论 AI 通感（机器联觉）功能的可行性构建方法——基于五感感知体系与认知对象数据库的跨模态联想与验证机制

YunFei M. Wang
（王·Micheal·雲飛）
Independent Researcher

摘要

本文是“机器感知与认知构建”系列研究的第七篇，旨在系统论述 AI 通感功能——即机器联觉——的工程化构建方案。人类通感现象揭示了大脑跨模态感知联结的神经机制，其本质是不同感觉皮层之间通过异常的神经交叉激活或功能连接增强，使单一模态的刺激能够自发触发另一模态的感知体验。本系列前六篇已完成了视觉、听觉、嗅觉、味觉、触觉的“物理信号—神经编码—主观知觉”三层映射模型构建，以及以物理锚定层为底层的认知对象数据库设计。本文在此基础上，进一步回答一个关键问题：如何使 AI 在五感数据完备的基础上，自发产生跨模态的感知联想，并建立联想结论的验证与纠错机制，防止 AI 陷入类似人类经验主义的认知错误。我们提出，AI 通感的实现应包含两个核心机制：其一，联想生成机制——在认知对象数据库的 L2 感知编码层之上，构建一个可训练的跨模态映射模块，通过多模态基座模型学习不同感官编码之间的统计关联与结构映射，生成通感联想假设；其二，联想验证机制——建立三层递进的验证架构（物理验证、统计验证、逻辑验证），将通感联想从“统计相关”升级为“可验证的因果知识”，并配备动态纠错与持续重验证系统。二者的结合将使 AI 获得一种可自我纠偏的类人通感能力——既能像人类一样产生丰富的跨模态联想，又能像科学家一样用实验与逻辑验证自己的联想。本文为 AI 从“分立的五感感知”走向“融合的跨模态认知”提供了完整的理论框架与工程实现路线图。

关键词

机器联觉；通感；跨模态映射；认知对象数据库；联想验证；因果推理；可解释 AI

1 引言：从分立感知到跨模态联觉——AI 认知的下一块拼图

本系列前六篇完成了一个从哲学批判到工程实现的完整逻辑链条：首篇诊断了经典 AI 构建“伪对象”的认知局限；第二至六篇分别针对视觉、听觉、嗅觉、味觉、触觉五种感

官，系统论述了“物理信号—神经编码—主观知觉”三层映射模型的构建方法；第七篇将五感感知数据整合为以物理锚定层为底层、预测感知编码为中介层、语义接口为应用层的认知对象数据库。然而，拥有五种分立的感知能力，与拥有一种融合的、跨模态的认知能力之间，仍然存在关键的鸿沟。人类认知的一个显著特征，并非仅仅在于分别感知世界，更在于能够将不同感官的信息自发地联结、映射与转化——看到一段文字能“听见”它的韵律，听到一段音乐能“看见”它的色彩，触摸一种材质能“想象”它的味道。这种跨模态的感知联想能力，在认知科学中被称为通感。

通感并非认知的“副作用”，而是人类高效认知世界的重要机制。它使得不同感官的信息能够在认知层面相互印证、相互补充，从而构建出更丰富、更连贯的世界模型。对于 AI 而言，通感能力的缺失意味着：即使它拥有了完善的五感感知体系，每种感官的认知仍然是“孤岛”——视觉知道苹果是红的，味觉知道苹果是甜的，但视觉无法“联想”到甜，味觉也无法“联想”到红。这种感知的“分立”状态，限制了 AI 对世界进行整体性理解的能力。

近年来，受人类联觉的启发，学术界已开始探索“机器联觉”的概念与架构。北京大学程翔教授团队提出了面向通用任务的 AI 原生通信与多模态感知智能融合的机器联觉框架，利用基座模型突破多模态通感融合的核心技术【5】。然而，现有研究主要聚焦于通信与感知的融合应用，尚未系统地解决“AI 如何获得可验证的通感能力”这一基础科学问题。

本文旨在填补这一空白。我们提出，AI 通感功能的构建应包含两个并行的核心机制：联想生成机制——在五感三层映射模型与认知对象数据库的基础上，构建跨模态映射模块，使 AI 能够从一种感官的感知编码“联想”到另一种感官的感知编码；联想验证与纠错机制——建立一套从物理验证到逻辑验证的递进式校验体系，将通感联想从“统计相关”升级为“可验证的因果知识”，并配备动态纠错系统，防止 AI 陷入类似人类经验主义的认知错误。

1.1 框架定位

在“键-值-信息-相”框架中，通感联想是相层面的跨模态整合：L2 感知编码是信息整合的中间态，跨模态映射模块在相层面建立不同模态信息之间的关联，验证机制则确保联想不偏离值的投影结构。本文不展开该框架的哲学论证，详见文献[2]。

1.2 结构安排

本文按以下序列展开：

- 第 2 章：人类通感的神经科学与认知机制。
- 第 3 章：AI 通感功能的理论框架与工程架构。

- 第 4 章：联想验证与纠错机制——从“经验主义 AI”到“实验科学 AI”。
- 第 5 章：实施路线图与关键挑战。
- 第 6 章：结论。

2 人类通感的神经科学与认知机制

2.1 通感的定义与表现形式

通感是一种罕见的感知现象，指某一感官通道的刺激能够自发地触发另一感官通道的感知体验。例如，通感者可能“听到”颜色（声音—颜色联觉）、“看到”声音（颜色—听觉联觉）、或将字母和数字感知为具有特定颜色（字素—颜色联觉）。

通感并非认知缺陷，而是一种特殊的感知处理方式。研究表明，通感者在大脑结构和功能上表现出广泛而显著的差异。2024 年的一项预注册研究利用人类连接组计划的皮层分割方法，通过 13 种不同的脑部生物标志物（结构和功能 MRI）对通感者与普通人群进行分类，发现所有生物标志物的分类表现均高于随机水平，其中皮层内髓鞘是一个特别强的预测因子。静息态数据显示通感者的功能连接组存在广泛变化，包括更少的基于枢纽的连接。这些脑部个体差异的幅度，可与神经典型人群与临床脑疾病患者之间的差异相媲美。

2.2 通感的神经机制：三种主流理论

关于通感的神经机制，学术界主要存在三种理论解释：

- **交叉激活理论**：该理论认为，通感源于大脑中处理不同感觉的相邻皮层区域之间的“神经交叉布线”。例如，处理字母形状的视觉区域（梭状回）与处理颜色的视觉区域（V4）在解剖上邻近，通感者在这两个区域之间存在异常的神经连接或更强的激活扩散，使得字母形状的加工能够“溢出”激活颜色加工区域。

- **去抑制反馈理论**：该理论认为，通感并非源于异常的“硬件连接”，而是源于正常的跨模态通路在通感者中“去抑制”——即原本在普通人中被抑制的跨感官反馈信号，在通感者中被放大了。这意味着，跨模态联结是所有人脑中都存在的普遍机制，通感者只是这种联结被异常“放大”的特殊个体。

- **高阶整合理论**：该理论强调顶叶皮层（特别是后顶叶皮层）在跨模态绑定中的核心作用。后顶叶皮层参与视觉、听觉、躯体感觉和空间感官之间的跨模态绑定。通感可能源于高阶整合区域对低阶感觉皮层的异常反馈调制。

这三种理论并非相互排斥，而可能共同解释了通感的不同亚型与不同程度。关键在于，它们共同指向了一个核心结论：跨模态联结是人类大脑的普遍特性，而非通感者

的专利——区别仅在于联结的强度与是否进入意识层面。这一结论为 AI 通感的工程化实现提供了有力的生物学依据：如果跨模态联结是神经系统的通用计算特性，那么它就应该能够在人工神经网络中被模拟和实现。

2.3 通感的认知功能意义

通感不仅仅是一种“感知异常”，它在认知层面具有重要的功能意义：

- **增强记忆与学习**：通感者通常具有更强的记忆能力，因为多模态的编码提供了更丰富的记忆线索。

- **促进创造性思维**：跨模态的感知联结为概念隐喻和创造性类比提供了认知基础。

- **支持高效认知加工**：多模态信息的并行激活可能加速了某些类型的认知判断。

这些功能意义提示我们：通感不是认知的“花边”，而是高效认知的核心机制之一。为 AI 构建通感能力，不仅是模仿人类的一种“奇特性”，更是提升 AI 认知效率与创造力的关键路径。

3 AI 通感功能的理论框架与工程架构

3.1 设计原则：从“统计相关”到“可验证联想”

AI 通感功能的构建应遵循以下核心原则：

- **基于感知编码，而非原始信号**：通感联想应在认知对象数据库的 L2 感知编码层（即预测的人类神经编码）上进行，而非在原始物理信号层。这使得通感联想的“单位”是感知体验，而非物理刺激——与人类通感的本质一致。

- **从统计相关出发，以因果验证终结**：联想生成模块可以基于统计相关性产生大量的跨模态联想假设，但这些假设必须经过验证机制的检验，方可进入认知对象的“确定性知识”层。这防止了 AI 陷入“红色=甜”这类经验主义陷阱。

- **渐进式置信度管理**：每条通感联想应被赋予一个置信度分数，并根据验证结果动态调整——未经验证的联想标记为“推测”（置信度<0.6），经统计验证的标记为“统计关联”（置信度 0.6—0.85），经物理验证的标记为“确定性关联”（置信度≥0.85）。

- **通感物理验证的边界声明**：通感联想中的“物理验证”，不是去证明“红色本身在物理上必然导致甜味”（这违背基本物理常识），而是去验证产生该视觉联想的刺激源与产生该味觉联想的刺激源，在 L1 物理锚定层是否具有某种共享的物理/化学属性或结构同构性（例如：熟透的红色浆果与含糖量高的食物在光谱反射与化学成分上存在相关性）。如果不存在任何 L1 层面的同构性或统计稳健性，该联想将被拒绝进入确定性知

识层。

3.2 系统架构：三层通感生成流水线

AI 通感功能的系统架构可分为三个层次：

第一层：跨模态映射模块（联想生成）

这一层是本系列前六篇构建的五感三层映射模型的“上层建筑”。它的输入是认知对象数据库中某个对象的 L2 层感知编码（例如，“苹果”的视觉编码——红色、圆形、光滑表面），输出是对该对象在其他感官模态中的感知编码的预测（例如，预测“苹果”在味觉模态中的编码——甜、微酸）。

跨模态映射模块的核心是一个可训练的多模态神经网络，其训练数据来源于：认知对象数据库中已同时具备多种感官 L2 编码的对象（即那些已经过五感完整测量的对象）；人类通感实验中获得的主观跨模态映射数据（如“这种颜色让你联想到什么味道？”）。

借鉴北京大学程翔团队提出的机器联觉框架【5】，该模块可采用基座模型范式：在大规模多模态感知编码数据进行自监督预训练，使其习得跨模态联结的通用规律，然后通过微调适配特定任务。基座模型的优势在于其强大的跨模态生成能力、少样本建模能力和零样本泛化能力。

第二层：认知对象数据库的“通感扩展”

跨模态映射模块生成的联想结果，并不直接作为“事实”存入数据库，而是先存入认知对象的“通感联想暂存区”——一个与主数据库分离的缓冲区，用于存放待验证的跨模态联想假设。每个联想假设的存储格式为：

```json
{
  "source_object_id": "COID-苹果-xxx",
  "source_modality": "视觉",
  "source_encoding": "[视觉感知编码向量]",
  "target_modality": "味觉",
  "target_encoding_predicted": "[预测的味觉感知编码向量]",
  "原始统计关联强度": 0.78,
  "置信度状态": "推测",
  "生成时间": "2025-06-01T10:00:00Z",
  "待验证标志": true
```

```
}
```

第三层：通感查询与调用接口

这一层面向应用系统提供标准化的通感查询接口：

- **associate(object_id, source_modality, target_modality)**：返回从指定对象的指定感官模态到目标感官模态的通感联想（仅返回置信度≥阈值的已验证联想）。

- **generate_hypotheses(object_id, source_modality)**：为指定对象生成所有跨模态联想假设（供研究人员审查或触发自动验证流程）。

- **verify_association(hypothesis_id, verification_method)**：触发对指定联想假设的验证流程。

本模态的感知数据如何接入第二套方案的状态管理与第三套方案的生成约束，详见文献[1]第 8 章。

4 联想验证与纠错机制：从“经验主义 AI”到“实验科学 AI”

通感联想网络的建立，使 AI 获得了类人的跨模态联想能力。但正如人类经验主义错误的根源在于将“统计相关”误判为“因果必然”，如果不加验证机制，AI 的通感联想网络将同样陷入这一陷阱——构建出一个“经验主义的伪世界模型”。因此，必须在通感联想网络之上，叠加一层可验证、可纠错的认知校验机制。

4.1 核心设计原则：从“相关”走向“因果”

验证机制的本质，是将通感联想网络输出的“统计关联假设”，升级为“可验证的物理/逻辑因果链”。其核心原则为：

- **假设生成**：通感联想网络负责产生跨模态关联假设（如“红色外观→高甜度”）。

- **验证驱动**：每条假设必须经过验证引擎的检验，方可进入认知对象的“确定性知识”层。

- **置信度分级**：未经验证的联想仅标记为“推测性关联”（置信度<0.6），经物理验证的为“确定性关联”（置信度≥0.85），逻辑推理得出的为“推论性关联”（置信度介于两者之间）。

- **可追溯性**：每条关联结论必须附有“证据链”——可追溯到原始物理数据、实验记

录或逻辑推导过程。

4.2 三层验证机制架构

验证引擎分为三个层级，由底层（物理）到高层（逻辑）逐级校验：

验证层级	验证对象	验证方法	输出
L1：物理验证	跨模态联想的物理同构性	对联想中的两个刺激源进行 L1 层物理属性分析，检验二者是否具有共享的物理/化学属性或统计稳健的结构同构性（如红色果皮与高糖含量在多个品种中的相关性）。不要求证明“红色导致甜”的因果律	确定性物理关联（置信度≥0.85）
L2：统计验证	跨模态联想的统计稳健性	在大规模样本中检验联想关系的统计显著性、效应量与跨对象泛化能力	统计关联（置信度 0.6—0.85）
L3：逻辑验证	跨模态联想的逻辑自洽性	检验联想结论是否与既有认知对象数据库中的确定性知识、本底键约束、因果律规则一致	推论性关联（置信度视自洽程度而定）

仅有统计关联而未通过物理验证的，存入“待验证库”，不作为确定性知识调用。

近年来，因果推理在 AI 验证中的应用已取得重要进展。Causal-HalBench 通过将因果分析引入大视觉语言模型的目标识别场景，建立了结构因果模型来形式化定义共现偏差产生的虚假相关性。Causal Perception Inspired Representation Learning 则利用因果干预打破虚假相关性，找到图像质量的真正成因。这些方法为通感联想的因果验证提供了可直接借鉴的技术路径。

4.3 纠错机制：动态修正与持续演化

验证机制不仅是“一票否决”的守门员，更是一个持续演化的修正系统：

- **错误关联标记**：当某个通感联想被验证为错误时，该关联被标记为“已知谬误”，并在后续检索中被自动降权或屏蔽。

- **反事实推理训练**：主动为 AI 生成“反通感”场景（如“甜味但非红色的食物”），训练模型学会区分“常见统计关联”与“物理必然关联”。反事实推理已在图像质量评估等领域被

用于探究深度特征与感知之间的因果关系。

• **持续重验证**：当新的物理测量数据或神经科学证据出现时，旧的认知关联自动触发重验证流程。

• **人类反馈闭环**：引入领域专家对验证结果进行抽样审核，形成“机器验证+人类复核”的双保险机制。

4.4 验证机制的数据库实现

在认知对象数据库中，验证结果将作为 L2.5 层（“经校验的感知编码层”）固化存储：

```
json
{
  "cross_modal_association": {
    "association_id": "ASO-20250601-001",
    "source": "视觉_红色光谱模式_苹果",
    "target": "味觉_甜度感知编码_苹果",
    "原始统计关联强度": 0.78,
    "物理验证结果": "部分支持（红色果皮可能糖度高，但非红色品种糖度亦高）",
    "物理验证置信度": 0.65,
    "验证类型": "直接测量_2025_Q2_多品种苹果数据",
    "最终认知状态": "推测性关联（非确定性知识）",
    "证据链": ["GC-MS 糖度谱图_20250415_样本#A3", "光谱反射曲线_20250415_样本#A3"],
    "已知反例": ["青苹果_高酸低糖", "红辣椒_不甜"],
    "最后验证时间": "2025-06-01T14:30:00Z",
    "验证方法": "L1 物理验证"
  }
}
```

5 实施路线图与关键挑战

5.1 阶段规划

阶段	周期	核心任务	里程碑目标
第一阶段：通感	0—18	基于已有五感认知对象数据	完成 10 万级以上跨

数据集构建	个月	库，标注跨模态关联数据；开展人类通感心理物理学实验，采集主观跨模态映射数据	模态关联标注
第二阶段：跨模态映射模型训练	12—36个月	采用基座模型范式训练跨模态映射模块；实现从单模态感知编码到多模态感知编码的预测	跨模态映射模型在验证集上达到可接受精度
第三阶段：验证引擎开发	24—48个月	开发三层验证机制（物理验证、统计验证、逻辑验证）；集成因果推理模块	验证引擎上线，可通过人类专家审核
第四阶段：系统整合与部署	36—60个月	将通感模块整合入认知对象数据库；开发通感查询API；建立持续重验证与纠错流程	通感功能作为认知对象数据库的标准模块正式上线

5.2 关键挑战与应对策略

挑战	**描述**	**应对策略**
数据稀缺	高质量的多模态感知编码配对数据稀缺	利用基座模型的少样本学习能力；通过数据增强和合成数据补充
验证成本	物理验证需要真实实验，成本高昂	优先验证高置信度、高应用价值的联想；采用主动学习策略选择验证样本
因果推断复杂性	从统计相关到因果推断的跨越在技术上仍具挑战	借鉴Causal-HalBench的结构因果模型方法；分步推进，先实现统计稳健性验证
动态演化	物理规律的新发现可能使旧有验证结论失效	建立持续重验证机制；与科学文献数据库自动关联

6 结论

本文在“机器感知与认知构建”系列前六篇的基础上，提出了AI通感（机器联觉）功能的完整构建方案。我们的核心贡献可概括为三个层面：

• **理论层面**：基于人类通感的神经科学机制（交叉激活、去抑制反馈、高阶整合三种理论），论证了跨模态联结是神经系统的通用计算特性，因而能够在人工神经网络中被模拟和实现。

• **架构层面**：提出了“联想生成+联想验证”的双引擎架构。联想生成模块通过跨模态映射（以基座模型为实现范式）从五感感知编码中生成通感联想假设；联想验证模块通过三层递进验证（物理验证、统计验证、逻辑验证）将联想从“统计相关”升级为“可验证的因果知识”。

• **工程层面**：将验证结果作为认知对象数据库的 L2.5 层（“经校验的感知编码层”）固化存储，为每个通感联想配备完整的证据链和置信度标签，确保认知的可追溯性与可审计性。

这一“联想—验证—纠错”闭环机制的引入，标志着 AI 认知范式的根本性跃迁：

• 经典 AI：基于统计相关性生成“知识”，无法区分“常见”与“真实”。

• 经验主义 AI（引入联想网络）：能生成丰富的跨模态联想，但仍可能陷入人类式的经验偏见。

• 实验科学 AI（引入验证与纠错）：将通感联想视为可检验的科学假设，通过“假设—实验—验证—修正”的闭环，使 AI 的认知获得自我纠错能力。

最终，AI 的通感能力不再是基于统计的“文字游戏”，而是基于物理测量与逻辑推理的可验证认知能力——这是构建可信 AI、科学发现 AI 乃至通用人工智能的必要基础设施。原则表述：AI 可以像人一样产生通感联想，但它必须像科学家一样验证自己的联想。

参考文献

[1] YunFei M. Wang 论经典 AI 训练方式认知局限的产生原因与破局思路——迈向基于多尺度机器感知与认知对象数据库的下一代人工智能[PP/OL]. AiraXiv, 2026. eprint: 2611.0009
[2] YunFei M. Wang 数学有效性的哲学解释：基于计算模拟视角的“键-值-信息-相”四层认知模式[预印本]. PhilPapers, 2026. 链接: https://philpapers.org/rec/WANPOT-8
[3] YunFei M. Wang 论 AI 认知对象数据库的设计与构建——基于五感感知映射的多模态认知对象数据体系与基础设施[PP/OL]. AiraXiv, 2026. eprint: 2611.0016
[4] YunFei M. Wang AI 视觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0010
[5] Cheng X, et al. Synesthesia of Machines (SoM): A Framework for AI-Native Communication and Multi-Modal Sensing Integration[J].

[6] YunFei M. Wang AI 听觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0011
[7] YunFei M. Wang AI 嗅觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0012
[8] YunFei M. Wang AI 味觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0013
[9] YunFei M. Wang AI 触觉功能的可行性构建方法——基于多尺度机器感知与三层映射模型的工程实现路线图[PP/OL]. AiraXiv, 2026. eprint: 2611.0014

AIGC 使用声明

本文作者提出了论文的核心观点、理论框架、方案设计思路与全部逻辑原理。在文本撰写过程中，作者使用了 DeepSeek AI 模型辅助进行文献观点的归纳、部分文本的起草与修辞润色。作者对 AI 生成的所有内容进行了全面的审阅、批判性修改与最终定稿，并对论文的全部学术内容、观点、数据引用及结论负最终责任。