

2025 年度拟提名陕西省自然科学奖项目公示内容

一、项目名称

多媒体数据概要表示和高效压缩的理论与方法

二、提名者及提名意见

提名者：陕西省教育厅

提名意见：

该成果面向多媒体大数据战略资源，围绕“快速概览、精确寻的、高效压缩”等关键问题开展系统研究。成果揭示了多媒体数据中关键内容的表示机理，构建了多媒体数据的特征域重构理论体系，设计了内容自适应的多媒体数据概要表示方法；发现了多媒体大数据中感兴趣目标的多维度特征变化规律，构建了感兴趣目标的特征不变性表示理论，设计了感兴趣目标的高效筛选方法；揭示了多媒体数据分层压缩框架下率失真依赖机理，创建了不同层级率失真描述理论架构，提出基于率失真优化理论的量化参数确定方法。成果选题面向国家重大需求，创新性强，成果发表多媒体、图像处理领域顶级期刊和会议，得到同行广泛引用和高度评价，产生了重要学术影响并对相关研究起到引领和示范作用，对学科建设和经济社会发展有重要的指导作用。

成果材料齐全、规范，无知识产权纠纷，人员排序无争议，符合陕西省自然科学奖提名条件。

提名该项目为陕西省自然科学奖一等奖。

三、项目简介

在国家数智经济时代，多媒体数据已不仅是信息传播的核心载体，更逐渐成为驱动经济发展的核心生产要素。然而，其天生具备的“海量性”与“高冗余性”两大痛点，直接制约了从数据产生到落地应用的全链路效率。这也决定了多媒体数据的“概要表示”与“高效压缩”并非可选项，而是支撑多媒体数据实现价值转化的基础刚需。从学术与战略维度看，多媒体数据概要表示和高效压缩既是电子与通信、计算机科学、智能信息处理等交叉学科的国际研究前沿，更是国家大数据战略的重要引领点，以及经济民生领域（如数字媒体、远程医疗）的关键支撑点。在此背景下，面对海量多媒体数据，如何实现“快速概览、精确寻的、高效压缩”，已成为破解多媒体数据监管分析、存储利用难题，推动其转化为科技生产力的核心关键问题。2010 年以来，该项目通过探索关键内容表示机理、感兴趣目标多维特征变化规律以及率失真依赖机理等科学问题，完成内容级概要表示、目标级精确筛选以及数据级高效压缩等关键成果，构建了多媒体数据概要表示和高效压缩的理论与方法体系，为多媒体大数据的智能解析与存储提供核心技术支持。

该成果的重要科学发现具体如下：

1. 揭示了多媒体数据关键内容的表示机理，构建了关键内容对多媒体数据的特征域重构理论框架，设计了内容自适应的多媒体概要表示方法，将视频概要信息获取时间缩短为视频时长的 1/395，解决了多媒体数据的“快速概览”难题。
2. 发现了多媒体数据中感兴趣目标多维度特征变化规律，构建了目标图像的特征不变性表示理论体系，设计了多媒体数据中感兴趣目标的高效筛选方法，实现了复杂视觉场景下目标的精准表征与筛选，将无人机多遥感视角下弱小目标检测精度提升 5%以上，解决了多媒体数据的“精确寻的”难题。
3. 发现了单层/可分层视频压缩框架下码率与失真的参考依赖关系，创建了不同层级的率失真描述理论架构，提出基于率失真优化理论的量化参数确定方法，将国际标准同等视频质量的码率降低 4.87%，从理论上率先填补了相关领域技术空白。

相关理论研究成果发表在多媒体、图像处理领域国际顶级期刊和会议，出版专著 2 部，形成 10 项国际/国内行业标准提案，支撑完成人作为发明人与华为、OPPO 等公司共同申请国际/国内发明专利 260 项。7 篇代表性论文专著得到美国国家工程院院士 Richard Szeliski 教授、加拿大工程院院士 Xiao-Ping Zhang 教授、英国皇家学会院士 Jon Crowcroft 教授、美国国家发明家科学院院士 Sam Kwong 教授等 10 名院士和 36 位 IEEE/ACM Fellow 等国内外著名学者的积极赞扬，评价所提出方法是“最先进的性能（state-of-the-art performance）”、“有效的（work well）”、“优于（outperform）”、“新颖的（novel）”等。

相关研究支撑培养国家级青年人才 2 人，陕西省优秀博士学位论文 1 篇，培养学生获得 IEEE ISPACS 2017 最佳论文奖、IEEE MSA-TC 2024 最佳论文荣誉提名奖、IEEE WHISPERS 2023 目标跟踪挑战赛和遥感图像语义分割挑战赛冠军以及大学生创新创业大赛陕西省赛金奖等荣誉。成果支撑了西北工业大学入选国家双一流建设序列和“信息与通信工程”第五轮学科评估等级提升，进一步巩固学科在国内同类领域的优势地位；保障了陕西省信息获取与处理重点实验室获得 4 次优秀 1 次良好的省级评估，为后续开展前沿领域研究奠定了坚实基础。研究成果已经应用于上海卫星工程研究所、中国航天科技集团有限公司第一研究院研究发展中心等科研院所及 OPPO 广东移动通信有限公司等知名企业，为微小卫星研制、高超声速平台探测系统论证、无人机远程侦查探测、新一代视频编码国际标准等提供了高效解决方案，不但提高了多型航空航天平台光电探测数据获取的有效率，而且提升了我国企业在多媒体编数据解码等领域的国际话语权与技术竞争力。

四、客观评价

针对科学发现 1：国际知名学者评价所提出方法是“国际先进的”、“性能更好的”，提取的概要“能够保留原始多媒体数据的重要内容”且“具有精炼性”。

- IEEE Fellow、希腊亚里士多德大学 Ioannis Pitas 教授指出代表性论文 1 的方法是有竞争性的、先进的，使用真正的 L_0 稀疏范数，可以使概要具有精炼性：

“two competing, state-of-the-art static video summarization algorithms: OffMSR from [32] and RPCA-KFE from [7]”、“while in [32]... Minimum Sparse Reconstruction (MSR) attempts to indirectly minimize the actual L_0 sparsity norm”、“enforce summary conciseness through a sparsity constraint, as in [32]”。

- IEEE Fellow、俄亥俄州立大学 Yuan F. Zheng 教授评价代表性论文 1 通过将视频摘要问题表述为对原始视频进行最小稀疏重建处理大型视频数据：“Mei et al. [25] dealt with large video data via video summarization, and the problem is formulated as a minimum sparse reconstruction of the original video.”。
- 上海交通大学赵俊教授（CMBBE:IV、IJBI 副主编）评价代表性论文 2 平衡了全局特征和局部细节的代表性，相较于帧级稀疏表示的方法有显著优势：“Patch-based sparse representation [33] has been proven to outperform frame-level sparse representation due to its balancing the representativeness of global features and local details.”。

针对科学发现 2：国际知名学者评价所提出的方法是“效果好的”、“优于之前工作的”，并且直接使用提出的技术或将提出的算法作为基准模型进行比较分析。

- 英国爱丁堡皇家学会院士、哥斯拉大学 Philippe G Schyns 教授在其网络中直接使用代表性论文 3 提出的损失函数：“We used SymTriplet loss [106,107], a version of the triplet loss function (“Face-Net”).”。
- 英国爱丁堡龙比亚大学 Amir Hussain 教授（Cognitive Computation 创始主编）指出代表性论文 4 通过利用全局信息和局部几何信息来减轻旋转姿态的影响：“these methods (Lee et al., 2023, Mei et al., 2023, Yu et al., 2023, Yu et al., 2020) attempt to mitigate the impact of rotational pose on point clouds by leveraging both global information and local geometric information.”。
- 中国科学院读大学叶齐祥教授（国家杰青、Elsevier 高被引学者）评价代表性论文 4 提出了一种新颖的循环极坐标卷积层，解决特征学习中的旋转不变性问题：“To handle the problem of rotation invariance for feature learning, Mei et al. [22] proposed a novel cyclic polar coordinate convolutional layer.”

针对科学发现 3：国际知名学者评价所提出方法是“动态自适应的”，且“降低率失真优化总成本/更加高效的”。相关基础理论研究支撑第三完成人万帅教授与华为、OPPO 等公司联合申请/获批国际/国内发明专利 260 项，形成国际标准提案 10 项，对于提升增强了我国企业在多媒体数据编解码等领域的国际话语权与技术竞争力，打破国外技术专利壁垒具有重要意义。

- 美国国家发明家科学院/加拿大工程院院士、IEEE Fellow、岭南大学 Sam Kwong 教授指出代表性论文 5 降低率失真优化中的总成本：“In [14], the first two frames of the sequence were encoded multiple times to obtain necessary statistics for layer-level dependency factor to reduce the total cost in RDO.”。
- IEEE Fellow、电子科技大学朱策教授多次肯定了代表性论文 5 对 HEVC 分

层编码提出了动态自适应的 QPC 方法：“[7]-[9] proposed adaptive QP cascading (QPC) methods for HEVC...are determined dynamically.”。（[8]为代表性论文 5）

五、代表性论文专著目录（不超过 8 条，其中代表性论文不超过 5 篇，代表性专著不超过 3 部）

序号	论文专著 名称	刊名	作者	年卷页 码（xx 年 xx 卷 xx 页）	发表 时间	通讯 作者	第一 作者	国内作者	他引 总次 数	检索 数据 库	知识产 权是否 归国内 所有
1	Video summarization via minimum sparse reconstruction	Pattern Recognition	Shaohui Mei, Genliang Guan, Zhiyong Wang, Shuai Wan, Mingyi He, David Dagan Feng	2015 年 48 卷 522-533 页	2014-08-15	Shao hui Mei	Shao hui Mei	梅少辉, 万帅, 何明一			是
2	Patch based video summarization with block sparse representation	IEEE Transactions on Multimedia	Shaohui Mei, Mingyang Ma, Shuai Wan, Junhui Hou, Zhiyong Wang, David Dagan Feng	2021 年 31 卷 732-747 页	2020-04-16	Shao hui Mei	Shao hui Mei, Ming yang Ma	梅少辉, 马明阳, 万帅, 侯军辉			是
3	Tracking persons-of-interest via unsupervised representation adaptation	International Journal of Computer Vision	Shun Zhang, Jia-Bin Huang, Jongwoo Lim, Yihong Gong, Jinjun Wang, Narendra Ahuja, Ming-Hsuan Yang	2020 年 128 卷 96-120 页	2019-09-03	Ming - Hsuan Yang	Shun Zhan g	张顺, 龚怡宏, 王进军			是

4	Rotation-invariant feature learning via convolutional neural network with cyclic polar coordinates convolutional layer	IEEE Transactions on Geoscience and Remote Sensing	Shaohui Mei, Ruoqiao Jiang, Mingyang Ma, Chao Song	2023 年 61 卷 1-13 页	2023-01-02	Shao hui Mei	Shao hui Mei	梅少辉, 蒋若峤, 马明阳, 宋超			是
5	Rate-distortion-optimization-based quantization parameter cascading technique for random-access configuration in H.265/HEVC	IEEE Transactions on Circuits and Systems for Video Technology	Yanchao Gong, Shuai Wan, Kaifang Yang, Yuan Yang, Bo Li	2017 年 27 卷 1304-1312 页	2016-03-08		Yan hao Gong	公衍超, 万帅, 杨楷芳, 杨原, 李波			是
6	新一代高效视频编码 H.265/HEVC: 原理、标	电子工业出版社	万帅, 杨付正	2014 年	2014-12-01		万帅	万帅, 杨付正			是
7	新一代通用视频编码 H.266/VVC: 原理、标准与实现	电子工业出版社	万帅, 霍俊彦, 马彦卓, 杨付正	2022 年	2022-08-01		万帅	万帅, 霍俊彦, 马彦卓, 杨付正			是
合 计											是

六、主要完成人情况（不超过 6 人）

姓名	排名	行政职务	技术职称	工作单位	完成单位	对本项目贡献
梅少辉	1		教授	西北工业大学	西北工业大学	梅少辉是团队和成果负责人，是代表性论文专著【1,2,4】的第一作者或通信作者，对该成果的科学发现 1 和科学发现 2 做出了重要贡献。对科学发现 1，揭示了多媒体数据间线性/非线性表示机理，构建了面向多媒体概要表示的协同/稀疏重构理论体系；对科学发现 2，发现了多媒体大数据中感兴趣目标的多维度特征变化规律，构建了感兴趣目标的旋转不变性和尺度特征不变性表示理论体系。
马明阳	2		副教授	西北工业大学	西北工业大学	马明阳是代表性论文专著【2 和 4】的共同第一作者/作者，对该成果的科学发现 1 和科学发现 2 做出贡献。对科学发现点 1，构建了多媒体数据间长短期依赖关系模型，提出了基于时间相关性、时空同步性的内容自适应多媒体数据概要表示方法；对科学发现点 2，设计了双路多分支卷积神经网络模型和双中心损失函数，有效提升了感兴趣目标的旋转不变特征表示和提取能力。
万帅	3		教授	西北工业大学	西北工业大学	万帅是代表性论文专著【1,2,5】的作者以及代表性论文专著【6,7】的第一作者，对该成果的科学发现 1 和科学发现 3 做出贡献。对于科学发现 1，结合算法承载平台参数，提出了物理约束的远程多媒体图像自适应筛选传输方法；对于科学发现 3，揭示了单层/可分层视频压缩框架下率失真依赖机理，创建了不同层级率失真描述理论架构。
张顺	4		副教授	西北工业大学	西北工业大学	张顺是代表性论文专著【3】的第一作者，对该成果的科学发现 2 做出贡献。对科学发现 2，建立了高差异区分的目标特征学习理论和动态目标特征表示方法，设计了高辨别性特征学习神经网络，创立了高价值感兴趣目标样本挖掘策略、对称三元组损失和自监督模式训练方法，获得了适应感兴趣目标外观变化的辨别特征。

七、主要完成单位情况（不超过 3 个）

完成单位	排名	对本项目主要贡献（限 600 字）
西北工业大学	1	<p>“多媒体数据概要表示和高效压缩的理论与方法”是梅少辉教授团队近十年来持续研究完成的。该项目研究先后得到国家自然科学基金和 863 计划资助，学校为项目的立项论证、管理、实施、监督、推广等方面做了大量工作，为本项目的科技创新和推广应用提供了全方位的支持与保障。</p> <p>在科研资源支持上，学校依托深厚的“三航”学科背景，积极布局基础研究，为项目提供了大量专业设备与先进技术平台。学校出台的《西北工业大学关于加强基础研究工作的若干意见》等文件，从顶层设计层面为项目的研究提供了保障，助力研究团队在多媒体数据概要表示方面，运用高性能计算设备进行大数据分析与处理，在多媒体数据高效压缩方面，积极推进产学研合作，助力团队与华为、OPPO 等科技企业联合开展国际标准论证，为建立新方法、揭示新机理提供了技术支撑。</p> <p>在人才队伍方面，西北工业大学的信息与通信工程、计算机科学与技术等相关学科优势显著。围绕本项目的研究方向，为团队配备了一批具有家国情怀、科学家精神优秀专业人才。另一方面，学校将各类大平台、大项目、大成果等优势科研资源转化为教育教学的源头活水，提升学生的实践能力和创新思维，为本项目的理论和技术突破注入新鲜血液。</p>

八、完成人合作关系说明

一、第一完成人梅少辉教授与项目第二完成人马明阳副教授、第三完成人万帅教授，自 2015 年开展合作。马明阳于 2015 年 9 月至 2021 年 12 月在梅少辉和万帅教授的联合指导下，进行硕博连读学习，并围绕“视频摘要”开展博士论文研究；博士毕业后，加入梅少辉教授负责的科研团队。在此期间，马明阳参与了梅少辉教授主持的两项 863 项目以及国家自然科学基金面上项目，三人合著 20 余篇高水平期刊或会议论文（含代表性论文专著 2），并共同获得 IEEE ISPACS 2017 会议最佳论文奖。

二、第一完成人梅少辉教授与项目第三完成人万帅教授，自 2010 年起长期开展合作研究。依托 863 项目和国家自然科学基金，在多媒体数据概要表示方面，合著了 20 余篇高水平期刊或会议论文（含代表性论文专著 1、2），完成了“面向空间微小操控平台的图像获取与自主筛选处理技术”成果，并通过科技

成果鉴定；在多媒体数据高效压缩编码方面，双方开展专利合作，共同授权了多项发明专利。

三、第一完成人梅少辉教授与项目第四完成人张顺副教授，自 2016 年开展合作。张顺副教授自 2016 年进入西北工业大学电子信息学院工作，加入梅少辉教授负责的团队。双方在多媒体数据目标特征学习和表示方面开展联合研究，合著了 10 余篇高水平期刊及会议论文。