

2026 年度陕西省自然科学奖公示信息

项目名称		以人为中心的视觉内容跨域生成理论与方法		
提名者		陕西省知识产权局		
提名意见				
我单位认真审阅了该项目提名书及附件材料，确认全部材料真实有效，并按照要求对该项目的基本情况进行了公示，公示期间无异议。 该项目面向视觉内容跨域生成领域的基础科学问题，围绕“如何有效学习以人类视觉感知为约束的跨域映射函数，实现高质量、高保真、高可信的跨域视觉内容生成”这一关键科学问题开展了长期系统研究，构建了以人为中心的“表示—生成—评价”闭环理论体系。项目取得了多粒度感知驱动的视觉内容跨域对齐表征理论、人类认知语义引导的视觉内容跨域生成模型和人类视觉感知牵引的质量评价与反馈增强框架三项科学发现。三个科学发现相互衔接、层层递进，将人的感知、认知与评价贯穿跨域映射学习全过程，具有鲜明的原创性和系统性，为视觉内容跨域生成提供了新的理论基础和方法体系。 项目代表性成果发表后受到国内外同行广泛关注和高度评价，八篇代表性论文、专著谷歌学术被引 2580 次，单篇最高被引 931 次，成果被 60 多个国家和地区的学者引用，包括多位国内外院士及 IEEE/ACM/AAAI Fellow，相关理论与方法具有重要国际学术影响。部分成果已服务于国家重要任务和深空探测等重大需求，体现出良好的科学价值和应用前景。 根据《陕西省科学技术厅关于 2026 年度省科学技术奖提名工作的通知》，参照自然科学奖评定条件和评定标准，提名该项目参评陕西省科学技术一等奖。 提名该项目为陕西省自然科学奖一等奖。				
主要完成人情况				
排名	姓名	技术职称	工作单位	完成单位
1	高新波	教授	西安电子科技大学	西安电子科技大学
2	王楠楠	教授	西安电子科技大学	西安电子科技大学
3	李洁	教授	西安电子科技大学	西安电子科技大学
4	朱明瑞	副教授	西安电子科技大学	西安电子科技大学
5	李航宇	讲师	西安电子科技大学	西安电子科技大学
6	牛振兴	教授	西安电子科技大学	西安电子科技大学
主要完成单位				
排名	单位名称			
1	西安电子科技大学			
代表性论文（专著）目录				

序号	论文（专著） 名称/刊名/作者	年卷页 码（xx 年 xx 卷 xx 页）	发表 （出 版） 时间 （年 月 日）	通讯作 者（含 共同）	第一作 者（含 共同）	作者	论文署名单 位
1	Channel-Wise and Spatial Feature Modulation Network for Single Image Super-Resolution/ IEEE Transactions on Circuits and Systems for Video Technology/Yanting Hu, Jie Li, Yuanfei Huang, Xinbo Gao	2019 年 30 卷 3911- 3927 页	2019 年 05 月 07 日	Xinbo Gao	Yanting Hu	胡彦婷, 李 洁, 黄源飞, 高新波	西安电子科 技大学
2	VideoReTalking: Audio- based Lip Synchronization for Talking Head Video Editing In the Wild/ SIGGRAPH Asia/Kun Cheng, Xiaodong Cun, Yong Zhang, Menghan Xia, Fei Yin, Mingrui Zhu, Xuan Wang, Jue Wang, Nannan Wang	2022 年 30 卷 1-9 页	2022 年 11 月 30 日	Xiaodong Cun	Kun Cheng, Xiaodong Cun	程坤, 寸晓 东, 张勇, 夏 孟涵, 印飞, 朱明瑞, 王 璇, 王珏, 王 楠楠	西安电子科 技大学 腾讯 AI Lab 清华大学
3	Ordinal Regression with Multiple Output CNN for Age Estimation/IEEE Conference on Computer Vision and Pattern Recognition/ Zhenxing Niu, Mo Zhou, Le Wang, Xinbo Gao, Gang Hua	2016 年 01 卷 4920- 4928 页	2016 年 12 月 12 日	Xinbo Gao	Zhenxing Niu	牛振兴, 周 默, 王乐, 高 新波, 华刚	西安电子科 技大学 西安交通大 学 微软亚洲研 究院
4	HSME: Hypersphere Manifold Embedding for Visible Thermal Person Re- Identification/Proceedings of the AAAI Conference on Artificial Intelligence/ Yi Hao, Nannan Wang, Jie Li, Xinbo Gao	2019 年 33 卷 8385 - 8392 页	2019 年 07 月 17 日	Nannan Wang	Yi Hao	郝毅, 王楠 楠, 李洁, 高 新波	西安电子科 技大学

5	Adaptively Learning Facial Expression Representation via C-F Labels and Distillation/ IEEE Transactions on Image Processing/ Hangyu Li, Nannan Wang, Xinpeng Ding, Xi Yang, Xinbo Gao	2021 年 30 卷 2016- 2028 页	2021 年 01 月 13 日	Nannan Wang	Hangyu Li	李航宇, 王楠楠, 丁鑫棚, 杨曦, 高新波	西安电子科技大学
6	模糊聚类分析及其应用/ 西安电子科技大学出版社/高新波	2004 年 1- 212 页	2004 年 01 月 01 日	高新波	高新波	高新波	西安电子科技大学
7	视觉信息质量评价方法/ 西安电子科技大学出版社/高新波, 路文	2010 年 1- 167 页	2010 年 01 月 01 日	高新波	高新波	高新波, 路文	西安电子科技大学
8	人脸图像合成与识别/高等教育出版社/高新波, 王楠楠	2022 年 1- 273 页	2022 年 03 月 01 日	高新波	高新波	高新波, 王楠楠	西安电子科技大学

客观评价

项目研究成果得到 60 多个国家/地区学者包括 30 余位院士和 100 余位 IEEE/ACM/AAAI Fellow 的广泛引用和高度评价。八篇代表性论文、专著谷歌学术被引 2580 次，单篇最高被引 931 次。

1. 科学发现点[1]的代表性评价

IEEE/ACM/AAAS/IAPR/SPIE Fellow、美国国家工程院院士、密歇根州立大学杰出教授 Anil K. Jain 等引用代表性论文 3，采用了该工作提出的 AFAD 数据集实现其人口属性估计模型的训练与测试。（“... AFAD [41] are the datasets for training and testing demographic estimation models”）

AAAS/ASA/IMS Fellow、美国国家科学院院士、斯坦福大学教授 Wing Hung Wong 等引用了代表性论文 3，指出该工作中的序数回归策略在年龄估计任务中展现出重要且有效的作用，能够有效应对面部衰老过程中存在的非平稳性问题。（“Ordinal regression is important and effective ... used in age estimation [17], [18] to address the non-stationary property of facial aging process.”）

IEEE/ACM/AAAI/AIMBE/IAPR/SPIE Fellow、欧洲科学院院士、美国国家发明家科学院院士 Jiebo Luo 教授等引用了代表性论文 4，指出该工作利用全局和局部特征提取到有效的图像表示，并取得令人满意的性能表现。（“... [7] used global features or local features of the holistic image as shown in Fig. 1(a) to extract effective feature representations and achieved satisfactory performance.”）

IEEE/AAIA/ISCA Fellow、美国卡内基梅隆大学教授 Bhiksha Raj 等引用了代表性论文 4，指出该工作在超球面上解决行人重识别中的表征学习问题具有良好有效性，并强调了其在嵌入空间下的泛化性。（“learning a representation on the hypersphere is also shown generally useful [63] ... beneficial to generalizability.”）

2. 科学发现点[2]的代表性评价

IEEE Fellow、德国国家工程院院士、慕尼黑工业大学教授 Alois Knoll 等人引用了代表性论文 1，沿用了该工作提出的通道和空间特征调制网络，有效提取了图像中的信息特征。（“Hu et al. proposed a channel-wise and spatial feature modulation network for image super-resolution to capture more informative features [10]. Our case of modulation is consistent with that of [10] to capture informative features for high-quality image restoration.”）

IEEE/MICCAI Fellow、美国医学与生物工程院院士、耶鲁大学教授 James S. Duncan 等引用了代表性论文 1，指出该工作提出的空间-通道注意力机制有效提升了图像超分辨率性能。（“Hu et al. [36] recently also demonstrated that spatial-channel attention can boost the image super-resolution performance”）

IEEE Fellow、美国国家发明家科学院终身院士、达特茅斯学院教授 Hany Farid 等引用了代表性论文 2，指出该工作是广泛使用的深度伪造引擎，并基于此构建了全新的伪造数据集。（“We created lip-sync deepfakes for the controlled and Anderson Cooper datasets using VideoRetalking [7] and Wav2Lip [17], two widely used deepfake engines.”）

IEEE Life Fellow、香港理工大学名誉教授 Wan-Chi Siu 和 H. Anthony Chan 等引用了代表性论文 2，指出该工作是说话人脸生成的最好方法。（“State-of-The-Art models in talking face generation including ... and VideoReTalking [28].”）

3. 科学发现点[3]的代表性评价

IEEE/AAAS/IFAC Fellow、美国国家发明家科学院院士、新泽西理工学院杰出教授 MengChu Zhou 等引用了代表性论文 5，受到该工作的启发提出了课程模糊策略。（“Inspired by the Curriculum Learning (CL) [13], [14] of human educational wisdom, we propose a CL-based fuzzy strategy for adaptively tuning ...”）

IEEE/ACM/AAAC/BCS/DIRDI/ELLIS/ISCA Fellow、英国帝国理工学院教授 Björn W. Schuller 等引用了代表性论文 5，指出该工作是有竞争力的方法，实现了最先进的性能表现。（“a number of competitive FER methods [5], [6], [7], [8], [9] ... we compare the proposed method with state-of-the-art methods: KTN [9]”）

4. 国内外重要科技奖励

相关算法分别获计算机视觉国际顶级会议 ICCV 2019 受限条件下超分辨重建竞赛第一名、ECCV 2020 图像极限修复竞赛第一名。

5. 成果评价及应用

基于科学发现点[1]和[3]的所支撑研发的图像跨域生成系统应用于国家某部，配合完成了某项重要任务，取得了突出成效，获得党和国家领导人的重要批示；《视觉信息质量评价方法》获第三届中国大学出版社图书奖优秀学术著作奖二等奖；《人脸图像合成与识别》获教育部一华为优秀教学资源奖。科学发现点[2]中的视觉内容跨域生成算法已应用于嫦娥七号月球探测器科学载荷服务于深空探测图像压缩与重建，实现了算力在 12 Tops 功耗小于 20W 要求下的高保真度图像传输；科学发现点[2]中的通道与空间双重自适应调制算法，落地华为 Mate70 系列和 MateX6 旗舰产品的长焦人像增强，并获得华为火花奖。基于科学发现点[2]和[3]的质量评价与反馈增强框架所研发的 AI 画质增强算法及其 IP 核，成功应用于海信发布的中国首颗全自研 8K 120Hz AI 画质芯片，填补了国产化超高清画质芯片领域的技术空白。

项目简介

视觉内容跨域生成在突破视觉模态固有边界、实现“感知—理解—创造”智能闭环方面具有深远的科学意义。视觉域是指在视觉特征空间和语义空间中均保持内在一致性的视觉内容集合，涵盖空间域（不同分辨率）、光谱域（可见光、红外等）、属性域（年龄、表情、身份等）及时空域（视频等影像序列）等，其元素服从特定的高维概率分布。视觉跨域生成是指从源域内容变换为目标域内容的过程，其科学本质在于建立以人的视觉感知与认知特性为约束的跨域映射关系。不同视觉域间的高维异构性与语义多样性使跨域映射难以通过解析方法获得闭式解，而已有数据驱动方法忽视了“人”在感知、认知与评价中的关键作用，导致跨域生成内容与人类视觉体验之间存在显著鸿沟，难以兼顾精确性、自然性与可信性。

在国家自然科学基金重点项目等资助下，经过多年攻关，团队构建了以人为中心的视觉内容跨域表示、跨域生成与质量评价反馈增强的理论与方法体系，形成了“表示—生成—评价”闭环的理论框架。该框架以“关注于人”为感知基石，以“服务于人”为认知核心，以“受控于人”为评价保障，将人的感知、认知与评价能力贯穿于跨域映射学习的全程。聚焦“如何有效学习以人类视觉感知为约束的跨域映射函数，实现高质量、高保真、高可信的跨域视觉内容生成”这一关键科学问题，本项目取得了如下科学发现：

1. 多粒度感知驱动的视觉内容跨域对齐表征理论。发现了不同视觉域特征空间的内在关联性，揭示了超球面流形嵌入与人类多粒度感知之间的一致性规律，以及有序结构约束与人类对连续属性变化感知的同向激励机制。针对可见光与红外等光谱域之间因成像机理差异导致的域鸿沟，提出了跨域超球面流形嵌入方法，将不同域特征映射到共享流形空间，以域内鉴别性约束与域间对齐性约束进行联合优化，解决了传统方法因域鸿沟导致的表征坍塌问题。在此基础上，将上述流形表征思想拓展至年龄、表情等具有人类认知序关系的表观属性域，提出了基于有序回归的结构化表征模型，使属性的序关系显式编码于特征空间中。受人的多粒度感知特性启发，依托模糊聚类分析理论对高维空间中不确定关系的柔性建模能力，将光谱跨域对齐与语义跨属性结构化统一于同一表征框架中，形成了多粒度感知驱动的视觉内容跨域对齐表征理论，突破了传统单一域特征表示在跨域泛化能力上的固有瓶颈。美国国家科学院院士、斯坦福大学 Wing Hung Wong 教授等评价该工作中的序数回归策略在展现出重要且有效的作用，能够有效应对面部属性变化过程中存在的非平稳性问题。

2. 人类认知语义引导的视觉内容跨域生成模型。发现了人类语义先验对跨域内容生成的有效引导作用，揭示了通道-空间联合注意力机制与人类视觉系统多级特征筛选的一致性规律，提出了“特征调制—语义解耦—精细增强”的跨域生成范式。以跨分辨率生成为例，通过模拟视觉注意的通道选择性增强与空间聚焦特性，设计了通道与空间双重自适应调制网络，实现了低分辨率到高分辨率的精细像素重建；以数字人生成为例，将上述调制机制从像素推广至语义层面，通过“表情解耦—口型同步—身份增强”的多阶段框架，将音频语义信息跨域投射至视觉空间，在保持说话人身

份一致性的前提下实现自然的口型同步与表情动态。在此过程中，人脸图像的语义解耦与身份保真约束为跨域生成提供了基础性的理论支撑，使得分辨率跨域与视听跨域两种形态被统一于“以语义为引导、以人为对象”的生成框架之下，突破了传统模型在跨域语义一致性和人像保真度上的根本性局限。IEEE Fellow、德国国家工程院院士、慕尼黑工业大学 Alois Knoll 教授等沿用了该工作提出的通道和空间特征调制网络，有效提取了图像中的信息特征。

3. 人类视觉感知牵引的质量评价与反馈增强框架。发现了人类视觉感知质量与视觉内容特征分布之间的相关性规律，构建了融合结构统计分布、模糊语义量化与深度特征表征的多粒度质量评价范式，建立了从主观感知到客观计算的映射桥梁。在此基础上，进一步揭示了质量评价反馈对特征学习过程的正则化作用机理，以人脸属性分析为例，提出了融合粗-细粒度标签策略与情感蒸馏机制的自适应质量增强方法，以高质量样本特征为“教师”，通过知识蒸馏将判别性知识迁移至“学生”样本，将人类感知先验转化为可计算的损失函数，驱动特征学习的质量自适应优化。由此形成了“评价建模—质量感知—知识蒸馏—生成增强”的闭环增强理论，将人对视觉内容的粗粒度层级判断转化为细粒度结构化反馈，驱动跨域生成过程的持续优化，解决了视觉内容主客观质量评价不一致的难题，实现了视觉内容生成的迭代优化和质量提升。IEEE Fellow、美国国家发明家科学院院士、新泽西理工学院杰出教授 MengChu Zhou 等受到该工作的启发提出了课程模糊策略。

上述三个科学发现呈现递进支撑的关系：视觉内容跨域对齐表征理论从人的多粒度感知特性出发，为生成提供统一表征基础；视觉内容跨域生成模型将人的高层认知注入跨域映射，实现高质量跨域转换；质量评价与反馈增强理论以人的主观评价为基准，驱动生成任务迭代优化与生成质量持续提升。三者以“关注于人—服务于人—受控于人”为内在逻辑主线，形成了“表示—生成—评价”闭环理论体系，为视觉内容跨域生成提供了系统的理论基础与方法支撑，推动了计算机视觉从“技术最优”向“以人为中心”的范式演进。

八篇代表性论文专著谷歌学术被引 2580 次，单篇最高被引 931 次，受到 60 多个国家/地区学者的广泛引用和高度评价，包括 30 余位中/美/英/加等国院士和 100 余位 IEEE/ACM/AAAI Fellow。所提出的理论方法应用于国家某部重要任务并取得突出成效，获党和国家领导人重要批示；所研制的大模型端侧推理芯片，应用于嫦娥七号月球探测器科学载荷，服务深空探测图像压缩与重建。项目组入选科技部重点领域创新团队和教育部创新团队，其中 2 人入选国家高层次人才支持计划，1 人入选 IEEE Fellow。

完成人合作关系说明

“以人为中心的视觉内容跨域生成理论与方法”项目第一完成人高新波、第二完成人王楠楠、第三完成人李洁、第四完成人朱明瑞、第五完成人李航宇和第六完成人牛振兴的完成单位均为西安电子科技大学。

项目开展期间，第一完成人高新波提出科学发现点[1]，与第六完成人牛振兴合著代表性论文 3，与第二完成人王楠楠合著代表性论文 4；提出科学发现点[2]，与第三完成人李洁合著代表性论文 1；提出科学发现点[3]，与第二完成人王楠楠、第五完成人李航宇合著代表性论文 5。

第二完成人王楠楠在西安电子科技大学攻读博士学位期间开始参与本项目的研究工作，第一完成人高新波是其博士生导师。第二完成人王楠楠博士毕业留校任教后继续作为核心人员参与本项目。在该项目中，第一完成人高新波与第二完成人王楠楠合著代表性论文 2、4、5。

自该项目开始，第一完成人高新波和第三完成人李洁进行了卓有成效的合作研究，并组建了学术团队，共同指导博士、硕士研究生开展视觉内容跨域生成理论与方法方面的研究工作。在该项目中，第一完成人高新波与第三完成人李洁合著代表性论文 1、5。

第四完成人朱明瑞在西安电子科技大学攻读博士学位期间开始参与本项目的研究工作，第三完成人李洁是其博士指导教师，毕业留校任教后继续作为核心人员参与本项目。在该项目中，第二完成人王楠楠与第四完成人朱明瑞合著代表性论文 2。

第五完成人李航宇在西安电子科技大学攻读博士学位期间开始参与本项目的研究工作，第二完成人王楠楠是其博士指导教师，毕业留校任教后继续作为核心人员参与本项目。在该项目中，第一完成人高新波与第五完成人李航宇合著代表性论文 5。

第六完成人牛振兴在西安电子科技大学攻读博士学位期间开始参与本项目的研究工作，第一完成人高新波是其博士生指导教师。第六完成人牛振兴博士毕业留校任教后继续作为核心人员参与本项目。在该项目中，第一完成人高新波与第六完成人牛振兴合著代表性论文 3。