

# 关于“信息-知识-智能转换规律”的研究

钟义信

(北京邮电大学,北京 100876)

**摘要:** 人类以资源和资源转换能力为生.工业时代的表征性资源是能量,“能量转换与守恒定律”就成为整个近代科学技术的基石;信息时代的表征性资源是信息,“信息-知识-策略-行为的转换与统一理论”将成为引领信息时代科学技术发展的灵魂.本文将阐明信息、知识、策略、行为的概念及其转换机制,建立“信息-知识-策略-行为的转换与统一理论”.这里,信息理论将大大超越 Shannon 理论,知识理论将大大突破“知识工程”,同时,原来的结构主义学派、功能主义学派、行为主义学派将被综合成为一个有机的整体.信息-知识-策略-行为的转换与统一理论将为智能科学的进一步发展提供新的视野和活力.

**关键词:** 全信息理论; 知识理论; 信息-知识-策略-行为的转换与统一理论

**中图分类号:** TP18 **文献标识码:** A **文章编号:** 0372-2112 (2004) 04-0601-05

## A Study on Information-Knowledge-Intelligence Transformation

ZHONG Yi-xin

(Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** Resources are the bases of human life. Energy is the characteristic resource in Industrial Age and therefore, the Law of Energy Transform and Conservation became the solid foundation of science and technology in modern time. Information is the characteristic resource in Information Age and thus the Transformational and Unified Theory of Information, Knowledge, Strategy and its Execution will be the soul of science and technology development in Information Age. The concepts and the mechanism of information, knowledge, strategy and its execution will be explained and the unified theory of information, knowledge, strategy and its execution can thus be established within which the information theory has yet greatly gone beyond Shannon Theory, the knowledge theory has also dramatically extended from knowledge engineering, and the three existing approaches of intelligence theory, the neural networks approach, the expert systems approach, and the sensor-motor approach, are integrated into a uniform system. The unified theory will provide new vision and new vigor for the further development of intelligence science.

**Key words:** comprehensive information theory (CIT); knowledge theory; unified theory of information, knowledge, strategy and execution

### 1 引言

历史,常常在无意间“重演”.这就是“螺旋式上升”的发展逻辑.比如最现代化的智能理论研究,就重演了古印度寓言“盲人摸象”的情景.而且,这种摸索长达半世纪之久,确实令人感慨良深.

众所周知,长期以来,人工智能理论的研究存在三个不同的学派:始于1943年基于结构主义精神的神经网络学派,始于1956年基于功能主义精神的符号逻辑学派,以及始于上世纪80年代末基于行为主义精神的感知控制学派.三大学派各执一端,互不认可.其间,还发生过许多不快的争端:1969年, Minsky 和 Papert 出版的专著《Perceptron》断言人工神经网络前景不妙,导致人工神经网络的研究陷入近20年的“黑暗年

代”;而标志神经网络复兴的1987年第一届 IEEE International Conference on Neural Networks 则宣称“AI is dead. Long live Neural Networks”.此后不久,行为主义的代表人 Brooks 以他研制成功的“六脚虫”机器为据,声称智能“无需知识的表示和推理”.

这种鼎足三分的局面不能不促使人们认真思考:智能这只“大象”的本质究竟是什么?为什么会产生三大学派的分庭抗礼?它们之间的本质区别与联系究竟在哪里?有没有使三大学派殊途同归的可能?

针对这些问题,作者和他的同事们进行了长期的研究和探索,终于对这只困惑了人们半个世纪之久的“智能”大象获得了初步的总体认识.它的核心本质便是本文将要阐明的“信息-知识-策略-行为的转换与统一理论”.

以下各节将逐步揭开“智能大象”的神秘面纱.第二节首

先阐明“智能”的基本定义和相关概念,第三节给出作者关于全信息理论的基本结果,第四节介绍作者关于知识理论的研究进展,第五节介绍智能理论的“三分归一统”,第六节是全文的结论,提炼出“信息-知识-策略-行为的转换与统一理论”。

## 2 智能定义和相关概念<sup>[1]</sup>

时至今日,人们已经拥有许多不同的智能定义。例如,Webster 辞典认为智能是“理解和适应的能力(capacity for understanding and for other forms of adaptive behavior)”;Oxford 辞典则认为智能是“进行观察、学习、理解和认识的能力(power of seeing, learning, understanding, and knowing)”,等等。

这些说法各有一定道理,但它们都只是关于智能的“现象学描述”,而不是关于智能本质的揭示。

我们认为,智能,可以概略地定义为“认识问题和解决问题的能力”。这大概不会有什么争议,但需要深入。

进一步的分析可以发现:智能,是一种有目的的行为;没有目的,谈不上有智能。目的,通常又由一系列具体的目标组成;所有具体目标都为实现最终目的服务。有目的,才能在具体环境中发现问题和定义问题;有目的,才能确定解决问题的目标;有目标,才能找到解决问题达到目标的策略。这样,就可以把智能定义为:在总体目的的驱动下,面对任何给定的环境,发现(定义)问题、确定目标、解决问题、达到目标的能力。

但在确定环境、问题、目标之后,怎样才能解决问题达到目标?所谓“确定了环境、问题和目标”,就意味着获得了关于环境、问题和目标的信息。为了解决问题,就必须拥有合理的策略,而后者只能根据与环境、问题及目标相关的信息和知识制定出来。于是,智能的完整定义就可以表述为:

**定义 1** 智能,是在总体目的驱动下,面对任何给定的环境,发现(定义)问题、确定目标、获得问题-环境-目标的信息、把信息提炼为知识、把知识激活为合理的策略、在策略引导下解决问题(满足约束)达到目标的能力。

用符号表示就是:

$$Intel(Info, K, S, Exe | O, E; P, G) \quad (1)$$

式中的符号 *Intel*, *Info*, *K*, *S*, *Exe*, *O*, *E*, *P* 和 *G* 分别表示智能、信息、知识、策略、策略执行、目的、环境、问题和目标。式中表明,智能可以划分为两个有联系又有区别的层次:隐智能(式中竖杠后面的部分),显智能(式中竖杠前面的部分)。

隐智能是指:在总体目的的驱动下,面对任何给定的环境,发现问题、定义问题和确定目标的能力;显智能则是指:在此给定前提下,获得信息、把信息提炼为知识、把知识激活为策略、把策略转化成行为(解决问题)的能力。隐智能体现的是目的,显智能体现的是手段。

式(1)表示的智能是一个“从认知到行事”的完整过程,可以称为“广义智能”,它的标志是“在总体目的的驱动下,成功地解决问题,达到目标”;而其中的“认知与决策”则可以称为“狭义智能”,它的标志是“形成了求解问题的策略”。所以,策略是狭义智能的体现,并常常称为智能策略。不难看出,广义智能贯穿在人们认识世界和改造世界的全部活动过程中,狭义智能则主要在人的思维器官范围内展开。在很长时间内,人

们比较注意狭义智能,然而,狭义智能不可能独立存在,它是广义智能的一个部分。

式(1)是一个比较有意义的智能定义。这是因为它不仅说明了智能是怎样的一类现象(What),而且说明了智能是怎样生成的(How),以及为什么智能与目的、信息、知识有关(Why)。显然,这里所说的“智能”是以生物系统特别是高等生物系统——人类系统为原型的,因此也可以称为“人类智能”。

那么,什么是人工智能?

人工智能是泛指机器系统具有的智能。它的定义可以这样表述:

**定义 2** 人工智能,是在给定问题、环境、目标的前提下,机器获取相关的信息、把信息提炼为知识、把知识激活为策略、并在策略引导下满足约束解决问题达到目标的能力。

用符号表示就是

$$AI(Info, K, S, Exe | E, P, G) \quad (2)$$

其中, *E*, *P*, *G* 是由系统设计者给定的具体条件。

比较式(1)和式(2)可以发现,人工智能与人类智能之间,既有本质联系又有原则区别。它们的联系表现在:两者都具有显智能的特征,即都具有“获取信息、提炼知识、生成策略、解决问题、满足约束、达到目标”的能力。它们的区别主要表现在:人类智能是显智能和隐智能的统一体;人工智能则只具有显智能,而不具有隐智能的能力。这是因为,人,是具有自觉能力的系统,有自身的明确目的;而机器只有系统设计者所赋予的目标。这种区别决定了:机器只能在人的安排下发挥聪明助手的作用。

值得注意,“如何在目的的驱动下,在任意给定环境中发现问题、定义问题和确定目标”,这是人类智能独具的能力,也是人类“创造力”的先决条件和首要基础:发现问题,是成功的一半;不能发现问题,就不可能解决问题。由于篇幅所限,本文暂不讨论这一课题。

式(1)和式(2)两个定义清楚表明,无论是人类智能还是人工智能,获取信息、提炼知识、生成策略(狭义智能),都是关键的要素。它们是人类智能和人工智能系统的核心机制和灵魂。下面就分别来讨论这些要素。

## 3 全信息理论<sup>[2]</sup>

信息是智能的源头,一切智能都是通过对信息进行适当的加工形成的。没有信息,智能便是无源之水。目前流行的信息理论是 1948 年 Shannon 在 *BSTJ* 杂志发表的“通信的数学理论(Mathematical Theory of Communication)”,后人称之为信息论。众所周知,通信的技术本质是信息传输,它用随机信号波形来表示所传送的信息。只要在接收端把发送端发来的信号波形恢复出来,就完成了通信的任务。因此,Shannon 理论只关心信号波形,不关心内容和价值。

从智能研究的性质来看,Shannon 理论显然不能满足需要。这是因为,不了解信息的内容和价值就很难据此做出合理的决策,而做出合理决策是智能研究的基本要求。为此,需要对 Shannon 信息论进行大幅度的改造。

信息的概念相当复杂,但它最基本的层次是本体论信息

和认识论信息。

定义 3 事物的本体论信息,是关于“事物的运动状态及其变化方式”的直接表现,与观察主体的因素无关。

定义 4 认识论信息是“主体所感知的事物运动状态及其变化方式,包括这种状态方式的形式(称为语法信息)、含义(称为语义信息)和价值(称为语用信息)”。语法信息、语义信息、语用信息三者的全体,称为“全信息”。

由于通信和智能涉及信息的深度很不相同,面向通信的信息可以看作是本体论信息;然而面向智能研究的信息却必须是认识论信息。

改造后的信息理论称为“全信息理论”。这是因为,首先,全信息理论不仅研究随机型信息,也研究非随机型信息(如模糊信息);同时,不仅研究语法信息,还研究语义信息和语用信息,全面满足了智能研究对信息理论的要求。我们把“全信息理论”的英文表示为 Comprehensive Information Theory,意思是:“全信息理论”研究的信息不仅覆盖完全,而且具有“理解”的含义。这是“全信息理论”能够满足智能研究需要的根本原因。关于“全信息理论”的系统阐述,可以参见《信息科学原理》一书的第一、二和三章<sup>[2]</sup>。书中证明,Shannon 的信息测度(熵)公式是全信息测度公式的一种特例,如图 1 所示。

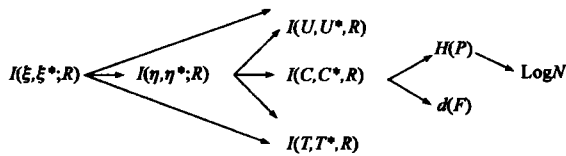


图 1 全信息测度与 Shannon 信息测度的关系

图中,  $I(\xi, \xi^*; R)$ ,  $I(\eta, \eta^*; R)$ ,  $I(C, C^*; R)$ ,  $I(T, T^*; R)$ ,  $I(U, U^*; R)$  分别是综合语用信息(即全信息)、综合语义信息、语法信息、语义信息、语用信息的测度公式,而  $H(P)$ ,  $d(F)$ ,  $\text{Log}N$  则分别是 Shannon 的信息熵公式、DeLuca-Termini 的模糊熵公式、Boltzmann-Ashby 的均匀分布熵公式。箭头的含义是:一种测度退化为另一种测度的方向。可见,Shannon 熵公式是全信息测度的退化情形,相应于各种退化过程的具体条件参看文<sup>[2]</sup>。

全信息理论的重要意义在于:通过分析语法信息,可以了解信息所反映的事物的结构;通过分析语义信息,可以了解信息所反映的事物的含义;通过分析语用信息,可以了解信息所反映的事物的价值。只有了解了信息的形式结构、内容含义和价值效用,才可能做出合理的、有智能的决策。

其实,全信息理论不仅为加工智能提供了必要的基础,而且对获取信息也具有重要的指导作用:人们在获取信息的时候,不能只考察信息的形式因素(语法信息),还必须考察信息的内容(语义信息)和价值(语用信息)。当前流行的搜索引擎给用户反馈的信息之所以严重缺乏针对性,问题就在于它们只利用了语法信息(关键词的形式匹配),没有利用语义信息和语用信息。搜索引擎(包括数据挖掘、知识发现、机器学习等等)的改进方向,也在于要设法有效地利用全信息。由此可以看出全信息理论在智能研究中的基础地位。

### 4 知识理论<sup>[3]</sup>

研究还发现,在很多情况下,信息加工的结果首先得到的是“知识”,而不是直接得到智能。智能是跟解决问题的目标联系在一起。因此,在没有明确目标之前,不可能直接处理智能的问题,但是却可以加工出知识,因为知识并不需要直接与目标相联系。这当然也可以有例外,后面还会提到这一点。

那么,什么是“知识”?最流行的知识定义是:

定义 5 知识是人们实践经验的结晶;经验,是有待确证的准知识。

这一定义不但表明了知识生成的重要途径——由经验升华而成知识,而且把经验也归入了知识的体系,这显然是合理的。不过,知识的生成不仅仅可以通过经验升华的途径来实现,也可以通过知识本身的推演来实现。因此,我们还需要更深刻的知识定义。

定义 6 所谓关于某类事物的“知识”,是人们关于这类事物的运动状态及其变化规律的描述,包括这种状态和规律的形式(形态性知识)、含义(内容性知识)和价值(效用性知识)。形态性知识、内容性知识、效用性知识三者一起构成了完整的知识。

比较定义 4 和定义 6 发现,除了“变化方式”换成了“变化规律”,全信息的定义和知识的定义几乎没有区别。这就可以理解:为什么我们说“知识是由信息提炼出来的抽象产物”,这个抽象的产物,就是由具体的“变化方式”提炼出来的“变化规律”。这就是知识与信息的本质联系与区别。不过,这里的信息应当是“全信息”。从 Shannon 信息只能获得统计的规律,只有从全信息才能获得可理解的规律。这里又一次看到了全信息理论的重要性。

由上面的讨论可以看出,由信息提炼知识的基本方法是归纳抽象:即把原始的具体的信息(事物运动状态及其变化方式)通过归纳处理,成为抽象的、具有一定普遍意义知识(事物运动状态及其变化规律)。当然,除了归纳抽象的基本方法之外,由信息加工提炼知识的过程也还存在其它的方法(如类比、联想、趋势外推、中值内插等等)。至于由原有知识推演出新知识的基本方法,则主要是演绎推理。但是,不管是由信息归纳的知识还是由演绎推演的知识,它们共同的特征都必然是“关于事物运动的状态及其变化的规律的描述”。

根据上面对于信息和知识定义的讨论,我们可以进一步把信息和知识的关系用图 2 来表示:

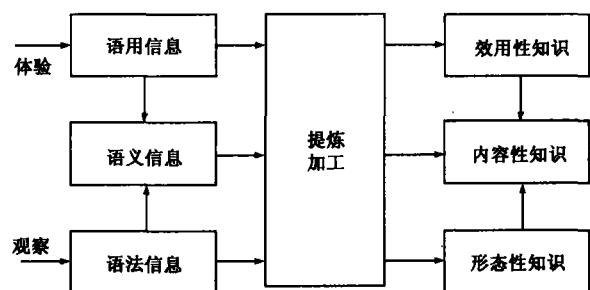


图 2 信息与知识的关系

如前所见,信息有语法信息、语义信息、语用信息之分,它们的有机整体称为“全信息”。与此相应,图中形态性知识、内容性知识、效用性知识的整体也可以称为“全知识”。图中还示出,在最原始的场合,语法信息可以通过直接观察(感知)获得,语用信息可以通过体验得到,语义信息则可以通过语法信息与语用信息的综合处理得到;知识的情形与此相似。相应的原则算法参见文[3]。

由定义6还可以引出另外一个重要的知识概念:常识。

**定义7** 常识是被普遍公认因而无需证明的知识。

这显然也是合理的:因为任何常识必然都是“关于某类事物运动状态及其变化规律的描述”,因而也是知识。只不过,由于这种“规律”是人所共知,司空见惯,习以为常,人们往往不把它当知识看待。但不管常识是怎样简单,它确实确实是知识,这是毫无疑问的。

图3示出了经验性知识、正规性知识、常识性知识的关系。

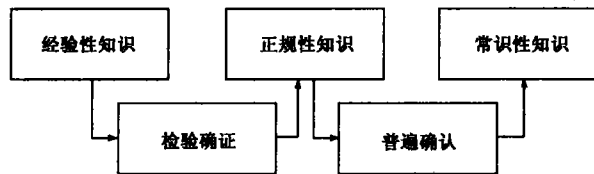


图3 经验、知识、常识的关系

总之,“知识”应当包括经验性知识、正规性知识、常识性知识。这是一个完整的知识的体系,不妨称之为“广义知识”。在尚未得到检验确证之前,知识应当是经验性知识;而知识一旦被普遍确认进入不言而喻状态的时候就成为常识。随着时间的不断向前推移,经验性知识将演变为正规性知识,后者将演变为常识。这是一个动态演进的过程。

除了知识生成的理论之外,知识理论还有另一方面的内容,这就是知识向智能的转化,称为知识的激活。我们把它归入下一节介绍。

## 5 综合智能理论<sup>[4]</sup>

前曾指出,知识激活成为(狭义)智能(体现为策略),需要有具体的待求解问题、环境约束条件和问题求解的目标。否则就会成为空洞的智能。但是,根据知识性质的不同,激活知识生成智能的方法也不同。

对于正规性知识而言,激活知识生成智能的基本过程是:在目标引导下,在给定环境条件的约束下,针对面临的问题,运用(知识库中已有的)相关知识进行推理,生成一个问题求解的可能策略;在策略作用下,问题的状态将发生相应的改变,根据问题新状态与目标状态之间的差异程度(效果),可以确定策略是否“既能满足客观约束又能达到主观目标”。这种智能策略的生成通常是一个动态演进和逐次逼近的过程,可以用图4来表示:

可以看出,这种智能策略生成的实质是在“问题-环境-目标”的共同约束下把知识映射为策略的过程,是基于功能主义的“专家系统”的机制。

对于经验性知识的情形,通常不存在事先设计好的知识库,也不是通过确定的知识推理来生成智能策略,而是利用某种经验性的结构(如神经网络),通过试验样本直接训练学习的方法来建立它的智能策略。它的工作过程一般是:

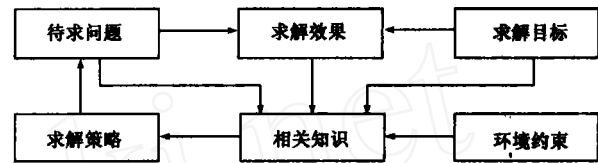


图4 规则性知识的激活算法

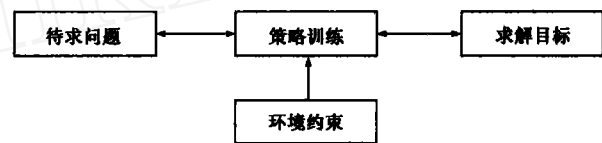


图5 经验性知识的激活模型

这种不利用先验知识库、不采用显式推理、而直接通过对某种经验结构进行样本训练的方法来建立智能策略的思路,是基于结构主义的“神经网络”的典型工作机制。

对于常识性知识的情形,由于这里的智能策略完全基于常识,因此,既不需要像神经网络那样把知识按照一定结构进行表示,也不需要像专家系统那样对知识进行推理;它所需要的只是“感知-控制”的直接过程:只要感知到了所面临的具体输入情况,就按照常识所确定的动作做出相应的反应。这就是基于行为主义的方式。

可见,从知识激活的机制看,人工智能现存的三大学派(基于规则性知识的功能主义学派、基于经验性知识的结构主义学派以及基于常识性知识的行为主义学派)正好构成了有机互补的人工智能理论体系,不妨称为“广义人工智能”。它们共同的机制都包含“知识的激活”,只是由于知识的性质不同,激活的具体方法不同而已。

顺便提及,意识也是介于信息与策略之间的过程,篇幅所限,从略。

## 6 结语:信息-知识-策略-行为的转换与统一理论

讨论表明:正像其它(物质和能量)资源可以转换一样,信息也是一类普遍存在的资源,可以通过相应的加工机制把它转换成为知识、策略和执行策略的行为,最终成为认知与行事的智能。信息是智能的源泉;智能是信息的归宿。这就是“信息-知识-策略-行为的转换与统一理论”,它是智能科学的核心。

具体地说,“信息-知识-策略-行为的转换与统一理论”包括以下要点:

第一,针对给定的问题、环境和目标,通过信息获取可以把相关的本体论信息转换为认识论信息(语法信息或全信息,依具体问题而定);

第二,通过加工提炼(归纳)可以把认识论信息转换成为知识;通过演绎又可以把老知识转换为新知识(经验性知识或正规性知识或常识性知识);

第三,针对给定的问题、环境和目标,可以把知识转换(激

活)成为求解问题的策略:把经验性知识转换成为结构主义的策略;把正规性知识转换成为功能主义的策略;把常识性知识转换成为行为主义的策略;

第四,通过控制和显示,可以把求解问题的策略转变为相应的行为,在满足环境约束的条件下达达到预定的目标,使问题得到解决。

可见,信息获取理论、全信息理论、知识理论、综合智能理论及控制理论之间相互贯通,构成了“信息-知识-策略-行为的转换与统一理论”,构成了智能科学的核心和精髓。

作者相信,“信息-知识-策略-行为的转换与统一理论”是信息科学与智能科学发展到今天必然要产生的结果。信息时代“信息-知识-策略-行为的转换与统一理论”的科学意义可以同工业时代“能量转换与守恒定律”相媲美,因为它们分别揭示了能量资源和信息资源加工转换的根本规律,为人们利用

能量和信息资源提供了理论的指导。自然,本文的结果仅仅是“信息-知识-策略-行为统一理论”的初步,更精彩的结果有待广大同仁共同努力去开发。

#### 参考文献:

- [1] 钟义信,等.智能理论与技术:人工智能与神经网络,邮电出版社,1992.
- [2] 钟义信.信息科学原理(第三版).北京邮电大学出版社,2002.
- [3] 钟义信.知识理论框架.中国工程科学,2000,9(2):50-64.
- [4] 钟义信.Unified Theory of Information, Knowledge and Intelligence. Chinese Journal of Electronics, 2003, 6.

#### 作者简介:

钟义信 男,1940年2月出生于江西赣州市,北京邮电大学教授,博士生导师。

(上接第 600 页)

- [8] K F Cheung, R J Marks II, L E Atlas. Synchronous vs. asynchronous behaviour of Hopfield's CAM neural net[J]. Appl Opt, 1987, 26(22): 4808-4813.
- [9] R P Lippman. An introduction to computing with neural nets[J]. IEEE ASSP Mag, Apr 1987, 4(2): 4-22.
- [10] N Farhat, D Psaltis, E Paek. Optical implementation of the neural model [J]. Appl Opt, 1985, 24(10): 1469-1475.
- [11] L E Atlas. Auditory coding in higher centers of the CNS[J]. IEEE Eng Med Biol Mag, June 1987, 29-32.
- [12] R J Marks II, L E Atlas. Geometrical interpretation of Hopfield's content addressable memory neural network [A]. In Proc. Northcon' 88 [C]. Seattle, WA, Vol. II Oct. 1988. 964-977.
- [13] Y S Abu-Mostafa, J M St. Jaques. Information capacity of the Hopfield model[J]. IEEE Trans IT, 1985, IT-31: 461.
- [14] R J McEliece, E C Posner, E R Rodemich, S S Venka-tesh. The capacity of the Hopfield associative memory[J]. IEEE Trans IT, 1987, IT-33: 461-482.
- [15] D E Rumelhart, J L McClelland, the PDP Research Group. Parallel Distributed Processing [M]. vols. I and II. Cambridge, MA: Bradford Books, 1986. 318-362.
- [16] D E Rumelhart, G E Hinton, R J Williams. Learning representations by back-propagation errors[J]. Nature, 1986, 323 (6088): 533-536.
- [17] C Stanfill, D Waltz. Toward memory-based reasoning[J]. Commun. ACM, Dec. 1986, 29: 1213-1228.
- [18] 王金根. 航空磁探系统中目标信号检测与磁性目标定位研究[D]. 博士学位论文. 武汉: 海军工程大学, 2001, 5.
- [19] 王金根, 林春生, 龚沈光. 基于交替投影神经网络的带限信号外推算法[J]. 电子学报, 2000, 28(10): 52-55.
- [20] 王金根, 龚沈光, 林春生, 唐劲飞, 刘胜道. 基于复交替投影神经网络的陷波器[J]. 数据采集与处理, 2001, 16(4): 440-445.
- [21] WANG Jin-gen, CHEN Shi-fu, GONG Shen-guang, CHEN Zhao-qian. An Extended Alternating Projection Neural Networks based weak-signal separation algorithm [A]. IEEE International Conference on Robotics, Intelligent Systems and Signal Processing [C]. Changsha, China, Oct. 2003. 554-558.