

# 多源数据融合算法综述

祁友杰, 王 琦

(中国航天科工集团 8511 研究所, 江苏 南京 210007)

**摘要:** 多源数据融合作为一种特殊的数据处理手段, 在目标识别领域得到了较大的重视和发展。在介绍多源数据融合的基本原理和功能模型的基础上, 对目前的多源数据融合算法进行了全面介绍, 从算法概念出发对其进行了分类, 分别为物理模型类、基于参数类和基于认识模型类, 并阐述了每类算法的特点以及相关算法的改进, 列举了目前国内外一些已经发表的重要算法, 为下一步多源融合的目标识别研究提供了一定的理论依据。

**关键词:** 多源; 多传感器; 信息融合; 数据处理

**中图分类号:** TN97 **文献标识码:** A

DOI:10.16328/j.htdz8511.2017.06.009

## Review of multi-source data fusion algorithm

Qi Youjie, Wang Qi

(No.8511 Research Institute of CASIC, Nanjing 210007, Jiangsu, China)

**Abstract:** As a special data processing method, multi-source data fusion is received great attention and development in the field of target recognition. The algorithm of multi-source data fusion is introduced integrally on the basis of its basic principle and function model, which is classified as physical model, based on parameter and based on cognitive model. The feature and improvement of each kind of algorithm are expounded. Some important algorithms which have been already published at home and abroad are listed provide theoretical basis for the next target recognition study of multi-source fusion.

**Key words:** multi-source; multi-sensor; information fusion; data processing

## 0 引言

信息融合是生物体所固有的一种本质特征, 是生物体环境感知和行为行动的基础, 其过程是生物体生存、进化和发展的基本能力要素, 它是人类和其他生物系统中普遍存在的一种基本功能。人类非常自然地利用这一能力把来自人类各传感器(眼、耳、鼻和四肢)的信息(景物、声音、气味和触觉)组合起来, 并用先验知识去估计、理解周围环境和正在发生的事件。无论在军事还是在非军事领域, 数据融合技术已经成为全球的研究热点之一。从军事应用的角度, 数据融合被定义为: 把来自多个传感器和信息源的数据和信息加以校准、联合、相关, 合并成统一的表示形式, 以获得精确的目标位置、状态估计、身份识别, 以及对战场态势和威胁的综合评估; 从非军事应用的角度来说, 数据融合是对多个传感器和信息源所提供的关于某一环境特征

的不完整信息加以综合, 以形成相对完整、一致的感知描述, 从而实现更加准确的识别和判断功能。

信息融合可以看成是一种形式框架, 其过程是用数学方法和技术工具综合不同源信息, 目的是得到高品质的有用信息。与单一信源独立处理相比, 信息融合的优势包括: 提高可探测性和可信度、扩大时空感知范围、降低推理模糊程度、改进探测精度等性能、增加目标特征维数、提高空间分辨率、增强系统容错能力等<sup>[1]</sup>。

## 1 数据融合模型

数据融合的初始模型分为: 像素级融合、特征级融合和决策级融合, 后来又被扩展成目标提炼、态势分析、威胁估计和过程精炼等四级<sup>[2]</sup>。四级融合模型通过动态监视融合处理过程, 优化资源和传感器管理, 实时反馈融合结果信息, 以使融合处理过程具有自适应性, 从而达到最佳融合效果。融合方法的研究是数据融合的重要研究内容之一, 与三级融合模型相对应的融合方法<sup>[3]</sup>有: 像素融合方法主要有加权平均法、选

收稿日期: 2017-06-24; 2017-10-15 修回。

作者简介: 祁友杰(1982-), 男, 博士, 主要研究方向为信号处理、图像处理。

举决策法、卡尔曼滤波法、数理统计法等;特征级融合方法主要有卡尔曼滤波法、模糊推理法、神经网络法、产生式规则法等;决策级融合方法主要有贝叶斯概率推理法、Dempster-Shafer 证据推理法(简称 D-S 方法)等。在实际应用中,所有的融合方法都必须面临着处理各种不确定信息的问题。

数据融合的一般功能模型对于设计融合系统结构,以及有效利用多传感器信息具有重要的指导意义。目前,国外最新的数据融合功能模型,是在美国 JDL 数据融合功能模型基础上的改进<sup>[2]</sup>,如图 1 所示。

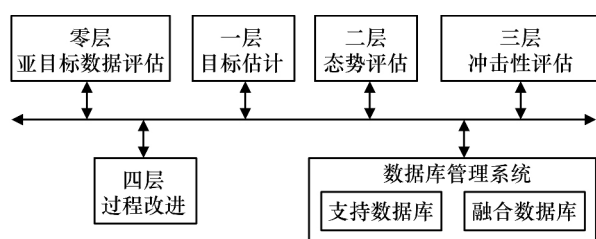


图 1 JDL 数据融合功能模型框图

该功能模型将数据融合分为 5 层:第零层——亚目标数据评估(Subobject Data Assessment):以像素或信号级上的数据关联和特征描述为基础,对具有可观测状态的信号、目标的估计和预测;第一层——目标估计(object Assessment):根据从观测到跟踪所建立的关系进行实体状态估计和预测,包括连续状态估计(如运动状态)和离散状态估计(如目标类型和身份);第二层——态势评估(Situation Assessment):实体之间相互关系的估计与预测,包括:火力结构与火力交叉关系,通信交互和人们之间的直接交互情况,战场周围环境等因素;第三层——冲击性评估(Impact Assessment):对参与者所制订的计划或预测行动结果的估计和预测,主要包括多个参与者行动计划间的相互影响(如估计出弱点以预测对某一计划制订者的威胁动作);第四层——过程改进(Process Improvement):它是一个资源管理成分,支持任务的自适应数据获取和处理。

从处理对象的层次上看,零层属于低级融合,它是经典检测理论的直接发展,是近十几年才开始的研究领域,目前绝大多数多传感器数据融合系统还不存在这一级,仍然保持集中式检测,而不是分布式检测,但分布式检测是未来的发展方向。第一和第二层属于中间层次,是最重要的两级,它们是进行威胁估计的前提和基础。实际上,融合本身发生在前两个层上,而态势评估和威胁估计只是在某种意义上与融合具有相似的含义。

## 2 多源数据融合算法的分类

多源数据融合涉及到多方面的理论和技术,并没有完全统一的算法能够适应所有的场景,所以在应用上,需要针对不同的应用背景选择相应的算法。按算法概念分类,主要分成三大类,分别为:物理模型类、基于参数类和基于认识模型类,如图 2 所示。

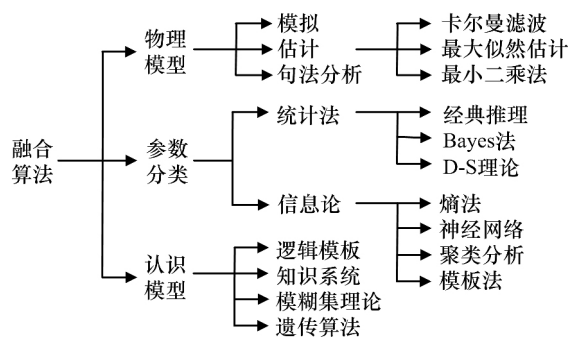


图 2 融合算法的概念性分类

### 2.1 物理模型类

此类算法是根据物理模型直接计算实体特征。预测一个实体特征的物理模型必须以被识别物体的物理特征为基础,而实际物理模型往往相当复杂,建立起来非常困难。尽管实际中很少使用这种方法,但在基础研究工作中却需要使用它。句法分析方法和估计理论法属于物理模型算法中的典型,句法分析方法开发了物理对象的语法和句法,而物理对象说明是分别从传感器数据分量(或原始信息)获得的,因此能对来自这个假定的结构关系的一个对象的表示进行汇集。估计理论法包括卡尔曼滤波算法、最大似然估计法及最小二乘法等。

### 2.2 基于参数类

参数分类算法是最常见,也是应用最广、研究最深入的一类算法。此类算法寻求一个标识说明且使之依赖于参数数据,在参数数据和一个标识说明间建立一个直接的映像。此类算法可进一步分为基于统计的算法和基于信息论技术的算法。

#### 2.2.1 基于统计的算法

统计算法主要包括经典推理法、Bayes 推理算法、基于 D-S 证据理论法等。经典推理:经典推理描述<sup>[4]</sup>在给出目标存在的假设条件下,所观测到的数据与标识相关的概率。其局限性是:仅能估计两个假设,即假设和与其相对的备择假设;当遇到多变量数据时,算法复杂性提高;需要一个先验密度函数的有效度,否则不能直接使用先验估计。Bayes 推理<sup>[5]</sup>:它解决了经典

推理方法的某些困难。它以最小风险代价为基本模型,在给定一个预先似然估计和附加证据(观测)条件下,能更新一个假设的似然函数。Bayes 推理的缺点是:定义先验似然比较困难;当有多个可能的假设和多个条件相关时,显得很复杂;要求对立的假设彼此互不相容;缺乏分配总不确定性的能力。D-S 证据理论<sup>[6]</sup>: D-S 证据理论是 Bayes 理论的广义扩展,它考虑了总不确定性程度。D-S 方法利用了概率区间和不确定区间来确定多证据下假设的似然函数,当所有的假设互不相容且完备时,Bayes 推理技术与 D-S 法产生相同的结果。D-S 法存在的问题是:不能有效地处理矛盾的证据;具有幂指数增长的计算量;推理链较长时,使用证据理论很不方便;D-S 组合规则具有组合灵敏性,但有时基本概率赋值一个很小的变化都可能导致结果很大的变化。

### 2.2.2 基于信息论的算法

在某些场合,多传感器数据融合目标识别并不需要用统计的方法直接模拟观测数据的随机形式,而是依赖于观测参数与目标身份之间的映射关系来对目标进行识别。这类方法称为基于信息论的融合识别算法,这类算法包括神经网络法<sup>[7-8]</sup>、熵理论法<sup>[9]</sup>、表决法<sup>[10]</sup>和聚类算法<sup>[9]</sup>等。

人工神经网络是在现代神经科学研究成果的基础上提出的。神经网络目标识别法是用非常简单的计算处理单元(神经元)进行互联构成的非线性网络系统,具有学习、记忆、计算能力以及各种智能识别处理能力,所有神经元可在没有外部同步信号作用的情况下执行大容量的并行计算。神经网络目标识别算法中最著名的是以自适应信号处理理论为基础发展起来的前向多层神经网络及其逆推学习(BP)算法。

熵法是基于融合系统的一种新技术。利用了事件发生的概率,反映了信息量的思想。它的原理是经常发生的事情熵最小,而不经常发生的事情熵最大,将其用于传感器数据融合过程中,就是要作出使熵极大的结论。熵法很有应用价值,尤其是对于实时性要求很高的系统,当准确的先验统计不可利用时,或者是从整个成本—效益观点来看,熵法都是很有吸引力的,国外已开始将熵法用于计算与假设有联系的信息内容的度量值。

表决法类似于日常生活中的投票选举,它包括布尔“与”、“或”处理,是多传感器数据融合目标识别算法中最简单的技术。它由每个传感器提供对被测对象状态的一个判断,然后由表决算法对这些判断进行搜索,以找到一个由半数以上传感器“同意”的判断(或采取

其他简单的判定规则),并宣布表决结果。也可采用加权方法、门限技术以及其他判定方法等,这种方法在没有准确的先验统计数据可利用时十分有用,特别是对于实时融合很有吸引力。

所谓聚类就是把大量的  $d$  维数据样本聚集成  $n$  个类,使同一类内样本的相似性最大,而不同类内样本的相似性最小。聚类分析法是一种启发式算法,在模式类数目不是精确知道的标志性应用中,这类方法很有效,它是按某种聚类准则将数据分组(聚类),并把每个数据组解释为相应的目标类。聚类分析可以作为一个单独的工具以发现数据库中数据分布的一些深入的、隐含的、有用的信息,并且概括出每一类的特点,或者把注意力放在某一个特定的类上以作进一步的分析。目前,已经提出的聚类算法有很多,主要有分裂法、分层设计法及基于网格的方法等。

### 2.3 基于认识模型类

基于认识模型的算法主要包括逻辑模板法、模糊集理论算法、遗传算法及知识系统法等。

逻辑模板法实质上是一种匹配识别的方法,它将系统的一个预先确定的模式(模板)与观测数据进行匹配,确定条件是否满足,从而进行推理。预先确定的模式中可以包含逻辑条件、模糊概念、观测数据以及用来定义一个模式的逻辑关系中的不确定性等。因此模板实质上是一种表示与逻辑关系进行匹配的综合参数模式方法<sup>[11]</sup>。决策模板法作为逻辑模板法的一类,是一种简单直观的决策层融合目标识别算法。经典的决策模板法没有充分利用各传感器对于不同类目标鉴别能力的先验信息,文献<sup>[12]</sup>提出利用传感器平均度量熵对决策模板法进行修正,合理度量多个传感器对不同类目标的分类鉴别能力,仿真结果表明改进的决策模板法能提高目标正确识别率。文献<sup>[13]</sup>针对经典的决策模板法不能反映传感器对于不同类目标的分类鉴别能力,不能适应待识别目标特征矢量起伏变化,同时没有保留训练样本全部信息等缺陷,提出了基于熵和  $K$  近邻方法的修正决策模板法,从而使得多传感器融合目标识别的性能得到很大的改善。

由于环境的复杂性、噪声干扰、识别系统的不稳定及采用不同识别算法等因素的影响,目标信号以及提取的特征参量信息存在不精确、不完整和不可靠性。此外,进行目标识别的判决方法上也存在一定的主观模糊性。模糊集理论为解决这类问题的强有力数学工具。模糊集理论中丰富的融合算子和决策规则为有效进行目标融合处理提供了必要的手段。近年来,国内外一些学者已开始将模糊集理论应用于目标身份确认

问题。如 Kewley<sup>[14]</sup> 基于电子侦察数据,使用模糊逻辑融合多传感器数据进行目标身份确认。Ludovic Roux 等<sup>[15]</sup> 基于模糊集的可能性理论,进行多谱卫星图像融合以解决图像分类问题。文献[16]通过运用模糊集理论中的模糊综合函数,建立了多传感器目标识别决策级融合模型,用可能性分布表征和描述传感器目标识别结果,讨论了基于模糊综合函数的目标识别融合算法,最后举例说明了该方法的有效性。文献[17]将模糊测度和模糊积分引入多传感器数据融合目标识别,给出了模糊积分应用于决策层数据融合目标识别的通用技术路线,并提供了信源重要程度的度量方法。文献[18]给出了一种多源多层次自适应变权的多传感器目标融合的模糊算法,提供了一种表达和处理不确定性信息的有效方法,充分利用了各传感器信息源之间的冗余性和互补性以及各信息源自身的可靠性信息,达到了较好的识别效果。不同类型传感器在相同的工作环境下具有不同的识别能力,模糊集理论算法考虑了信源的重要程度,反应了客观实际,融合识别结果优于 D-S 证据理论,并且相对于 D-S 证据理论,模糊积分中不涉及识别框架的论域问题,有利于增大识别框架,提高融合系统的实用性。

遗传算法是一个群体优化过程,它由一组初始值出发进行优化,优化过程就是这个群体不断繁衍、竞争和遗传、变异的过程<sup>[19]</sup>。文献[20]将遗传算法与神经网络目标分类器结合,通过识别结果的反馈信息,控制遗传算法的遗传进化方向实现特征优化,为克服遗传算法的未成熟收敛问题,提出了相关选择与自适应遗传算子相结合的改进遗传算法,实验结果验证了方法的有效性。

知识是对某些客观对象的认识,并通过属性来刻画这种认识。如果能够给出某个对象集合的属性特征,则该对象集合表达了一种完全知识<sup>[21]</sup>。文献[22]以 PSM 和本体为核心,构筑了知识系统建模框架,将应用领域知识系统结构要素分为需求本体、概念本体、方法本体、控制知识和解释知识。文献[22]还将概念本体和方法本体融入多主体建模中,探讨了知识本体维护和管理机理及系统结构,为知识系统建模集成探索了一条有效的途径,也为基于知识系统的多传感器数据融合目标识别算法研究提供了清晰有效的思路。

#### 4 结束语

随着传感器技术、数据处理技术、计算机软硬件技术及人工智能技术等相关技术的发展,多传感器数据融合目标识别必将成为未来目标识别研究的重点和热

点。本文在介绍了多源数据融合目标识别原理及其算法理论基础的前提下,从算法概念分类方面,对多传感器数据融合的目标识别算法进行了较为详尽的总结,为下一步多传感器数据融合目标识别研究提供了一定的理论依据。■

#### 参考文献:

- [1] 潘泉,王增福,梁彦,等. 信息融合理论的基本方法与进展[J]. 控制理论与应用, 2012,29(10):1234—1244.
- [2] Hall DL, Llinas J. An introduction to multisensor data fusion[J]. Proc. of the IEEE, 1997,85(1):6—23.
- [3] 康耀红. 数据融合理论与应用[M]. 西安:西安电子科技大学出版社, 1997.
- [4] 张新曼,韩九强.基于矩阵分析的一种不确定性推理的数据融合方法[J]. 西安交通大学学报, 2002, 36(8): 851—854.
- [5] 王俊林,张剑云. 采用 Bayes 多传感器数据融合方法进行目标识别[J]. 传感器技术, 2005,24(10):86—88.
- [6] Dempster AP. Upper and lower probabilities induced by a multivalued mapping[J]. Annals of Mathematical Statistics, 1967,38(2):325—339.
- [7] 付耀文. 雷达目标融合识别研究[D]. 长沙:国防科技大学, 2003.
- [8] Barbera HM, Skarmeta AG, Izquierdo MZ. Neural networks for sonar and infrared sensors fusion[C]// Proc. of the third international confer. on information fusion, 2002:3—4.
- [9] Rizvi SA, Nasrabadi NM. Fusion techniques for automatic target recognition [C] // Applied Imagery Pattern Recognition Workshop, 2003:201—207.
- [10] 徐小琴. 多传感器数据融合目标识别算法综述[J]. 红外与激光工程, 2006, 35:321—328.
- [11] 杨应雷,周金和,王川潮. 加权决策模板业务感知算法[J]. 计算机工程与应用, 2015,53(2):118—123.
- [12] 张翼,朱玉鹏. 决策模板法在决策层融合目标识别中的应用和改进[J]. 电光与控制, 2005, 12(6):16—19.
- [13] 帅军. 基于熵和 KNN 的决策模板法在目标识别中的应用[J]. 湖南文理学院学报, 2005,17(1):18—21.
- [14] Kewley DJ. Notes on the use of D-S and fuzzy reasoning to fuse identity attribute data[M]. Salisbury, S.A.; Surveillance Research Laboratory, 1992.
- [15] Roux L, Desachy J. Information fusion for supervised classification in a satellite image[C]// International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium, 1995:1119—1125.
- [16] 杨建勋,史朝辉. 基于模糊综合函数的目标识别融合算法研究[J]. 雷达技术, 2004(33): 11—13.

- [17] 刘永祥, 黎湘, 庄钊文. 基于 Choquet 模糊积分的决策层信息融合目标识别[J]. 电子与信息学报, 2003, 25(5): 695—699.
- [18] 刘源, 谢维信. 多传感器图像模糊融合算法在图像识别中的应用[J]. 西安电子科技大学学报, 2000, 27(1): 5—8.
- [19] 李君. 智能学习的目标识别算法研究[J]. 红外月刊, 2003, 2(2): 11—32.
- [20] 牛丽红, 倪国强. 多传感器目标识别系统的特征优化方法[J]. 光学技术, 2005, 31(3): 420—426.
- [21] 张文修, 徐宗本. 知识系统与认知分析[J]. 系统工程理论与实践, 2002, 10(10): 118—122.
- [22] 朱欣娟, 兰壮丽, 刘风华. 知识系统建模框架研究[J]. 西安工程科技学院学报, 2004, 18(1): 48—54.

## 征稿简则

1.《航天电子对抗》由中国航天科工集团有限公司主管、中国航天科工集团有限公司 8511 研究所主办, 主要报道与导弹武器系统、空间系统相关的雷达、制导、通讯、导航、C4ISR 等电子设备和系统总体的电子、光电及信息攻防对抗技术。来稿应切合本刊报道内容, 具备科学性、创新性和逻辑性, 要求论点明确、结构严谨、论据可靠、文字和算式推导简练无误, 有较高的理论水平和实用价值。稿件篇幅以 6000 字以内(含摘要、图表、公式及参考文献)为宜, 最多不超过 8000 字。

2. 来稿应包括: 题名、作者姓名、单位、摘要和关键词(3~8 个), 并提供上述内容的英译文; 文章的中图分类号、文献标识码; 引言、正文、结论或结束语、参考文献和作者简介(包括通讯地址和联系方式); 如果投稿文章为国家自然科学基金等基金项目资助课题或为省、部委重点课题, 获奖课题等请予以注明(提供正式全名和编号)。具体为:

1) 题名是能反映文章中心内容的恰当的、简明的词语的逻辑组合, 一般不超过 20 个汉字, 应避免使用不常见的外文缩写。英文题名的第 1 个词的首字母大写, 其余均小写。

2) 作者不超过 5 人, 署名与排序应协商一致, 如有变动请提前通知。

例: 方萍, 文绍川

Fang Ping, Wen Shaochuan

3) 作者单位用中文全称标注, 其英译文应统一、规范, 实词首字母均大写。

例: (中国航天科工集团 8511 研究所, 江苏 南京 210007)

(No.8511 Research Institute of CASIC, Nanjing 210007, Jiangsu, China)

4) 摘要应完整、高度地概括出文章的目的、方法、结果及结论; 采用第三人称, 句型要力求简单, 慎用长句, 字数以 200 字左右为宜; 应排除常识性内容, 避免与题名简单重复; 其英译文应完整、准确, 尽量使用被动式。

5) 作者简介的著录规范为: 姓名(出生年—), 性别, 职称, 学历, 研究方向, 其它。并注明可靠的通讯地址和联系方式(电话、手机等), 如有变更, 请尽快通知编辑部。

6) 来稿中图表应精选, 切勿过大(一般不超过 8cm 宽)。图表中的字符和数据应准确无误, 且与正文一致, 图表序号、图表题、图表正文应以中文标注, 曲线图请注明纵横坐标的变量名、单位和刻度值。

7) 应使用法定计量单位, 一般不使用“公斤”、“英里”等单位, 单位符号应为正体; 量符号必须使用斜体字母, 矢量、张量和矩阵符号为黑斜体。编写公式应尽量使用公式编辑器。

8) 参考文献应是公开发表过的文献, 发表年限应尽量在最近 10 年之内, 按引用的先后顺序排序, 作者姓名一律姓前名后。

3. 请使用电子邮箱投稿(编辑部 E-mail: HTDZ8511@126.com), 稿件须为 Word 格式, 主题请写为“新投稿”, 修改稿主题请写为“修回(文章编号)”。

4. 本刊收稿后 2 个月左右通知作者是否刊登, 在此期间切勿一稿多投。已通知录用的文章必须在要求的时限内修回, 如不能按时修回请提前向编辑部予以说明, 否则视为作者撤稿。请自留底稿, 来稿一律不返还。

5. 来稿必须经过保密审查, 并附有单位盖章的保密审查单, 文责自负。本刊对录用稿有修改权, 所发表文章版权归编辑部。本刊已被封面所示的数据库全文收录, 作者的著作权使用费和稿酬一次性付给, 如作者不同意文章被收录, 请在来稿时申明, 本刊将做适当处理。