

# 人机共融智能

於志文 郭 斌  
西北工业大学

关键词：人机共融智能 群物交互特征 群体行为特征

## 人机共融智能的前身

一方面，随着普适计算和物联网技术的发展，大量智能物体 (smart objects) 不断涌现，另一方面，人工智能技术，特别是机器人技术的发展，又促使大量不同用途的机器人出现。人类将与大量智能物体、机器人在一起生活。

从计算系统的角度看，传统的计算系统中将不断融入人类和社会特征，形成由人类、机器和智能物体共同组成的感知计算系统——人机物融合系统。

人机物融合系统面临很多方面的挑战，其中人类和机器智能的融合成为关键挑战之一。人工智能技术在近年来不断取得突破，成为国际前沿研究热点，并被提升为国家战略。2017年7月国务院发布了《新一代人工智能发展规划》，这为探索人机物融合环境下人类和机器智能的深度融合创造了前所未有的机会。

自从1956年人工智能概念被提出以来，研究学者就在不断地思考人与智能化的计算设备之间的关系。同年，计算机学科的先驱与控制理论的专家们就提出了“智能增强”(intelligence amplification)<sup>[1]</sup>的概念。智能增强是指通过有效地使用信息技术来增强人的智能，其主要用于扩展人类处理信息的能力。时至今日，这一概念仍然发挥着有效的作用，计算器、个人电脑和智能手机，以及能够自动处理用户数据的计算机软件等设备都属于增强智能的范畴。此后，随着计算机和人交互融合的深入，计算机科学家利克莱德(Licklider)于1960年进一步提出了“人机共

生”(man-computer symbiosis)<sup>[2]</sup>的概念，指出人类的大脑将会与计算机等设备紧密耦合(tightly coupled)，并且这种耦合将形成一种我们前所未见的合作关系，其将以我们现在无法描述和构想的形式进行思考和分析数据。然而，受限于当时计算机硬件技术的发展，这些人机智能的融合方式都仅仅停留在论文严谨的推导和证明中，并没有具体的硬件为当时的学者提供人机智能混合的具体实现。

20世纪70年代，专家系统(expert system)在医疗经济学领域盛行，人与机器的智能混合初现端倪。然而，这类系统仅仅是人类推理知识的计算机表达形式，智能混合还显得比较浅显，即人将数据输入给计算机，计算机返回处理结果，人再通过自己的智能进行最终的决策。在这个过程中，人类始终扮演着推理的角色，而计算机担任着数据处理和分析的职责。

## 什么是人机共融智能

随着计算机软硬件技术的发展，人工智能进一步繁荣，20世纪60年代由科学家们所提出的人机融合形态成为可能，并呈现在真实世界中。然而，新的时代催生出了新的需求，我们需要重新审视人与计算机智能的关系，并构建出新型的人机智能混合形态。在已有研究基础上，本文结合人机物融合系统这一新的背景，提出“人机共融智能”概念。人机共融智能定义为：利用人类智能和机器智能的差异性和互补性，通过个体智能融合、群体

智能融合、智能共同演进等，实现人类和机器智能的共融共生，完成复杂的感知和计算任务。

人机共融智能的关键特性包括：

**个体智能融合** 机器的优势在于快速、低成本地对信息进行存储、比较、排序和检索，人脑的优势在于联想、推理、分析和归纳。针对复杂任务，巧妙利用人的识别、推理能力<sup>[3,4]</sup>，实现人机协作增强感知与计算，发挥二者的互补优势。

**群体智能融合** 除人脑智能外，人机共融智能重点强调群体智能，尤其是隐式智能，通过利用群体行为特征、结构特征及交互特征等在特征和决策层面与机器智能进行融合<sup>[5]</sup>，实现智能增强。

**智能共同演进** 未来我们不希望看到机器智能不断增长，而人类智能停滞不前乃至衰退，人机共融智能的目标是人类智能和机器智能互相适应，彼此支持，相互促进，实现智能的共同演进和优化。

## 与目前一些术语的区别与联系

人类和机器智能融合成为近年来学术界和企业界关注的热点，目前已经存在一些相关术语，如人机混合智能、群体智能、群智感知计算等，但人机共融智能与这些术语既有区别，又有联系。

**人机混合智能**<sup>[6]</sup> 强调人类智能和机器智能的互补，实现复杂问题和场景下的协作决策。包括人在回路的系统决策，以及面向机器智能不足的人类协作两个层面。其强调的是显式人类智能参与，是面向特定问题的，而人机共融智能同时关注显式智能和隐式智能，更大的不同是智能的共同演进。

**群体智能**<sup>[7]</sup> 是一种通过群体智慧来解决问题的方式，强调人类群体间协同互补以及机器在辅助人协作方面的作用。我们强调人类智能和机器智能的深度融合，群体智能重点突出群体隐式智能的作用。

**群智感知计算**<sup>[8]</sup> 以大量普通用户使用的移动设备作为基本感知单元，通过物联网、移动互联网协作，实现感知任务分发与数据收集，进而完成大规模、复杂的社会与城市感知任务，侧重于感知过程中人的参与，人机共融智能扩展为感知计算过

程，并且强调智能的共同演进。

## 研究挑战

人机共融智能面临的研究挑战有：精准感知、深度画像、临境自然交互、人机协同感知、人机融合计算和人机智能演进。

**精准感知** 指基于传统接触式感知技术与新兴的非接触式感知技术对人的行为及周围环境的高精度感知。例如人及环境实物的高精度(厘米、毫米级)的位置感知，高精度动作行为(手势、指形、唇形、力度等)感知，生理信号(呼吸、心跳、脉搏、血压等)感知等。机器精准感知人和环境是人机共融智能的基础，利用先进感知技术让机器了解周围的物理环境和环境中的人，为后续高级智能提供具有语义的数据输入。

**深度画像** 人机共融需要机器对人有全面深入的了解，人物深度画像指利用高精度感知技术和社交网络对在物理世界中的行为和网络空间中的交互行为进行融合分析，从而推断出人物的特征和生活规律，如健康状况、性格、兴趣、爱好、情绪等。主要的研究挑战是通过原始感知数据，结合社交网络数据，提取用户高级语义。人类行为的多面性、善变性、演化性是深度画像的难点。

**临境自然交互** 人机共融需要让人类通过更加自由/自然的方式与机器进行交互。临境自然交互从两个侧面描述了人机交互的形式：(1)临境，即身临其境，人机的边界将变得模糊。用户意识不到计算设备的存在，需要交互的信息将被突出展示，用户将“浸入”到需要的信息中，并与之进行交互。(2)自然，即交互的方式是自由的，不会受到硬件设备的限制。人类将通过自然的行为方式与计算机进行交互，例如手势、语音和触摸等。

**人机协同感知** 协同感知作为信息采集的一种方式，其内涵是汇集来自不同设备、不同数据源以及不同感知实体的信息，并通过处理这些多元的数据信息来更加全面地感知物理世界，为人类提供精准和智能的服务。人机协同感知将人类作为感知

节点,通过融合人类的智能,来提升传统基于机器设备的感知能力,实现人机的优势互补,从而提高感知的效能。人机协同感知面临两个技术难点:(1)如何实现跨空间的数据协同感知;(2)如何完成人与机器感知能力的互补增强。

**人机融合计算** 人机融合计算是指人与机器通过显式或隐式的融合范式,达到人机智能的协作与增强。显式人机融合计算中,人按照任务要求有意识地参与,将识别、联想、推理能力融入计算任务中。隐式人机融合计算仅靠行为习惯无意识参与,将人群无意识表现出的行为规律作为智能用于求解问题。人机融合计算的挑战在于如何对计算任务进行分割,然后给人和机器分配各自擅长的子任务,并确定执行子任务的顺序(串行或并行),以及如何对计算结果进行融合。

**人机智能演进** 人机智能演进的目标是发挥人和机器的优势互补,促进人的智能和机器智能的共同进步。从机器的角度,以人的知识作为输入指导机器,使得其自身的智能通过不断迭代,变得更加智能和高效,交互式遗传算法正是这一思想的典型体现<sup>[9]</sup>。机器自身亦可以利用机器之间的相互协作,借助机器提供的反馈,通过博弈的方式,强化机器的智能,从而实现机器智能的自我演进。应用强化学习的AlphaGo便是通过机器之间的相互对弈,完成了自身对弈能力的不断更新换代。反过来从人的角度,随着机器智能的提升,人也可以通过机器的反馈而受到启发,从而丰富自身的经验和知识,提高认知能力。例如,人们可以从AlphaGo、微软小冰等学习到从未想到的棋招和诗句。随着类脑计算和芯片技术的发展,在不久的将来,便会出现可植入式人工智能芯片,实现深度人机智能演进,满足许

多特定需求,如大脑先天发育不良、神经系统受损的病人,有特殊作战需求的军人等。人机智能演进需要研究如何实现人类智能与机器智能的共同学习,以及如何实现具备人机相互协作与促进特征的人机智能共同演进方法。

## 初步实践

我们近期完成的两个研究工作就是应用了人机共融智能的思路,主要是融合隐式群体智能与机器智能,完成复杂感知认知任务。

**利用群物交互特征对拍照图像进行分类和语义标注**<sup>[10]</sup> 海报是一种很好的宣传形式。电影海报、招聘广告、学术报告等海报形式,为人们出行、工作、学习、购物等提供了便利。然而,物理世界的海报往往又受到覆盖范围、传播速度等影响,为信息共享和传播带来了局限。基于此,我们设计和实现了FlierMeet系统(如图1所示),对分布在城市社区的公共信息实现了跨空间转发、分类和共享。FlierMeet通过群体参与实现用户感兴趣海报信息的采集和在线共享,通过文本和群物交互(crowd-object interaction)特征对海报进行分类(学术海报、广告等)和语义(热度、专业度、群组兴趣度)标注。首先,通过OCR识别技术提取图像中的文本信息,并进行海报分类;然后,从多维群体-感知对象交互数据中提取群物交互特征(群体智能),包括人-感知

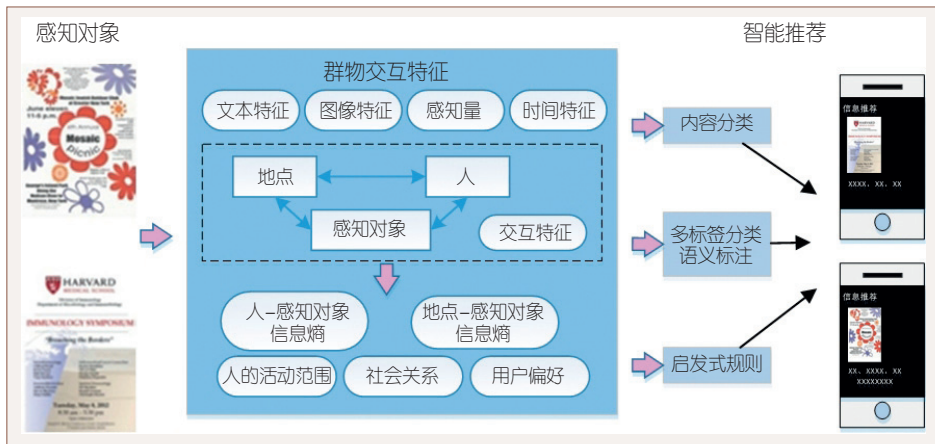


图1 基于群物交互特征的图像语义理解

对象-地点间交互信息熵、用户偏好、社会关系等，提出基于多标签分类与启发式规则融合的方法实现图像的语义标注。以往在处理图像类感知数据时，多采用基于内容挖掘的方法；而 FlierMeet 则通过多侧面的群物交互特征来对图像进行语义标注，利用隐式群体智能实现了高效图像语义理解。

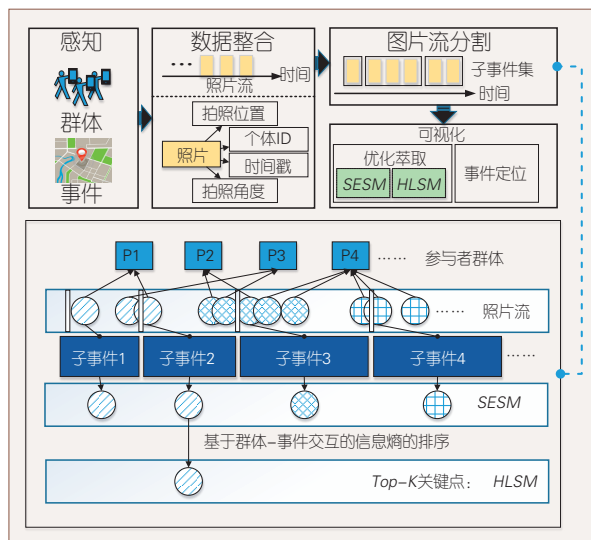


图2 基于群体行为规律的物理事件感知与多粒度呈现

基于群体行为特征的感知数据萃取<sup>[11]</sup>为及时感知城市热点事件（热门活动、公共安全事件）的发生和发展，我们研究并提出了视觉群智事件感知系统 InstantSense（如图2所示）。该系统汇聚机会社群中不同成员提供的数据，利用个体智能和群体智能萃取出高效用数据，实现了对事件发展过程的智能呈现。首先，因为同一区域可能有多个事件发生，先根据群体感知节点感知事件发生时的位置和角度对事件的发生地点进行定位。将拍摄照片时可能覆盖到的地理区域划分为网格，使用正态分布（假设事件区域通常占据照片的中间部位）计算每个网格的定位权重，通过融合群体智能（群体贡献照片的元数据信息，如拍摄位置、方向等）计算网格的累计权重，权重越大则越有可能是事件发生的地点。其次，通过个体/群体智能实现对事件的分割，即将照片流根据事件发展划分为不同子事件。个体智能体现在个体对拍照时机的选择，一般

在事件发生变化时进行拍摄；群体智能则通过群体拍摄行为体现事件的重要时刻和不同侧面的感知。第三，InstantSense不但得到低冗余、高覆盖的事件全面感知结果，还会根据子事件对目击者的吸引力进行排序，选择 Top-K 照片构成事件感知的关键点。InstantSense 通过个体/群体照片拍摄行为规律实现对事件的语义分割，进而从两个粒度对事件进行呈现。一是针对多维感知信息（远近、角度等）对子事件进行全方位呈现，二是根据群体拍摄信息熵进行关键子事件的发现。

## 结语

人机物融合正在成为新一代计算技术的重要特征与主要趋势，人类和机器智能的融合是其关键。人机智能融合不仅仅是智能在物理上一时的混合，而是要让这种融合产生化学反应，迸发出更多更强的智慧，通过不同粒度智能的深度交融、共同演进等，实现人类和机器智能的共融共生。研究人员已经做了一些初步尝试，但距离真正的人机共融智能，还有很长一段路要走，很多理论和技术挑战需要去克服。人机共融智能，任重而道远！



### 於志文

CCF 杰出会员、理事、CCF 普适计算专委会副主任，CCCF 前编委。西北工业大学教授。主要研究方向为普适计算、移动感知等。  
zhiwenyu@nwpu.edu.cn



### 郭斌

CCF 高级会员。西北工业大学教授。主要研究方向为普适计算和群智感知计算。  
guob@nwpu.edu.cn

## 参考文献

- [1] Ashby W R. An Introduction to Cybernetics[J]. Chapman & Hall Ltd, 1961.
- [2] Licklider J C R. Man-Computer Symbiosis[J]. IRE

- Transactions on Human Factors in Electronics, 1960 (1): 4-11.
- [3] Bragg J., Weld D. S., et al. Crowdsourcing multi-label classification for taxonomy creation[C]// Proceedings of First AAAI conference on Human Computation and Crowdsourcing, 2013.
- [4] Franklin M. J., Kossmann D., Kraska T., et al. Crowddb: answering queries with crowdsourcing[C]//Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data (SIGMOD11), 2011: 61-72.
- [5] Guo B., Wang Z., Yu Z., et al. Mobile Crowd Sensing and Computing: The Review of an Emerging Human-Powered Sensing Paradigm[J]. ACM Computing Surveys, 2015, 48(1):7.
- [6] Zheng N., Liu Z., Ren P., et al. Hybrid-augmented intelligence: collaboration and cognition[J]. Frontiers of Information Technology & Electronic Engineering, 2017, 18(2):153-179.
- [7] Li W., Wu W., Wang H., et al. Crowd Intelligence in AI 2.0 Era[J]. Frontiers of Information Technology & Electronic Engineering, 2017, 18(2):15-43.
- [8] 刘云浩. 群智感知计算 [J]. 中国计算机学会通讯, 2012; 8(10): 38-41.
- [9] Michelucci P., and Dickinson J. L., The power of crowds[J] Science 2015;351 (6268):32-33.
- [10] Guo B., Chen H., Yu Z., et al. FlierMeet: A Mobile Crowdsensing System for Cross-Space Public Information Reposting, Tagging, and Sharing[J]. IEEE Transactions on Mobile Computing, 2015;14(10) 2020-2033.
- [11] Chen H., Guo B., Yu Z., Han Q.. Toward Real-time and Cooperative Mobile Visual Sensing and Sharing[C]// Proceedings of the 35th IEEE International Conference on Computer Communications (INFOCOM 2016), 2016: 10-15.