线考核以数量效率为代表的分选效果是不实际的；实际精煤产率即为精煤产量与入洗原煤量之比的百分数；而理论产率则指的是入洗原煤中小于分选密度的比例，测定方法为：用规定密度级的密度液对入洗原煤代表性样品进行浮沉试验。

错配物含量是指：跳汰机各产品中不符合各自分选密度的部分占各自总量的百分数。由于精煤产品的灰分即能代表其中包含的中煤甚至矸石含量并且精煤灰分为产品合格与否的标志性指标，因此，错配物含量主要指矸石中小于矸石分选密度的含量和中煤中小于中煤分选密度的含量，也即“矸石带煤”和“中煤带煤”。

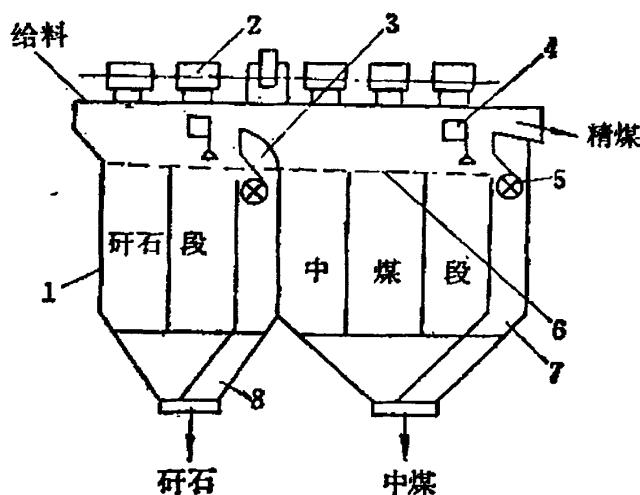
错配物含量或者“矸石带煤”和“中煤带煤”这项指标与数量效率是一一对应的，但是却比数量效率易于考核并且非常直观，更是现场日常关注的指标，在控制系统中也比较方便地检测，因此，本次研究对分选效果的考核即以“矸石带煤”和“中煤带煤”为目标。

影响跳汰机分选效果的因素众多，但归结起来主要为三个方面，即跳汰机的结构特征、原煤物理性质以及操作或控制水平。

2.1 跳汰机的结构特征

即使对同一种原煤、同样的操作制度，跳汰机结构不同，直接影响着分选效果，而不管是否采用了智能控制抑或智能控制的作用如何。

众多中外制造商、研究机构，在跳汰机的结构完善方面进行了不懈的努力，形成了各具特色的跳汰机类型。但尽管跳汰机的种类很多，其结构的主要部分都包括：机体、筛板、风阀、溢流堰、排料装置、检测控制装置等，以图 2—1 为例说明。



1-机体；2-风阀；3-溢流堰；4-浮标传感器；5-排料装置；6-筛板；7,8-矸石中煤排料道

图 2—1 跳汰机结构图

figure 2—1 jig structure

2.1.1 跳汰机的机体

跳汰机机体是由厚钢板焊接而成，沿跳汰机纵向可分为单段、两段或多段，一般为两段。在每段又分成2个或3个隔室，每个隔室都设有风阀和筛下顶水管。在顺煤流方向各段的末端均设有排除重产物的闸门和通道，每段的长度均已系列化并可根据原料性质和产品质量要求选定。

跳汰机的带孔筛板与机体联合构成床层的分层空间，筛板承托床层，筛孔的大小和冲水的大小又控制着透筛排料速度和重产物床层的水平移动速度，因此，要求筛板具有一定的刚性、耐磨性、适当的开孔率和坚固耐用性，并须有一定的倾角和孔形便于比重大的小颗粒物料透筛。筛板合理的倾角和孔形可克服上升水流的阻力，使床层顺利移动又不易堵塞筛孔或便于清理。

2.1.2 跳汰机的空气室

国内外主要跳汰机代表品牌其空气室设置各具特点：

- 1、德国 Batac 跳汰机空气室的位置由原来的在跳汰室两侧各设一个空气室，改为在跳汰室中间设一个空气室（图 2—2），并适当加宽了空气室的宽度并降低了进气口，改善了空气室内液位的稳定性。末煤 Batac 跳汰机筛下物采用了钢管导出的方式，既防止了能量扩散，又可以避免扩散的能量破坏排料的平衡；
- 2、日本的筛下空气室式跳汰机每一个分室内有两个空气室，基本位于跳汰室的中间；
- 3、中国的 X 型跳汰机空气室在跳汰室的中间（图 2—3），空气室的下方设有流线型的导流板，可减少能量扩散和产生涡流；
- 4、中国的 SKT 跳汰机采用了单格室漏斗形组合式机体，各室侧面设一个空气室（图 2—4），洗水呈无干扰 U 形振荡，可减少能量损失；上升水流沿导流板扩散产生一前冲力，有利于输送物料前行，提高处理能力；各室底口收缩成漏斗状通过钢管将透筛物导入排料端，既可阻尼脉动水流向下扩散损失能量，还可大大地减少机体自身重量和容水量。

上述几种机体结构各具特点，但目的都是通过优化设计，使脉动水流平稳、均匀，

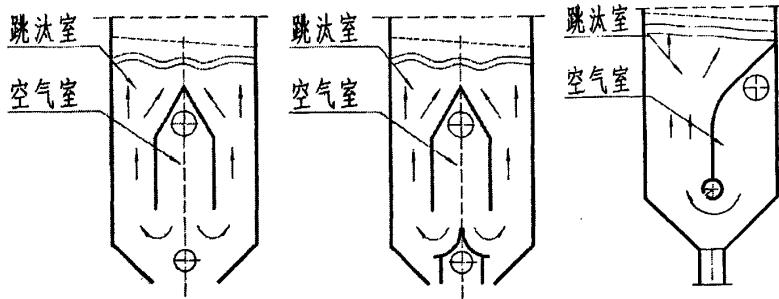


图 2-2 Batac 跳汰机

figure 2-2 Batac jig

图 2-3 X型跳汰机

figure 2-3 X jig

图 2-4 SKT 跳汰机

figure 2-4 SKT jig

以提高物料的整体分层效果，同时减少动力消耗，并减轻机体的体积和重量。

2.1.3 跳汰机的风阀

风阀是跳汰机的重要部件，是跳汰周期的产生源，其功能是周期性地使空气室与压缩空气和大气相连或隔绝，以在跳汰室形成脉动水流。风阀的结构和工作周期对水流在跳汰机中的脉动特性有很大影响，不仅直接影响跳汰机的分层效果，同时对跳汰机的生产能力影响也很大，风阀控制是跳汰机控制的核心之一。

长期以来，围绕着风阀结构及控制方式，专家们进行了大量的研究和改进^[4]。老式跳汰机所采用的电动旋转阀因打开速度慢且难于调整而被数控气动风阀所取代，而数控气动风阀又经历了多次的更新换代。

一、气动盖板阀

德国 Batac 跳汰机在上世纪 70 年代应用了数控气动盖板阀技术，随之我国的一些跳汰机上也采用了这种风阀（图 2—5）。

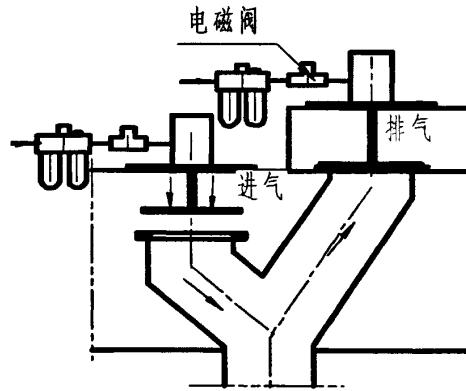


图 2-5 气动盖板阀

figure 2-5 air-operated cover board valve

数控盖板阀结构简单，打开速度快，能够获得良好的工艺效果。其不足之处是盖板上作用有数百公斤的背压，打开风阀必须克服背压，动力消耗大。此外盖板关闭时和管口频繁撞击，易损坏橡胶垫和阀体。

二、气动蝶阀

近年来德国 Batac、澳大利亚 Apic 和我国 SKT 等跳汰机还采用了蝶阀。蝶阀开关省力，但传动机构较复杂，可靠性有待提高，目前尚未大量推广应用。

三、气动滑阀

为克服盖板阀的不足，煤科总院唐山分院在 SKT 跳汰机上采用了数控气动滑阀，滑阀运动无背压、无撞击，不易损坏。但因阀芯和阀套间隙较小，调整不当时洗水进入排气阀会阻塞阀芯灵活运动。为此，SKT 跳汰机又将圆柱形滑阀改进为圆锥形滑阀，当阀芯运动时和阀套产生较大的间隙，可减小摩擦阻力，防止煤泥水阻塞，使用效果明显好于圆柱形滑阀。

四、无背压软接触盖板阀

为解决以上几种风阀的不足，2002 年唐山分院研制成功了无背压软接触盖板阀专利技术。该风阀和原有的盖板阀结构相似，但巧妙地改变了进风方向，将原来阻碍盖板打开的背压转化成了推动盖板打开的动力，缩短了风阀的打开时间并加快了开启速度。该阀还将过去的实心缓冲橡胶垫改成了弧形橡胶圈，盖板关闭时和弧形橡胶圈呈软接触而无硬性撞击，风压的作用恰好促使橡胶圈贴紧盖板免予漏风。这一改进可谓一举多得，它兼备了原有几种风阀的优点，又避免了它们的缺点：打开无背压、关闭无撞击，运动无磨损，省力节能，不易损坏。现在装配该新型风阀的 SKT 跳汰机已大量推广应用，使用效果十分理想。

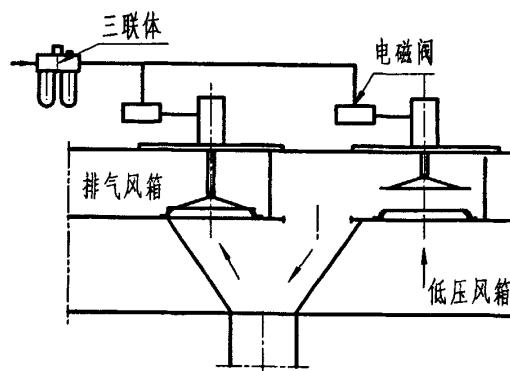


图 2—6 无背压软接触盖板阀

figure 2—6 none back pressure & soft-contact cover board valve

五、多室共用风阀

数控气动风阀能改善分选效果已得到一致认可，但风阀零部件多，系统复杂，维护量较大，此外还需配用一台功率约 40kw 的空压机。为此，唐山分院于 1999 年开发出了多室共用风阀专利技术，将常规两段五室跳汰机的 10 个阀简化为 4 个（每段进、排气阀各一个），同时，将 5 套气源三联体改为一套集中净化加油装置。通过以上改进，使风阀的故障率、易损件数量、维护量、高压风用量均下降了 70% 以上，空压机功率减少近 30kw。现在，中小型的 SKT 跳汰机随机自带一台小型空压机即可使用，用户不再需要专门配备大空压机，既节省了资金，又方便了使用。

六、多频脉动风阀

跳汰机矸石段和中煤段物料的粒度组成、密度组成及可选性差异很大。矸石段物料粒度大，密度高，应采用低频高振幅来提供较大的颗粒换位空间。而中煤段物料粒度小，密度小，需要高频低振幅提高分层效率并避免过度透筛。因此跳汰机矸石段和中煤段应采用不同的风水制度和跳汰制度，以适合各自的分选要求。但以往跳汰机的矸石段和中煤段只能采用同一个频率，不能分别调整。对此，煤科总院唐山分院又开发了多频脉动风阀，它可分别调节矸石段和中煤段的脉动频率和振幅，满足它们各自的分选需要。

七、复振风阀

为了提高分选效果并降低分选下限，各国还把注意力放在了研究复振变波跳汰机上。德国和美国采用的方法相类似，都是利用两套风阀分别作为主振阀和辅振阀，在主振阀产生低频脉动的同时辅振阀高频进风，叠加在主脉动上。因主振、辅振压力不同，故用两个风源分别向两个阀供风；日本采用的是二次进风方式，即在一个脉动周期将一次进风变成二次进风，先用风压 $0.39\text{Kg}/\text{cm}^2$ 进一次风，使床层迅速起振到一定高度，当床层回落时再用风压 $0.24\text{ Kg}/\text{cm}^2$ 进一次风，将正弦波形变成梯形波来延长床层的松散期。我国的煤科总院唐山分院研制的复振跳汰机，采用主、辅两套风阀，共用一个风源，它既可以产生叠加振动，也可实现二次进风。目前装备复振风阀的 SKT 跳汰机已经应用于多家选煤厂。现场反馈信息表明：使用复振风阀后床层分层状况明显改善，提高了精煤产率，降低了矸石及中煤的带煤损失，且有效节省了顶水用量。

2.1.4 跳汰机的排料机构

排料机构也是跳汰机的重要部件之一，其作用是将按密度分层后的物料，准确、及时和连续地排出重产品，并保证床层稳定、保证产品质量和跳汰机的处理能力。性能好的排料机构，可使跳汰机得到较好的分选效果和较高的生产率。跳汰机各段的轻产物是通过水平水流的输送，随水流一起经过溢流堰排出的，各段的重产物，由筛上末端排料机构和透筛排料两部分排出。块煤和末煤的排料方式不同，块煤和不分级原煤的重产物是以筛上末端排料机构排出为主；末煤重产物则以透筛排料为主；煤泥重产物几乎全随透筛物排出。

目前跳汰机的排料机构结构型式多种多样，大体上可以分为末煤、块煤和不分级煤三种排料方式。

德国 Batac 跳汰机针对末煤和块煤采用了两种不同的排料结构。末煤跳汰机采用液压闸门来调节排料口的大小（图 2—7），而块煤跳汰机则是用液压缸调节筛板倾角来调节排料口的大小（图 2—8）。

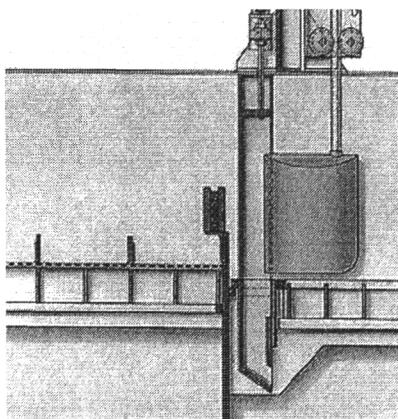


图 2—7 Batac 末煤排料机构

figure 2—7 Batac fine coal discharge framework
discharge framework

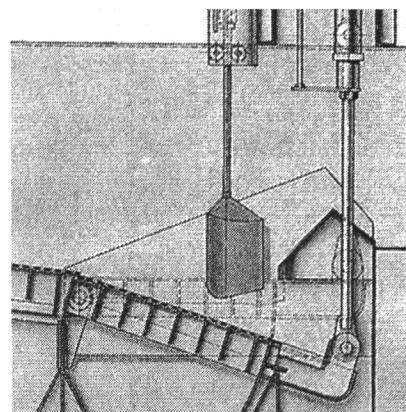


图 2—8 Batac 块煤排料机构

figure 2—7 Batac coarse coal
discharge framework

波兰的 BOSS-2000 型跳汰机，采用了排料闸门和溢流堰互动的排料方式，动力源是伺服马达，调整伺服马达的静态和动态工作参数确定产品的排出量。Apic 跳汰机的排料方式是在溢流堰和闸门内部设置脉动冲孔隔板，使闸门区域到床层末端能够连续不断地脉动，这样闸门在不断运动时，开口也在随着脉动开闭，一是不会被堵塞，二是大粒度的物料也会顺利通过。

中国选煤厂基本都采用不分级混合洗选，跳汰机入料粒度范围大，而且往往筛分

破碎环节把关不严，超粒现象普遍存在，这就要求排料机构必须对宽粒级的物料具有广泛的适应性。

中国目前常用的跳汰机有 LTX 型、LTG 型、X 型、SKT 型等。LTX 和 LTG 型跳汰机是早期设计的机型，均采用溢流堰、直闸门、排料轮和护板组成的排料机构。由于排料轮距下料口太近，排料轮上的物料易受脉动水流作用挤开护板而自行流出，使排料轮难以控制排料量。此外，排料轮放在排料道内容易被大块物料卡住，因此这种结构一直未能得到良好使用。

X 型跳汰机采用液压拖板的排料方式（图 2—9）。

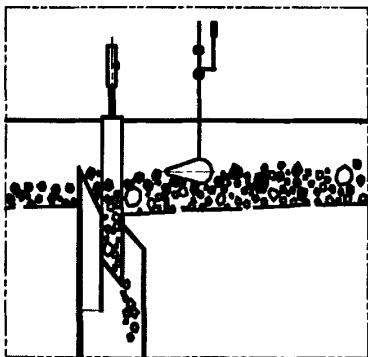


图 2—9 X 型跳汰机排料机构

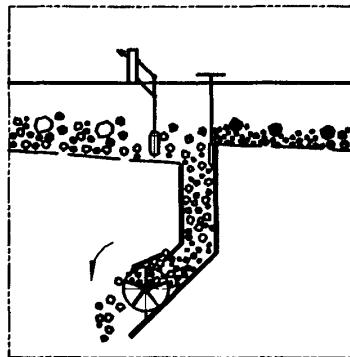


图 2—10 SKT 跳汰机排料机构

figure 2—9 discharge framework of X jig figure 2—10 discharge framework of SKT jig

该结构和 Batac 末煤跳汰机相似，是通过液压缸推拉拖板来调节排料口的大小，以此控制排料量。因末煤粒级窄，排料速度受粒度变化影响小，排料口大小与排料量基本上对应，因此这种结构较适合用于末煤跳汰机。但由于块煤或不分级煤的流动性受粒度变化影响大，流速不均匀，排料口大小与排料量不呈线性对应，故采用拖板排料易造成排料量忽大忽小，不连续、不稳定。

为提高对各种不同煤质的适应性，SKT 跳汰机采用了无溢流堰深仓式稳静排料方式（图 2—10）。这种方式有以下几项优点：

- 1、取消溢流堰可防止分好层的物料撞击和翻越溢流堰造成二次混杂；
- 2、增设活动溢流堤可灵活调节各段底流床层厚度；
- 3、加深料仓可阻止水流在排料道中上下窜动而影响排料稳定；
- 4、排料轮设在排料道底口下方可减小排料轮长度，增大过料断面，使大粒物料不易卡轮。
- 5、排料轮强制性主动排料，其转速可无级调整并和排料量呈线性对应，跳汰机

可以连续、稳定、准确地控制排料量，产品质量易于保障。

综上所述，排料是跳汰机的一个重要环节，跳汰机的排料系统应该具有以下功能：

- (1) 灵敏准确地检测出重产物床层高度；
- (2) 控制和保护好前期的分层状态；
- (3) 保证排料过程的连续进行。

为完成以上功能，跳汰机的自动排料装置应由检测和控制两部分组成。检测部分的任务是检测重产物状况，现在广泛使用的是浮标检测。控制部分的具体要求是：排放速度要与床层的运动速度相适应，排料时不产生瞬时局部塌陷和瞬时排空现象，避免破坏床层的稳定，保证排料过程的连续进行，而且一定要做到排出的重产物是从床层的最底层连续地排出的，以保证排料的纯净。

2.1.5 跳汰机结构的可调参数

跳汰机结构的可调参数主要是筛板倾角，跳汰机筛板倾斜安装的作用在于调整重产物的排放速度和床层按密度分选时间。筛板安装倾角的选择应与床层按密度分层的速度相适应，对于矸石含量多中间煤含量少的可选性好的原料煤，特别是黄铁矿较多的原料煤，床层按密度分层的速度快。矸石段筛板倾角应大些，以加快床层水平移动速度。因为在矸石含量多的情况下，筛板倾角小，床层下部的矸石层运动速度慢，势必影响跳汰机的生产能力。但是，对矸石含量少中间煤含量多的可选性差的原料煤，筛板倾角就应小一些，因为倾角太大，床层不均匀，增加透筛损失，另外倾角大，床层水平移动速度快，缩短了分选时间，影响分选效果，甚至造成排料口堆积。前者协调量，后者协调质，两者都不容忽视。所以，筛板倾角应视原料煤性质而定，而表征筛板倾角合适的特征是床层厚度均匀，这可在跳汰机正常生产时满载停机进行测试。

2.2 原煤性质

入洗原料煤性质不同是影响跳汰机分选效果的一个重要方面，必须在控制系统中予以考虑。原煤性质一般包括以下几个方面：

- (1) 煤的可选性

可选性好的煤在分选过程中，颗粒按密度分层的过程进行得明显且较快，因而分选指标好，生产能力也较大；对可选性差的煤，虽然延长分选时间，颗粒也不能很好

地按密度分离，产品中混入非规定密度物（错配物）较多。

（2）原煤的粒度组成

原煤的粒度组成决定跳汰机筛板上床层的水力阻力，决定水流下降时细小颗粒通过大颗粒间的空隙而落入筛板下的条件，同时也决定上升水流将细小的沉矸冲到床层上部与浮煤一起从溢流口排出的可能性。在跳汰机内颗粒按密度分层的过程中，较大颗粒进行的快且明显，因而块煤跳汰机有较大的单位生产能力。块煤跳汰机的分选指标一般比末煤跳汰机好。

（3）煤的机械强度

煤及其伴生矿物质的机械强度对于跳汰分选有一定的影响，当分选含有坚硬而不易浸润的硬煤时，其分选效果较好。反之，当分选含有较多易泥化矿物的煤时，容易在分选过程中产生许多高灰分细泥，使洗水浓度增加，以致精煤受到污染，灰分增高，产品脱水困难，分选效果变坏。

（4）颗粒形状

颗粒的形状对分选同样有影响，特别是扁平形状。扁平的矸石颗粒在横向下降时，会受到上冲水流的反向作用力而浮至床层上部，与精煤一起排出，从而增加了精煤的灰分，在铺有长石床层的跳汰机内，扁平的矸石及中煤不易通过筛孔。

2.3 跳汰机的操作变量

2.3.1 跳汰过程

在跳汰机中水流运动包括两部分：垂直升降的变速脉动水流和水平水流。前者对颗粒按密度分层起主要作用，后者对颗粒按密度分层也有影响，但主要是起润湿和运输的作用；润湿是为了防止物料进入水中后结团，运输是负责将分层之后居于上层的低密度物料冲带走，使其越过溢流堰排出机外。

跳汰机筛板承托的全部物料即为跳汰床层，由被选物料给入跳汰机内落到筛板上形成。在给料的同时，从跳汰机下部周期性给入上下交变的水流，此垂直变速水流透过筛孔进入床层，促使物料按照密度差异在这种脉动水流中完成分层。

当水流上升时，床层被冲起，呈松散及悬浮的状态。此时，床层中的矿粒，按其自身的特性（密度、粒度和形状），彼此作相对运动，开始进行分层。在水流已停止上升，但还没有转为下降水流之前，由于惯性力的作用，矿粒仍在运动，床层继续松

散、分层；水流转为下降时，床层逐渐紧密，但分层仍在继续；当全部矿粒落回筛面，彼此之间已丧失相对运动的可能时，则分层作用基本停止。此时，只有那些密度较高、粒度很细的矿粒，穿过床层中大块物料的间隙，仍在向下运动，这种行为也可看成是分层现象的继续。下降水流结束时，床层完全紧密，分层便暂告终止。水流每完成一次周期性变化所用的时间即为跳汰周期。在一个跳汰周期内，床层经历了从紧密到松散分层、到再紧密的过程，床层内的各种颗粒得到了分选。

在经历多个跳汰周期之后，分层逐渐完善，床层从上到下密度由低到高排列，低密度物料（精煤）集中在床层上部，高密度物料（矸石）集中在床层下部，二者之间为中煤。

高密度的矸石从矸石排料口排出，其余的物料随横向水流进入中煤段，并在中煤段将其分成中煤和精煤，中煤从中煤排料口排出，精煤则从跳汰机末端的精煤溢流堰排出，从而得到跳汰洗选的三种产物。

跳汰过程流程可用下图 2—11 表示：

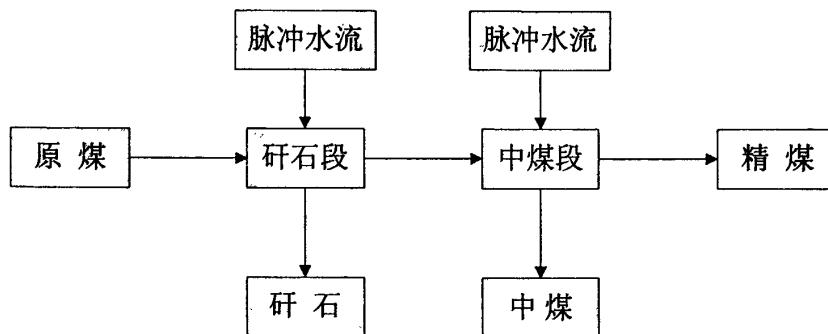


图 2—11 跳汰过程流程图

figure 2—11 flow chart of jig process

因此，物料在跳汰过程中之所以能分层，起主要作用的内因，是矿粒自身的性质，但能让分层得以实现的客观条件，则是垂直升降的交变水流。

2.3.2 跳汰周期分析

为了研究颗粒在跳汰床层中的分层过程，可以从一个跳汰周期内脉动水流的运动过程分析入手。

一个脉动周期分为上升初、上升末、下降初、下降末四个阶段。脉动水流的上升初期包括上升水速从零升至最大时整个上升过程的加速阶段。在上升过程中，跳汰床层具有最大的上举力。由于颗粒自身所受压力的不同，这时低于分选密度的颗粒，包括粒度大的低密度颗粒，从上升初期就开始持续地上升，而密度高的重颗粒的上升过

程比轻密度颗粒晚一些，同时上升速度也小一些。在此过程中，应尽可能保证高密度颗粒在筛板上产生一些位移，这样有利于床层的松散及重颗粒在筛面的运动和排放，设定的关键即为进气期的长短。

脉动水流的上升末期，即上升水速由最大逐渐减小至零的阶段。在此过程中，脉动水流的上升力逐渐减小，水流的加速度方向也由向上转为向下；床层中颗粒的上升速度也逐渐降低，密度和粒度相对较大的颗粒也会从上升转为下降，此阶段床层的底部也将会回到紧密状态，但此时一部分粒度和密度相对较小的颗粒仍处于上升状态，较上层的床层仍会处于松散状态。

脉动水流的下降初期，即水流的加速下降阶段，在此过程中，水流速度和水流加速度的方向都向下，所有颗粒都先后转为下降，床层的各部分都会逐渐回到紧密状态。

脉动水流的下降末期，即水流的减速下降阶段，在此过程中，床层大都仍处在紧密状态。虽然粗颗粒已不可能随水流继续向下运动，但由于吸啜作用的存在，细颗粒仍可穿过床层中的缝隙继续向下运动。

在跳汰过程中床层的松散是由于在脉动水流作用下床层中性质不同的颗粒运动状况间的差异引起的，颗粒在床层中运动的基本方程如式(2-1)所示：

$$\frac{dv}{dt} = \frac{1}{2\rho_s + \rho} [2(\rho_s - \rho)g + \frac{3\rho}{\varepsilon} \times a \pm \frac{1.34\mu^2}{d\varepsilon^{4.6}}] \quad (2-1)$$

其中： v ——颗粒运动速度； ρ ——介质的密度；

ρ_s ——颗粒的密度； ε ——床层的松散度；

μ ——水流运动速度； a ——水流运动加速度；

g ——重力加速度； d ——颗粒粒度。

其中床层松散度定义为整个床层中孔隙体积占床层体积的百分数。床层松散是水流脉动作用的结果，理论上，脉动水流特性（如速度、加速度）直接制约着床层的松散状况。

从式(2-1)的颗粒运动方程可以看出，使床层向上运动的作用力是水流上升速度以及上升期的加速度或下降期的减加速度产生的上举力。床层的松散与颗粒的上升运动就是靠这两种作用力。床层何时开始松散由上升水速与方向向上的水流加速度（在水流的下降期则是负加速度）两者的综合效应所决定。在一般情况下，床层开始松散的时间发生在脉动水流的上升初期，因为这段时期水速与加速度都是向上，在上升末期与下降初期床层的松散度达到最大，随着上升水速的减小，特别是下降水速的不断增加，床层逐渐恢复紧密。

对上述影响分选的四个阶段的分析后可以看出，上升水流是影响分层的重要因素。由于水流加速度对颗粒运动的影响只与颗粒的密度有关，与颗粒的粒度无关；而水速对颗粒运动的影响则不但与颗粒的密度有关，也与颗粒的粒度有关，因此在上升

初期,应尽量利用水流加速度对分层产生的作用。在水流脉动的上升初期与下降末期,脉动水流的加速度方向向上,在上升期它可以使低密度颗粒获得比高密度颗粒更快的上升运动;在下降期则可以使低密度颗粒获得比高密度颗粒更大的减速下降运动。因此在这两个阶段希望脉动水流具有较大的加速度。反之,在脉动水流的上升末期与下降初期,水流的加速度将会促使低密度颗粒比高密度颗粒更快的下降,对按密度分层不利。所以希望在这两个阶段脉动水流具有较小的加速度。所以比较理想的脉动水流速度波形应是两头陡(有较大加速度),中段平缓(有较小加速度)的不对称波形。

水流速度对完全按密度分层是不利的,它对分层的影响具体表现在:在上升期,水速会使细颗粒上升得比粗粒更快,可能会使一部分高密度细粒污染到上层中去;但在下降期,水速会使细颗粒下降得比粗粒快,特别是当床层处于较紧密状态时,吸啜分选作用将有利于将上升期被冲到上层去的一些高密度细颗粒重新回落到下层来,这样可在一定程度上克服由水流速度带来的负面影响。

从上面可以看出脉动水流直接影响着跳汰床层的状态,而脉动水流直接受风阀周期的控制。所以要得到好的跳汰床层的分层状态,就要合理地控制风阀周期。

在跳汰过程中,跳汰周期特性的基本形式有三种,即间断上升介质流、间断下降介质流、呈升降突变介质流。跳汰周期的特性在一定程度上决定了跳汰分层的效果,并间接地体现了跳汰机的主要结构特征。但在结构一定或者结构改进已经日臻完善的情况下,跳汰机的操作变量便是影响分选效果的决定性因素并且始终将是研究的重点。

2.3.3 跳汰机的操作变量

跳汰机的控制系统必须在分析清楚对分选效果的影响因素后再进行设计,这些因素可以归纳为如下几个方面。

(1) 跳汰频率

跳汰频率的合理选择能解决跳汰机分离效果与单机生产能力之间的矛盾。实践证明:低频率大振幅,床层松散度较大,分层较快,单机生产能力较大,但按密度分选所受颗粒的粒度、形状以及水流速度等影响较大,对风、水、给料等变化较敏感,操作较困难,所以只适应于块煤或易选煤的分选;而高频率小振幅跳汰机工作较稳定,按密度分选所受上述影响因素较小,故操作容易,产品品质较好,但因松散度减小,分层速度减慢,单机生产能力亦降低,故适应于末煤和较难选煤。

(2) 风阀周期

在一个跳汰周期内，风阀的进气期、膨胀期和排气期的作用是：进气期使床层产生位移，膨胀期使床层松散分层，排气期使细物料透筛。因此，风阀进气期、膨胀期和排气期的长短（或面积）应根据入选物料的性质来选择，选择不当，必将影响分选产品的品质和产率。在一个跳汰周期内，加长其一，必缩短其二，过长过短都不利于床层分选。风阀周期是控制系统的主要任务之一，控制系统必须能适应上述变化。

（3）给煤量

入选煤的数、质量能够稳定在适当范围内，是跳汰选煤获得良好分选效果的因素之一。当其他操作因素一定时，给煤量过大，床层变得致密，破坏了按密度分层的条件；反之，给煤量过小，床层变空容易被搅乱，同样破坏了跳汰过程，并且还降低跳汰机的处理能力。因此，选煤司机或一个控制系统，必须严格地保证跳汰机的合适的处理能力，不使入选煤的供给量有较大的变动。通常情况下，当原煤性质变动时，一般是以调整跳汰机的其它因素来保证要求的分选精度，从而尽可能稳定给煤量。

（4）产物的排放量

产物排放应及时、合理，特别是重产物的排放速度应与分层速度及高密度床层的移动速度相适应。必须保证床层的层次，应尽量保持排料口垂直闸门的开口高度或尽量保持排料轮的旋转速度，稳定矸石层和中煤层在筛板上的固定厚度。排放过多，会增加矸石（或中煤）中的带煤量，并使床层过薄、紊乱，不利于分选；排放过少会使重产物床层加厚，在同样风、水、给料量的条件下，床层跳不起来、振幅变小、颗粒换位空间不够、分层效果变差并导致一段矸石排不完，二段精煤受污染，分选效果同样不好。排料闸门不可不调，也不可不按实际乱调，操作或控制时，应根据原料煤中矸石和中间煤含量的变化，对排料闸门开口高度进行适当的调整。

（5）风、水量

风是松散床层的主要因素，是跳汰选煤的动力源。风量过大，床层易混乱，浮煤损失大，回收率低，床层移动速度快，分选时间缩短；风量过小，床层紧密，细颗粒透筛效果差，床层移动速度慢。用水量也应适宜，水量过大，床层显得无根，在水流下降期，颗粒分层受粒度影响较大，细矸不易透筛排出；由于水速快，以致分选时间短，分选效果差，还会给后续煤泥水处理带来不便。水量过小，床层过于紧密，颗粒间无置换空间，床层整起整落且移动速度慢，造成分选效果差，生产能力降低。

（6）床层厚度

跳汰机出厂时的结构特性基本决定了床层的厚度，但仍可具有可调节性，可以从可调整溢流堰的高度、原煤给料量的大小、风水大小、排料机构负荷大小等几方面来调控。

从床层的松散条件看，床层的厚度应该尽可能地小些，因为，床层过厚，在风压和风量不足时，上冲水流难以保证颗粒上升到必要的高度，也就难以保证床层具有必要的松散度，以致破坏了颗粒按密度分选的过程。但床层过薄，吸啜作用过强，轻产物中透筛损失增加，物料压不住床层还会导致床层紊乱，不易分层。这个合适高度取决于在整个筛面上形成清晰而稳定的分离床层，它与原煤的性质、筛板的宽度、筛孔尺寸、频率以及振幅是相互制约的。

原料煤中末矸多，床层可以薄些，使其易透筛排出；原料煤中块矸多，重产物层的孔隙度就大，为避免低密度物透筛损失，床层应该厚些。原煤可选性对床层厚度的影响是：中间煤含量多时，床层就应该厚些，因为，中间煤含量多，精煤被污染的程度严重，如果溢流堰低则床层薄，振幅势必增大，易使部分中间煤混入精煤中。

筛板宽度和筛孔尺寸对床层厚度的影响是：筛面越宽，沿宽度的脉动水流和给料越难均匀，因此要求床层厚些；筛孔尺寸较大时，为避免物料透筛损失，床层应厚些；反之宜薄些。

床层厚度与频率振幅的关系一般是：薄床层采用高频率、低振幅的跳汰制度，厚床层则应采用低频率高振幅的跳汰制度。

从影响跳汰机分选效果的种种因素分析看，这些因素都构成了需控变量，并且变量之间存在交互作用，在实际的跳汰机自动控制过程中，必须把这些变量进行梳理，抓住主要矛盾，才能完成这一复杂控制系统。

如果把上述每一个变量均纳入控制系统，一是太复杂，尤其是考虑变量之间的交互作用的存在，更使复杂程度加大；二是没有必要，因为在实际生产中，一些因素是相对固定的，例如溢流堰高度、风压、水压等。

从跳汰机分选过程可以看出，整个分选过程可以概括为两部分：分层和排料。床层的分层好坏是分选效果的前提，分好层的物料的准确、及时排出是必要条件，同时排料的好坏又对分层产生影响，而床层的分层好坏主要来源于脉动水流的作用，即风阀周期的合适选择。到目前为止，跳汰机的测控主要对象就是针对上述两方面。

2.4 跳汰机的测控方式

跳汰机的控制可概括分为排料和分层的控制,如果一台跳汰机既实现了分层的自动控制又实现了排料的自动控制,则可称为实现了整机控制。分层的控制是给跳汰机以合理的风水制度,使原煤按照密度较好地分开。风水制度有风阀周期、顶水、跳汰频率几个控制变量,其中顶水和跳汰频率在生产过程中变化不大,主要是风阀周期的控制。跳汰机的排料是跳汰机按比重分选的最后实现,直接决定了洗选产品的质量和洗选效率,排料控制是跳汰机控制的一个重要环节。排料的控制应在分层状态比较好的情况下,根据检测到的物料分层状态具体情况,控制系统给出控制信号控制排料闸门的开度,使物料排出。

2.4.1 传统方式下分层状态检测

对于跳汰过程状态和运行参数的测试技术的研究,国外开始于五十年代末,其中风压和垂直水流位移的测试已有较成功的经验;松散过程的测试经历了一段时间的研究,未有成熟的应用报道;水流脉动速度的测试近几年也有一定程度的进展,已取得了丰富的经验数据,从本质上说,也不成为技术难点;跳汰床层分层状况的检测目前仍然是手工探杆探测和浮标探测共存的状况,可以说停留在了目前的自由浮标水平上,实际上已经限制了跳汰控制水平的提高。

在目前业内应用的跳汰机自控系统中,普遍的方法是利用浮标传感器对跳汰床层的重物料层进行检测,探知重产物床层厚度,以此为依据对排料进行控制,目标是保持这一层厚度的稳定。浮标替代了手工探杆所探测的部分工作,即使巴达克跳汰机也是如此。究其内在结构,浮标是利用重力漂浮原理,模仿同密度的物料在跳汰床层内的运动,浮标传感器比探杆探测的主要优势在于它是连续测量,并且通过机—电转换,将床层重物料层厚度信号通过接口送入计算机,能较为准确地反映重物料层的厚度,完成跳汰机床层重产物厚度自动检测和排放。综合而论,浮标传感器有代替人工探杆探测跳汰重物料层之功绩,但仍仅是对床层分层状况的局部测量,并且受其自身的机械性能影响而会影响其精度,不能提供确切的床层分层信息。

目前,出厂的跳汰机大都随机携带浮标,用以在线测量重产物床层厚度,但有很大的局限性。为使分层效果达到最佳状态、精确控制精煤灰分、提高数量效率、减少

精煤流失，研制可靠的床层检测手段，成为跳汰机自动控制发展的先决条件。

2.4.2 传统方式下跳汰机的控制

目前，跳汰机自动控制的主要内容仍为自动排料，即通过对排料轮转速的实时控制，维持重产物床层厚度的稳定，目的就是达到跳汰机按给定密度进行分选。重产物床层厚度一般采用浮标检测，煤质变化时浮标的比重由人工调节配重，调节的结果由在线精煤灰分仪反映，或由人工采样化验精煤灰分，再通知操作司机，实际上仍是由人工来反复调节配重、摸索合适分选密度的过程。

控制系统一般采用一定值闭环调节：

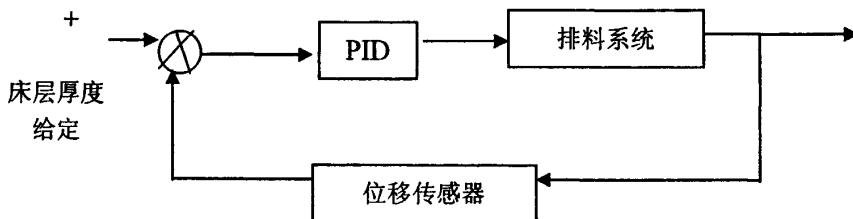


图 2-12 跳汰机 PID 自动排料系统

figure 2-12 jig PID control loop of discharge

控制核心是利用计算机实现 PID 控制，虽然采用的是 PID 的差分方程式，严格讲是一种断续方式，但由于采样周期相对于被控对象的时间常数小得多，因此也可近似于连续控制方式。

2.4.3 γ 射线跳汰床层密度探测器

一、 γ 射线的工业应用

工业部门因多种不同的目的已广泛使用了同位素放射源的各种射线，包括各种过程自动控制与最优化、质量控制等。

在大多数应用中，同位素技术尚不可能被现有可更替的技术所取代，因为它在目前尚是最有效、最经济的选择。

在生产过程控制领域，采用 ^{241}Am 、 ^{137}Cs 和 ^{60}Co 等放射源制成的同位素仪表，如密度计、料位计和核子秤，已经大量用于在线非破坏性测量。同位素示踪剂如 ^{127}Xe 、 $^{99}\text{Tc}^+$ 、 ^{85}Kr 、 ^{203}Hg 、 ^{82}Br 和 I 被用来进行液体流动测量和泄漏探测。

作为工业应用实例，法国在 1995 年总计有约 5000 个行业使用了 10000 台以上的同位素仪表。其涉及的工业部门包括：石油和天然气、采矿和矿物原料的输送、核燃料循环、钢铁、冶金、化学、塑料、民用建筑、宇航、造纸和卷烟工业等。我国近几年来同位素的应用领域也越来越大，在煤炭行业，使用较多的是管道密度计、核子称、灰分仪、料位计等，但同位素测量跳汰床层密度的应用不仅在国内，甚至在国外也未见应用报道。作者参加太原理工大学机电所课题组的 γ 射线床层密度探测器研制全过程，并成功实现了工业应用。

二、 γ 射线密度测量原理

当 γ 射线穿过物质时，由于射线与物质的三种相互作用：光电效应、康普顿效应和电子对效应，其强度受到衰减。对于窄束平行的射线，它遵从如下的指数衰减规律：

$$I = I_0 \exp(-\mu_s \rho d) \quad (2-2)$$

式中， μ_s 为质量吸收系数 (cm^2/g)， ρ 、 d 为被穿透物质的密度 (g/cm^3) 和厚度 (cm)， I_0 、 I 为吸收前后 γ 射线的强度。

$$\mu_s = \tau_s + \sigma_s + X_s \quad (2-3)$$

τ_s 、 σ_s 、 X_s 为光电效应、康普顿效应和电子对效应引起的质量吸收系数，大小与 γ 射线的能量 E_γ 有关。当 $E_\gamma < 0.3 \text{ MeV}$ 时，光电效应起主要作用； $E_\gamma > 1.02 \text{ MeV}$ 时，电子对效应起主要作用；当强度在 $0.3 \text{ MeV} < E_\gamma < 1.2 \text{ MeV}$ 时，康普顿效应占主要地位。在实际过程中，三种作用同时存在，因此对总的质量吸收系数都有贡献。实验发现，对中等原子序数的材料，当强度在 $0.3 \text{ MeV} < E_\gamma < 2 \text{ MeV}$ 时，总质量吸收系数就与化学成份无关（除少量重元素外）。此时穿过吸收体的射线强度仅仅取决于该吸收体的厚度与密度。换言之，选取合适的射线能量，可以使 γ 射线强度的衰减只与 ρ 、 d 有关，其它影响因素可以忽略。

目前工业使用的 γ 射线密度计就是根据这一基本原理而研制的。一定强度和特定能量的 γ 射线，在距离 d 一定的情况下，射线强度的衰减随被测物质的密度而变，

密度大衰减大；反之，衰减则小。对上式取对数，可知物质的密度是射线强度的对数关系。由此通过测量射线强度强弱的变化，可以实现对物质（跳汰床层）密度的连续监测。

三、跳汰机床层内密度的表述

第一，煤的物理性质及其密度组成

煤是成分和结构都很复杂的物质，它可看成由碳质、矿物质和水组成。碳质的平均原子序数 $Z \geq 6$ ，属于低原子序数物质；矿物质的 $Z \geq 12$ ，属于高原子序数物质。密度是煤的主要物理性质之一，凡涉及到煤的体积和质量关系的工作，都使用煤的密度。

第二，真相对密度

煤的真相对密度，是指在 20℃时煤的质量与同温度、同体积（不包括煤的内外表面孔隙）水的质量之比，用 TRD20/20 表示，一般用煤的真密度来表征高变质的无烟煤。

第三，视密度

也叫容积密度或假密度，一般的测定方法有两种，一是掏槽法：在煤层中掏出一个有规则的槽，量出它的体积并称出它的重量，换算出单位体积的重量；二是在实验室利用阿基米德原理，用水或水银通过置换法测定。

跳汰机床层内某一层的密度一般指容积密度即假密度，因为床层在正常工作过程中始终为水所浸泡，煤与煤的缝隙中为水所占据，利用射线在线测量的密度值是该层煤和水的平均密度。

选煤厂生产过程中对各产品的密度检测主要是用浮沉法，即用氯化锌配置一定密度的溶液，将待测的煤或矸石置于其中，捞出浮在上部的即为小于该密度液密度的部分，沉在底部的即为大于该密度液密度的部分。这种方法只能区分各个密度区间的物料量，如-1.3 的量和+1.3 的量等，也是一种近似估计测量，不能确定一个煤样的真、假密度，因为不可能配置无限多密度的密度液供实验，仅可用于指导生产，不可能用于在线控制。

四、跳汰床层密度探测器

跳汰床层密度探测器由辐射源和接收器、前置放大器、信号调理器与整形电路、滤波电路、门控电路等组成，辐射源和接收器（含前置放大器）分别安装在两个密闭

容器中，两者之间间隔一定的距离，置于床层内，位于排料口附近，检测待排料部分。射源被置于铅罐内，铅罐壁厚按国家辐射防护标准由专门的核辐射研究所设计，以满足安全需要，探测器在跳汰机床层内的位置如下示意图。

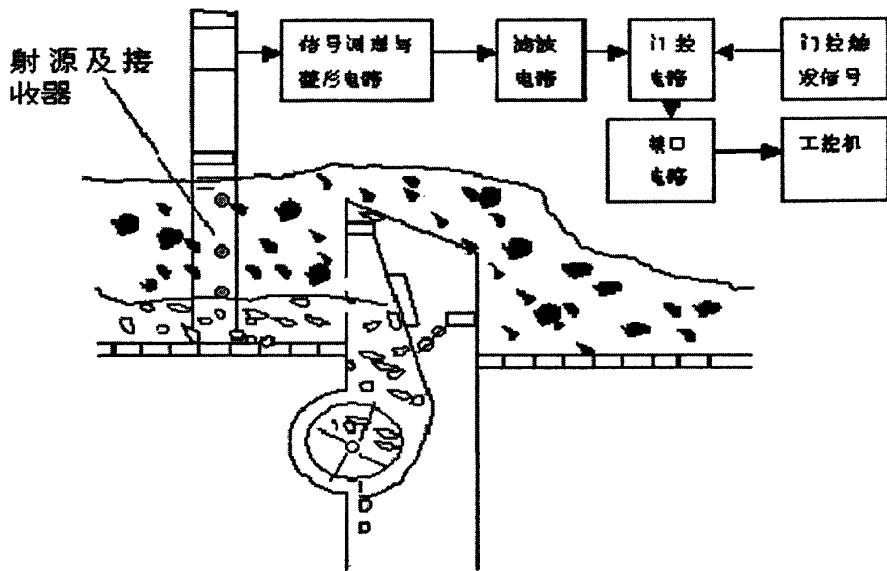


图 2—13 密度探测器示意图

figure 2—13 sketch map of density detector

五、 γ 射线探测器实验系统与试验设计

实验包括两部分：一是实验室实验，二是现场实验。实验室实验主要在太原理工大学机械电子工程研究所进行，目的是用邢台矿现场采集的各密度级煤样，测试放射性射源对煤炭密度的衰减特性，从而确定跳汰床层分层状态测试部分的参数；现场实验在邢台矿选煤厂进行，主要工作是通过现场跳汰床层采样分析，完善工业实际应用的床层分层状态（分层密度）探测系统。

重复性试验结果表明：跳汰机床层密度分布存在乘幂规律，考虑实验误差，规律性很强；而且试验的重复性很好， γ 射线密度探测器可以用有限的几个置于床层不同的层位内，即可完整表述该密度分布；完全可以满足跳汰床层密度检测需要。即使考虑实际操作误差因素，也可以认定：同位素探测器可以测量跳汰机床层的真密度。只要检测到了床层内某点的的密度，不管是散密度还是真密度，即可确定该点的灰分值，这样只要测出某层散密度，即可算出该层内+1.8 的含量，在实际控制系统中，这一点

非常珍贵，因为只要算出下部矸石层散密度，并计算出+1.8含量，即可知道最终产品矸石中的带煤损失，控制住这一指标，可非常有效地控制数量效率。

同样，只要测出中煤段某层散密度，即可算出该层内1.4-1.8的含量，在实际控制系统中，这一点同样非常珍贵，因为只要算出下部中煤层散密度，并计算出1.4-1.8含量，即可知道最终产品中煤中的非中煤所占的比重，即错配物含量。控制住这一指标，可非常有效地控制数量效率。

可以知道，错配物含量的多少就是数量效率的高低，知道了床层内沿高度的密度分布也就等于知道了床层下部待排料的错配物含量，在此基础上在进行风、水、排料的控制，依据就非常强了。

因此，跳汰机仅有的两段：矸石段和中煤段，利用 γ 射线测出床层内各层的散密度，即可十分明确而可靠地确定床层的密度分布状况，并可通过该数据的分析计算，来控制跳汰机的操作，来达到稳定和提高洗选数量效率的目的。

本文通过实验室和现场工业试验取得实时数据，对跳汰过程中床层的分层状况进行分析研究。

2.4.4 现场基于床层密度探测的跳汰机排料系统的模糊控制

跳汰机通过排料来实现最终产品（矸石、中煤、精煤）的分离，在物料分层状态一定的情况下，排料控制的优劣是影响精煤灰分与产率的主要矛盾。排料过程中，稳定床层是一个很重要的控制指标，若因排料造成床层不稳，会严重影响分离精度，因此，排料控制方法研究意义重大。

跳汰机排料控制具有一定的特殊性：传统意义上是要追求床层厚度的稳定，这是一个动态“稳定”量，精确意义上的稳定点是不要求的，因此，基于 γ 射线密度探测器探测床层密度上的排料系统，最能准确反映床层的波动。在控制系统的综合特性中，响应速度是主要矛盾，稳态误差是次要矛盾。因此需要一种合适的控制系统。而模糊控制最大特征是，它能将操作者或专家的控制经验和知识表示成语言变量描述的控制规则，然后用这些规则去控制系统。模糊控制器是一类非线性控制器。模糊控制器特别适用于数学模型未知的、复杂的非线性系统的控制。

在“跳汰过程物料分层状态检测”研究成果的基础上，现场开发了一套模糊控制系统，并在此系统基础上进行必要的试验、数据采集及本文的后续分析研究。

一、模糊控制的基本原理及组成

1) 模糊控制基本原理

最基本的模糊控制系统结构如图 2—14 所示。图中 R 为设定值，代表 γ 射线密度探测器最底部位置处床层平均密度给定值， y 为系统的输出值，代表 γ 射线密度探测器最底部实际测量值，它们都是清晰量。从图中可以看到模糊控制器的输入量是系统的偏差量 e ，它是确定数值的清晰量。通过模糊化处理，用模糊语言变量 E 来描述偏差。模糊推理输出 U 是模糊量，在系统中要实施控制时，模糊量 U 要进行清晰化处理，得到可操作的确定值 μ ，这就是模糊控制器的输出值，通过 μ 的调整控制作用，使偏差尽量小。

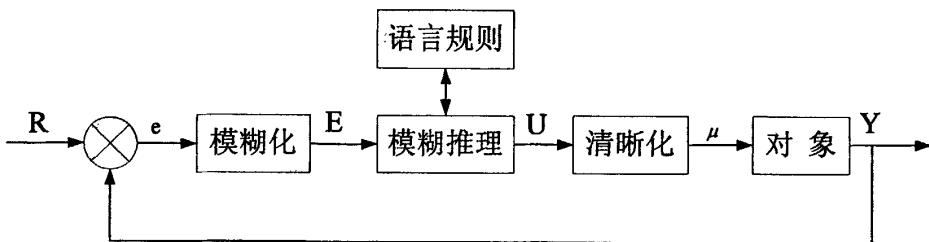


图 2—14 模糊控制器原理图

figure 2—14 schematic diagram of fuzzy controller

2) 模糊控制器的组成

模糊控制器的基本组成如图 2—15 所示。它含有模糊化接口、规则库、模糊推理和清晰化接口。模糊控制的核心是包含语言规则的规则库和模糊推理。模糊推理就是一种模糊变换，它将输入变量模糊集变换为输出变量的模糊集，实现论域的转换。规则库是由若干条语言控制规则组成，它是模糊化、模糊推理和清晰化的语言变量及其赋值的总则。

输入变量是系统实测值与给定值之间的差，输出变量是系统的实时控制修正变量。模糊化接口将清晰量模糊化具有两个作用：(1) 将基本论域转化到语言变量的论域；(2) 求出对应语言变量的隶属度。清晰化接口是将模糊的集合转变为可执行的清晰变量。

一般的模糊控制器可分为单变量模糊控制器和多变量模糊控制器，单变量模糊控制器就是一个或多个输入对应一个输出的模糊控制器，多变量模糊控制器是多输入对应多输出的模糊控制器。本文所研究的排料系统是偏差和偏差率两输入对应一个输出

控制量的单变量模糊控制器。

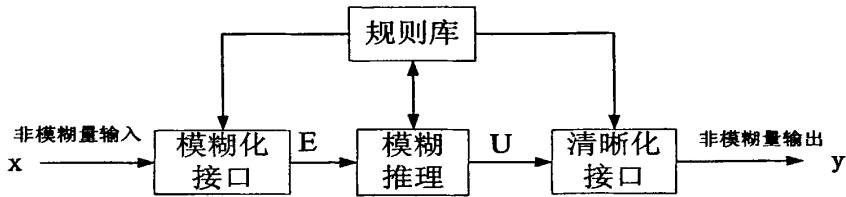


图 2—15 模糊控制器的基本组成

figure 2—15 ultimate composition of fuzzy controller

二、排料系统模糊控制系统的设计

跳汰机排料，必须是一个连续的过程，尽管限于床层移动速度只能在 10^{-1}m/s 数量级上，但鉴于一般跳汰机都在处理着成百上千吨/小时的原煤并且从系统稳定出发，所研究的排料模糊控制器还是要保证高的响应速度，并寻求小的稳态误差域。采用带自调整因子的双输入单输出模糊控制器，如图 2—16 所示，图中， H_i 为实测床层密度， H_s 为床层密度设定值； e 和 \dot{e} 为实际的床层密度偏差及偏差变化率。控制器模型结构所涉及的 3 个语言变量的模糊集合是偏差 $\tilde{E}(e)$ 、偏差变化率 $\tilde{E}_c(\dot{e})$ 和输出控制量 \tilde{U} 。

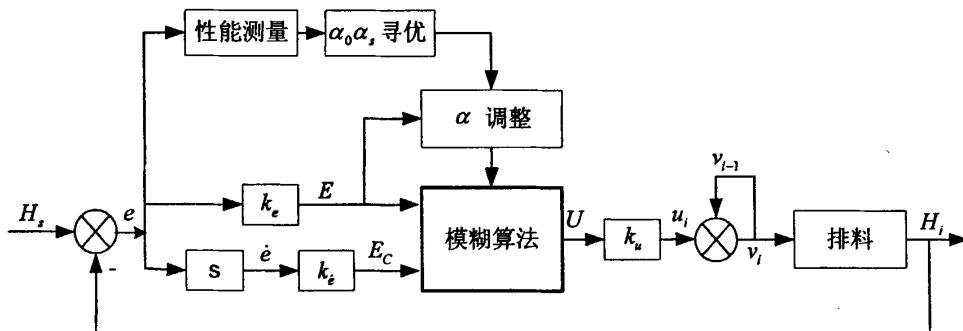


图 2—16 带自调整因子的模糊控制器框图

figure 2—16 fuzzy controller with self-adjust gene

设选定偏差 $\tilde{E}(e)$ 的离散论域为 $\{-6, -5, \dots, -0, +0, \dots, +5, +6\}$ 偏差变化率 $\tilde{E}_c(\dot{e})$ 和 \tilde{U} 的论域为 $\{-6, -5, \dots, 0, \dots, 5, 6\}$ ，并取 $\tilde{E}_c(\dot{e})$ 和 \tilde{U} 分别有 7 个

语言变量值为 PB、PM、PS、Z0、NS、NM、NB； $\tilde{E}(e)$ 8 个语言变量值为 PB、PM、PS、PO、NO、NS、NM、NB；各模糊集合对应的隶属度函数取 GAUSS 函数。

设偏差 e 的基本论域为 $[-15, 15]$ ，则偏差量化因子 $k_e = 2/5$ ，根据隶属度函数确定出偏差模糊子集并建立语言变量 $\tilde{E}(e)$ 的赋值表 2-1；同时设偏差变化率 \dot{e} 和 u 的基本论域为 $[-10, +10]$ ，量化因子 $k_{\dot{e}}$ 和 k_u 为 $5/3$ 。同样得到 $\tilde{E}_c(\dot{e})$ 的赋值表 2-2

表 2-1 语言变量 $\tilde{E}(e)$ 赋值表table2-1 Language variable $\tilde{E}(e)$ endued list

\tilde{E}	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
NL	1.0	0.8	0.7	0.4	0.1	0	0	0	0	0	0	0	0
NM	0.2	0.7	1.0	0.7	0.3	0	0	0	0	0	0	0	0
NS	0	0.1	0.3	0.7	1.0	0.2	0.2	0	0	0	0	0	0
NO	0	0	0	0	0.1	1.0	1.0	0	0	0	0	0	0
PO	0	0	0	0	0	0.1	1.0	0.6	0.1	0	0	0	0
PS	0	0	0	0	0	0	0.2	0.7	1.0	0.7	0.3	0.1	0
PM	0	0	0	0	0	0	0	0	0.2	0.7	1.0	0.7	0.3
PL	0	0	0	0	0	0	0	0	0.1	0.4	0.7	0.8	1.0

表 2-2 语言变量 $\tilde{E}_c(\dot{e})$ 赋值表table2-2 Language variable $\tilde{E}_c(\dot{e})$ endued list

\dot{e}	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
NL	1	0.8	0.4	0.1	0	0	0	0	0	0	0	0	0
NM	0.2	0.7	1.0	0.7	0.2	0	0	0	0	0	0	0	0
NS	0	0	0.2	0.7	1.0	0.9	0	0	0	0	0	0	0
ZO	0	0	0	0	0	0.5	1.0	0.5	0	0	0	0	0
PS	0	0	0	0	0	0	0	0.9	1.0	0.7	0.2	0	0
PM	0	0	0	0	0	0	0	0	0.2	0.7	1.0	0.7	0.2
PL	0	0	0	0	0	0	0	0	0	0.1	0.4	0.8	1.0

一般的模糊控制规则算法为：

$$\tilde{U} = (\tilde{E} + \tilde{E}_c)/2 \quad (2-4)$$

且规则固定不变。这些规则都是人工经验并带有主观因素，使控制规则往往在某种程度上显得精度不高或不完善，并且当对象的动态特性发生变化，或受到随机干扰的影响时，都会影响到模糊控制的效果。因此需要对控制规则和查询表不断及时地进行修正。如果在上式的基础上引入一个可调整因子即 α （或称加权系数）就可以对规则进

行调整即：

$$\tilde{U} = -[\alpha \tilde{E} + (1-\alpha) \tilde{E}_c] \quad (2-5)$$

其中 $\alpha \in [0, 1]$ 为在线调整。对于 $X \in \mathbb{R}$, 正态模糊数 $\tilde{E} \in X$ 取自系统偏差的模糊量化

$\tilde{E} = \langle k_e e(t) \rangle$, 正态模糊数 $\tilde{E}_c \in X$ 取自系统偏差变化率的模糊量化

$\tilde{E}_c = \langle k_e \dot{e}(t) \rangle$, $\tilde{U} \in Y$ ($Y \in \mathbb{R}$) 仍为正态模糊数。

α 的调整规则为：

$$\alpha = \frac{1}{N} (\alpha_s - \alpha_0) |E| + \alpha_0 \quad (2-6)$$

式中, α_0 和 α_s 为 α 的上下限, $0 \leq \alpha_0 \leq \alpha_s \leq 1$; N 取 $|E|_{\max}$ 。

由式(1)知, α 表征了按偏差大小自动调整偏差和偏差变化对控制作用的权重, α 选取得当与否对系统响应速度、超调量及鲁棒性有很大影响。如 α 取小值时, 虽无超调, 但响应速度太慢。为此, 对 α_0 和 α_s 进行寻优。

定义 X 为模糊控制误差矢量:

$$X = [x(t_s) \ x(t_s+1) \ \dots \ x(t_f)]^T \quad t_s = 1, \dots, t_f - 1$$

式中 x 为床层厚度误差标量值, t_s 和 t_f 分别表示调整期间的离散时间起点和终点。

定义 J 为控制系统性能指标:

$$J = \|X\| / (t_f + 1 - t_s)^{1/2} \quad (2-7)$$

式中 $\|\cdot\|$ 为欧范数。令 $\partial J / \partial X = 0$, 得到优化控制量。

当输出语言变量论域上的模糊子集 \tilde{U}_1 确定后, 由论域 Y 上的每个元素

y_i ($i = 1, 2, \dots, n$) 对输出模糊子集 \tilde{U}_1 的隶属度 $\mu_c(y_i)$ 取加权平均, 将得到模糊控制系统的非模糊判决结果:

$$y_0 = \frac{\sum_{i=1}^n y_i \mu_c(y_i)}{\sum_{i=1}^n \mu_c(y_i)} \quad (2-8)$$

控制器输出精确值 $u_1 = k_u \bullet y_0$. 当 v_i 、 v_{i-1} 分别表示第 i 和 $i-1$ 时刻的闸板开度时, 有:

$$v_i = v_{i-1} + u_i$$

2.5 本章小结

跳汰分选的核心是按密度差异进行物料的分层和分离，从跳汰机结构上看，以唐山分院研制的 SKT 跳汰机为代表，在风室结构、风阀构造、排料机构等方面进行了大量改进工作，日臻完善；从跳汰分选过程看，分选效果、生产能力受诸多因素共同影响。

而分选效果体现在分层和排料两个过程中，两个过程同时发生，共同作用，互相影响，涉及到的跳汰机控制变量众多，是一个复杂、时变、非线性系统。

跳汰床层密度探测是跳汰机理研究和跳汰控制等所有问题分析的基础， γ 射线密度探测器实现了跳汰床层在线测量，再加上通过大量实验确定的不同工况下床层密度分布规律，于是有限的几个探测器置于床层不同的层位内，即可完整表述床层的密度分布^[25]，这是跳汰床层密度检测的重大突破。并且，只要检测到了床层内某点的密度，即可间接测量出该层内 $+1.8\text{ g/cm}^3$ 密度级含量，而对于跳汰机两段的重产物层位的密度探测即是最终产品矸石、中煤的带煤损失，也即错配物含量，也即数量效率。

我们研发的基于跳汰床层密度在线探测的模糊排料控制系统，一改传统的浮标探测，利用 γ 射线密度探测器实测床层密度，与床层密度设定值、排料闸板开度一起根据模糊规则，实行模糊推理输出控制，效果显著。本文在现场大量采样、试验的基础上，利用该控制系统实时收集系统信息数据，为进一步分析研究奠定了基础。

但是，这种控制方式仍然没有把分层和排料这两个过程完全地结合起来考虑，对两者之间的相互作用也未能充分体现，尤其是对入洗原煤性质改变情况下如何根据床层分层状况进行操作变量的适时调整仍没有得到解决，利用人工神经网络的映射功能进行分析、研究显得尤为必要。作者采用人工神经网络对上述问题进行分析、研究，提出了解决方法。

第三章 基于人工神经网络的跳汰过程状态识别

人工神经网络（Artificial Neural Network）是一种全新的模拟人类智能的技术，它无需建立物理模型和进行人工干预，而是通过神经元建立起输入、输出状态间的映射关系。通过训练学习，它能自动调节各神经元之间的结合强度，使网络能正确映射其输入输出关系。

人工神经网络具有的对任意非线性对象的良好逼近能力，给为非线性辨识、非线性建模提供了新的方法，跳汰控制系统的问题有可能得到较好的解决。作者使用神经网络中应用最广泛的基于误差反向传播算法(BP 网络)和多层前向径向基函数网络(RBF 网络)，对跳汰控制系统的工业实验数据进行分析。

3.1 现场实验及样本数据

在正常生产的条件下，每一小时左右人工采集某矿 307#跳汰机入洗原煤、矸石各 2—3kg，分别做密度为 $1.5\text{g}/\text{cm}^3$ 、 $1.8\text{g}/\text{cm}^3$ 的浮沉试验，以确定入洗原煤性质以及矸石带煤量的大小。

考虑矸石样品在排料斗采集与矸石排料门采集间存在一定的运输距离，矸石带煤量与其对应的风水量及床层在线密度在时间上有一定的差距，所以采用示踪法确定其时间差，从而使样本中各个数据在时间上相对应。具体做法是将带漆矸石放在排料门，当排料时开始记时；当在采集点看到带漆矸石时停止记时，这一段时间即为人工修正的时间差。部分样本数据如附录 1 所示。

表中所列数据中，床层的三个层位的密度值 (FL1,FL2,FL3)、排料门的开度值 (GATE1L) 和风阀周期 (JQ1,PZ1,... ...) 等参数由计算机系统自动记录，而矸石带煤 (WASTE) 为人工测定的矸石中密度小于 1.8 部分的含量，RC 表示入洗原煤中密度大于 1.8 部分的含量，和原煤灰分正相关，人工测定，代表入洗原煤性质的变化，实际控制系统可采用原煤在线灰分仪的数据。

从附录 1 可以看出，样本中包含了对分层起主要作用的控制变量、分层状态数据、原煤性质数据、排料状态数据，这些数据成为跳汰系统分层和排料两个过程的完整信息集合。

要使人工神经网络具有良好的泛化能力，合理的样本选择是其必要条件。首先通

过观察采集来的样本数据，对于样本密集的地方，由于外在的因素的干扰有可能出现矛盾数据，进行人工删减。另外通过观察删除了一些由于实验因素而导致的明显异常数据。最后选出了 600 组样本数据作为神经网络的训练和测试。

尽管对于跳汰机这样复杂的设备，如果没有足够多的训练样本，就不足以表征所研究问题的基本特性。但是可以使网络通过合理的学习机制和以后的应用过程中不断补充的典型样本采集而不断在线修正，使网络逐渐提高泛化能力。

3.2 基于 BP 网络的模式识别

BP 网络的误差反馈特性解决了隐含层引入以后的学习问题。其实质是采用梯度下降法使权值的改变总是朝着误差变小的方向改进，最终达到最小误差。网络实质上实现了一个从输入到输出的映射功能，而数学理论已证明它具有实现任何复杂非线性映射的功能^[29]。由于其结构简单，可塑性强，使得它在函数逼近，模式识别和信息分类等领域得到了广泛的应用，而且具有一定的推广、概括能力和自学习的能力。BP 网络目前是采用最多，理论发展最完善，应用最广泛，最为成熟的训练算法之一。但是 BP 网络的收敛速度比较慢，由于采用了梯度下降法，就有使网络容易陷入局部极小值的可能，而且网络的泛化能力比较差。而遗传算法具有全局随机搜索能力，所以用遗传算法优化神经网络的初始权值和阈值，使其减小陷入局部极小值的可能性。

3.2.1 研石带煤 BP 网络拓扑结构的设计

BP 神经网络是一种前向网络。典型的 BP 网络是三层网络，由输入层、隐含层和输出层组成，见图 3—1。同层节点间无关联，异层神经元之间前向连接。复杂的 BP 网络是由多个隐含层组成的。网络结构的设计包括确定网络的隐含层数、隐含层节点数及输入输出节点数。

作者首先利用 BP 网络对跳汰机研石段研石带煤与分层状态、排料之间的关系进行识别：给定网络输入层为 5 个输入节点（三个层位的密度值，排料门的开度值，原煤

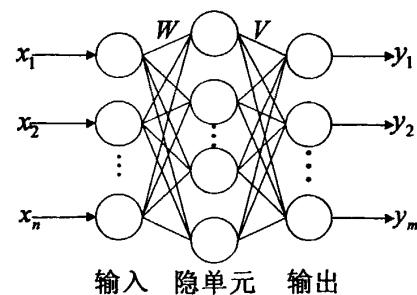


图 3—1 BP 网络的拓扑结构
figure 3—1 topology structure of BP

中+1.8 部分含量即灰分值代表），输出层 1 个输出节点（矸石带煤值）。

在把数据输入神经网络前需要对输入样本数据作必要的预处理，主要考虑归一化：使原始输入数据 P 分布在 $[-1, 1]$ 之间，归一化后的数据用 P_n 表示，公式为：

$$P_n = 2 * (P - P_{\min}) / (P_{\max} - P_{\min}) - 1 \quad (3-1)$$

归一化的原因有：避免具有不同物理意义和量纲的输入变量不能平等使用，BP 中常采用 sigmoid 函数作为转移函数，归一化能够防止净输入绝对值过大引起的神经元输出饱和现象，保证输出数据中数值小的不被吞食，另外还可以提高神经网络的训练效率。

3.2.2 隐含层函数的选择

BP 网络常用的非线性传递函数有以下几种：

(1) Sigmoid 函数

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3-2)$$

(2) Tansig 函数。

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3-3)$$

(3) purelin 函数。

$$f(x) = x \quad (4-4)$$

由 Kolmogorov 定理^[20]知，采用 S 形函数作为节点作用函数就可使神经网络模型逼近任何连续函数。所以用 sigmoid 函数作为隐含层激活函数。中间层用 S 型函数是分类用的，使网络有非线性映射能力，而输出层用线性型函数，这样即可不会导致输出数据被压缩。

3.2.3 BP 网络算法

网络学习的指导思想是，对网络权值 (ω_{ij}, T_{ij}) 修正，使误差函数 E 沿负梯度方向下降。BP 网络三层节点表示为，输入节点： x_j ，隐节点： y_i ，输出节点： O_l 。

输入节点与隐节点间的权值为 ω_{ij} , 隐节点与输出节点间的网络权值为 T_{il} , 当输出节点的期望输出为 t_l 时, BP 模型的计算公式如下:

1) 隐含层节点的输出:

$$y_i = f(\sum_j \omega_{ij} x_j - \theta_i) \quad (3-5)$$

2) 输出节点的计算输出:

$$O_l = f(\sum_j T_{il} y_i - \theta_l) \quad (3-6)$$

3) 输出节点的误差公式:

$$\begin{aligned} E &= \frac{1}{2} \sum_l (t_l - O_l)^2 = \frac{1}{2} \sum_l (t_l - f(\sum_i T_{il} y_i - \theta_l))^2 \\ &= \frac{1}{2} \sum_l t_l - f(\sum_i T_{il} f(\omega_{ij} x_j - \theta_i) - \theta_l))^2 \end{aligned} \quad (3-7)$$

式中: t_{pi}, O_{pi} 分别为期望输出和网络的计算输出。

BP 算法实质上是梯度最快速下降法的一种应用结果。所谓梯度最快速下降法, 指在点 x 处按 $f(x)$ 的负梯度方向进行搜索的一种优化方法。

因此, BP 算法可小结为: 设网络有 m 层, 第一层为输入层, 第 m 层为输出层, 有 s 个输入样本

1) 选定初始权值 W ;

2) 重复下述过程直到网络收敛。

a) 对于任何一个输入样本, 计算网络的实际输出即网络中任意一个节点的输出值(正向传播过程)。

b) 从第 m 层到输入层对各节点反向计算误差信号(反向传播过程)。

c) 修正权值: $\omega_{ij}(k+1) = \omega_{ij}(k) + \Delta\omega_{ij}$

3.2.4 隐含层数目的选择

根据“Kolmogorov 神经网络映射存在定理”^[19, 20, 21], 它证明了: 一个三层前向馈入式神经网络, 当隐含层神经元数目足够多时, 可以以任意精度逼近任何一个具有有限间断点的非线性函数。然而这仅是一个存在定理, 它并未指明是否能够使该网络学习到正确的权值。在 Kolmogorov 的工作的基础上, Hecht-Nielson 及其他人也证明:

具有正确权值的三层后向传递神经网络能按照指定精度实现任意平方可积函数，但是，应用现有的学习规则仍不能保证一定能学到正确的权值。鲍立威，何敏，沈平引证了 Kolmogorov 定理，并得出了四层网络才可任意逼近任何连续函数的结论。而且采用严格的数学证明说明了在三层网络中，隐节点数不能任选，它其实恰好是对应函数展开式的项数，可见，隐含层节点的数目是很固定的。这又说明了三层 BP 网络用于学习及记忆时，为了记住较多样本并能逼近一个函数，需要相当数量的隐含层节点，但是理论及实践都证明这样的三层网络对样本数据携带的噪声的鲁棒性极差，且计算效率随隐含层节点数的增加下降很大。由上面论述可见：三层网络很难说最好。

所以虽然三层神经网络可以拟合任何非线性函数，但是并不代表三层就是最好的选择。相反很多复杂问题用四层网络效果表现更好。但是四层网络比三层网络更容易陷入局部最小，由于实验数据少又会出现过度拟合现象，从而使网络的泛化能力变差。

考虑前面所论述，为了使网络有良好的泛化能力，先考虑设一个隐含层。而一般经验表明：采用两个隐含层时，如在第一个隐含层设置较多的隐节点而第二个隐含层设置较少的隐节点，则有利于改善多层前馈网的性能。此外考虑对于有些实际问题，采用双隐含层所需要的隐节点总数可能少于单隐含层所需的隐节点数，所以，当一个隐含层的节点数很多仍不能改善网络性能时，再考虑增加一个隐含层。

3.2.5 隐含层节点数的选择

在一般情况下，神经网络的训练能力差时，预测能力也差，并且一定程度上，随训练能力的提高，预测能力也提高。但这种趋势有一个极限，当达到此极限时，随训练能力的提高，预测能力反而下降，即出现所谓“过拟合”现象。

“过拟合”现象是网络隐含层节点过多的必然结果，它的出现影响了网络的泛化能力，在满足精度的要求下，逼近函数的阶数应该越少越好，低阶逼近可以有效防止“过拟合”现象，从而提高网络的预测能力，反映到多层前向神经网络中，就是在精度满足的要求下，网络的隐节点数越少越好。但是隐含层神经元数目如果太少网络就无法进行训练，如果刚好能够进行训练则网络的鲁棒性差，抗噪声能力不强，不能辨识以前没有遇到的模式。而如果网络隐含层的神经元数目太多就会需要大量的训练样本，而且伴随而来的是为训练而耗费大量时间和内存，甚至出现过拟合。

因此，关于隐含层的节点数的选择是个复杂的问题，对于不同的解决问题，节点

数是不相同的，在实际应用中，还没有一套成熟的理论方法来确定网络的隐节点，隐节点的确定基本上依赖经验，主要是采用递增或递减的试探方法来确定。但有下面几个经验公式值得参考：其中 m 为隐含层节点数， n 为输入节点数， o 为输出节点数， a 为1-10之间的常数，特别情况可以大于10。

$$1) \quad m = \sqrt{n+o} + a \quad (3-8)$$

$$2) \quad m = \log 2^n \quad (3-9)$$

$$3) \quad m = \sqrt{n \cdot o} \quad (3-10)$$

参考上面的经验公式，作者选定隐含层节点数的范围为3到13之间。如果在此区间网络的性能不能满足要求时，再考虑增加一个隐含层。

而为了提高神经网络的泛化能力，针对过拟合问题采用以下解决方法：

1) 正则化方法：

正则化方法是通过修改神经网络的训练性能函数来提高其推广能力，一般情况下神经网络采用均方误差函数 mse 作为训练性能函数，函数如下：

$$mse = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2 \quad (3-11)$$

式中 e_i ——预测值与实际值的误差；

N ——样本对的个数。

在正则化方法中，网络的性能函数经过改进变为如下形式：

$$msereg = \gamma \cdot mse + (1 - \gamma) msw \quad (3-12)$$

其中 γ 为比例系数， msw 为所有网络权值平方和的平均值，即

$$msw = \frac{1}{N} \sum_{j=1}^N (w_j)^2 \quad (3-13)$$

可见，采用新的性能指标函数后，可以保证网络训练误差尽可能小的情况下，还要使网络具有较小的权值，即：使得网络的有效权值尽可能少，这实际上相当于自动缩小了网络的规模。

但是常规的正则化方法通常很难确定比例系数 γ 的大小，当 γ 太大时网络可能会出现过拟合，若 γ 太小，则网络将无法进行训练。最好的策略是能自动地设置调整参数，

一种较好的方法是基于贝叶斯决策规则的调整方法即贝叶斯正则化方法^[29, 34]。该方法假定网络的权值和偏置值是具有特定分布的随机变量，可以用统计方法来估计调整参数，从而使网络不会产生过拟合。在 matlab 的 NNTool 中，贝叶斯正则化方法是通过 trainbr 函数来实现的。它依据 Levenberg-Marquardt 优化理论对网络的权值和阈值进行调整。

2) 提前停止法:

提前停止法是避免产生过拟合、提高网络的泛化能力的另外一种有效的方法。在该方法中，训练样本集被分为训练集，验证集和测试集，其中测试集是可选的。训练集用来对神经网络进行训练，验证集用于在对神经网络训练的同时监控网络的训练进程。在训练的初始阶段，验证集形成的验证误差通常会随着网络的训练误差的减小而减小，但是当网络出现过度训练时，验证误差就会逐渐增大，当验证误差增大到一定程度时，网络训练就会提前停止。这时训练函数会返回当验证误差取最小值时的网络对象。其中测试集形成的测试误差在网络训练时没有被使用。

以上两种方法能消除过拟合现象，但是不能选择最合适隐含层神经元数目，无法获得最佳的网络结构。而其中的贝叶斯正则化方法只用于小规模网络的函数拟合或函数逼近问题，不适用于解决模式分类问题，而且其收敛速度一般也比较慢。

另外对于样本的划分，最简单的技术就是把有待训练的样本分为两个互不相交的集合，即训练集合和检验集合，若两个集合中的向量的数量均很多，这样检验集合所得出的抽样误差率比较高。这里，把 $2 / 3$ 的可变向量归入训练集合，而另外 $1 / 3$ 则归入检验集合。对于交叉检验方法，即把有待训练的样本分成 k 个互不相交的子集。然后，选其中一个子集作为检验集合而另外 $k-1$ 个子集作为训练集合。这样做 k 次，每次选不同的一个子集作为检验集合而剩下的作为训练集合。

为了得到好的网络结构，综合考虑上面提到的各种方法，作者采取下面的办法：

根据经验公式首先把隐含层的节点数限制在3到13之间，采用试凑法从小到大逐渐增加隐含层数目。其中神经网络的初始权值和阈值采用遗传算法优化得到，遗传算法的终止条件设为最大进化代数为200。对于每个隐含层数目，采用提前终止方法。把数据分为4个互不相交的子集，选其中一个子集作为验证集合，一个集合作为测试集合，而另外两个子集作为训练集合，这样有12种不同的组合方法。记录每次训练终止时的训练步长，训练达到的目标误差值，和测试误差的相对大小的平均值。对于每个隐含

层都训练12次，然后计算12次训练步长，取目标误差值和测试误差的相对大小的平均值。

当取得的测试误差的平均值随着隐含层的增加开始增加时，训练结束，其中的测试误差的平均值最小时对应的隐含层节点数即为设计的神经网络的隐含层节点数。

经过仿真，当隐含层节点数为3，4，5时网络不能收敛，当隐含层节点数大于5时网络能够收敛，其收敛情况整理如下表3—1所示：

表3—1 神经网络仿真结果

table 3—1 emulation results of ANN

隐含层节点数	6	7	8	9
训练步长	58.2	42.4	49.5	35.8
训练结果mse	0.0258251	0.0266238	0.0198324	0.0122568
测试误差	21.3873%	11.2505%	16.6426%	20.8577%

从上表中可以看出，当隐含层节点数为7时，测试误差达到了比较小的程度而且训练误差也很小了，同时随着隐含层节点数的增加测试误差反而增加了，所以采用隐含层节点数为7的神经网络结构。

3.3 遗传算法与神经网络的结合

最近十多年来，人工神经网络，尤其是BP神经网络获得了极其广泛的应用。它有效地克服了多层网络无法解决的非线性分类问题的缺陷，但BP算法属于梯度下降算法，从本质上来说属于局部寻优算法，在存在较多局部极小的情况下很容易陷入局部极小点，而且要求误差函数必须可导，受网络结构限制等缺点。针对这些问题，国内外研究人员进行了不懈的努力，提出了很多改进算法和新的算法，但到目前为止，前向网络的训练问题还远没有得到解决。

而遗传算法具有全局随机搜索能力，它能够在复杂的、多峰值的、不可微的大矢量空间中迅速有效地寻找到全局最优解，能够弥补传统神经网络学习算法存在的不足，另外，由于适应度函数无需可导，因此进化学习算法可适用的神经元(激活函数)类型更为广泛。同时由于遗传算法使用简单，鲁棒性强的特点，用遗传算法进化神经网络无疑具有重要的意义。

所以为了解决BP网络易于陷入局部极小值的现象，本文用遗传算法来优化BP网

络的初始权值和阈值。有两种方法可以优化，一种是直接用遗传算法优化神经网络的权值和阈值，另一种是先用遗传算法优化 BP 网络的初始权值和阈值，把网络的权值和阈值限定在一个合适的范围内，然后用神经网络的梯度下降法在这个范围内快速找到最小值。

由于遗传算法也有其自身的缺点，如计算量大、训练速度慢等，它擅长全局搜索而局部搜索能力不足，研究发现，遗传算法可以用极快的速度达到最优解的 90%左右^[94, 95]，但要达到真正的最优解还要花费很长的时间。这样如果用第一种方法进行优化，优化的速度将会很慢。而采用第二种优化方法把遗传算法和 BP 算法结合起来，形成 GA-BP 混合训练算法，以遗传算法优化 BP 网络的初始权值和阈值，再由 BP 算法按负梯度方向修正网络权值及阈值，这样神经网络可以快速找到极小点。所以采用后一种方法进行网络训练，只要遗传算法设计得当就可以避免 BP 网络易陷入局部极小问题，达到优化网络目的。

3.3.1 遗传算法分析

遗传算法(Genetic Algorithm, GA)是近些年发展起来的一种自适应全局优化概率搜索算法，是一种模拟生物界自然选择和自然遗传机制的随机化搜索算法，它借用了生物遗传学的观点，通过自然选择、遗传、变异等作用机制，实现各个个体的适应性的提高。这一点体现了自然界中“物竞天择、适者生存”进化过程。其主要特点是群体搜索策略和群体中个体之间的信息交换，搜索不依赖于梯度信息。在实现机制上，它又是一种离散动力学系统，在给定初始群体和遗传操作的前提下，通过迭代实现群体的进化。它尤其适用于处理传统搜索方法难于解决的复杂和非线性问题，可广泛用于组合优化、机器学习、自适应控制、规划设计和人工生命等领域^[17]。

遗传算法研究的目的，一是抽取和解释自然系统的自适应过程，二是设计具有自然系统机理的人工系统。1975 年，Holland 在《自然界与人工系统中的适应性》一书中，正式提出遗传算法的概念和理论框架。当前，遗传算法与进化策略、进化程序设计共同构成了“进化计算”的主干，而且它们之间的边界也被打破了。

遗传算法是一种由一个“染色体群”通过“自然选择”的机制转化成另一个“染色体群”的方法。目前关于遗传算法的定义说法不一。但是，所有称为“遗传算法”的方法都至少具有如下特征：染色体群，依据适应度进行的选择，通过交叉产生新的子代，

随机产生的的子代突变现象等等。

在遗传算法中，染色体对应的是数据或数组，在标准遗传算法中，通常是由一维的串结构数据来表现的。串上的各个位置对应上述的基因座，而各个位置上所取的值对应上述的等位基因。遗传算法处理的是染色体，或叫基因型个体。一定数量的个体组成了群体，也叫集团。群体中个体的数目称为群体的大小，也叫群体规模。而个体对环境的适应程度叫做适应度。此外，遗传算法的执行包含两个必须的数据转换操作，一个是表现型到基因型的转换，另一个是基因型到表现型的转换。前者是把搜索空间中的参数或解转换成遗传空间中的染色体或个体，即编码操作；后者是前者的一个相反操作，即译码操作。

每个染色体可以看作搜索空间中的一个点，代表了一个候选解。遗传算法对染色体群进行处理，不断地用一个新的染色体群替换原来的染色体群，也就是不断地尝试新的候选解。为测试候选解的可行性，需要一个适应度函数来检查每一个候选解的适应性。这个函数给出每一个候选解的适应性的度量，也就是这个候选解对于所给定问题的有效性的度量。

3.3.2 遗传算法的实现

遗传算法是一种群体型操作，该操作以群体中的所有个体为对象。最简单的遗传算法中最常用的算子有如下几种：

(1) 选择算子(selection/reproduction)：选择算子从群体中按某一概率成对选择个体，某个体 x_i 被选择的概率 P_i 与其适应度值成正比。

(2) 交叉算子(Crossover)：交叉算子将被选中的两个个体的基因链按概率 P_c 进行交叉，生成两个新的个体，交叉位置是随机的。其中 P_c 是一个系统参数。例如，两个染色体分别为10000100和11111111，随机选定的位置是第三位，进行交叉的结果是10011111和11100100。交叉操作粗略地模拟了两个单倍体生物进行繁殖时染色体的配对过程。

(3) 变异算子(Mutation)：变异算子将新个体的基因链的各位按概率 p_m 进行变异，对二值基因链(0, 1 编码)来说即是取反。例如，二进制串 01000100 在第二位上发生变异，变成 00000100。变异可能发生在符号串的任意位上。通常发生变异的概率很低。

上述各种算子的实现是多种多样的，而且许多新的算子正在不断地提出，以改进 GA 的某些性能。

再次，采用串行运算的遗传算法(Sequential Genetic Algorithm, SGA)，按如下过程进行：

- (1) 对待解决问题进行编码；
- (2) 随机初始化群体 $X(0) = (X_1, X_2, \dots, X_n)$ ；
- (3) 对当前群体 $X(t)$ 中每个个体 X_i 计算其适应度 $F(X_i)$ ，适应度表示了该个体的性能好坏；
- (4) 应用选择算子产生中间代 $X_r(t)$ ；
- (5) 对 $X_r(t)$ 应用其它的算子，产生新一代群体 $X(t+1)$ ，这些算子的目的在于扩展有限个体的覆盖面，体现全局搜索的思想；
- (6) $t = t + 1$ ；若遗传代数达到允许值或其他收敛条件以满足时，停止，否则转到步骤 3。

3.3.3 遗传算法优化神经网络权值及阈值的步骤

一、编码方式分析：

编码是应用遗传算法首先要解决的问题，也是设计遗传算法的一个关键步骤。编码方法除了决定个体的染色体排列形式之外，还决定了个体从搜索空间的基因型变换到解空间的表现型时的解码方法，编码方法也影响到交叉算子、变异算子等遗传算子的运算方法。所以，编码方法在很大程度上决定了如何进行群体的遗传进化运算以及遗传进化运算的效率。

用遗传算法优化 BP 网络的初始权值和阈值，首先要考虑网络权值和阈值的编码。编码原则：

针对跳汰系统具体应用问题，如何设计一种完美的编码方案是重点课题，也是遗传算法的一个重要研究方向。可以说目前还没有一套既严密又完整的指导理论及评价准则能够帮助我们设计编码方案。

作为参考，De Jong 曾提出了两条操作性较强的实用编码原则(又称为编码规则)^[17]：

- 编码原则一(有意义积木块编码原则)：应使用能易于产生与所求问题相关的且

具有低阶、短定义长度模式的编码方案。

- 编码原则二(最小字符集编码原则): 应使用能使问题得到自然表示或描述的具
最小编码字符集的编码方案。

由于遗传算法应用的广泛性, 迄今为止人们已经提出了许多种不同的编码方法。

总的来说, 这些编码方法可以分为三大类:

1、二进制编码方法:

二进制编码方法是遗传算法中最常用的一种编码方法, 它使用的编码符号集 是由二进制符号 0 和 1 所组成的二值符号集{0, 1}, 它所构成的个体基因型是一个二进制编码符号串。

二进制编码方法的优点:

- 编码、解码操作简单易行;
- 交叉、变异等遗传操作便于实现;
- 符合最小字符集编码原则;
- 便于利用模式定理对算法进行理论分析。

格雷码编码方法

格雷码编码方法是二进制编码方法的一种变形。其连续的两个整数所对应的编码值之间仅仅只有一个码位是不相同的, 其余码位都完全相同。

格雷码编码优点:

- 便于提高遗传算法的局部搜索能力;
- 交叉、变异等遗传操作便于实现;
- 符合最小字符集编码原则;
- 便于利用模式定理对算法进行理论分析。

二进制编码的缺点:

- 二进制编码存在着连续函数离散化时的映射误差。个体编码串的长度较短时, 可能达不到精度要求; 个体编码串的长度较长时, 虽然能提高编码精度, 但却会使遗传算法的搜索空间急剧扩大。
- 二进制编码不便于反映所求问题的特定知识, 这样也就不方便开发针对问题专门知识的遗传运算算子, 人们在一些经典优化算法的研究中所总结出的一些宝贵经验也就无法在这里加以利用, 也不便于处理非平凡约束条件。

2、浮点数编码方法:

个体的每个基因值用某一范围内的一个浮点数来表示；个体的编码长度等于其决策变量的个数。因为这种编码方法使用的是决策变量的真实值，所以浮点数编码方法也叫做真值编码方法。

在浮点数编码方法中，必须保证基因值在给定的区间限制范围内；遗传算法中所使用的交叉、变异等遗传算子也必须保证其运算结果所产生的新个体的基因值也在这个区间限制范围内。当用多个字节来表示一个基因值时，交叉运算必须在两个基因的分界字节处进行，而不能在某个基因的中间字节分隔处进行。

浮点数编码方法的优点：

- 适合于在遗传算法中表示范围较大的数；
- 适合于精度要求较高的遗传算法；
- 便于较大空间的遗传搜索；
- 改善了遗传算法的计算复杂性，提高了运算效率；
- 便于遗传算法与经典优化方法的混合使用；
- 便于设计针对问题的专门知识的知识型遗传算子；
- 便于处理复杂的决策变量约束条件。

浮点数编码方法的缺点：

- 有些问题需要专门的遗传操作算子。

3、符号编码方法:

个体染色体编码串中的基因值取自一个无数值含义、而只有代码含义的符号集。

这个符号集可以是一个字母表，如 $\{A, B, C, D \dots\}$ ；

也可以是一个数字序号表，如 $\{1, 2, 3, 4, 5 \dots\}$ ；

还可以是一个代码表，如 $\{A_1, A_2, A_3, A_4 \dots\}$ 等等。

符号编码的主要优点：

- 符合有意义积木块编码原则。
- 便于在遗传算法中利用所求解问题的专门知识。
- 便于遗传算法与相关近似算法之间的混合使用。

但对于使用符号编码方法的遗传算法，一般需要认真设计交叉、变异等遗传运算

的操作方法，以满足问题的各种约束要求，这样才能提高算法的搜索性能。

综上考虑，由于对跳汰系统神经网络样本值而言，优化的神经网络的初始权值和阈值是实数，二进制编码存在着离散化时的映射误差，而且有精度要求和搜索空间之间的矛盾，个体编码串的长度较短时，可能达不到精度要求。由于优化的目标总体数目比较多，个体编码串的长度较长时，每个个体的二进制编码的符号串长度就会很长，大的搜索空间寻优肯定会使得遗传算法的运行性能相当差，甚至可能无法进行下去。符号编码方法一般用于解决一些专门的问题。而浮点数编码是连续参数优化问题的自然描述，个体的编码长度等于其决策变量的个数，有比较高的精度，便于较大的搜索空间，对于一些复杂的问题也有较高的运算效率，而且便于和其他搜索技术相结合。

所以采用浮点数编码的方法进行具体编码：

设输入层有 m 个节点，隐含层有 s 个节点，输出层有 n 个节点。这样对 BP 网络的权值（共有 $m * s + s * n$ 个权值）编码如下：

$$W_{11}, W_{21}, \dots, W_{ji}, \dots, W_{sm}, \dots, W_{sj}, \dots, W_{ns}$$

其中 W_{ji} 为输入层第 i 个节点到隐含层第 j 个节点的连接权。 W_{sj} 为隐含层第 j 个节点到输出层第 s 个节点的连接权。

对阈值 b 进行编码，有隐含层和输出层共有 $s+n$ 个阈值，编码如下：

$$B_1, B_2, \dots, B_{s+n}$$

最后得到的遗传算法的种群个体由权值和阈值向量组成。表示如下：

$$W_{11}, W_{21}, \dots, W_{ji}, \dots, W_{sm}, \dots, W_{sj}, \dots, W_{ns}, B_1, B_2, \dots, B_{s+n}$$

遗传算法编码长度为 $m * s + s * n + s + n$ 。

二、遗传算法的初始化：

首先对 BP 网络初始化。其中包括网络的输入，输出向量的生成，网络的输入层，隐含层，输出层神经元节点的个数和遗传算法编码长度的确定。

1、适应度函数的确定：

遗传算法的一个特点是它仅使用所求问题的目标函数值就可得到下一步的有关搜索信息。而对目标函数值的使用是通过评价个体的适应度来体现的。适应度函数是评价个体优良的标准，优秀的个体适应度值大。评价个体适应度的一般过程是：

(1) 对个体编码串进行解码处理后，可得到个体的表现型；

(2)由个体的表现型可计算出对应个体的目标函数值;

(3)根据最优化问题的类型,由目标函数值按一定的转换规则求出个体的适应度。

这里设计的神经网络是以均方误差函数 mse 最小为目标函数,

$$mse = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2 \quad (3-14)$$

这样目标函数总是取非负值,但是遗传算法要把问题转化为求目标函数最大值的优化问题,故可直接用总的训练样本的均方误差函数 mse 的和的倒数作为个体的适应度。适应度 f 即为:

$$f = \frac{1}{\sum_{i=1}^N (e_i)^2} = \frac{1}{\sum_{i=1}^N (t_i - a_i)^2} \quad (3-15)$$

2、初始群体的设定:

遗传算法中由众多的个体组成了群体,遗传操作是对众多个体同时进行的。群体中所含的个体数目称为群体规模。群体规模越大,群体中个体的多样性越高,算法陷入局部解的危险越小。所以,从考虑群体多样性出发,群体规模应该大一些。但是,群体规模太大,其适应度评估次数就会增加,从而影响算法的效能,另一方面,群体规模太小,会使遗传算法的搜索空间分布范围有限,因而搜索有可能停止在未成熟阶段,引起未成熟收敛现象,必须保持群体的多样性,即群体规模不能太小。通常种群规模为 20 到 100 之间,这里取其平均值选定初始种群为 60,根据隐含层节点数目的增加可以适当增加种群中个体的数目。由于遗传算法是一种全局最优化方法因此初始群体的选择对算法的收敛性不会产生太大的影响,所以采用最简单的随机方法产生初始群体。由于神经网络的权值和阈值应该是比较小的实数,所以把遗传算法编码的范围定在 -1 到 1 之间。随机产生初始种群 $P = [X_1, X_2, \dots, X_n]$, 对任一 X , 为一个神经网络的权值向量和一个阈值向量组成。

3、遗传操作:

遗传操作是模拟生物基因遗传的操作。通过编码组成初始群体后,遗传操作的任务就是对群体的个体按照它们对环境的适应的程度进行的操作,从而实现优胜劣汰的进化过程。从优化搜索的角度而言,遗传操作可使问题的解,一代一代的逐渐优化,从而逐渐逼近最优解。遗传操作包括三个基本的遗传算子:选择操作,交叉操作,变

异操作。这三个遗传算子的操作都是在随机扰动情况下进行的。遗传操作是随机化操作，但是这种随机化操作和传统的随机搜索方法是有区别的。遗传操作进行的是高效有向的搜索而不是如一般随机搜索方法所进行的无向搜索。其效果和编码方法、群体大小、初始群体以及适应度函数的设定密切相关。三个基本遗传算子的操作方法和操作策略随具体求解的问题不同而异。和个体的编码方式直接有关。下面针对本文中的问题选择合适的遗传算子。

4、选择算子：

选择是从群体中选择优胜的个体，淘汰劣质个体的操作。其目的是把优化的个体直接遗传到下一代或通过配对交叉产生新的个体再遗传到下一代。选择操作是建立在群体中个体的适应度评估基础上的。常见的选择算子如下：

适应度比例选择：

适应度比例选择是目前遗传算法中最基本的选择方法。也叫赌轮选择。该方法中各个个体被选中的概率与其适应度大小成正比。设群体大小为 M ，个体 i 的适应度为 F_i ，则个体 i 被选中的概率 P_{is} 为：

$$P_{is} = F_i / \sum F_i \quad (i=1,2,\dots, M) \quad (3-16)$$

显然，概率 P_{is} 反映了个体 i 的适应度在整体的个体适应度总和中所占的比例。适应度越高的个体被选中的概率也越大，适应度越低的个体被选中的概率也越小。但是由于随机操作的原因，这种选择方法的误差比较大，有时甚至连适应度较高的个体也选择不上。

5、分级选择或称排序选择：

在遗传算法中个体适应度数值上的差别有时会很大，尤其是在算法的早期这种差别更是悬殊。因此，个别特优个体会多次被选中进行复制，经过几代后它们在群体中数目愈来愈多，冲淡了群体的多样性。因此，人们提出分级的概念，用连续渐变的分级代替数值悬殊的适应度。

操作方法如下：

设群体中有 m 个个体，将它们按降序方法依次排序，规定个体优劣的等级依次为：
1, 2, 3, ..., i , ..., M 。

采用线性分级，使各个个体被选中的可能性有如下线性关系：

$$p(i) = q - (i-1)*d \quad (3-17)$$

其中， q —— 最优个体被选中的概率；

d —— 相邻个体被选中概率之差。

上述线性关系使 $p(i)$ 构成等差级数，即：

$$q, q-d, q-2d, \dots, q-i*d, \dots, q-(M-1)*d$$

由于概率定义要求 $\sum_{i=1}^M p(i)=1$ ，按级数求和，有：

$$M \cdot q - \frac{M(M-1)d}{2} = 1 \quad (3-16)$$

$$\text{即 } q = \frac{(M-1)d}{2} + \frac{1}{M} \quad (3-17)$$

分级选择能够消除个体适应度差别悬殊时的影响，代替适应度的缩放技术。但是却抹杀个体适应度的实际差别，未能充分运用遗传信息。

6、竞技选择法：

这种选择法通过相互竞争，优胜者成为下一代的个体。在每一代群体中，每次都随机选择 k 个个体构成一个小群体，然后从这 k 个个体中确定性地取适应度最大的个体复制，进入下一代群体。被复制后的个体仍返回父代群体中，参加下一次 k 个个体的随机选择。这种随机选择重复 M 次，产生 M 个下一代个体。

在竞技选择法中，选择的力度取决于 k 值的大小。 k 值愈大，每次选出的优胜者具有很高的适应度。反之， k 值愈小，优胜者的适应度或高或低，随机性很强。

最优保存策略：

在遗传算法的运行过程中，通过对个体进行交叉、变异等遗传操作而不断地产生出新的个体。虽然随着群体的进化过程会产生出越来越多的优良个体，但由于选择、交叉、变异等遗传操作的随机性，它们也有可能破坏掉当前群体中适应度最好的个体。希望适应度最好的个体要尽可能地保留到下一代群体中。为达到这个目的，可以使用最优保存策略进化模型来进行优胜劣汰操作。

其基本思想是当前群体中适应度最高的个体不参与交叉运算和变异运算，而是用它来替换掉本代群体中经过交叉、变异等遗传操作后所产生的适应度最低的个体。

最优保存策略可视为选择操作的一部分。该策略的实施可保证迄今为止所得到的最优个体不会被交叉、变异等遗传运算所破坏，它是遗传算法收敛性的一个重要保证条件。但是它容易使得某个局部最优个体不易被淘汰掉反而快速扩散，从而使得算法的全局搜索能力不强。

这里用遗传算法优化神经网络的权值和阈值的初始值，是利用了遗传算法的全局搜索能力，避免其出现早熟现象。而分级选择算法的在遗传早期能避免由于个体适应度数值上的差别有时会很大时个别特优个体会多次被选中进行复制，而使遗传算法早熟现象的产生。所以选择分级选择算子。

7、交叉算子的选择：

在生物的自然进化过程中，两个同源染色体通过交配而重组，形成新的染色体，从而产生出新的个体或物种。交配重组是生物遗传和进化过程中的一个主要环节。同样，遗传算法中起核心作用的是遗传操作的交叉算子。交叉操作是指把两个父代个体的部分结构加以替换重组而生成新个体的操作。通过交叉，遗传算法的搜索能力得以飞跃提高。

交叉算子的设计和实现与所研究的问题密切相关，一般要求它既不要太多地破坏个体编码串中表示优良性状的优良模式，又要能够有效地产生出一些较好的新个体模式。另外，交叉算子的设计要和个体编码设计统一考虑。其设计包括如何确定交叉点的位置和如何进行部分基因交换两方面的内容。

算术交叉：

算术交叉是由两个个体的线性组合而产生出两个新的个体。为了能够进行线性组合运算，算术交叉的操作对象一般是由浮点数编码所表示的个体。

算术交叉产生的新个体如下：

$$X_A^{t+1} = \alpha X_B^t + (1 - \alpha) X_A^t \quad (3-18)$$

$$X_B^{t+1} = \alpha X_A^t + (1 - \alpha) X_B^t \quad (3-19)$$

式中： x 为个体； α 为一 0 到 1 的参数，它可以是一个常数，此时所进行的交叉运算称为均匀算术交叉；它也可以是一个由进化代数所决定的变量，此时所进行的交叉运算即为非均匀算术交叉。

算术交叉的主要操作过程

1. 确定两个个体进行线性组合时的系数。
2. 依据上式生成两个新的个体。

纵上所述，本文采用浮点数编码方法，而算术交叉对浮点数编码方法操作比较方便，能够进行线性组合运算，在一定程度上能够产生出性能更优的新个体，可以使算法有较好的稳定性。所以采用算术交叉运算。其中的参数 α 每次交叉运算前随机生成。

8、变异算子的选择：

变异运算是对群体中的个体串的某些基因座上的基因值作变动。尽管只是产生新个体的辅助方法，但它也是必不可少的一个运算步骤，因为它决定了遗传算法的局部搜索能力。在遗传算法中使用变异算子主要有以下两个目的：一是使遗传算法的局部的随机搜索能力。当遗传算法通过交叉算子已接近最优解邻域时，利用变异算子的这种局部随机搜索能力可以加速向最优解收敛。二是使遗传算法可以维持群体的多样性，防止出现早熟现象。交叉算子与变异算子的相互配合，共同完成对搜索空间的全局搜索和局部搜索，从而使得遗传算法能够以良好的搜索性能完成最优化问题的寻优过程。

常用的变异方法如下：

基本位变异：

基本位变异操作是指对个体编码串中以变异概率随机指定的某一位或某几位基因座上的基因值作变异运算。基本位变异操作改变的只是个体编码串中的个别几个基因座上的基因值，并且变异发生的概率也比较小，所以其发挥的作用比较慢，作用的效果也不明显。

均匀变异：

均匀变异操作指分别用符合某一范围内均匀分布的随机数，以某一较小的概率来替换个体编码串中各个基因座上的原有基因值。

均匀变异的具体操作过程如下：

1. 依次指定个体编码串中的每个基因座为变异点。
2. 对每一个变异点，以变异概率从对应基因的取值范围内取一随机数来替代原有基因值。

均匀变异操作特别适合应用于遗传算法的初期运行阶段，它使得搜索点可以在整个搜索空间内自由地移动，从而可以增加群体的多样性，使算法处理更多的模式。

边界变异：

边界变异是均匀变异操作的一个变形算法。在进行边界变异操作时，随机地取基因座的二个对应边界基因值之一去替代原有基因值。

当变量的取值范围特别宽，并且无其他约束条件时，边界变异会带来不好的作用。但它特别适用于最优点位于或接近于可行解的边界时的一类问题。

非均匀变异：

均匀变异操作取某一范围内均匀分布的随机数来替换原有基因值，可使得个体在搜索空间内自由移动。但另一方面，它却便于对某一重点区域进行局部搜索。而非均匀变异不取均匀分布的随机数去替换原有的基因值，而是对原有基因值作一随机扰动，以扰动后的结果作为变异后的新基因值。对每个基因座都以相同的概率进行变异运算之后，相当于整个解向量在解空间中作了一个轻微的变动。非均匀变异的具体操作过程与均匀变异相类似，但它重点搜索原个体附近的微小区域。

非均匀变异可使得遗传算法在其初始运行阶段进行均匀随机搜索，而在其后期运行阶段进行局部搜索，所以它产生的新基因值比均匀变异所产生的基因值更接近于原有基因值。故随着遗传算法的运行，非均匀变异就使得最优解的搜索过程更加集中在某一最有希望的重点区域中。

高斯变异：

高斯变异操作是改进遗传算法对重点搜索区域的局部搜索性能的另一种变异操作方法。

高斯变异操作是在进行变异操作时，用符合均值为 u 、方差为 σ^2 的正态分布的一个随机数来替换原有基因值。

高斯变异的具体操作过程与均匀变异类似。具体实现高斯变异时，符合正态分布的随机数可由一些符合均匀分布的随机数利用公式来近似产生。

从上面的讨论可以看出：基本位变异操作发生的概率也比较小，发挥的作用比较慢，作用的效果不明显。边界变异，非均匀变异和高斯变异主要是改进遗传算法对重点搜索区域的局部搜索性能的变异操作方法。而这里希望遗传算法在运算要有全局的搜索能力，尽量使初期运行阶段的群体要有多样性。所以选用均匀变异操作。

9、终止条件：

遗传算法是一个反复迭代的过程，每次迭代期间，要执行适应度计算，复制，交

叉，变异等操作，直到满足终止条件。

使遗传算法终止的方法有三种：

(1) 规定最大迭代次数 N 。一旦遗传算法达到了规定最大迭代次数 N ，则停止操作，输出结果。

(2) 规定最小的偏差 δ 。对于适应度目标已知的遗传算法，可以用最小偏差 δ 制定终止条件，即：

$$|f_{\max} - f^*| \leq \delta \quad (3-20)$$

式中 f^* 为已知的适应度目标， f_{\max} 为每代最大的适应度。

(3) 观察适应度的变化趋势。在遗传算法的初期，最优个体的适应度以及群体的平均适应度都较小，而随着操作，适应度值增加。到了遗传算法后期，这种增加已趋缓和或停止，一旦这种增加停止，即终止遗传算法。

10、实数编码交叉概率和突变概率的确定：

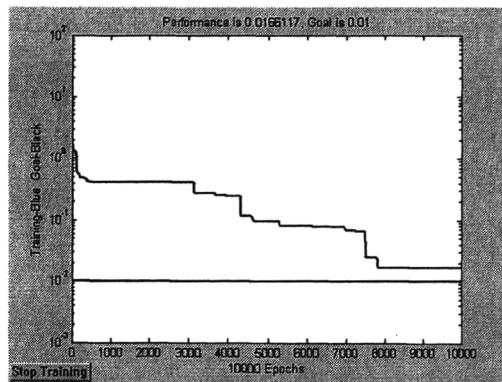
当选择策略确定后影响遗传算法收敛性的主要因素是交叉和变异算子。交叉和变异算子对于遗传算法收敛性的影响在编码方式确定后通过交叉率和变异概率来体现。较大的交叉概率可增强遗传算法开辟新的搜索区域的能力，但高性能的模式遭到破坏的可能性增大，若交叉概率太低，遗传算法搜索可能陷入迟钝状态。变异概率对于保持群体结构多样性，防止过早收敛是一种重要手段。变异概率过大将引起遗传算法的振荡，过小则影响各进化迭代搜索区域的范围。由此可知交叉率和变异率对于遗传算法的收敛性有重要影响。对于交叉概率和突变概率的选择还没有一定的标准，目前常用的参数范围是交叉概率为 0.4—0.99，变异概率范围为 0.005—0.1，为了增强遗传算法的搜索能力，选择较大的交叉概率，为了使遗传算法稳定选取较小的变异概率。最后选定交叉概率为 0.7，变异概率为 0.03。

3.4 遗传算法与神经网络结合的仿真结果

神经网络结构确定后，对于矸石带煤模式识别用遗传算法优化神经网络的初始权值和阈值和不用遗传算法优化神经网络的初始权值和阈值的matlab仿真结果的对比：

对于不用遗传算法优化神经网络的初始权值和阈值即随机给定网络的初始权值和阈值：由于不同的初始权值和阈值可能会导致完全不同的训练结果，所以训练了 10 次

网络，设训练最大步长 10000 次。在训练神经网络时由于不能过于追求较小的误差训练目标，以免过小的误差训练目标值会使网络陷入局部最小点，即过拟合状态，同时也会增加训练时间，甚至导致无法收敛，因此误差训练目标值设为 0.01，完全可以满足跳汰机的控制需要。实际运行发现，由于神经网络的初始权值和阈值是随机选取的，十次仿真结果中有两次是收敛的。部分仿真结果如下：其中的 elapsed_time 为训练网络所需要的时间，横坐标为训练步数，纵坐标为训练误差。

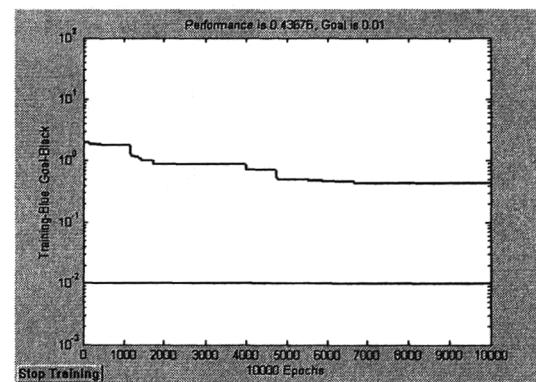


训练时间 elapsed_time = 113.3530

图3—2 随机给定神经网络初始权值和阈值训练结果1

figure 3—2 train result 1 with random beginning

weights and threshold

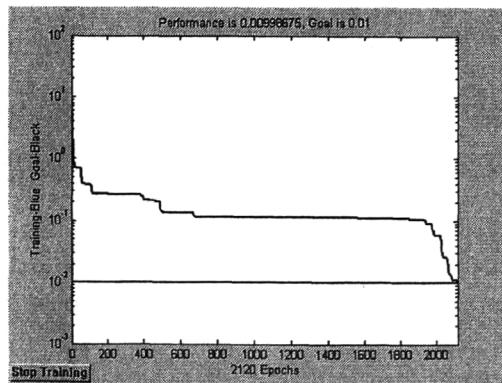


训练时间 elapsed_time = 167.0300

图3—3 随机给定神经网络初始权值和阈值训练结果2

figure 3—3 train result 2 with random beginning

weights and threshold

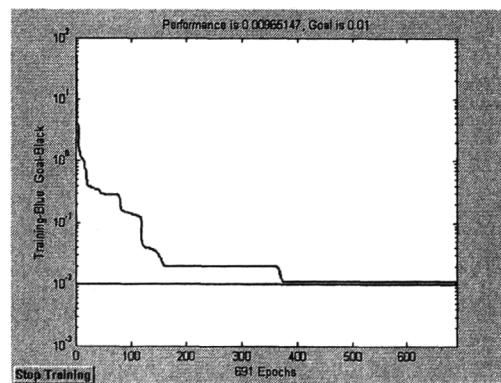


训练时间 elapsed_time = 34.0390

图3—4 随机给定神经网络初始权值和阈值训练结果3

figure 3—4 train result 3 with random beginning

weights and threshold



训练时间 elapsed_time = 11.5960

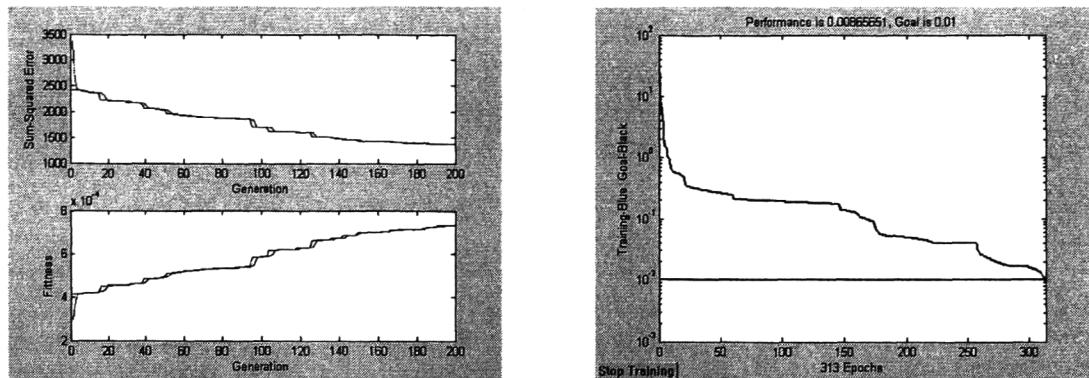
图3—5 随机给定神经网络初始权值和阈值训练结果4

figure 3—5 train result 4 with random beginning

weights and threshold

在采用遗传算法优化神经网络的初始权值和阈值后再用神经网络训练的方法训练网络时，共训练了10次网络，发现有9次网络能较好收敛。下面是部分训练结果。其中的 elapsed_time 为训练网络所需要的时间，横坐标分别为遗传代数和训练步数，纵坐

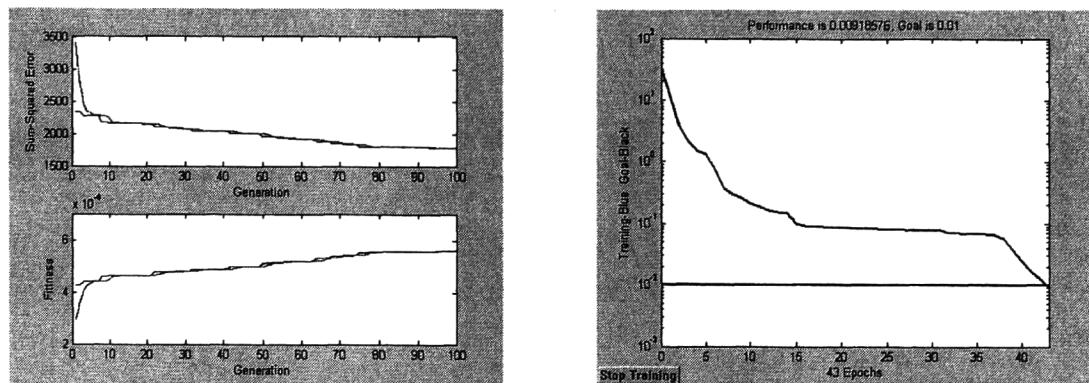
标为训练误差。



训练时间 elapsed_time = 125.7480

图 3—6 遗传最大迭代次数为 200 时训练结果

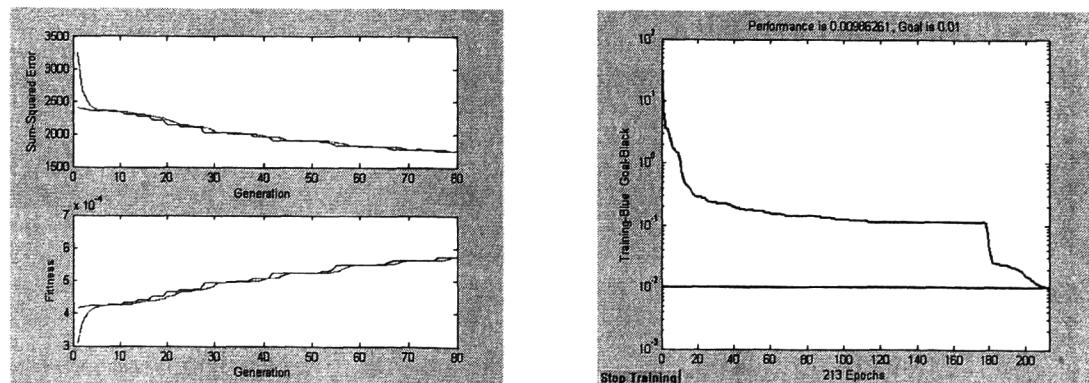
figure 3—6 train results with max inherit generation 200



训练时间 elapsed_time = 18.7070

图 3—7 遗传最大迭代次数为 100 时训练结果

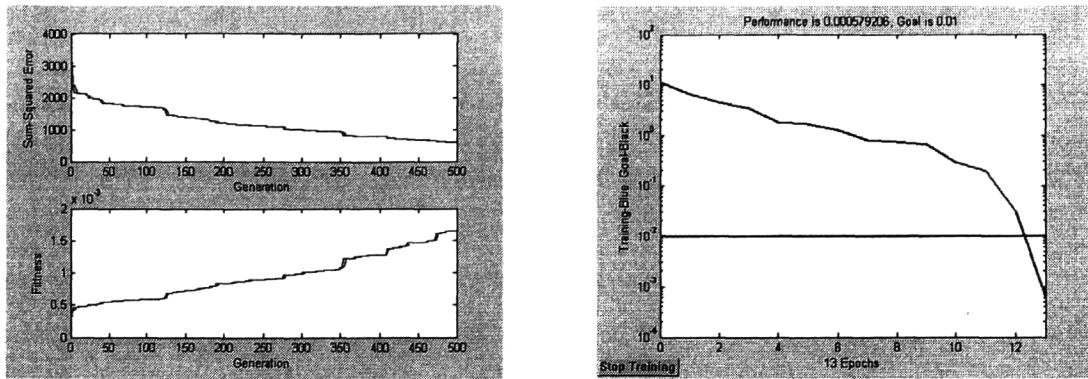
figure 3—7 train results with max inherit generation 100



训练时间 elapsed_time = 29.9030

图 3—8：遗传最大迭代次数为 80 时训练结果

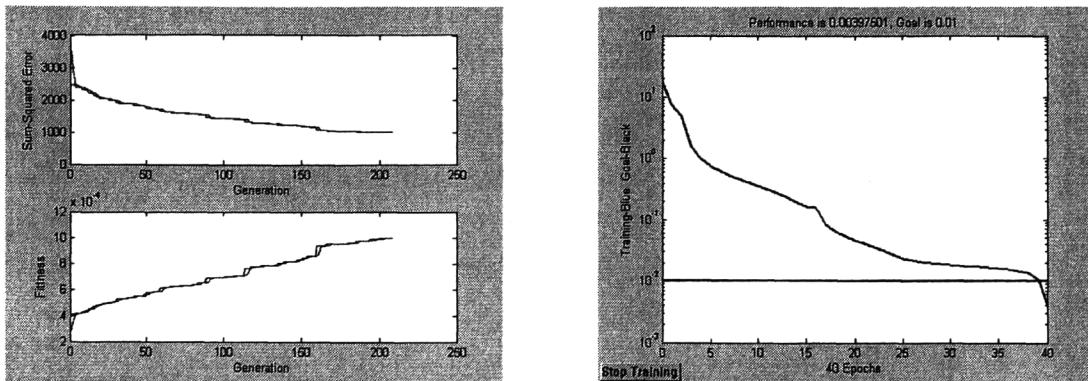
figure 3—8 train results with max inherit generation 80



训练时间 elapsed_time = 219.4950

图3—9 遗传最大迭代次数为500时训练结果

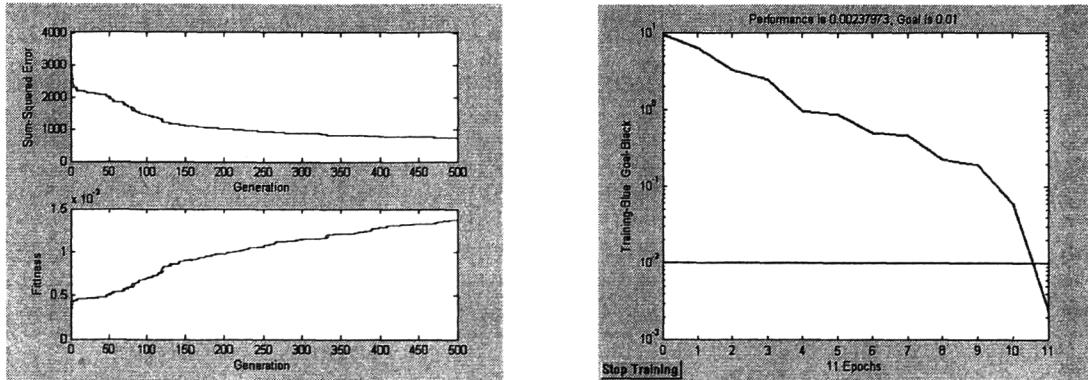
figure 3—9 train results with max inherit generation 500



训练时间 elapsed_time = 101.6100

图3—10 遗传最大迭代次数为500且最大适应度值为0.0001时训练结果

figure 3—10 train results with max inherit generation 500 & max adaptability is 0.0001



训练时间 elapsed_time = 208.5740

图3—11：遗传最大代数为500时训练结果

figure 3—11 train results with max inherit generation 500

通过比较发现：采用遗传算法优化神经网络的初始权值和阈值，使神经网络能比较好地摆脱局部极小，但是网络的计算量变大了。

而随机选择神经网络的初始权值和阈值，网络很容易陷入局部极小，但是如果网络能收敛，训练时间还是比较短的。上面网络收敛时，迭代步数为2120时，仿真时间为34.0390秒。当迭代步数为691时，仿真时间为11.5960秒。

采用遗传算法和神经网络相结合的算法时，仿真时间和选择的遗传算法最大迭代代数有很大的关系，而和遗传算法终止时最终的适应度的大小关系不大。

如果选择了较小的迭代代数，就会使神经网络的迭代步数变大，最终的仿真时间也就比较长同时可能会陷入局部极小，如果选择了较大的遗传算法最大迭代代数，神经网络的迭代步数就会变小，但是遗传算法仿真时间就变得比较长。

如选择遗传最大代数为500时，仿真时间为208.5740秒，选择遗传最大代数为80时，仿真时间为29.9030秒。但是当选择的遗传最大迭代代数为100时，仿真时间为18.7070秒，仿真时间达到了比较小的值。

这表明如果选择了遗传算法合适的最大迭代代数，就会使训练达到比较好的效果。

3.5 基于 RBF 网络的矸石带煤模式识别

BP神经网络虽然是用途最广泛的一种网络，但网络的学习算法存在训练速度慢、由于利用梯度下降算法，在学习阶段可能会陷入局部极小，全局搜索能力差，虽然有一些方法，如上面的遗传算法，虽然可以在某种条件下避免局部最小，但一般都需要巨大的计算量，从而极大地限制了实时应用。而且网络存在难以确定隐含层和隐节点数等缺点。

从理论上而言，RBF网络和BP网络一样可以近似逼近任何连续的非线性函数，两者的主要区别在于各使用不同的作用函数，BP网络中的隐含层节点使用的是sigmoid函数，其函数值在输入空间中无限大的范围内为非零值，而RBF神经网络的作用函数则是局部的，是一种典型的局部逼近神经网络，网络的每个隐含层神经元传递函数都构成了拟合平面的一个基函数，网络也由此而得名。

由于二者的构造本质不同，RBF网络与BP网络相比通常规模较大，但学习速度较快，并且网络的函数逼近能力、模式识别与分类能力都优于后者，但它是在局部范围内而言的。

RBF网络是以函数逼近理论为基础而构造的一类前向网络。这类网络的学习等价于在多维空间中寻找训练数据的最佳拟合平面。即对输入空间的某个局部区域，只有少数几个权值影响网络的输出。而全局逼近神经网络对于每对输入输出数据对，网络的每一个权值均需要调整，从而导致全局逼近网络学习速度很慢，这个缺点对于控制来说常常是不可忽视的。RBF神经网络对于每个输入输出对，只有少量的权值需要调整，从而使局部逼近网络具有学习速度快的优点。

3.5.1 RBF 网络的结构

RBF神经网络通常是一种前向网络。也是由输入层、隐含层和输出层组成，见图4—12。输入层只传递信号到隐含层，隐含层节点由像高斯函数那样的辐射状作用函数构成，隐含层节点中的作用函数(基函数)对输入信号将在局部产生响应，当输入信号靠近基函数的中央范围时，隐含层节点将产生较大的输出，因此这种网络具有良好的局部逼近能力。

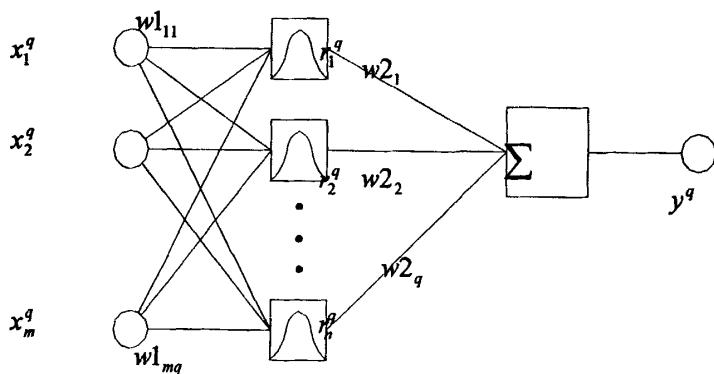


图3—12 径向基网络结构

figure 3—12 RBF NN structure

对于输入层，只负责信息的传输，其输入与输出相同；对于隐含层，每个神经元将自己和输入层神经元相连的连接权值矢量 $w_{1,i}$ （也称为第 i 个隐含层神经元的基函数中心）与输入矢量 X^q [表示第 q 个输入矢量， $X^q = (x_1^q, x_2^q, \dots, x_j^q, \dots, x_m^q)$]之间的距离乘以本身的偏置值 $b_{1,i}$ 作为自己的输入，见图3—13。

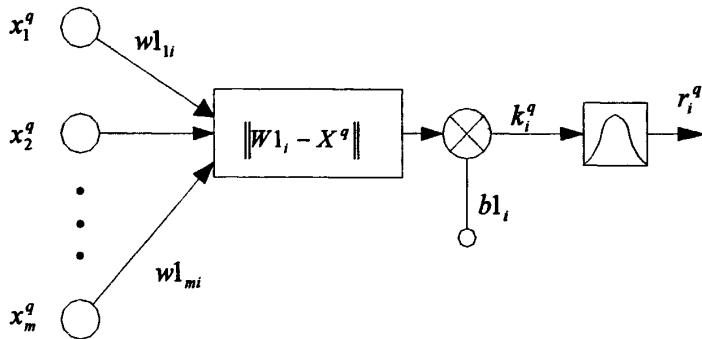


图 3-13 RBF 网络隐含层神经元的输入与输出示意图
figure 3-13 inputs & outputs of RBF of hidden layer neuron

对应输入层第 q 个输入产生的隐含层第 i 个神经元的输入为 k_i^q :

$$k_i^q = \sqrt{\sum_j (w_{1ji} - x_j^q)^2} \times b_{1i} \quad (3-20)$$

隐含层第 i 个神经元的输入经过高斯函数的变换产生的输出为 r_i^q :

$$r_i^q = e^{-(k_i^q)^2} = e^{-\left(\sqrt{\sum_j (w_{1ji} - x_j^q)^2} \times b_{1i}\right)^2} = e^{-(\|w_{1i} - X^q\| \times b_{1i})^2} \quad (3-21)$$

径向基函数的偏置值 $b1$ 可以调节函数的灵敏度，但实际工作中常用另一参数 C (称为扩展常数)。 $b1$ 和 C 的关系在实际应用中有多种确定方法，在MATLAB神经网络函数中 $b1$ 和 C 的关系设置为 $b1_i = 0.8326/C_i$ ，此时隐含层神经元的输出变为:

$$r_i^q = e^{-\frac{(\|w_{1i} - X^q\| \times 0.8326)^2}{C_i}} = e^{-0.8326^2 \left(\frac{\|w_{1i} - X^q\|}{C_i}\right)^2} \quad (3-22)$$

RBF网络的输出节点一般是线性的，即输出层的传递函数为线性函数。输出单元对隐节点输出进行线性加权组合，整个网络执行的是通过非线性基函数的线性组合、从输入空间到输出空间的非线性变换。对应输入层第 q 个输入产生的输出层神经元输出 y^q 为:

$$y^q = \sum_{i=1}^n r_i^q \times w_{2i} \quad (3-23)$$

RBF网络的训练分为2步：第1步为非监督式学习，为了训练输入层与隐含层间的权值 W_1 和基函数的宽度，是一个非线性优化的问题，其求解方法较复杂，目前可选用的学习方式较多，主要有随机选取RBF中心(直接计算法)、无监督学习选取RBF中心(k-均值聚类法)、正交最小二乘法(OLS)等。第2步为监督式学习为了训练隐含层与输出层间的权值 W_2 。是一个线性优化问题，学习速度较快，且有唯一确定的解。

3.5.2 RBF 网络设计

作者采用精确设计函数newrbe和普通设计函数newrb两个函数来创建径向基神经网络。

newrbe函数可用来精确设计径向基神经网络，所谓精确，是指该函数生成的网络对于训练样本数据达到了零误差。函数newrbe的调用形式是

```
net = newrbe(P, T, Spread)
```

其中，P为输入向量；T为期望输出；Spread为径向基函数的分布系数扩展常数，代表径向基函数的宽度，缺省值为1.0。

而newrb函数的调用形式是

```
net=newrb(P, T, GOAL, Spread, MN, DF)
```

其中，P为输入向量；T为期望输出；GOAL为目标误差，MN表示最大神经元个数，其最大值为训练样本个数，DF表示迭代过程的显示频率，Spread为径向基函数的分布系数。

newrbe函数在建立网络时生成的隐含层神经元个数与训练样本数据的个数相同，隐含层神经元的扩展常数为 $0.8632/Spread$ 。

newrb函数利用迭代方法设计径向基神经网络，采用误差纠正RBF学习算法，是对RBF网络的中心、基函数的宽度和输出单元的权值同时从样本中进行监督学习。训练过程中，开始时网络径向基层(即隐含层)的神经元个数为零，通过检查输出误差使网络自动增加神经元，在每次迭代中，网络首先进行仿真并找到对应于最大输出误差的输入样本矢量，然后径向基层添加一个神经元并把权值设为该输入矢量，最后再修改线性层(输出层)的权值以达到最小误差。并检查新网络的误差，重复此过程直到平方和误差下降到目标误差以下或隐含层神经元个数达到最大值时迭代停止。

由上述可知当训练样本数据过多时，利用newrbe函数生成的网络会过于庞大，从而使网络的实用性变差。所以使用newrb函数对网络进行训练，随机抽取200组数据作

为测试数据，其余的400组数据作为训练数据。

newrb 函数的 RBF 网络结构设计中最重要的是目标误差值和径向基函数的分布的取值，它直接影响到网络的拟合和泛化能力。

网络的误差目标是使网络的实际输出和目标输出在训练样本上允许相差的程度。误差太大，网络训练不完全；误差太小、可能出现过度拟合现象，而使其泛化能力太差。前面提到一般的工程问题误差训练目标值选到 $0.01 \sim 10e-5$ 即可，采用缺省的 Spread 值，改变误差值，使误差值从 $10e-5$ 到 0.01 之间变化，经过仿真发现当误差值取 0.004 时，网络的测试误差也达到了满意的结果。所以首先选定网络的误差目标值为 0.004 。

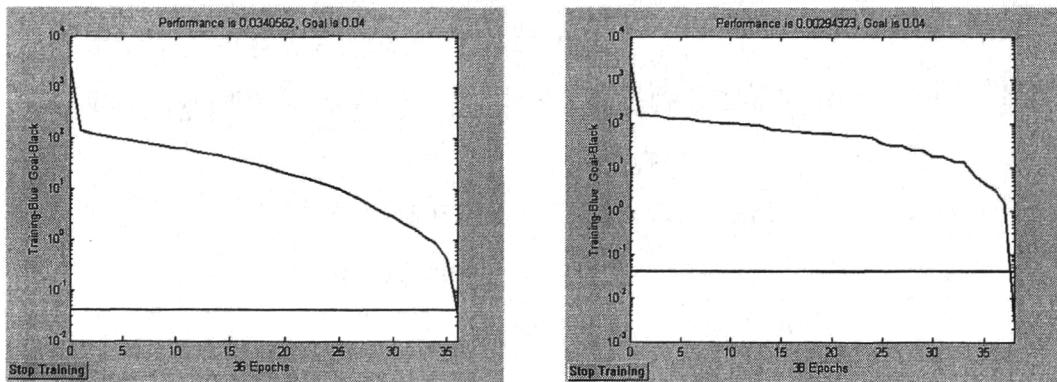
表 3—2 RBF 神经网络仿真结果

table 3—2 emulation results of RBF ANN

Spread	隐含层节点数	训练误差	测试误差
0.2	39	0.000340562	24.3873%
0.4	32	6.62643e-029	21.6153%
0.6	38	1.84593e-028	22.3032%
0.8	37	0.00294323	19.1516%
1.0	37	0.0242084	15.7699%
1.2	37	0.0145366	12.8585%
1.4	36	1.66844e-027	10.4620%
1.6	37	7.13998e-026	13.0194%
1.8	36	2.75581e-026	16.5181%
2.0	37	2.30799e-025	18.3671%
2.2	37	4.06876e-025	20.0030%
2.4	37	0.000462797	23.9882%
2.6	36	0.0196457	24.2263%
2.8	38	2.42864e-024	24.2498e%
3.0	36	8.38925e-02	24.3395%
4.0	36	4.63922e-022	23.7964%
5.0	37	5.80587e-021	24.9158%

然后确定 Spread 参数的值: Spread 越大, 需要非常多神经元来适应这种快速变化, 如果 Spread 过小, 需要许多神经元来适应这种缓慢变化。因此, 需要对 Spread 进行尝试, 以确定最优值。

对 Spread 参数采用从 1 为中心以 0.2 递增和递减用实验的方法确定 Spread 参数。表 3—2 是 RBF 神经网络仿真结果, 图 3—14 是 RBF 网络的部分训练结果。



训练时间 elapsed_time = 1.6020 训练时间 elapsed_time = 1.5130

图 3—14: RBF 网络的部分训练结果

figure 3—14 part train results of RBF NN

从上面的表, 可以看出, 当 Spread 参数的值为 1.4 时, 训练误差达到了 **1.66844e-027**, 同时测试误差是仿真的最小值 **10.4620%**。所以选定 Spread 参数的值为 **1.4**。

通过比较 BP 网络和 RBF 网络的仿真结果, 可以看出: 对于跳汰机的矸石段床层分选模式识别, RBF 网络的学习速度比 BP 网络快, 同时测试误差也比 BP 网络要小些, 尽管网络的隐含层结点数目比较大即网络的规模比较大, 但可以达到比较理想的结果, 测试误差可以达到 10.46%。

3. 6 本章小结

以床层的三个层位的密度值、入洗原煤性质、排料门的开度值和风阀周期等参数作为输入, 而把矸石带煤作为输出, 构建了 BP 和 RBF 人工神经网络。

选用 600 组样本数据用作神经网络的训练和测试: 把数据分为 4 个互不相交的子集, 选其中一个子集作为验证集合, 一个子集作为测试集合, 另外两个子集作为训练集合。

用遗传算法优化 BP 神经网络的初始权值和阈值，遗传算法的终止条件设为最大进化代数 200；对于每个隐含层数目，采用提前终止方法；确定神经网络隐含层节点数为 7；采用浮点数编码的方法；选用均匀变异操作。仿真结果为：

当选择的遗传最大迭代代数为 100 时，仿真时间最小，为 18.7070 秒，而对 RBF 网络，当 Spread 参数的值为 1.4 时，训练误差 $1.66844e-027$ ，同时测试误差最小。

通过比较 BP 网络和 RBF 网络的仿真结果，可以看出：对于跳汰机研石段床层模式识别，RBF 网络的学习速度比 BP 网络快，同时测试误差也比 BP 网络要小些，尽管网络的隐含层结点数目比较大，但可以达到比较理想的结果，测试误差可以达到 10.46%。

本章的研究说明，人工神经网络可以用来构建跳汰机分层、排料一体化模型，该模型可以是 BP 网络也可以是 RBF 网络，并且模型中包含了跳汰机分层和排料的所有信息：诸如风阀周期、排料门开度、床层分层状态、入洗原煤性质等，同时，模型的输出直接是研石带煤这一最终控制目标，即对于跳汰机这一复杂系统而言，人工神经网络模型可以通过简单的实验来获得并能表现出较好的功能。

然而，可以看出，尽管采用了遗传算法优化 BP 神经网络的初始权值和阈值等方法，测试误差仍高达 10% 以上，如果不尽可能降低测试误差，势必影响其使用效果。

逐渐提高网络泛化能力、降低测试误差，尽管可以在以后的应用过程中通过采集典型样本不断对模型在线修正实现，但是毕竟样本的采集具有一定的困难和复杂性，如何提高小样本情况下模型的适应性成为必然的研究对象，在此情况下，研究支持向量机在跳汰系统模式识别中的应用可能及效果就显得尤为必要了。

第四章 基于支持向量机的跳汰过程状态识别

鉴于利用GA+ANN对跳汰机床层状态识别存在误差大(10%)、网络构建复杂和需要大量的训练样本等问题。特别是后者，生产过程中要获取大量样本需要很长的时间，在这段时间内过程参数(例如入洗原煤性质等)可能发生很大变化，所取得的样本就有很大差别，不可能全部用作训练样本，这就给上述神经网络的应用产生困难。于是，作者考虑当出现这种情况时，采用只需要有限样本的基于支持向量机跳汰过程状态识别方法。

支持向量机方法及其应用于跳汰过程辨识的几个主要优点：

1. 它是专门针对有限样本情况的，其目标是得到现有信息下的最优解而不仅仅是样本数趋于无穷大时的最优值；

跳汰过程辨识不可能采集数千个甚至更多的样本，原因之一是采集过程复杂、大量散状的煤炭样品其代表性就是一个非常值得探究的问题；原因之二是即使跳汰机控制参数不变，入洗原煤却在时时刻刻改变着，采集的样本就必须加入入洗原煤性质这样的参数，使样品的复杂性大大增加。

而支持向量机的第一个优点正好适应了这种情况。

2. 算法最终将转化成为一个二次型寻优问题，从理论上说，得到的将是全局最优点，解决了在神经网络方法中无法避免的局部极值问题；

局部极值问题在跳汰过程中的表现一方面表现在控制变量适应性差，不能像有经验的人工操作那样可以进行大幅度的调整；另一方面表现在对目标函数(矸石带煤)反映的准确性上面，即：矸石带煤在某种情况下是有超常值出现的，例如，在出现机械故障甚至人为故障致使床层出现排空时即是如此。支持向量机的二次型寻优问题，即可有效避免局部极值的发生。

3. 算法将实际问题通过非线性变换转换到高维的特征空间(Feature Space)，在高维空间中构造线性判别函数来实现原空间中的非线性判别函数，特殊性质能保证机器有较好的推广能力，同时它巧妙地解决了维数问题，其算法复杂度与样本维数无关；在SVM方法中，只要定义不同的内积函数，就可以实现多项式逼近、贝叶斯分类器、径向基函数(Radial Basic Function或RBF)方法、多层感知器网络等许多现有学习算法。

这一特点的有效应用，应该可以使跳汰系统辨识推广能力得到提高，从而使智能控制得以实现。

4.1 统计学习理论的基本思想

统计学习理论的基本思想就是通过让机器对大量信息的认知，进行内部学习、分析归类、记忆存储和随时更新等处理过程，达到做出预期输出的目的，基本结构框图见图 4—1。

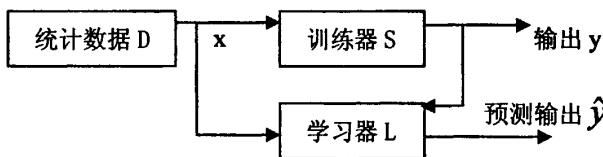


图 4—1 统计学习理论基本结构框图

figure 4—1 basic structure of statistical learning theory

统计学习理论是研究小样本统计估计和预测的理论，其核心内容主要包括四个方面：

- (1) 经验风险最小化准则下统计学习一致性的条件；
- (2) 在这些条件下关于统计学习方法推广性的界；
- (3) 在这些界的基础上建立的小样本归纳推理准则；
- (4) 实现新的准则的实际算法。

其中，最有指导性的理论结果是推广性的界，与此相关的一个核心概念是VC维(VC Dimension)。它是描述函数集或学习机器的复杂性或者说是学习能力(Capacity of the machine)的一个重要指标，在此概念基础上发展出了一系列关于统计学习的一致性(Consistency)、收敛速度、推广性能(Generalization Performance)等重要结论^[81]。统计学习理论是建立在一套较坚实的理论基础之上的，为解决有限样本学习问题提供了一个统一的框架。它能将很多现有方法纳入其中，有望帮助解决比如神经网络结构选择问题、局部极小点问题等。

4.1.1 VC 维

在模式识别方法中 VC 维的直观定义是：对于一个指示函数集，如果存在 h 个样本

能够被函数集里的函数按照所有可能的 2^h 种形式分开，则称函数集能够被 h 个样本打散。函数集的 VC 维就是能够打散的最大样本数目 h 。如果对任意的样本数，总有函数能打散它们，则函数集的 VC 维就是无穷大。有界实值函数的 VC 维可以通过使用阈值将函数转化成指示函数来定义。它描述了组成学习模型的函数集合的容量，也就是说刻化了函数集合的学习能力。VC 维越大，函数集合越大，其相应的学习能力就越强。

VC 维是统计学习理论的核心概念，它是目前为止对函数集学习性能的最好描述指标。一般而言，VC 维越大，学习机器的学习能力就越强，但学习机器也越复杂(容量越大)。遗憾的是，目前还没有通用的关于计算任意函数集的 VC 维理论，只有对一些特殊函数集的 VC 维可以准确知道。例如： n 维实数空间中线性分类器和线性实函数的 VC 维是 $n + 1$ 。而对于一些比较复杂的学习机器(比如神经网络)，其 VC 维除了与函数集的选择有关外，通常也受学习算法的影响。因此要确定其 VC 维将更加困难。对于给定的学习函数集，如何用理论或实验的方法计算其 VC 维是当前统计学习理论研究中有待解决的一个问题。

4.1.2 推广性的界与结构风险最小化原理

对于各种类型的函数集，统计学习理论系统地研究了经验风险和实际风险之间的关系，也即推广性的界。它是分析学习机器的性能和发展新的学习算法的重要基础。根据统计学习理论中关于函数集推广性界的结论，假设输出变量 Y 与输入变量 X 之间存在某种对应的依赖关系，即为未知概率分布 $P(X, Y)$ ， $P(X, Y)$ 反映了某种知识。学习问题可以概括为：根据 l 个独立同分布的观测样本，学习到一个假设 $H = f(x, w)$ 作为预测函数，其中 w 是广义参数。它对 $P(X, Y)$ 的期望风险 $R(w)$ 是(即统计学习的实际风险)：

$$R(w) = \frac{1}{2} \int |y - f(x, w)| dP(X, Y) \quad (4-1)$$

而 $E[(y_i - f(x_i, w))] = \sum_{i=1}^l [(y_i - f(x_i, w))] \cdot P(X, Y = y_i)$ 对观测样本上产生的风险

$R_{emp}(w)$ 被称为经验风险(学习的训练误差)：

$$R_{emp}(w) = \frac{1}{2l} \sum_{i=1}^l |y_i - f(x_i, w)| \quad (4-3)$$

首先 $R_{emp}(w)$ 和 $R(w)$ 都是 w 的函数，传统概率论中的定理只说明了（在一定条件下）当样本趋于无穷多时 $R_{emp}(w)$ 将在概率意义上趋近于 $R(w)$ ，却没有保证使 $R_{emp}(w)$ 最小的点也能够使 $R(w)$ 最小。

根据统计学习理论中关于函数集的推广性的界的结论，对于两类分类问题中的指示函数集 $f(x, w)$ 的所有函数（当然也包括使经验风险最小的函数），经验风险 $R_{emp}(w)$ 和实际风险 $R(w)$ 之间至少以不下于 $1 - \eta$ ($0 \leq \eta \leq 1$) 的概率存在这样的关系：

$$R(w) \leq R_{emp}(w) + \phi(h/l) \quad (4-4)$$

$$\phi(h/l) = \sqrt{\frac{h(\ln(2l/h+1) - \ln(\eta/4))}{l}} \quad (4-5)$$

其中 h 是函数 $H = f(x, w)$ 的 VC 维， l 是样本数。

4.1.3 结构风险最小化基本想法

如果要求风险最小，就需要不等式（4—4）中的两项相互权衡，共同趋于极小；另外，在获得的学习模型经验风险最小的同时，还需要使 VC 维 h 值尽可能小，这样对未来样本才会有较好的推广能力，这就是结构风险最小化准则的基本思想。

一般的学习方法（如神经网络）是基于 $R_{emp}(w)$ 最小，满足对已有训练数据的最佳拟和，在理论上可以通过增加算法（如神经网络）的规模使得 $R_{emp}(w)$ 不断降低以至为 0。

但是，这样就使得算法（神经网络）的复杂度增加，VC 维 h 增加，从而 $\phi(h/l)$ 增大，导致实际风险 $R(w)$ 增加，实际是学习算法的过度拟合。

4.2 支持向量机基本方法

支持向量机（Support Vector Machine或SVM）方法是建立在统计学习理论的VC维理论和结构风险最小原理基础上的学习方法。根据结构风险最小化原则，SVM通过最优超平面的构造，在固定学习机经验风险的条件下最小化VC维h。

SVM是从线性可分情况下的最优分类面发展而来的，基本思想可用图4—2的二维情况加以说明。图中，三角形点和圆形点代表两类样本，H为分类线，H₁、H₂分别为过各类中离分类线最近的样本且平行于分类线的直线，它们之间的距离叫做分类间隔（margin）。最优分类线即是要求分类线不但能将两类正确分开（训练错误率为0），而且使分类间隔最大。

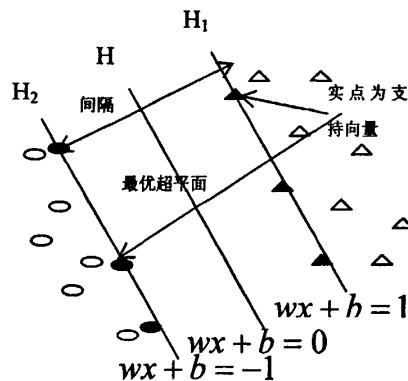


图 4—2 最优超平面

figure 4—2 optimum super plane

设有n个线性可分样本 x_i 及其所属类别 y_i : $(x_i, y_i), x \in R^d, y_i \in \{+1, -1\}, i=1, \dots, n$

超平面方程 $w \cdot x + b = 0$, w 为权向量, b 为阈值, 能将两类样本正确区分, 并使分类间隔最大的优化问题表示如下:

在式:

$$y_i[(w \cdot x_i) + b] - 1 \geq 0, i=1, \dots, n \quad (4-6)$$

的约束下求

$$\varphi(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w \cdot w) \quad (4-7)$$

的最小值。满足式(4—6)且使 $\frac{1}{2}\|w\|^2$ 最小的分类面即最优分类面, H_1 、 H_2 上的训练样本点即支持向量。

利用Lagrange 优化方法可以把上述最优分类面问题转化为其对偶问题, 即: 在约束条件

$$\sum_{i=1}^n y_i \alpha_i = 0 , \quad (4-8)$$

和 $\alpha_i \geq 0 \quad i=1, \dots, n \quad (4-9)$

下对 α_i 求解下列函数的最大值:

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (4-10)$$

α_i 为原问题中与每个约束条件(4—8)对应的Lagrange 乘子。这是一个不等式约束下二次函数寻优的问题, 存在唯一解。容易证明, 解中将只有一部分(通常是少部分) α_i 不为零, 对应的样本就是支持向量。解上述问题后得到的最优分类函数是

$$f(x) = \text{sgn}\{(w \cdot x) + b\} = \text{sgn}\{\sum_{i=1}^n \alpha_i^* y_i (x_i \cdot x) + b^*\} \quad (4-11)$$

式中的求和实际上只对支持向量进行, b^* 是分类阈值, 可以用任一个支持向量(满足(4—8)中的等号)求得, 或通过两类中任意一对支持向量取中值求得。

对非线性问题, 可以通过非线性变换转化为某个高维空间中的线性问题, 在变换空间求最优分类面。这种变换可能比较复杂, 因此这种思路在一般情况下不易实现。但是注意到, 在上面的对偶问题中, 不论是寻优目标函数(4—10)还是分类函数(4—11)都只涉及训练样本之间的内积运算(x_i, x_j)。那么, 设有非线性映射 $\phi: R^d \rightarrow H$ 将输入空间的样本映射到高维(可能是无穷维)的特征空间 H 中。当在特征空间 H 中构造最优超平面时, 训练算法仅使用空间中的点积, 即 $\phi(x_i) \cdot \phi(x_j)$, 而没有单独的 $\phi(x_i)$ 出现。因此, 如果能够找到一个函数 K 使得 $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$, 这样, 在高维空间实际上就只需进行内积运算, 而这种内积运算是可以用原空间中的函数实现的, 甚至没有必要知道变换 ϕ 的形式。根据泛函的有关理论, 只要一种核函数 $K(x_i, x_j)$ 满足Mercer

条件，它就对应某一变换空间中的内积。

因此，在最优分类面中采用适当的内积函数 $K(x_i, x_j)$ 就可以实现某一非线性变换后的线性分类，而计算复杂度却没有增加，此时目标函数（4—10）变为：

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \quad (4-12)$$

而相应的分类函数也变为

$$f(x) = \text{sgn}\left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b^*\right) \quad (4-13)$$

即为支持向量机。

这一特点提供了可能导致的“维数灾难”问题^[81]的方法：在构造判别函数时，不是对输入空间的样本作非线性变换，然后在特征空间中求解；而是先在输入空间比较向量（例图4—1最优分类面如求点积或是某种距离），对结果再作非线性变换。这样，大的工作量将在输入空间而不是在高维特征空间中完成。SVM分类函数形式上类似于一个神经网络，输出是中间节点的线性组合，每个中间节点对应一个支持向量（如图4—3所示）。

函数 K 称为点积的卷积核函数，它可以看作在样本之间定义的一种距离。

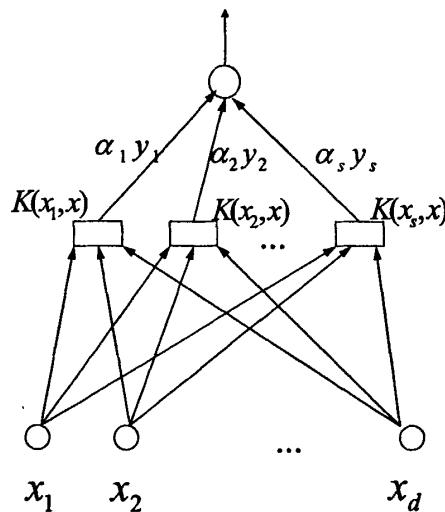


图 4—3 支持向量机示意图

figure 4—3 sketch map of SVM

其中权值为: $\alpha_i y_i$, 输入向量为: $X = (x_1, x_2, x_3, \dots, x_d)$, 输出为:

$$y = \text{sgn}(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b).$$

其中核函数在支持向量机中起了重要作用, 它将高维特征空间中的非线性运算转换为原输入空间的核函数计算, 避免了“维数灾难”和确定非线性函数的形式和参数等复杂问题。不过核函数的选择却并不是容易的一件事, 根据 Mercer 定律, 核函数是满足 Mercer 条件的任意对称函数。常用的核函数有线性核函数、多项式核函数、径向基核函数和 Sigmoid 核函数四种。

(1) 线性内核:

$$K(x_i, x_j) = x_i \cdot x_j \quad (4-14)$$

(2) 径向基函数内核

$$K(x_i, x_j) = \exp\left\{-\frac{|x_i - x_j|^2}{\sigma^2}\right\} \quad (4-15)$$

当采用径向基核函数时, 支持向量机的实现等价为一个径向基函数网络, 但这里每一个基函数中心对应于一个支持向量, 它们及输出权系数均是由算法自动确定的。

(3) 多项式内核

$$K(x_i, x_j) = [(x_i \cdot x_j) + 1]^q \quad (4-16)$$

此时得到的支持向量机是一个 q 阶多项式分类器。其中 q 是由用户决定的参数。

(4) S 形内核

$$K(x_i, x_j) = \tanh(v(x_i \cdot x_j) + c) \quad (4-17)$$

此时得到的支持向量机则等价为一个单隐含层前向网络, 但隐含层结点数、各层权系数均由算法自动确定, 并且能达到最优分类效果。特别是对高维样本有积极意义。

显然, 上面的方法在保证训练样本全部被正确分类, 即经验风险 R_{emp} 为 0 的前提下,

通过最大化分类间隔来获得最好的推广性能。如果希望在经验风险和推广性能之间求得某种均衡，可以通过引入正的松弛因子 ξ_i 来允许错分样本的存在。这时，约束(4—8)变为

$$y_i[(w \cdot x_i) + b] - 1 + \xi_i \geq 0, i = 1, \dots, n \quad (4-18)$$

而在目标——最小化 $\frac{1}{2} \|w\|^2$ 中加入惩罚项 $C \sum_{i=1}^n \xi_i$ ，这样，wolf对偶问题可以

写成在

$$\sum_{i=1}^n y_i \alpha_i = 0$$

和 $0 \leq \alpha_i \leq C, i = 1, \dots, n$ 条件下，

使

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \quad (4-19)$$

取最大值。

这就是SVM方法的最一般的表述。为了方便后面的陈述，把对偶问题的最优解做一些推导。

定义：

$$\omega(\alpha) = \sum_i \alpha_i y_i \phi(x_i) \quad (4-20)$$

$$F_i = \omega(\alpha) \cdot \phi(x_i) - y_i = \sum_j \alpha_j y_j K(x_i, x_j) - y_i \quad (4-21)$$

对偶问题的Lagrange 函数可以写成：

$$L = \frac{1}{2} \omega(\alpha) \cdot \omega(\alpha) - \sum_i \alpha_i - \sum_i \delta_i \alpha_i + \sum_i u_i (\alpha_i - C) - \beta \sum_i \alpha_i y_i \quad (4-22)$$

KKT条件为：

$$\frac{\partial L}{\partial \alpha_i} = (F_i - \beta)y_i - \delta_i + u_i = 0 \quad (4-23)$$

$$\delta_i \alpha_i = 0 \quad \text{且} \quad \delta_i \geq 0$$

$$u_i(\alpha_i - C) = 0 \quad \forall i$$

最优化求解得到的 α_i 中, α_i 可能是: $\alpha_i = 0$, $0 \leq \alpha_i \leq C$, $\alpha_i = C$ 。

若 $\alpha_i = 0$ 时, 则 $\delta_i \geq 0$ $u_i = 0$ 由此 $(F_i - \beta)y_i \geq 0$

若 $0 \leq \alpha_i \leq C$ 时, 则 $\delta_i = 0$ $u_i = 0$ 由此 $(F_i - \beta)y_i = 0$

若 $\alpha_i = C$ 时, 则 $\delta_i = 0$ $u_i \geq 0$ 由此 $(F_i - \beta)y_i \leq 0$

由于 KKT 条件是最优解应满足的充要条件, 所以目前提出的一些算法几乎都是以是否违反 KKT 条件作为迭代策略的准则。

4.3 支持向量机算法

尽管 SVM 方法有较好的理论基础和现在已经在一些领域的应用中表现出来了优秀的推广性能, 但是标准支持向量机算法在计算上存在着一些问题, 如训练算法速度慢、算法复杂而难以实现以及检测阶段运算量大等。传统的利用标准二次型优化技术解决对偶问题的方法可能是训练算法慢的主要原因: SVM 方法需要计算和存储核函数矩阵, 当样本点数目较大时, 需要很大的内存, 例如, 当样本点数目超过 4000 时, 存储核函数矩阵需要多达 128 兆内存; 另外, SVM 在二次型寻优过程中要进行大量的矩阵运算, 多数情况下, 寻优算法是占用算法时间的主要部分。

近年来人们针对方法本身的特点提出了许多算法来解决对偶寻优问题。大多数算法的一个共同的思想就是循环迭代: 将原问题分解成为若干子问题, 按照某种迭代策略, 通过反复求解子问题, 最终使结果收敛到原问题的最优解。

同时, 关于 SVM 算法本身的改进和算法的实际应用, 都陆续提了出来。目前支持向量机的变形算法主要有 C-SVM 系列、V-SVM 系列、One-class SVM、 ε -SVM 系列、WSVM (weighted SVM) 和 LS-SVM (least-square SVM) 等算法。这些变形算法主要是通过增加函数项, 变量或系数等方法使公式变形, 产生出各种有某一方面优势或者一定应用范围的算法。下面对几种算法进行必要的探讨。

一、 ε -Support Vector Regression (ε -SVR) 算法

设有数据 $\{(x_1, z_1), \dots, (x_l, z_l)\}$, 其中 $x_i \in R^n$ 为输出向量, $z_i \in R^l$ 为输出向量, 则标准的 ε -SVR 算法形式为:

$$\text{在 } \omega^T \phi(x_i) + b - z_i \leq \varepsilon + \xi_i ,$$

$$z_i - \omega^T \phi(x_i) - b \leq \varepsilon + \xi_i^*,$$

$\xi_i^*, \xi_i \geq 0, i = 1, \dots, l$ 条件下

$$\text{求 } \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l \xi_i + C \sum_{i=1}^l \xi_i^* \text{ 的最小值。}$$

采用拉格朗日乘子法求解这个具有线性约束的二次规划问题，得到的对偶问题为：

$$\sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0,$$

$0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, l$ 条件下使

$$\frac{1}{2} (\alpha - \alpha^*)^T Q (\alpha - \alpha^*) + \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l z_i (\alpha_i - \alpha_i^*) \text{ 取最小值。}$$

$$\text{这里 } Q_{ij} = K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$$

决策函数变为：

$$\sum_{i=1}^l (-\alpha_i + \alpha_i^*) K(x_i, x) + b \quad (4-24)$$

二、 ν -Support Vector Regression (ν -SVR) 算法

设有数据 $\{(x_1, z_1), \dots, (x_l, z_l)\}$ ，其中 $x_i \in R^n$ 为输出向量， $z_i \in R^1$ 为输出向量，则标准的 ε -SVR 算法形式为：

$$\text{在 } \omega^T \phi(x_i) + b - z_i \leq \varepsilon + \xi_i, \quad ,$$

$$z_i - (\omega^T \phi(x_i) - b) \leq \varepsilon + \xi_i^*, \quad ,$$

$\xi_i^*, \xi_i \geq 0, i = 1, \dots, l \quad \varepsilon \geq 0$ 条件下

$$\text{使 } \frac{1}{2} \|\omega\|^2 + C \left(\nu \varepsilon + \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*) \right) \text{ 取最小值。}$$

其对偶问题为：

$$\text{在 } e^T (\alpha - \alpha^*) = 0, e^T (\alpha + \alpha^*) \leq C\nu,$$

$0 \leq \alpha_i, \alpha_i^* \leq C/l, i = 1, \dots, l$ 条件下

$$\text{使 } \frac{1}{2}(\alpha - \alpha^*)^T Q(\alpha - \alpha^*) + z^T(\alpha_i - \alpha^*) \quad \text{取最小值。}$$

这里，可以认为把 C/l 看成参数 C ，这样问题就变为在

$$e^T(\alpha - \alpha^*) = 0, \quad e^T(\alpha + \alpha^*) \leq Clv,$$

$$0 \leq \alpha_i, \quad \alpha_i^* \leq C, \quad i=1, \dots, l \quad \text{条件下}$$

$$\text{使 } \frac{1}{2}(\alpha - \alpha^*)^T Q(\alpha - \alpha^*) + z^T(\alpha_i - \alpha^*) \quad \text{取最小值。}$$

决策函数同 $\varepsilon-SVR$ 一样为：

$$\sum_{i=1}^l (-\alpha_i + \alpha_i^*) K(x_i, x) + b \quad (4-25)$$

三、最小二乘支持向量机 (least-square SVM) 算法

最小二乘支持向量机是基于支持向量机方法的一种改进算法。它是支持向量机在二次损失函数下的一种形式，通过构造损失函数将原支持向量机中算法的二次寻优变为求解线性方程，其求解速度较快，在函数估计、逼近和系统建模中得到了广泛应用。

在最小二乘支持向量机算法中，优化指标采用了平方项，只有等式约束，而没有不等式约束，从而推出不同的一系列的等式约束，其问题表示为在

$$y_i(\omega^T \phi(x_i) + b) = 1 - \xi_i \quad i=1, \dots, l \text{ 的约束下}$$

$$\text{使 } \frac{1}{2}\|\omega\|^2 + \frac{1}{2}r \sum_{i=1}^l \xi_i^2 \quad \text{取最小值。}$$

可得到线性方程组

$$\begin{bmatrix} 0 & y^T \\ y & Q + r^{-1}I \end{bmatrix}_{(l+1) \times (l+1)} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ e \end{bmatrix} \quad (4-26)$$

其中 $e \in R^{(l)}$ 为元素为 1 的向量， $I \in R^{(l \times l)}$ 为单位阵， $a = [a_1, a_2, \dots, a_l]^T \in R^{(l)}$ ，

$y = [y_1, y_2, \dots, y_l]^T \in R^{(l)}$ ， $Q = [q_{ij}]_{l \times l}$ ， $q_{ij} = y_i y_j K(x_i, x_j)$ 。式 (4-26) 用最小二乘即可解。在最小二乘支持向量机算法中，将二次规划问题转变成线性方程组的求解，简化了计算复杂性。

4.4 基于支持向量机的跳汰过程状态识别

采用支持向量机求解模式识别问题时需要选择一个合适的核函数。尽管只要满足 Mercer 条件的函数在理论上都可选为核函数，但是不同的核函数，其仿真的结果可能完全不同。因此，针对跳汰过程状态识别这一特定问题，选择什么样的核函数是至关重要的。而且，即使选择了某一类核函数，其相应的参数（如多项式的阶次、径向基函数的宽度）对仿真结果也有影响。因此，分别用线性函数，多项式函数，径向基函数和 S 形函数作为支持向量机的核函数对数据进行仿真，以探求最合适的核函数。

本文把前述 600 组数据分为两份，其中一份为 400 组，用来对支持向量机进行训练，另外的 200 组数据用来检验训练的结果。

4.4.1 基于 ε -SVR 算法和 ν -SVR 算法的仿真

对于 ε -SVR 方法采用不同的核函数，选取不同的参数对跳汰过程进行仿真比较。用线性函数作为支持向量机核函数时，改变损失函数 ε 的值、允许的终止判据的值和参数 C 的值，发现这 3 个参数对训练误差和测试误差的影响都很小，训练误差为 2.734%，测试数据的相对误差为 11.23%。

表 4-1 采用多项式核函数预测结果

table 4-1 predict results of polynomial kernel

Q 的值	1	2	3	4	5
训练误差	2.872%	1.356%	3.642%	2.546%	3.634%
测试误差	14.57%	8.717%	11.43%	18.57%	20.35%

用多项式函数作为支持向量机核函数时，改变损失函数 ε 的值、允许的终止判据的值和参数 C 的值，发现这 3 个参数对训练误差和测试误差的影响也都很小，当改变 q 的值，从 1 逐渐增加，发现随着 q 的增加训练误差逐渐减小，但增加到 2 以后测试误差开始增加了（见表 4-1 所示）。

用径向基函数作为支持向量机核函数时，改变 σ^2 的值，仿真结果和 σ^2 值关系不大，但随着参数 C 的增加，训练误差有所减少，增加到 3 以上时测试误差开始增加。

表 4—2 采用径向基核函数预测结果

table 4—2 predict results of RBF kernel

C 的值	1	2	3	4	5
训练误差	1.576%	0.7147%	1.234%	3.557%	4.352%
测试误差	11.12%	8.08%	7.143%	10.42%	13.52%

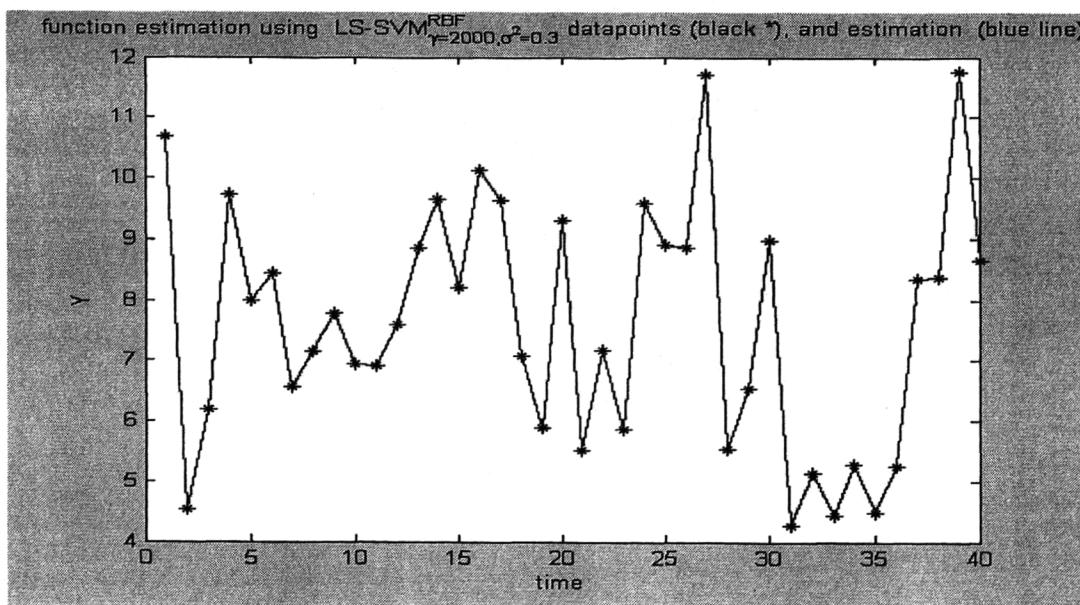
采用 S 型内核时发现测试误差和训练误差都比较大，在 20% 左右。

采用 ν -SVR 算法时也采用不同的核函数，发现仿真结果和 ε -SVR 算法有相似的结果。

4.4.2 基于最小二乘支持向量机(Least-Square SVM)算法的仿真

对于最小二乘SVM算法，采用不同核函数进行仿真，发现采用线性函数、多项式函数和S形函数作为支持向量机核函数时，训练误差和测试误差都比较大，而采用径向基函数作为支持向量机核函数时训练误差比较小，而且测试误差更小。

从图 4—4 可以看出训练误差已经很小了，图中 elapsed_time 为 matlab 训练网络所需要的时间，秒。表 4—3 是预测结果。



预测时间 elapsed_time = 0.7343

图 4—4 采用径向基核函数预测结果

figure 4—4 predict results of RBF kernel

表 4—3 采用径向基核函数预测结果

table 4—3 predict results of RBF kernel

σ^2 的值	0.1	0.2	0.3	0.4	0.5
测试误差	8.743%	8.027%	7.658%	7.724%	8.308%

由上述分析可以看出，采用神经网络为跳汰机建模时，需要根据经验设计网络的结构，而用支持向量机建模时，可以自动确定支持向量机模型的结构。而且不会陷入局部极小。通过仿真比较，神经网络的测试误差最小达到了 **10.46%**。而支持向量机通过选取合适的核函数或方法会使测试误差达到 **7.143%**。在本文的问题上，支持向量机表现出了更加优秀的性能。

4.5 本章小结

建立在坚实的数学基础之上的统计学习理论开始受到越来越广泛重视的今天，鉴于利用GA+ANN对跳汰机床层状态识别仍存在10%的误差和网络构建复杂、样本采集困难等缺点，以及以统计学习理论为基础的支持向量基的许多优异性能，研究SVM在跳汰床层状态识别上应用显得意义重大。

把已取得的跳汰过程的600组数据分为两份，其中一份为400组，用来对支持向量机进行训练，另外的200组数据用来检验训练的结果。采用 ε -SVR 算法时：

用线性函数作为支持向量机核函数时，仿真显示：训练误差为 **2.734%**，测试误差为 **11.23%**。

用多项式函数作为支持向量机核函数时，仿真显示：训练误差为 **1.356%**，测试误差为 **8.717%**。

用径向基函数作为支持向量机核函数时，仿真显示：训练误差为 **1.234%**，测试误差为 **7.143%**。

采用 S 形内核时发现测试误差和训练误差都比较大，在 20% 左右。

采用径向基函数作为支持向量机核函数时训练误差比较小，而且测试误差更小。

采用神经网络为跳汰机建模时，需要根据经验设计网络的结构，而用支持向量机建模时，可以自动确定支持向量机模型的结构，而且不会陷入局部极小。通过仿真比较，神经网络的测试误差最小达到 **10.46%**，而支持向量机通过选取合适核函数或方法会使测试误差达到 **7.143%**。在本文的问题上，支持向量机表现出了更加优秀的性能。

第五章 基于神经网络预测控制的跳汰风阀控制研究

5.1 MPC 应用现状及跳汰预测模型构建思路

模型预测控制 MPC(Model Predictive Control)，对于预测模型与被控对象实际输出之间的偏差，采用反馈校正环节进行修正，即在每一采样时刻，用实测偏差对模型预测的未来输出进行修正，并按修正后的预测输出进行滚动优化，计算控制规律。

线性系统的预测控制问题已解决，但对于非线性系统，由于被控对象的非线性结构未知，且不能充分描述，因此预测模型对其输出难以做出精确的预报，从而可能导致控制失败。

本文前面提到的神经网络和支持向量机方法都具有极强的非线性逼近能力等优点，因此，神经网络和支持向量机方法用于预测控制系统可以很好解决非线性系统预测控制。

基于模型的预测控制研究方向主要有如下两方面：首先是利用神经网络或支持向量机充分逼近任意非线性函数，学习和适应不确定性系统的动态特性，以及并行处理算法进行快速运算等能力，建立模型作为预测模型。

此外，基于神经网络解耦的多变量系统广义预测控制，本文采用为了消除多变量系统各控制变量间的相互约束和耦合影响，应用基于神经网络的静态解耦和动态解耦。

正如前面所述，预测控制不是某一独立体系理论的产物，而是在工业实践中不断发展和完善起来的一类新型应用计算机的控制算法。随着预测控制的发展，应用范围不断扩大。基于神经网络的预测控制在复杂工业过程控制中发展起来，并且取得了许多成功的应用。

陈增强和袁著祉等从预测方式、控制律求解方法和典型应用等几个角度对基于神经网络的非线性预测控制做了综述，对其中的关键技术做了深入阐述，并指明了今后的发展方向。见文献[97]。

李彦琴在其硕士学位论文中，针对卫星姿态和轨道控制的可靠性和稳态精度的要求越来越高的现状以及该领域中被控对象结构复杂，非线性严重等特点，提出了神经网络和预测控制相结合的方法，对卫星姿态仿真系统进行了建模与控制。见文献[96]

陈兵，王华介绍了电弧炉计算机智能控制系统的优点与方法，并着重讨论了神经

网络预估模型的设计，包括遗传算法在建立神经网络模型的应用和神经网络模型实时学习的要求，基于此基础对于多步预估进行了讨论和研究。通过实际应用证实该系统具有效率高，电能消耗低，电极消耗和耐火材料消耗少等一系列优点。见文献[98]。

休斯敦大学化学工程系的 Michael Nikolaou 在他的论文“模型预测控制：理论与实际应用的综合”中，从 MPC 的发展、理论分析、目前应用情况以及未来的需求出发，提出性能更优异的 MPC 以及更多理论的支持是众多过程控制的急切需要。见文献[38]

柴天佑院士针对钢球磨中储式制粉系统具有多变量强耦合、非线性以及生产工况变化大的特性，建立了钢球磨中储式制粉系统智能解耦控制仿真系统。该软件集成了钢球磨中储式制粉系统的仿真模型以及针对钢球磨中储式制粉系统而设计的多变量智能解耦控制技术，为钢球磨中储式制粉系统自动控制问题的深入研究提供了有力的实验工具。见文献[99]。

此外神经网络预测控制在热电厂和太阳能电厂都有成功应用的实例。这些成功的应用实践表明结合神经网络与预测控制的优势而形成的神经网络预测控制在复杂非线性系统中具有广泛的应用前景，具有较大的发展空间。

通过上述章节的数据分析可以看出，跳汰过程作为一个高度非线性的过程，过程参数相互影响、共同作用、系统直接输出可观测性差的特点，可以实现跳汰最终输出一矸石带煤（I 段）的较小误差预测。因此，本文通过神经网络实现跳汰过程的预测控制。同时，鉴于 BP 网络的网络结构比径向基函数网络的网络结构小，作者采用 BP 网络建立跳汰机的预测模型。

5.2 预测控制的基本原理

预测控制最基本的算法有模型算法控制、动态矩阵控制和广义预测控制，其他算法都是以它们为基础发展起来的。作者将从这三种预测控制算法和基于神经网络的预测控制入手，研究跳汰过程控制。

5.2.1 模型算法控制分析

模型算法控制（简称 MAC）采用基于对象脉冲响应的非参数模型进行描述，主要流程包括四部分：（1）输出预测的内部模型；（2）通过反馈校正纠正模型不确定性和外部对模型的干扰；（3）预测输出跟踪参考轨迹进行滚动优化；（4）通过控制逼近参考

轨迹。

模型算法控制分为单步模型算法控制和多步模型算法控制

一、单步模型算法控制：

单步模型算法控制的基本原理如图 5—1 所示。 $u(k)$ 为当前时刻的输入，这里包括了跳汰床层内的三密度值、排料门的开度、风阀周期、入洗原煤性质等跳汰过程变量， $y(k)$ 为当前时刻的输出亦即矸石带煤， $y_m(k)$ 为内部模型的输出， $y_m(k+1)$ 为内部模型对被控对象的预测输出， $y_p(k+1)$ 是 $y_m(k+1)$ 经过校正的预测输出， $y_r(k+1)$ 表示 $k+1$ 时刻的参考值。

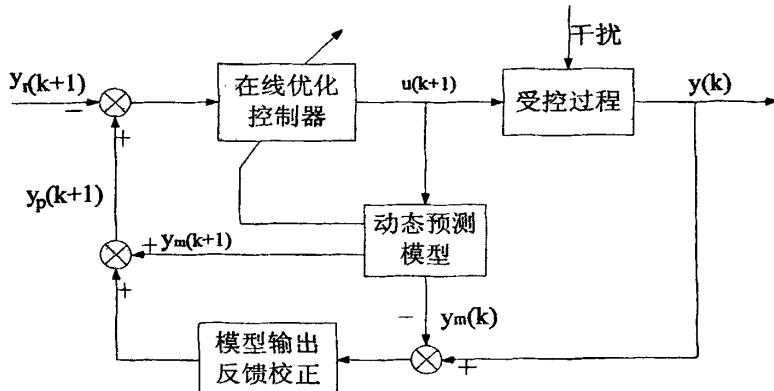


图 5—1 单步模型控制基本原理图

figure 5—1 basic schematic diagram of single step model control

单步模型算法控制是应用过去时刻输入和当前时刻输入共同作用对下一时刻进行预测，而过去时刻是已知量，当前时刻输入是待求量。假设系统的脉冲响应模型为：

$$y_m(k+1) = c_1 u(k) + c_2 u(k+1) + \cdots + c_N u(k-N+1) \quad (5-1)$$

其中 $c_i (i=1, \dots, N)$ 是脉冲序列，如图(5—2)所示。 $u(k)$ 是当前输入； $u(k), \dots, u(k-n+1)$ 是过去时刻输入；N 是脉冲模型的长度。



图 5—2 脉冲响应及模型 figure 5—2 pulse response model

由于在模型中存在干扰与误差，而且尽管采用适当的控制可以使模型的预测输出 $y_m(k+i)$ 接近 $k+i$ 时刻的输出期望值 $y_r(k+i)$ ，但还是不能保证系统未来的实际输出 $y(k+i)$ 接近 $y_r(k+i)$ 。为了能更好地估计误差，算出比较准确的控制量，应当对输出预测进行校正。但是由于模型的误差和未来时刻的噪音、干扰一般不容易测量，所以通常都是采用前一时刻预测值的偏差来近似校正。按这种方法，输出预测的校正值为

$$e(k) = y(k) - y_m(k)$$

作为反馈校正量，可得下一时刻修正后的预测输出为

$$y_p(k+1) = y_m(k+1) + he(k) \quad (5-2)$$

式中 h 为反馈校正系数。

参考轨迹在设定值发生跃变时，若要求输出迅速跟踪这一变化，往往需要施加大幅度的控制变量。这在工程上常常是很难办到的；即使能够办到，也往往会导致输出变化不稳。所以在预测控制中一般都设定一条从当前时刻实际输出 $y(k)$ 出发的指数曲线作为参考轨迹曲线，使输出由当前值逐步地过渡到设定值。

令设定值为 $y_r(k)$ ，它可以是常量，也可以是某个时间函数。通常采用参考轨迹是一阶指数形式，即：

$$y_r(k+i) = q_r y(k) + (1 - q_r) y_s \quad (i=1,2,\dots) \quad (5-3)$$

式中 $q_r = e^{-T_0/\tau}$ 且 $0 < q_r < 1$ ，这里 T_0 为采样周期， τ 为参考轨迹的时间常数。对于单步模型算法， i 取 1。可以看出，参考轨迹的时间常数 τ 越大， q_r 的值越大，系统的柔性越好，鲁棒性越好，但是控制速度却减慢，所以需要根据实际情况来确定 τ 的值。

有了预测输出和参考轨迹，就可建立 k 时刻滚动优化的性能指标：

$$J = \alpha[y_p(k+1) - y_r(k+1)]^2 + \beta u^2 \quad (5-4)$$

式中 α, β 分别为跟踪误差和控制量的加权系数。它们用来平衡能量与跟踪误差，式中第一项最小化可保证预测输出与参考轨迹最接近，第二项最小化保证跟踪参考轨迹所需的能量最小。跟踪误差与所需能量是一对矛盾。当 α 一定时，增大 β 可使控制量减小，但跟踪误差将会增大；反之，减小 β 可使控制量增大，跟踪误差减小。但为了保

证良好的跟踪特性， β 不能取值太大。

为了使性能指标达到最小，令 $\partial J / \partial u(k) = 0$ ，便可得到最小控制量：

$$u(k) = [y_r(k+1) - he(k)] / [c_1 + \frac{\lambda}{\alpha c_1}] \quad (5-5)$$

二、多步模型算法控制

多步模型算法和单步模型算法基本原理相同，多步预测同样是利用脉冲模型得到 N 步预测。它的基本结构为：

$$Y_m(k+1) = CU(k) + FU(k-1) \quad (5-6)$$

其中 $Y_m(k+1) = [y_m(k+1), y_m(k+2), \dots, y_m(k+N)]^T$ ， N 为预测长度。

$U(k) = [u(k), u(k+1), \dots, u(k+M-1)]^T$ 为当前和将来时刻的未知控制量， M 为控制时域长度。 $U(k-1) = [u(k-P+1), u(k-P+2), \dots, u(k-1)]^T$ 为过去时刻的已知控制向量， C 和 F 为脉冲系数，其形式如下：

$$C = \begin{pmatrix} c_1 & & & & \\ c_2 & c_1 & & & \\ \vdots & \vdots & & & \\ c_m & c_{m-1} & \cdots & c_1 & \\ \vdots & \vdots & & \vdots & \\ c_N & c_{N-1} & & \sum_{j=1}^{N-m+1} c_j & \end{pmatrix}_{N \times M}$$

$$F = \begin{pmatrix} c_p & c_{p-1} & \cdots & c_{p-N+1} & \cdots & c_2 \\ c_p & \cdots & c_{p-N+2} & \cdots & c_3 & \\ \ddots & & \vdots & & \vdots & \\ c_p & \cdots & c_{N+1} & & & \end{pmatrix}_{N \times (p-1)}$$

由 (5-6) 式可知， N 步预测模型输出 $Y_m(k+1)$ 是有两部分组成，第一部分 $CU(k)$ 含有当前时刻和将来时刻未知控制量，第二部分 $FU(k-1)$ 由过去已知量构成。

与单步模型算法控制相同，由于建模误差和不确定性的影响，多步模型算法控制的实际输出与模型输出存在误差即 $e(k) = y(k) - y_m(k)$ ，对 N 步模型输出进行反馈校正，即可得到 N 部的预测输出

$$Y_N(k+1) = Y_m(k+1) + he(k) = CU(k) + FU(k-1) + he(k) \quad (5-7)$$

式中 $h = (h_1, h_2, \dots, h_3)^T$ 在大多数情况下 $h_i = 1$

其参考轨迹 $y_r(k+i)$ 的选取仍与单步模型预测控制相同，由（5-4）式可得到 K 时刻的优化性能指标：

$$J = \sum_{i=1}^N q_i (y_N(k+i) - y_r(k+i))^2 + \sum_{j=1}^M \lambda_j (u(k+j-1))^2 \quad (5-8)$$

为了方便下面公式推导设

$$Q = \begin{pmatrix} q_1 & & & \\ & q_2 & & \\ & & \ddots & \\ & & & q_N \end{pmatrix}; \quad \Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_3 \end{pmatrix}; \quad Y_r = \begin{pmatrix} y_r(k+1) \\ y_r(k+2) \\ \vdots \\ y_r(k+3) \end{pmatrix}$$

因此性能指标可写成如下形式：

$$J = (Y_p(k+1) - Y_r(k+1))^T Q (Y_p(k+1) - Y_r(k+1)) + U^T(k) \Lambda U(k) \quad (5-9)$$

令 $\partial J / \partial U(k) = 0$ ，可得到如下优化控制率：

$$U(k) = (C^T Q C + \Lambda)^{-1} C^T Q (Y_r(k+1) - FU(K-1) - he(k)) \quad (5-10)$$

上述多步模型算法的基本原理可用图（5—3）来表示。从当前时刻实际输出 $y(t)$ 出发，利用设定值，在预测时域 $k+1$ 到 $k+n$ 内确定一条参考轨迹，通过确定控制 $u(k), u(k+1), \dots, u(k+M-1)$ ，使整个预测时域内预测输出对参考轨迹误差最小。

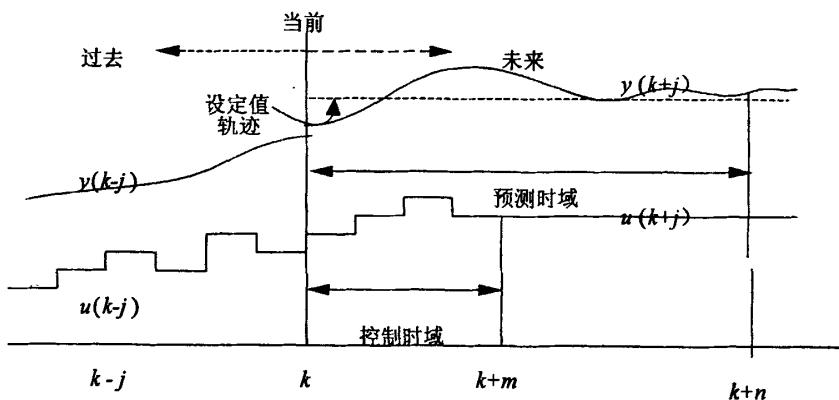


图 5—3 多步模型算法的基本原理图

figure 5—3 schematic diagram of multistep model arithmetic

在以上的理论推导过程中不难得出模型算法控制是有一定的实用范围的。以得到模型算法控制应用条件是：（1）系统线性或近似线性；（2）过程和控制器都稳定。

5.2.2 动态矩阵控制分析

一、动态矩阵控制基本原理：

动态矩阵控制(*DMC*)是一种基于对象阶跃响应模型的模型预测控制算法。该算法的预测模型是基于非参数模型，采用在工程上易于获取的对象阶跃响应作模型，直接在输入处施加阶跃信号，得到阶跃响应曲线。根据阶跃响应曲线，可以测出曲线在采样时刻 $k=1,2,\dots,n$ 的幅值 a_k 。对于这样一个对象，它在 k 时刻的输出是 k 时刻以前所有的输入增量产生的，即：

$$y(k) = a_1 \Delta u(k-1) + a_2 \Delta u(k-2) + \dots + a_n \Delta u(k-n) = \sum_{i=1}^n a_i \Delta u(k-i) \quad (5-11)$$

其中， $\Delta u(k-i) = u(k-i) - u(k-i-1)$ 为 $k-i$ 时刻的控制增量。

DMC 中的预测模型就采用上述阶跃响应表达式，但为了和对象的真正表达式加以区别，预测模型的输出通常记为

$$y_M(k) = \sum_{i=1}^n \hat{a}_i \Delta u(k-i) \quad (5-12)$$

式中， \hat{a}_i 为预测计算所用阶跃响应系数； $y_M(k)$ 为预测输出。此式所涉及的是当前 k 时刻，如果要控制 k 时刻以后 M 步，那么当前及未来时刻的控制增量序列为 $\Delta u(k), \Delta u(k+1), \Delta u(k+2), \dots, \Delta u(k+M-1)$ ，未来时刻的输出预测值为

$$y_M(k+1) = \hat{a}_1 \Delta u(k) + \hat{a}_2 \Delta u(k-1) + \dots$$

$$y_M(k+2) = \hat{a}_1 \Delta u(k+1) + \hat{a}_2 \Delta u(k) + \dots + \hat{a}_3 \Delta u(k-1) + \dots$$

⋮

$$y_M(k+M) = \hat{a}_1 \Delta u(k+M-1) + \dots + \hat{a}_M(k) + \dots + \hat{a}_{M+1} \Delta u(k-1) + \dots$$

⋮

$$y_M(k+p) = \hat{a}_{p-M+1} \Delta u(k+M-1) + \dots + \hat{a}_p(k) + \dots + \hat{a}_{p+1} \Delta u(k-1) + \dots$$

记 $k+1$ 至 $k+p$ 时刻的预测输出向量 y_M 和 k 至 $k+m-1$ 时刻的控制增量向量 Δu 为

$$Y_M(k+1) = [y_M(k+1), y_M(k+2), \dots, y_M(k+p)]^T$$

$$\Delta U(k) = [\Delta u(k), \Delta u(k+1), \dots, \Delta u(k+M-1)]^T$$

那么输出预测值记为

$$Y_M(k+1) = A\Delta U(k) + A_0\Delta U(k-1) \quad (5-13)$$

式中, A 为 $P \times M$ 维矩阵, A_0 为 $P \times (N-1)$ 阶矩阵, 均由阶跃响应系数构成,

$$A = \begin{bmatrix} \hat{a}_1 & & & & \\ \hat{a}_2 & \hat{a}_1 & & & \\ \vdots & \vdots & \ddots & & \\ \hat{a}_M & \hat{a}_{M-1} & \vdots & \hat{a}_1 & \\ \vdots & \vdots & \vdots & \vdots & \\ \hat{a}_p & \hat{a}_{p-1} & \cdots & \hat{a}_{p-M+1} & \end{bmatrix}$$

$$A_0 = \begin{bmatrix} \hat{a}_n - \hat{a}_{n-1} & \hat{a}_{n-1} - \hat{a}_{n-2} & \cdots & \hat{a}_{n-p+1} - \hat{a}_{n-p} & \cdots & \hat{a}_3 - \hat{a}_2 & \hat{a}_2 \\ & \hat{a}_n - \hat{a}_{n-1} & \cdots & \hat{a}_{n-p} - \hat{a}_{n-p-1} & \cdots & \hat{a}_4 - \hat{a}_3 & \hat{a}_3 \\ & & \ddots & \vdots & \vdots & \vdots & \vdots \\ & & & \hat{a}_n - \hat{a}_{n-1} & \cdots & \hat{a}_{p+2} - \hat{a}_{p+1} & \hat{a}_{p+1} \end{bmatrix}$$

A 由对象动态响应系数构成, 称为动态矩阵。显然, $A\Delta U$ 代表当前及未来时刻的控制对输出的影响, 而 $A_0\Delta U(k-1)$ 则代表过去的控制所产生的输出。

与前面介绍的模型算法类似, 动态矩阵控制在建模时存在误差和干扰, 所以应当对输出预测进行校正, 输出的校正值为:

$$e(k) = y(k) - y_M(k)$$

可得到 M 步的预测输出为:

$$Y_p(k+1) = Y_M(k+1) + he(k) = A\Delta U(k+1) + A_0\Delta U(k-1) + he(k) \quad (5-14)$$

$$\text{其中 } Y_p(k+1) = [y_p(k+1), y_p(k+2), \dots, y_p(k+p)]^T; \quad h = [h_1, h_2, \dots, h_p]^T$$

给定如下参考轨迹:

$$Y_r = [y_r(k+1), y_r(k+2), \dots, y_r(k+P)]^T$$

根据 $Y_p(k+1)$ 和 $Y_r(k+1)$, 可以得到二次性能指标

$$\begin{aligned}
J &= (Y_p(k+1) - Y_r(k+1))^T Q (Y_p(k+1) - Y_r(k+1)) + \Delta U^T \Lambda \Delta U \\
&= (A \Delta U(k+1) + A_0 U(k-1) + h e(k) - Y_r(k+1))^T \bullet Q \bullet \\
&\quad (A \Delta U(k+1) + A_0 U(k-1) + h e(k) - Y_r(k+1)) + \Delta U^T \Lambda \Delta U \quad (5-15)
\end{aligned}$$

令 $\partial J / \partial \Delta U(k) = 0$, 可求得如下优化控制率

$$\Delta U(k) = (A Q A^T + \Lambda)^{-1} A Q (Y_r(k+1) - A_0 U(k-1) - h e(k))$$

这里 Q 和 Λ 分别为 $P \times P$ 维和 $M \times M$ 维加权阵。

二、动态矩阵控制参数的选择

(1) 预测步程 P

预测步程 P 表示对从 k 时刻起对输出 $y_m(k+i)$ 逼近期望值 $y_r(k+i)$ ($i=1, 2, \dots, P$) 的步数。它是滚动优化的关键。

为使滚动优化真实存在, 预测步程 P 应该包括对象的真实动态部分, 也就是说应把当前控制影响较大的所有响应包括在内。一般 P 选为近似等于过程的上升时间。对于有时延的系统, P 必须选得超过对象脉冲响应(或阶跃响应)的时延部分, 并覆盖对象的重要动态响应部分。

预测步程 P 的选择, 直接影响控制的稳定性和快速性。下面咱们进行具体讨论。当是 P 取值足够小, 最小为 $P=1$, 这时就变为单步内选择控制 $\Delta u(k)$, 达到输出 $y(k+1)$ 跟踪输入 $y_r(k+1)$, 即 $y(k+1) = y_r(k+1)$ 。在这种情况下, 如果模型准确, 它可使对象输出跟踪预定输出期望值, 即实现一步控制。但对模型失配或有干扰情况, 和对有时延及非最小相位系统, 上述一步跟踪目标无法实现, 且有可能导致系统不稳定。当保持有限的控制步程 L , 而把 P 取得充分大, 这样虽然是动态优化, 但实际上输出已接近于稳态优化。此时系统的动态响应接近于对象的固有特性, 这样动态响应很慢不具有适时滚动优化。此外, 大的 P 还会使控制矩阵的阶次显著增高, 增加计算时间。总结上述两种极端情况, 前者虽然快速性好, 但稳定性和鲁棒性较差; 后者虽然稳定性好, 但动态响应慢, 且增加了计算时间, 降低了系统的实时性。 p 的这两种取法都是不可取的。实际选择时, 可在上述两者中间取值, 使系统既能获得所期望的稳定鲁棒性, 又能具有所要求的动态快速性。

(2) 控制步程 M

控制步程 M 在优化性能指标中表示要计算和确定的未来控制量改变的数目。由于优化主要针对未来 P 个时刻的输出预测误差进行的，它们最多只受到 P 个控制增量的影响，所以 $M \leq P$ 。

在 P 已经确定的情况下， M 的选取通常考虑以下情况，一般 M 选得越小则越难保证输出紧跟期望值变化，反映在性能指标中效果也越差。例如，若取 $M = 1$ ，则意味着只用一步控制增量 $\Delta u(k)$ 使系统在以后的输出 $k+1, \dots, k+p$ 时刻都能跟踪期望值变化，显然，对于复杂动态过程，这是不可能的。为了改善跟踪性能，就要用增加控制步数 M 来提高系统的控制能力，使输出误差尽可能小，也就是说，把 P 个点的输出误差优化，要求由给出的 M 控制变量来分担。但也不是 M 越大越好， M 越大，控制的机动性越强，可提高控制的灵敏度，但系统的稳定性和鲁棒性随之而下降。为提高系统的稳定性和鲁棒性，又要求 M 选得小一些，因为 M 越小，远程跟踪控制能力虽然有所削弱，但可以得到一个稳定的控制。总之， M 的选择应兼顾快速性和稳定性，综合平衡考虑。此外，当控制步程 M 增大时，控制矩阵的维数也增加，计算控制参数的时间迅速增加，会降低系统的实时性。

还必须指出，通过许多仿真研究表明，在许多情况下， M 和 P 在性能指标中起着类似相反的作用，既增加 M 与减小 P 有着类似的控制效果。所以，为简便，设计时可以先根据对象的动态特性初选 M ，然后再根据仿真和调试结果确定 P ，以节省调试时间。

(3) 模型长度 N

为使模型较好地描述对象的动态特性，一般要求 \hat{h}_i 在 $i > n$ 之后约等于 0。或者要求 \hat{h}_n 与输出测量误差有相同的数量级。对于一般特性不太复杂的过程， $N = 30 \sim 50$ 即可满足要求。缩短采样周期，虽能使模型更加准确，但 N 的增大会增加计算量；甚至会因为 \hat{h}_i 太小使控制量过强，或者导致病态矩阵出现。对于较慢的或特性复杂的过程，必须适当加大 N ，但仍以不造成矩阵计算困难及不影响实时计算速度为限。

综上所述， P 、 M 和 N 的取值都应根据实际情况合理取值，一般情况下应使 P 、 M 和 N 保持 $N \geq P \geq M$ 。

5.2.3 广义预测控制分析

广义预测控制是以最小方差为基础采用最小化参数模型的一种预测控制，它的这些特点决定了在预测和滚动优化过程中，计算量明显比较少。

广义预测控制采用的模型一般为 CARIMA(Controlled Auto-Regressive Integrated moving Average)模型，即受控自回归积分滑动平均模型，其具体模型为：

$$A(z^{-1})y(k) = B(z^{-1})u(k-1) + C(z^{-1})\xi(k)/\Delta \quad (5-21)$$

式中参数代表的意义是：

$\Delta = 1 - z^{-1}$ 为差分算子， ξ 为零均值不相关的白噪音向量，

$$A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_{na} z^{-na} = 1 + \sum_{i=1}^{na} a_i z^{-ni}$$

$$B(z^{-1}) = 1 + b_1 z^{-1} + \cdots + b_{nb} z^{-nb} = 1 + \sum_{i=1}^{nb} b_i z^{-ni}$$

$$C(z^{-1}) = 1 + c_1 z^{-1} + \cdots + c_{nc} z^{-nc} = 1 + \sum_{i=1}^{nc} c_i z^{-ni}$$

为了公式推导暂令 $C(z^{-1}) = 1$ ，对于 $C(z^{-1}) \neq 1$ 的情况可以仿照此法推导。

求解下面 Diophantine 方程，便可得到 P 步预测输出：

$$1 = E_j(z^{-1})A(z^{-1})\Delta + z^{-1}F_j(z^{-1}) \quad (5-22)$$

$$E_j(z^{-1})B(z^{-1}) = G_j(z^{-1}) + z^{-1}H_j(z^{-1}) \quad (5-23)$$

其中 $j = 1, \dots, n$ ，并且

$$E_j(z^{-1}) = e_0 + e_1 z^{-1} + \cdots + e_{j-1} z^{-j+1}$$

$$F_j(z^{-1}) = f_0 + f_1 z^{-1} + \cdots + f_{j-1} z^{-j+1}$$

$$G_j(z^{-1}) = g_0 + g_1 z^{-1} + \cdots + g_{j-1} z^{-j+1}$$

$$H_j(z^{-1}) = h_0 + h_1 z^{-1} + \cdots + h_{j-1} z^{-j+1}$$

其中 $E_j(z^{-1})$ 表达式中 e_0, e_1, \dots, e_{j-1} 是 $j = 1, \dots, n$ 时由 (4-22) 和 (4-23) 求出 $E_j(z^{-1})$ 的解，同理 $F_j(z^{-1}), G_j(z^{-1}), H_j(z^{-1})$ 表达式中系数所代表的意义与 e_0, e_1, \dots, e_{j-1} 相同。

为了书写方便，在下面的推导过程中将后移算子 z^{-1} 省略。则 P 步预测输出可表示为：

$$Y_p(k+1) = G\Delta U(k) + Fy(k) + H\Delta u(k-1) + E \quad (5-24)$$

其中 $Y_p(k+1) = [y_p(k+1), y_p(k+2), \dots, y_p(k+p)]^T$

$$\Delta U(k) = [\Delta u(k), \Delta u(k+1), \dots, \Delta u(k+m-1)]^T$$

$$F = [F_1, F_2, \dots, F_p]^T; \quad H = [H_1, H_2, \dots, H_p]^T$$

$$E = [E_1 \xi(k+1), E_2 \xi(k+2), \dots, E_p \xi(k+p)]^T$$

$$G = \begin{pmatrix} g_0 & & & \\ g_1 & g_0 & & \\ \vdots & \vdots & \ddots & \\ g_{m-1} & g_{m-2} & \cdots & g_0 \\ g_m & g_{m-1} & \cdots & g_1 \\ \vdots & \vdots & \cdots & \vdots \\ g_{p-1} & g_{p-2} & \cdots & g_{p-m} \end{pmatrix}$$

设参考轨迹为：

$$Y_r(k+1) = [y_r(k+1), y_r(k+2), \dots, y_r(k+p)]^T$$

广义预测控制的目标同样也是使预测输出对参考轨迹的跟踪误差最小，由此可确定如下性能指标：

$$J = \psi \{(Y_p(k+1) - Y_r(k+1))^T (Y_p(k+1) - Y_r(k+1)) + \Lambda U^T(k)U(k)\} \quad (5-25)$$

式中 ψ 表示求期望值， Λ 为加权因子，令 $\partial J / \partial \Delta U(k) = 0$ ，便可以得到优化控制量

$$\Delta U(k) = (G^T G + \Lambda I)^{-1} G^T (Y_r(k+1) - Fy(k) - H\Delta U^T(k-1)) \quad (5-26)$$

5.3 神经网络预测控制

5.3.1 神经网络在预测控制中的应用

由前面讨论可知，预测模型、滚动优化和反馈校正是预测控制算法的本质特征。由前面的预测控制理论可知，线性系统的预测控制问题已得到了解决，但这些算法的

运算负荷相当大，不利于其在速快系统中应用。另外由于外界的不确定性干扰，引起系统参数的变化，从而导致系统的真实模型与理论模型的失配，较大的模型失配使得系统的动态特性和控制质量变差，甚至造成系统的不稳定。为此，作者采用支持向量机建立系统模型，利用神经网络模型来提高其预测精度。

神经网络预测控制系统的结构框图如图（5—4）所示，为未来下一个时刻即下一步的预测过程，而多步预测与此类似。图中预测模型采用神经网络构成， $y_r(k)$ 为未来输出的期望值， $u(k+1)$ 为经过优化的控制量， $y(k)$ 为输出量， $y_m(k+1)$ 为预测下一时刻的输出结果。

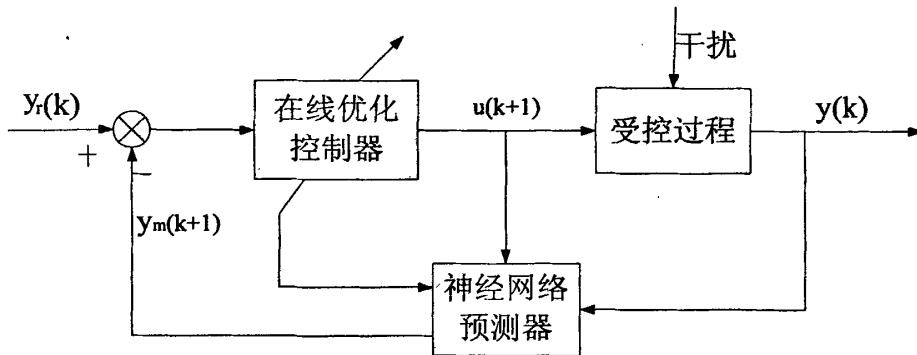


图 5—4 神经网络预测控制系统结构

figure 5—4 system structure of ANN predict control

首先对用来描述跳汰对象的动态行为的预测模型，根据系统当前的输入输出信息，预报系统未来的输出值。在线性系统的情况下，可利用叠加原理，并利用系统的阶跃响应、脉冲响应或者状态空间模型等，预测系统的未来输出。由于跳汰系统的机理结构未知，不能充分描述，作者采用神经网络来预测其模型，而将其用于 MPC 之中。

其次是神经网络的滚动优化，神经网络预测控制器的性能指标为

$$J = \sum_{j=r}^p e^2(k+j) + \beta \sum_{j=1}^m \Delta u^2(k+j-1) \quad (5-27)$$

$$e(k+j) = y_r(k+j) - y_p(k+j) \quad (5-28)$$

式中 $e(k+j)$ 为未来时刻期望值与未来时刻预测值之间的误差， p 称为最大输出预报区间，表明了待优化的将来输出跟踪范围； m 是控制长度，表明了要纳入考虑的未来控制的范围； β 是加权因子，表示对控制能量变化的重视程度。

有了未来输出的期望值 $y_r(k+j)$ 及优化性能指标函数 J 就可以求出未来的控制量，根据控制长度，在每个采样点，算出的未来范围的控制量有一定的时间长度，因为控制是一步一步向前推进的，所以施加的控制信号只能是计算出来的第一个控制量 $u(k)$ 。在下一时刻，这些量一般与此之前的优化的未来控制量一致，因此必须重新优化，即所谓的滚动优化。

再次就是反馈校正环节，反馈校正环节可以提高预报的精确值，其差值 $y(k) - y_m(k)$ 经过滤波器后作用于控制器。

经过上面的分析可知神经网络预测控制的算法过程：

(1) 由控制要求获取未来的期望输出序列值，即输入参考轨迹：

$$y_r(k+j) \quad (j = k, k+1, \dots, m)$$

(2) 预先测量得到被控对象的开环输入输出数据，把它们作为训练神经网络辨识器的样本，对神经网络进行离线训练，得到令人满意的神经网络模型。

(3) 利用神经网络模型产生预报输出 $y_m(k+j)$ ，经模型偏差 $e(k)$ 校正后得到系统的预测输出 $y_p(k+j)$ ，即 $y_p(k+j) = y_m(k+j) + he(k)$ ； h 为反馈校正系数。

(4) 计算与未来时刻期望值的误差 $y_r(k+j) - y_p(k+j)$ 。

(5) 极小化性能指标函数 J ，获得控制的最优控制量 $u(k+j-1)$

(6) 将第一控制量 $u(k)$ 作用于系统，返回 (1)。

由以上分析可以得出，神经网络预测控制的实质就是利用作为对象辨识模型的神经网络产生预测信号，然后利用优化算法求出控制矢量，从而实现对变参数、不确定性系统的跳汰过程预测控制。

5.3.2 基于 BP 神经网络的预测控制

BP 神经网络预测控制的基本原理是根据对象的输入输出，利用 BP 神经网络来构造对象的预测模型，以预测控制算法作为优化手段来实现系统的控制。所以 BP 神经网络预测控制的结构如图 5—4 所示，神经网络预测控制器由 BP 网络构造。预测控制算法中需要有一个描述对象动态行为，具有预测功能，即根据系统的历史信息和选定的未来输入，预测未来输出的预测模型。

基于神经网络设计预测控制系统主要分两部分进行：一部分是BP网络预测器的设计；另一部分是优化器的设计。

神经网络预测模型（预测器）通过训练一个BP神经网络来实现对未来时刻输出的预测。BP神经网络预测模型如图5—5所示，三层结构，有一个输入层，输入节点个数为 $n+m+1$, n, m 表示 k 时刻以前输入和输出阶次，有一个隐含层和一个输出层。其中作为预测器，输入层的输入量为当前时刻以及以前时刻的输出结果和经过不断优化的控制量。

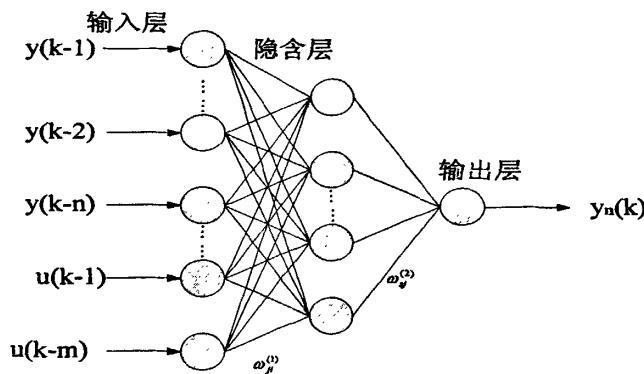


图 5—5 BP 神经网络预测模型

figure 5—5 predict model of ANN

设非线性系统可由下面数学模型表示：

$$y(k) = g(y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-m)) \quad (5-29)$$

其中 $g(\bullet)$ 为未知的非线性函数，由此式可推导出多步预测模型 $y_m(k+j)$ ：

一步预测模型可以表示为：

$$y_m(k+1) = g(y(k), \dots, y(k-n+1), u(k+1), \dots, u(k-m+1))$$

P步预测模型可以表示为：

$$y_m(k+p) = g(y_m(k+p-1), \dots, y_m(k-n+p), u(k+1), \dots, u(k-m+1))$$

显然上式使用了 k 时刻以后的预测值 $y_m(k+j)$ ，对 k 时刻和 k 时刻以前的预测值，可以用其真实值来代替，所以多步预测模型可以表示为：

$$y_m(k+j) = g(y_m(k+j-1), \dots, y_m(k-n+j), u(k+j), \dots, u(k-m+j))$$

其中： $j=1, 2, \dots, p$ (5-30)

以上公式中 $g(\bullet)$ 是未知非线性函数，公式右端的输入和公式左端的输出，很难用具体的数学公式表示，为此可用公式左右两端数据来训练 BP 网络使其在要求的精度下逼近函数 $g(\bullet)$ 。

综上所述，建立跳汰神经网络预测模型可分为两步：第一步，以对象的输入输出数据为样本，训练具有一定精度的神经网络模型，使网络的输出 $y_m(k)$ 尽量逼近被控对象的输出 $y(k)$ ，也就是说对被控对象进行模型辨识。第二步，求出多步预测值 $y_m(k+1)$ ，用 BP 网络对被控对象进行模型辨识。

在控制系统工作时，由于 BP 算法计算量比较大，对控制的实时性有一定的影响，在实际实施过程中，可采取如下措施加以改进。

- (1)、系统运行前先对 BP 网络进行离线训练，以减少在线学习时间；
- (2)、在预测过程中需要实时检验，当预测误差较大时，需要离线再训练网络，否则不做调整，以减少计算量。

5.3.3 神经网络预测控制中控制量的计算

在神经网络预测控制中，控制量是通过优化器计算得到，优化计算是建立在神经网络预测模型的基础之上的。要获得最优控制律，必须使性能指标函数 J 趋于最小。本文研究的对象—跳汰机研石段研石带煤为多变量非线性系统，为了能找到最优控制律，采用前面训练好的 BP 网络为多变量函数，使用蚂蚁算法优化连续函数得到最优控制量。

一、蚂蚁算法的提出：

蚂蚁算法是一种新颖的通用型的随机全局搜索算法，它吸收了昆虫王国中蚂蚁的行为特性：蚂蚁有能力在没有任何可见提示下找出从其巢穴至食物源的最短路径，并能随环境的变化而变化，适应性地搜索新的路径，产生新的选择，也就是说蚂蚁具有找到蚁巢与食物之间的最短路径的能力。这种能力是靠其在所经过的路径上留下一种挥发性分泌物 Pheromone(称为信息素，该物质随着时间的推移会逐渐挥发消失)来实现的，使得一定范围内的其他蚂蚁能够察觉到并影响其行为。当一些路径上通过的蚂蚁越来越多时，其留下的信息素轨迹(Trail)也越来越多，以致信息素强度增大(当然，随着时间的推移会逐渐减弱)，后来蚂蚁选择该路径的概率也越高，从而更增加了该路径的

信息素强度，这种选择过程被称之为蚂蚁的自催化行为(Auto catalytic Behavior)，蚂蚁群体就是靠着这种内部的生物协同机制逐渐形成一条最短路线的。受到蚂蚁群体的这种启发，近几年人们提出蚂蚁算法，并将其用在多方面的研究中，而且取得了比较好的结果。在本文中，主要研究，利用蚂蚁算法优化连续非线性多变量函数。

二、蚂蚁算法的基本原理和简单算法推导：

从图 5—6 真实蚁群寻找最优路径的原理分析：

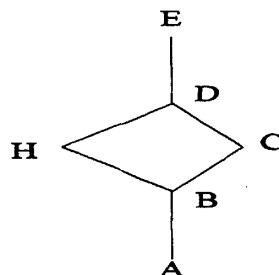


图 5—6 蚂蚁算法的基本原理图

figure 5—6 schematic diagram of ant arithmetic

设 $DH=HB=1$, $DC=CB=0.5$ 。假定在每个离散的等时间间隔: $t=0, 1, 2, 3\dots\dots$, 有 30 只蚂蚁从 A 到达 B, 同时有 30 只蚂蚁从 E 到 D, 每只蚂蚁的速度为 m 并且, 每有一只蚂蚁经过时, 在时间 t 留下信息素密度为 1。蚂蚁在选择路径时, 那些多数蚂蚁曾经选择过的路径(也就是具有更高信息素密度的路径), 被再次选中的可能性最大。

当 $t=0$ 时, 没有信息素, 有 30 只蚂蚁分别在 B 和 D 蚂蚁走哪条道路是完全随机的。假设, 在每个点上蚂蚁将有 15 只经过 H, 另外 15 只经过 C。

当 $t=1$ 时有 30 只蚂蚁从 A 到 B, 它们发现指向 H 道路上的信息素密度是 15, 是由从 B 出发的蚂蚁留下的; 指向 C 道路上的信息素密度是 30, 其中 15 是由 B 出发蚂蚁留下, 另外 15 是从 D 出发经过 C 已经到达 B 的蚂蚁留下。因此, 选择经过 C 到 D 的可能性就更大, 从 E 出发到 D 的 30 只蚂蚁也面临着同样的选择, 由此产生一个正反馈过程, 选择经过 C 的蚂蚁越来越多, 直到所有的蚂蚁都选择这条道路。蚁群算法就是利用蚂蚁的这一特性, 解决最优化问题。

算法的实现:

蚂蚁算法自从诞生以来, 第一个应用的是旅行商问题, 之后对蚂蚁算法的一些改进、评价其算法的优劣性主要是在旅行商问题上进行的。而且经过长期实践和时间的考验, 蚂蚁算法显示了其独特的求解力和其他算法难以比拟的搜索能力。蚂蚁算法的

优点：

- (1) 蚂蚁算法在优化过程中，采用的是大规模并行优化计算，所以蚂蚁算法有较强的全局优化能力；
- (2) 蚂蚁算法是一种全域按概率随机搜索，不易陷入局部优化；
- (3) 有较强的鲁棒性，可以用于多方面研究。
- (4) 蚂蚁算法很容易与多种启发式算法结合，以改善算法的性能。

在本文的研究中，利用蚂蚁算法作为研石带煤预测控制中的控制器，可以实现本文的目的即研石带煤最少时得到最优控制量。

三、蚂蚁算法对神经网络预测控制的优化器的改进

在神经网络预测控制中，优化器的设计正处于一个发展的阶段，单一的预测可以采用大量的数据以图 5—5 的方法进行网络训练，而预测与控制的结合增加了由理论到实践的难度。由于神经网络是对非线性不确定数学模型的不断学习逼近，所以对于从预测结果得到控制量的优化器的设计比较困难。到目前为止，Matlab 中的预测控制器也只实现了单输入单输出的控制量优化，但本文中研究的研石带煤是一个多输入单输出的系统，所以 Matlab 中的预测控制器在本文已不实用。为此作者将预测控制器中的预测器由 BP 网络设计，优化器用训练好的 BP 网络作为可调用函数、由蚂蚁算法将其优化得到控制量。

5.4 神经网络预测控制在跳汰风阀控制中的应用

矸石作为废弃物，矸石中所携带的煤就成为绝对损失。跳汰机研石段中的矸石带煤是跳汰研究的主要控制对象和主要代表，因此，降低矸石带煤可以从改造跳汰机结构和改造控制过程入手。跳汰机结构改造日趋完善，改造效果上升空间不大，而不管结构改造如何有效，也不能改造其“在一个重力加速度下完成分选”的实质，而跳汰控制过程的改造耗资小，方法多，易执行，是本文研究的内容，也是控制工程的一个重要领域。

5.4.1 跳汰风阀控制思路

跳汰机分层主要是受垂直升降的交变水流影响。而水位的波动状态将随着风阀参数的变化而变化，不合理的风阀参数将会导致形成不良的床层。因此，合理的控制风

阀是进行良好分层的必要条件。

由于对跳汰机这样复杂的生产过程不能建立起比较精确的数学模型，因此，上述几章通过 BP 人工神经网络、RBF 网络、支持向量机建立了跳汰系统研石带煤的软测量，测试误差可以达到 7.143%~10.46%，正是这项软测量构成了跳汰机模型预测控制的基础。

在应用神经网络预测控制之前，首先需要再分析跳汰机参数间的关系。在研石段，研石带煤直接由分层好坏和排料的好坏来决定，而分层又由风水量来控制。对分层来讲，能改变分层的作用力来源于产生垂直方向脉动水流的风、水量。因此，控制思路以及仿真路线为：

采用神经网络预测器对研石带煤下一时刻进行预测，将预测值与实际值进行比较再经过优化器对风阀控制量进行在线优化。其中预测器的输入量为分层情况、排料量（分层情况由三密度直接体现）和当前时刻研石带煤量，输出量为下一时刻研石带煤量；预测值与期望值比较后通过优化器对风阀周期进行优化。仿真采用前述采集的 600 组样本数据，其中当前时刻的研石带煤即为人工实测值，实际应用的控制系统中可以由研石在线灰分仪测得，期望值则根据当前一段时间内入洗原煤情况预先给定。控制过程流程如图 5—7 所示：

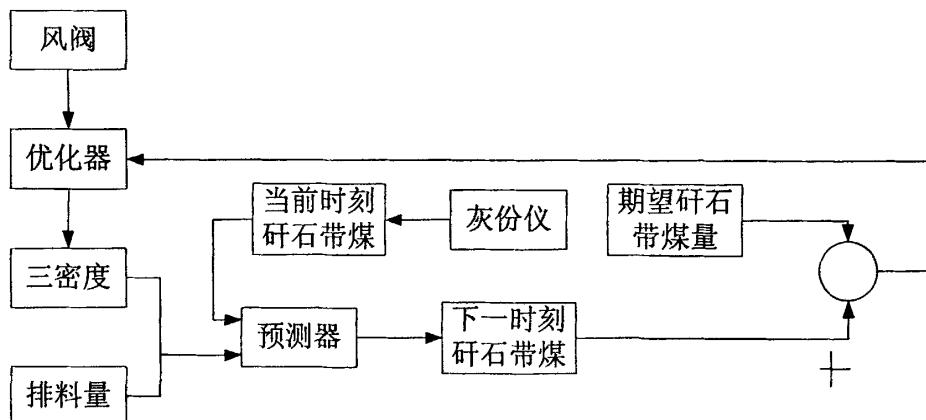


图 5—7 研石带煤预测控制流程图

figure 5—7 flow chart of coal lost in waste predict control

5.4.2 跳汰风阀神经网络预测控制的设计

在跳汰机的研石段，一般为两个跳汰室，实验所用属于旧有的跳汰机机型，每室

各布置一个进、排气风阀。这里，考察风阀的控制如何影响分层以及矸石带煤，而排料控制由模糊控制器完成，顶水阀门开度设为定值。那么，可控制变量为进气周期(JQ)、膨胀周期(PZ)和排气周期(PQ)，每室各三个可控变量。以前面的理论为铺垫，以矸石带煤为目标的预测控制，也分为预测器和优化器。

一、矸石带煤预测器的设计

由训练好的BP神经网络作为预测器，预测下一时刻的矸石带煤。BP神经网络的结构如图5—5所示，其参数选择以及优化算法在第三章已经阐述，将矸石段风阀的六个周期量($JQ1(k)$ $PZ1(k)$ $PQ1(k)$ $JQ2(k)$ $PZ2(k)$ $PQ2(k)$)、矸石段三密度和排料量以及前一时刻的矸石带煤量($Waste(k-1)$)作为训练样本的输入量，将当前时刻的矸石带煤($Waste(k)$)作为训练样本的输出量，对网络进行训练。下图为矸石带煤神经网络预测器的性能变化曲线，可以看到用BP神经网络预测精度较高，可以作为预测器。

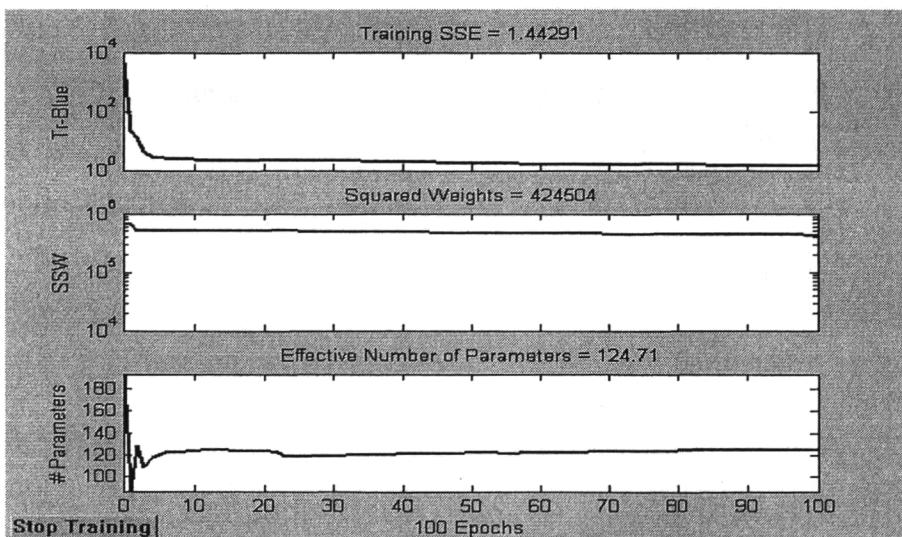


图 5—8 神经网络预测器的性能变化曲线图

figure 5—8 performance of ANN predictor

二、矸石带煤优化器的设计

在优化过程中，由于大部分算法或需要精确的数学模型、或是单输入单输出非线性，因此不确定数学模型的系统优化还难以实现。针对这种情况，将前面训练好的BP神经网络作为描述矸石带煤系统的连续空间函数，利用蚂蚁算法优化连续空间函数作为优化器得到的最优函数。

根据实际经验，跳汰机的风阀控制具有一定的取值范围($JQ1 \in [34, 39]$, $PZ1 \in [21, 28]$, $PQ1 \in [16, 30]$, $JQ2 \in [35, 39]$, $PZ2 \in [22, 28]$, $PQ2 \in [16, 25]$)，所以描述的系统函数是有界的连续函数。用蚂蚁算法优化这个有界连续函数作为矸石带煤优化器。

利用设计好的预测器和优化器对矸石带煤进行仿真计算可得到附录 2、附录 3 的结果。

分析结果可以发现：矸石带煤量明显减少，而且分层情况也明显变好，即代表分层情况的三密度中上密度减少，下密度增加。说明了风阀的预测控制对矸石带煤起到了明显的控制作用。

跳汰风阀的控制是跳汰控制的核心之一，风阀周期是物料得以分选的总动力源。以矸石带煤为目标函数，通过神经网络的预测控制，可以实现较好的控制效果；针对跳汰机的整体复杂控制，即多变量复杂系统难以建立预测模型的问题，采用神经网络建立被控对象的预测模型和基于神经网络预测的蚂蚁算法优化控制，实现了以矸石带煤为目标的风阀控制，也即实现跳汰机的效率控制，解决了跳汰机这一线性时不变系统一直无法解决的控制问题。

5. 5 本章小结

预测控制，注重的是模型功能，而不是结构形式，过程的描述可以通过简单的实验获得，不需要深入了解过程的内部机理。因此，预测控制算法改变了现代控制理论对模型结构较严格的要求，更着眼于根据功能要求，按最方便途径建立多样性模型。这一点对于诸如跳汰机这种理论严重落后于实践的系统尤为适用。以预测模型为基础，采用在线滚动优化性能指标和反馈校正的策略，来克服受控对象建模误差和不确定因素的影响，有效弥补了现代控制理论对复杂系统所无法避免的不足之处。

在有限时域内滚动优化，可在线反复进行优化计算跳汰床层的时变性，虽然在每一时刻只能得到全局的次优解，然而却能使模型失配、畸变、干扰等引起的不确定性及时得到补偿，从而得到较好的动态控制性能。

基于神经网络的预测控制在复杂工业过程控制中发展起来，并已取得了许多成功的应用。神经网络和支持向量机方法都具有极强的非线形逼近能力等优点，因此，将其用于跳汰床层预测控制系统是必要的尝试。

跳汰机矸石段中的矸石带煤作为跳汰控制研究的主要对象和主要代表，确定控制思路为：采用对矸石带煤下一时刻进行预测，利用事先离线训练好的神经网络预测器（定期对其进行检查校正，校正方法就是将近期积累的合理数据对神经网络预测器再进行离线训练）得到下一时刻矸石带煤量，据此控制风阀的周期。

主要工作与创新点

本文在对跳汰机分选机理进行深入研究的基础上，以某矿洗煤厂 307#跳汰机为研究对象，以 γ 射线密度探测器在线探测跳汰床层分层效果为基础，通过现场采集试验样本，以“矸石带煤”这一最终目标为考核对象，分别采用人工神经网络、支持向量机等方法对影响跳汰机分选效果的诸因素进行了模型建立，并利用 BP 网络对跳汰机进行了预测控制研究，达到了满意的效果。

本文的主要工作：

一、从中国能源战略出发，从洁净能源的需求出发，选煤发展任重道远；在研究了跳汰理论、过程控制研究的历史与现状后，认为：利用智能控制技术改善跳汰分选效果条件成熟、意义重大。

二、对跳汰机的可控变量进行了分析，提出了在矸石段以“矸石带煤”为控制目标，并以此来控制风阀和排料，以最大限度地在保证精煤质量的前提下提高精煤回收率。

三、在对BP人工神经网络进行分析的基础上，实现了基于人工神经网络的跳汰过程状态识别，并对BP神经网络和RBF神经网络进行了比较。针对BP神经网络的一些缺点，采取了具体措施，并提出了一套跳汰过程状态识别BP神经网络结构参数的选择的方法。

四、将遗传算法和BP算法相结合，进行神经网络的混合训练。通过利用遗传算法对初始权值分布进行优化，在解空间中定位出一个较好的搜索空间，然后采用BP算法在这个小的解空间中搜索出最优解，仿真表明神经网络能比较好地摆脱局部极小。

五、采用支持向量机对跳汰过程状态进行了识别，并与BP神经网络和RBF神经网络的识别结果进行了比较。结果表明：支持向量机有比较好的对比效果。

六、将神经网络预测控制应用在非线性、大滞后、多变量的跳汰控制中，并利用蚂蚁算法处理连续的非线性、多变量函数，解决了多变量、非线性、不确定数学模型系统难以实现优化的问题，实现了跳汰机风阀的有效控制。

主要创新点：

一、以 γ 射线密度探测器在线探测跳汰床层分层效果为基础，通过大量的现场采样、实验、分析、整理，建立了多达600条记录的跳汰机综合试验样本，该样本记录了从入洗原煤性质、风阀周期参数、排料参数等影响分选效果的主要因素，成为分析跳汰效果

及其控制的重要依据。

二、利用人工神经网络建立了跳汰机分层、排料一体化模型，彻底摆脱了目前跳汰控制不能将二者合二为一进行考虑的问题，该模型兼顾了各种控制变量对跳汰分选效果的作用，使跳汰机控制研究成为一个整体。鉴于样本采集困难，采用支持向量机比人工神经网络表现出了更加优异的性能。

三、以矸石带煤为目标函数，通过对MPC的分析，设计了跳汰风阀的预测控制系统；仿真结果显示，矸石带煤有了明显的降低。

本文的不足与展望：

一、神经网络训练的数据还比较少，有待今后工作中进一步丰富。

二、神经网络的理论还有待于完善，进而提高神经网络的泛化能力。

三、遗传算法各参数的选取一般而言仅凭经验，规则并不统一，对现有的遗传算法还需做进一步的研究和改进，以克服其不足。蚂蚁算法在函数连续空间中的优化理论还处于发展阶段，在多维变量函数优化过程中计算比较缓慢，冗余量较大，还需在理论上进一步完善。

四、跳汰机的分层和排料神经网络一体化模型对跳汰机理的诠释还有待进一步研究。

人工神经网络以及支持向量机在跳汰系统过程识别中表现出了比较好的性能，随着现场采集样本的不断增加以及控制理论的不断完善，跳汰系统的控制性能必将得到极大的提高，跳汰机分选效果也必将得到更大的改善，其在国民经济中的地位将更加突出。

附录

附录 1 跳汰机部分样本数据

附录 2. 优化前各参数值

附录 3. 优化后参数值

附录 1 跳汰机部分样本数据

appendix 1. Part sample data of jig

序号	Date	Time	上密度 FL1 g/cm ³	中密度 FL2 g/cm ³	下密度 FL3 g/cm ³	GATE1L	WASTE	1进 JQ1 1# PZ1	1# PQ1	2进 JQ2 2# PZ2	2# PQ2	周期 ZQ	原煤 RC
1	1025	12:59:10	1.522672	1.680465	1.987829	4	10.69	39	24	23	39	24	23
2		13:59:23	1.502059	1.670439	1.944577	16	4.52	39	24	21	39	24	21
3		15:19:51	1.511351	1.696173	2.059883	20	6.19	37	24	23	39	24	21
4		16:14:23	1.518703	1.720433	2.05783	56	9.71	39	24	25	39	24	25
5		17:14:03	1.558959	1.773396	2.296125	24	7.99	39	24	25	39	24	25
6		18:14:43	1.558415	1.749625	2.132596	16	8.43	39	24	25	39	24	25
7		19:14:22	1.521999	1.746932	2.249266	16	6.57	39	24	25	39	24	25
8		20:15:01	1.518424	1.707758	2.365479	52	7.14	39	24	25	39	24	25
9		21:14:41	1.518535	1.670648	2.015541	36	7.77	39	24	23	39	24	23
10	1026	3:44:07	1.518703	1.693338	2.047215	48	6.93	39	22	23	39	22	23
11		4:45:00	1.534376	1.698601	1.923626	20	6.91	39	22	23	39	22	23
12		6:44:50	1.506765	1.690264	1.978489	20	7.29	39	25	23	39	22	23
13		13:14:37	1.542259	1.692991	1.863141	12	8.00	39	25	23	39	22	23
14		14:14:16	1.495482	1.683781	1.949976	16	2.28	39	25	23	39	22	23
15		16:15:01	1.511571	1.761304	2.155515	20	7.35	39	25	23	39	22	23
16		17:14:40	1.499161	1.721591	2.103016	16	8.33	39	25	23	39	22	23
17		18:14:20	1.540044	1.742314	2.063217	20	9.86	39	25	23	39	22	23
18		19:14:54	1.516924	1.732139	2.12845	8	11.34	39	25	23	39	22	23
19		20:14:27	1.518146	1.710018	2.153442	12	8.31	39	25	23	39	22	23
20		21:14:59	1.494421	1.690917	2.060517	12	8.81	39	25	23	39	22	23
21		22:19:29	1.522785	1.725199	2.097722	20	10.20	39	25	23	39	22	23
22	1027	12:39:15	1.53622	1.675585	1.872353	12	10.05	39	23	23	39	22	22

附录 2. 优化前各参数值

appendix 2. parameters value before optimum

序号	上密度 FL1 g/cm ³	中密度 FL2 g/cm ³	下密度 FL3 g/cm ³	WASTE %	1 进 JQ1 %	1 膨 PZ1 %	1 排 PQ1 %	2 进 JQ2 %	2 膨 PZ2 %	2 排 PQ2 %
1	1.522	1.680	1.987	10.69	39	24	23	39	24	23
2	1.502	1.670	1.944	4.52	39	24	21	39	24	21
3	1.511	1.696	2.059	6.19	37	24	23	39	24	21
4	1.518	1.720	2.057	9.71	39	24	25	39	24	25
5	1.558	1.773	2.296	7.99	39	24	25	39	24	25
6	1.558	1.749	2.132	8.43	39	24	25	39	24	25
7	1.521	1.746	2.249	6.57	39	24	25	39	24	25
8	1.518	1.707	2.365	7.14	39	24	25	39	24	25
9	1.518	1.670	2.015	7.77	39	24	23	39	24	23
10	1.518	1.693	2.047	6.93	39	22	23	39	22	23
11	1.534	1.698	1.923	6.91	39	22	23	39	22	23
12	1.506	1.690	1.978	7.29	39	25	23	39	22	23
13	1.542	1.692	1.863	8.00	39	25	23	39	22	23
14	1.495	1.683	1.949	2.28	39	25	23	39	22	23
15	1.511	1.761	2.155	7.35	39	25	23	39	22	23
16	1.499	1.721	2.103	8.33	39	25	23	39	22	23
17	1.540	1.742	2.063	9.86	39	25	23	39	22	23
18	1.516	1.732	2.128	11.34	39	25	23	39	22	23
19	1.518	1.710	2.153	8.31	39	25	23	39	22	23
20	1.494	1.690	2.060	8.81	39	25	23	39	22	23
21	1.522	1.725	2.097	10.20	39	25	23	39	22	23
22	1.536	1.675	1.873	10.05	39	23	23	39	22	22

附录 3. 优化后参数值

appendix 3. parameters value after optimum

序号	上密度 FL1 g/cm ³	中密度 FL2 g/cm ³	下密度 FL3 g/cm ³	WASTE %	1 进 JQ1 %	1 膨 PZ1 %	1 排 PQ1 %	2 进 JQ2 %	2 膨 PZ2 %	2 排 PQ2 %
1	1.504	1.667	1.997	8.73	36	26	17	37	23	22
2	1.502	1.699	1.944	4.03	36	24	21	36	23	22
3	1.497	1.709	2.107	5.70	37	24	24	36	23	22
4	1.520	1.710	2.053	8.91	36	25	24	39	23	22
5	1.508	1.732	2.003	8.06	36	24	25	39	23	24
6	1.531	1.749	2.232	6.43	36	25	25	37	24	24
7	1.520	1.693	2.278	6.57	36	24	23	39	24	24
8	1.518	1.707	2.300	7.14	36	23	25	39	23	24
9	1.517	1.683	2.213	7.01	36	22	24	36	24	23
10	1.526	1.692	2.047	5.36	36	22	22	39	23	23
11	1.499	1.742	2.135	5.29	36	22	21	39	24	24
12	1.506	1.692	1.999	7.29	38	25	23	39	22	23
13	1.552	1.692	1.863	7.32	38	24	21	39	22	22
14	1.495	1.738	1.938	5.63	39	25	23	38	22	23
15	1.500	1.764	2.237	7.35	37	25	20	39	22	23
16	1.492	1.719	2.130	8.00	38	25	23	36	22	22
17	1.540	1.764	2.063	7.86	38	26	17	39	23	23
18	1.496	1.741	1.986	10.59	37	23	23	36	22	23
19	1.511	1.694	2.153	8.31	37	22	19	39	22	21
20	1.494	1.690	2.060	8.71	36	24	23	39	22	23
21	1.522	1.731	2.113	5.31	36	24	20	39	22	23
22	1.522	1.726	1.872	11.20	36	24	23	36	22	22

参考文献

- [1] 张世奎, 我国煤炭资源保障程度与合理开发, 中国国土资源报, 2004-02-25
- [2] 吴式瑜, 王美丽, 煤炭在中国能源的地位, 煤炭加工与综合利用, 2006, 5
- [3] 吴式瑜, 叶大武, 马剑, 中国选煤的发展, 煤炭加工与综合利用, 2006, 5
- [4] 杨康, 跳汰选煤评述 选煤技术, 2006, 10
- [5] 谢广元, 选矿学 中国矿业大学出版社, 2001, 8
- [6] 张建民, 王涛, 王忠礼, 智能控制原理及应用 冶金工业出版社, 2003, 2
- [7] 蔡自兴, 智能控制 第2版 电子工业出版社, 2004, 8
- [8] 戴忠达, 自动控制理论基础 清华大学出版社, 1991
- [9] 黄乾贵, 张艳, 人工智能的发展现状与展望, 煤矿机械, 2004, 2
- [10] 朱金波, 用人工神经网络确定跳汰分选指标及最佳操作, 中国矿业大学学报, 1999, 2
- [11] 钱立全, 跳汰床层松散度的灰色预测及神经网络控制方法, 矿业工程, 2004, 5
- [12] 贾廷贵, 刘剑针, 智能选煤的探讨与分析, 中国洁净煤技术, 2004, 2
- [13] 周明, 孙树栋, 遗传算法原理与应用 国防工业出版社, 1999
- [14] 刘勇, 康立山, 非数值并行算法, 科学出版社, 1995
- [15] 余雪丽, 孙承意, 冯秀芳等, 神经网络与实例学习 中国铁道出版社, 1996
- [16] 云庆夏, 进化算法, 冶金工业出版社, 2000
- [17] 周明, 孙树栋, 遗传算法原理及应用, 国防工业出版社, 1999
- [18] 蔡自兴, 智能控制, 电子工业出版社, 2004
- [19] 袁曾任, 人工神经元网络及其应用, 清华大学出版社, 1999
- [20] 高隽, 人工神经网络原理及仿真实例, 机械工业出版社, 2003, 7
- [21] 袁曾任, 人工神经元网络及其应用, 清华大学出版社, 1999
- [22] 许东, 吴铮, 基于 MATLAB 6, X 的系统分析与设计—神经网络, 第2版, 西安电子科技大学出版社, 2002, 9
- [23] 符东旭, 跳汰机自动控制的途径, 煤炭自动化, 1998, (1): 21-23
- [24] K, J, Astrom, J, J, Anton, K, E, Arzeu, Expert Control, Automation, Vol,

22, No3, P277-283, 1996

[25] 符东旭, 熊诗波, 跳汰机床层密度分布规律的研究, 选煤技术, 2001, (1): 13-17

[26] Bennett S, A brief history automatic control, IEEE Control Systems, 1996, 16(3):17-25

[27] Nielsen R H. Kolmogorow's Mapping Neural Network Existence Theorem. San Diego: Proc. IEEE First International Conference on Neural Networks. 1987, 3: 11-14

[28] H, White, et al, Artificial Neural Networks: Approximation and Learning Theory, Cambridge, Mass: Blackwell, 1992, 16—18

[29] 焦李成, 神经网络的应用与实现, 西安电子科技大学出版社, 1995

[30] 张建明, 智能控制原理及应用, 冶金工业出版社, 2003

[31] Sun Zhihui, Li Guochen, Hu Wen, Design of fresh-keeping system of intelligent pressure reducing and air conditioning based on fuzzy control, 2004 CIGR international Conference

[32] 汪镭, 吴启迪, 蚁群算法在连续空间寻优问题求解中的应用, 控制与决策, 2003(1)

[33] Dorigo M, Gambardella L, Ant colony system: A cooperative learning approach to the traveling salesman problem, IEEE Trans, Evolutionary Computation, 1997, 1(1):53-66

[34] 薛定宇, 控制系统计算机辅助设计—MATLAB 语言应用, 清华大学出版社, 1996, 6

[35] 王骥程, 祝和云, 化工过程控制工程, 第 2 版, 北京: 化学工业出版社, 1991

[36] 徐立鸿, 预测控制方法及其应用, 上海: 同济大学出版社, 1999

[37] Dawson, C, W, Wilby, R, An artificial neural network approach to rainfall runoff modeling, Hydro Science , 1998, 43(1):47~66

[38] Michael Nikolaou, Model Predictive Controllers: A Critical Synthesis of Theory and Industrial Needs , Chemical Engineering Dept., University of Houston

[39] Yamazak T, An improved algorithm for a self-adjusting Controller and its Experimented Analysis, Ph, D, Thesis, QMC, Univ of London, 1982, 9

[40] B, Wayne Bequette, Model Predictive Control, 2000, 9

[41] J, M, Maciejowski, Predictive Control with Constraints, 2000, 10, 14

- [42] Mark J, Willis Ming, T Tham, Advanced Process Control, Chemical and process engineering
- [43] Richalet J, Model Predictive Heuristic Control: Application to Industrial Processes. Automatica. 1978, 14(5), 413-428
- [44] Rouhani R, Mehra R.K., Model Algorithmic Control (MAC), Basic Theoretical Properties. Automatica. 1982, 18(4), 401-414
- [45] Garcia C E, Morari M. Internal Model Control. A Unifying Review and Some New Results. Ind. Eng. Chem. Pro. Des., 1982, 21(2), 308-323
- [46] 何振亚, 神经智能-认知科学中若干重大问题的研究, 湖南科学出版社, 1997
- [47] 张晓强, 在全国煤炭工作会议上的讲话, 国家发展和改革委员会副主任
2005-7-21
- [48] 濮洪九, 在煤炭工业经济运行座谈会上的讲话 2004, 中国能源, 2001, (6)
- [49] Rossiter J A, Rossner, J R and Kouvaritakis B, Infinite horizon stable predictive control, IEEE Trans, Automat, Control, 1996, 41 (10):1522-1527
- [50] Rawlings, J B and Muske K B, The stability of constrained receding horizon control [J], IEEE Trans, Automatic Control, 1993, 38 (II):1512- 1516
- [51] Konvaritakis B, Rossitcr ,JA and Chang AOT, stable generalized predictive control[J], Proc, IEE, Part D, 1992, 139(5):349-362
- [52] Chen H, Stability and Robustness Consideration in nonlinear model Predictive Control [M] Duesseldorf: VDI Verlag; 1997
- [53] 席裕庚, 关于预测控制的进一步思考, 控制理论与应用, 1994, 12(2) : 219-221
- [54] 钱颂迪, 运筹学, 第 2 版, 北京: 清华大学出版社, 1991, 413-473,
- [55] 魏权龄等, 数学规划引论, 北京: 北京航空航天大学出版社, 1991, 413-473
- [56] 席裕庚, 李康, 工业过程有约束多目标多自由度优化控制的可行性分析, 控制理论与应用, 1995, 12(5) :590-596
- [57] 宣家骥, 线性不等式组的相容条件和目标规划算法, 系统工程理论与实践, 1985, 5(4) : 10-7
- [58] Meng Fanqin, Wang Yaocai, Intelligent control and maintenance management integrated system based on mutiagents for coal preparation plant, J. China Univ. of Ming &

Tech(English Edition), jun. 2006, vol. 16, No.2

[59] Kuznetsov A, Optimization in Constrained Predictive Control, Proc, Of the OPERNICUS Summer School on Identification and Optimization, Prague, The Czech Republic, 1995

[60] 方崇智, 萧德云编著, 过程辨识, 北京: 清华大学出版社, 1988, 162

[61] 刘良宏, 袁渭康, 固定床反应器的控制, 化工学报, 1998, 47(6):727-724

[62] Scokaert P O , Rawlings J B, Infinite horizon linear quadratic with constraints[A], 13th IFAC World Congress[C], San Francisco, 1996, M:109- 114

[63] Bamberger W, Isermann R, Adaptive On-Line Steady-State Optimization of Slow Dynamic Processes, Automatic, 1978, 14(3):223-230

[64] 刘漫丹, 大纯滞后过程的自适应模糊逻辑控制系统及其应用研究, 浙江大学博士学位论文, 2000, 88-98

[65] Economou C G, Morari M, Internal Model Control: Extension to Nonlinear Systems, Ind, Eng, Chem., Res, 1986, 25(2): 403-411

[66] Li W C, Biegler L T, Process Control Strategies for Constrained Nonlinear Systems, Ind, Eng, Chem., Res, 1988, 27(8): 1421 — 1433

[67] Li WC, Biegler L T, Multistep Newton Type Control Strategies for Constrained Nonlinear Process, Chem., Eng, Res, Des, 1989, 67(6):562-577

[68] Patwardhans S C, Madhavan K P, Nonlinear Model Predictive Control Using Second Order Model Approximation, Ind, Eng, Chem., Res, 1993, 32: 334

[69] 席裕庚, 王凡, 非线性系统预测控制的多模型方法[J], 自动化学报, 1996, 22(4) : 456-461

[70] Hernandez E, Arkun Y, Control of Nonlinear Systems Using Polynomial ARMA Model, J, 1993, 39(3):446-460

[71] Zhu X, Seborg D E, Nonlinear Predictive Control Based on Hammer stein Models, Proc, Of PSE, 1994

[72] Clarke D W , Mohtadi C, Tuffs P S, Generalized predictive control: II, Extension and interpretations [J], Automatic, 1987, 23 (2) : 149~ 160

[73] Zhang Y, Mehta P G, Feed forward and Feedback Adaptive Control Design and

- Analysis [C], Beijing: Proceedings of 14th IFAC World Congress, 1999, I: 409~414
- [74] Fa Jenglin, Mu Pingchen, Wavelet neural network control for linear ultrasonic motor drive via adaptive sliding-mode technique, IEEE Transactions on Ultrasonics, June 2003, vol.50, No.6
- [75] 侯忠生, 非参数模型及其自适应控制理论, 科学出版社 , 1999
- [76] 舒迪前, 预测控制系统及其应用, 北京: 机械工业出版社, 1996
- [77] R, Toag, A Retrospective View of Fuzzy Control System, Fuzzy Sets and System, 1984, 2:199~210
- [78] B, R, Gaiaee&M, L, G, Shaw, From Fuzzy Logic to Expert Systems, Automatic Sciences, 1988, (1) :5-16
- [79] 鲍新福, 自调整比例因子模型控制器, 自动化学报, 1987, (2) : 129—133
- [80] 李洪光, 工程模糊数学方法及应用, 天津: 天津科学技术出版社, 1993, 10
- [81] 许建华, 统计学习理论, 电子工业出版社
- [82] 龙开照, 汪培应, Fuzzy 控制规则自调整问题模糊数学, 1982; (2)
- [83] 李士勇, 糊控制和智能控制理论与应用, 哈尔滨: 哈尔滨工业大学出版社, 1990
- [84] 诸静等著, 模糊控制原理与应用, 北京: 机械工业出版社, L995
- [85] Barambones O, Etxebarria V, Robust neural control for robotic manipulators, Automatic, 2002, 38(2):235-242
- [86] 贺剑锋, 模糊控制的新进展, 控制理论与应用, 1994, 11(2) : 129-136
- [87] Clarke D W, e1, Generalized Predictive Control-Part I and Part II, Automatic, 1987, 23(2):137-160
- [88] 梁坚, 杨永臻, 多模态自适应模糊控制器, 重庆大学学报, 1995, 18(6) :82-87
- [89] 白丽, 杨永臻, 分层型多变量自适应模糊控制算法的研究, 重庆大学学报, 1995, 18(2) : 27-32
- [90] Zhao Yingkai, Wang Yiling, Intelligence control of nonlinear system based on Matlab simulation and the real-time control platform, Proceedings of the 4# World Congress on Intelligent Control and Automation, June 10-14, 2002
- [91] Marek B. Zaremba, Gerard Morel, Integration and control of intelligence in distributed manufacturing, Journal of Intelligent Manufacturing, 14, 25-42,2003

- [92] 李少远, 席裕庚, 陈增强, 智能控制的新进展, 控制与决策, 2000, 15 (1): 1-5
- [93] 姚昆亮, 跳汰分层准重液机理探讨, 选煤技术, 1999, 第3期
- [94] 潘正君, 康立山, 陈毓屏, 演化计算 [M], 北京: 清华大学出版社, 1998,
- [95] 金菊良, 杨晓华, 丁晶, 基于实数编码的加速遗传算法 [J], 四川大学学报, 2000, 32(4): 20-24,
- [95] “TKX—1型跳汰机自动控制系统的研究”通过鉴定, 选煤技术, 1999, 第六期
- [96] 李彦琴, RBF型神经网络预测控制在卫星姿态仿真系统中的应用技术研究与试验, 硕士学位论文, 北京化工大学, 2003
- [97] 陈增强, 袁著祉, 张燕, 基于神经网络的非线性预测控制综述, 控制工程, 2002年04期
- [98] 陈兵, 王华, 基于神经网络的单步和多步电弧炉模型的研究与应用, 现代电子技术, 2003年02期
- [99] 岳恒, 周晓杰, 王昕, 柴天佑, 白焰, 钢球磨中储式制粉系统智能解耦控制仿真系统, 系统仿真学报, 2002 Vol.14 No.10 P.1340-1343

攻读博士学位期间发表的学术论文目录

- [1] 符东旭, 煤泥重介工艺的应用效果及作用分析, 中国煤炭, 2002, 第 12 期, 第 38 页—第 40 页
- [2] 符东旭, 用递归人工神经网络实时检测跳汰过程, 中国煤炭, 2004, 第 12 期, 第 56 页—第 58 页
- [3] 王国炎 左乙山 杨洁明 符东旭 熊诗波,《选煤新技术的研究与应用》第七章, 煤炭工业出版社, 北京, 2005 年, 共 482 页
- [4] 符东旭, 跳汰机矸石段过程状态的智能识别, 中国煤炭, 2005, 第 10 期, 第 51 页—第 53 页
- [5] 陶有俊, 符东旭, 陶东平, 刘炯天, 赵跃民, 复合物理力场中细粒煤脱硫研究, 煤炭学报, 2005, 第 4 期, 第 493 页—第 496 页, EI 收录, 收录号 05399388723
- [6] 邱长贵 , 陈平涛 , 符东旭, 煤泥水药剂自动添加优化控制在邢台矿选煤厂的应用, 河北煤炭, 2005, 第 3 期, 第 42 页—第 43 页
- [7] 符东旭, 熊小晋, 遗传算法和人工神经网络在跳汰过程状态识别中的应用, 振动、测试与诊断, 2007, 第 02 期, EI 收录, 收录号 073110727759
- [8] 符东旭, 陈兴辉, 基于小波与径向基神经网络的轴承故障诊断, 振动、测试与诊断, 2007 年增刊。
- [9] 符东旭, 熊小晋, 基于人工神经网络的跳汰模型研究, 振动、测试与诊断, 2006 年增刊, EI 收录, 收录号 070110349440

攻读博士学位期间科研情况

1. 发明专利: 跳汰洗选床层密度分布探测装置, 第三发明人, 2008 年。
2. 鉴定: 老年无烟煤重介洗选关键技术研究, 河南省科技厅鉴定, 国际领先水平, 技术第一负责人, 2007 年。

致谢

本文是在尊敬的导师熊诗波教授悉心指导下完成的，从论文的选题、构思到撰写及修改完成都凝聚了导师大量的心血和汗水。导师的渊博、思辨、严谨以及实事求是的科学态度、孜孜不倦的工作作风和对学生至真至诚的关怀，学生将终生铭记，引以为楷模。在此论文完成之际，谨向导师致以最崇高的敬意和最衷心的感谢！

感谢师母金老师在生活上的多方照顾。

感谢程绗老师的大力协助。

感谢师兄李文英、梁义维、杨洁明博士给予的多方面的支持和帮助。

感谢同学熊小燕、魏晋宏及陈兴辉、周成召硕士给予的帮助。

感谢河北金牛能源股份有限公司的领导、同事给予的帮助和理解。

感谢所有关心、帮助本论文研究的老师、同学和亲友，特别是我的父母亲和妻子多年来对我的学业和生活的关心、支持和帮助。