



师资队伍

Teaching staff

学术科研

您所在的位置：[教师科研](#) > [师资队伍](#)

详细资料

Details



金琴

金琴，教授，博士生导师。主要研究领域包括多媒体智能计算，人机交互。获2016年IBM SUR Award；2018-2020 CVPR ActivityNet Dense Video Captioning Challenge蝉联冠军，2016-2017 ACM Multimedia (Video to Language) Grand Challenge蝉联...

电话：010-62510042

个人主页：www.jin-qin.com

电子邮箱：qjin@ruc.edu.cn

[更多](#)

教育经历

1999.08 - 2007.01, 美国卡内基梅隆大学, 博士

1991.09 - 1999.01, 清华大学, 本科、硕士

工作经历

2013.01 - 今, 中国人民大学, 教授

2012.04 - 2012.12, IBM中国研究院, 研究员

2008.06 - 2012.03, 美国卡内基梅隆大学, 研究员, 教师

2007.02 - 2008.05, 美国卡内基梅隆大学, 博士后

研究方向

多媒体智能计算, 人机交互

讲授课程

多媒体技术 (Multimedia Technologies)

言语信息处理 (Spoken Language Processing)

科研项目

国家自然科学基金 (2021.01-2024.12)

国家自然科学基金 (2018.01-2021.12)

北京市自然科学基金 (2019.01-2021.12)

北京市储备课题基金 (2018-2020)

国家重点研发计划项目 (2016.07-2020.07)

北京市自然科学基金 (2014.01-2016.12)

中国人民大学科学研究基金 (2014.01-2016.12)

教育部留学回国人员科研启动基金 (2015.01-2016.12)

科研成果

- Shizhe Chen, Qin Jin, Peng Wang, Qi Wu. Say As You Wish: Fine-Grained Control of Image Caption Generation With Abstract Scene Graphs. CVPR, 2020.
- Shizhe Chen, Yida Zhao, Qin Jin, Qi Wu. Fine-Grained Video-Text Retrieval With Hierarchical Graph Reasoning. CVPR, 2020.
- Jia Chen, Qin Jin. Better Captioning With Sequence-Level Exploration. CVPR, 2020.
- Sipeng Zheng, Shizhe Chen, Qin Jin. Skeleton-based Interactive Graph Network for Human Object Interaction Detection. ICME, 2020.
- Shizhe Chen, Qin Jin, Alexandar Hauptmann. Unsupervised Bilingual Lexicon Induction from Mono-lingual Multimodal Data. AACL, 2019.
- Jingjun Liang, Shizhe Chen, Jinming Zhao, Qin Jin, Haibo Liu, Li Lu. Cross-culture Multimodal Emotion Recognition with Adversarial Learning. ICASSP, 2019.
- Shizhe Chen, Yuqing Song, Yida Zhao, Qin Jin, Zhaoyang Zeng, Bei Liu, Jianlong Fu, Alexander Hauptmann. ActivityNet 2019 Task 3: Exploring Contexts for Dense Captioning Events in Video. CVPR 2019, ActivityNet Large Scale Activity Recognition Challenge.
- Shizhe Chen, Qin Jin, Jianlong Fu. From Words to Sentences: A Progressive Learning Approach for Zero-resource Machine Translation with Visual Pivots. IJCAI, 2019.
- Shizhe Chen, Qin Jin, Jia Chen, Alexander G. Hauptmann. Generating Video Descriptions With Latent Topic Guidance. IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 21, NO. 9, SEPTEMBER 2019.
- Jinming Zhao, Shizhe Chen, Jingjun Liang, Qin Jin. Speech Emotion Recognition in Dyadic Dialogues. INTERSPEECH, 2019.
- Yuqing Song, Shizhe Chen, Qin Jin. Unpaired Cross-lingual Image Caption Generation with Self-Supervised Rewards. ACM Multimedia, 2019.
- Sipeng Zheng, Shizhe Chen, Qin Jin. Visual Relation Detection with Multi-Level Attention. ACM Multimedia, 2019.
- Shizhe Chen, Bei Liu, Jianlong Fu, Ruihua Song, Qin Jin, Pingping Lin, Xiaoyu Qi, Chunting Wang, Jin Zhou. Neural Storyboard Artist: Visualizing Stories with Coherent Image Sequences. ACM Multimedia, 2019.

- Sipeng Zheng, Xiangyu Chen, Shizhe Chen, Qin Jin. Relation Understanding in Videos. ACM Multimedia, Grand Challenge: Relation Understanding in Videos, 2019.
- Jinming Zhao, Ruichen Li, Jingjun Liang, Qin Jin. Adversarial Domain Adaption for Multi-Cultural Dimensional Emotion Recognition in Dyadic Interactions. AVEC, 2019.
- Shizhe Chen, Yida Zhao, Yuqing Song, Qin Jin, Qi Wu. Integrating Temporal and Spatial Attentions for VATEX Video Captioning Challenge 2019. ICCV, VATEX Video Captioning Challenge 2019.
- Weiying Wang, Yongcheng Wang, Shizhe Chen, Qin Jin. YouMakeup: A Large-Scale Domain-Specific Multimodal Dataset for Fine-Grained Semantic Comprehension. EMNLP, 2019.
- Yuqing Song, Yida Zhao, Shizhe Chen, Qin Jin. RUC_AIM3 at TRECVID 2019: Video to Text. NIST TRECVID, 2019.
- Jingjun Liang, Shizhe Chen, Qin Jin. Semi-supervised Multimodal Emotion Recognition With Improved Wasserstein GANs. APSIPA ASC, 2019.
- Shizhe Chen, Yuqing Song, Yida Zhao, Qin Jin, Alexandar Hauptmann. RUC+CMU: System Report for Dense Captioning Events in Videos. CVPR ActivityNet Large Scale Activity Recognition Challenge, 2018.
- Shizhe Chen, Jia Chen, Qin Jin, Alexandar Hauptmann. Class-aware Self-Attention for Audio Event Recognition. ACM International Conference on Multimedia Retrieval (ICMR), 2018. (Best Paper Runner-up)
- Jinming Zhao, Shizhe Chen, Qin Jin. Multimodal Dimensional and Continuous Emotion Recognition in Dyadic Video Interactions. Pacific-Rim Conference on Multimedia (PCM), 2018.
- Xiaozhu Lin, Qin Jin, Shizhe Chen, Yuqing Song, Yida Zhao. iMakeup: Makeup Instructional Video Dataset for Fine-grained Dense Video Captioning. Pacific-Rim Conference on Multimedia (PCM), 2018.
- Jinming Zhao, Ruichen Li, Shizhe Chen, Qin Jin. Multi-modal Multi-cultural Dimensional Continuous Emotion Recognition in Dyadic Interactions. ACM Multimedia Audio-Visual Emotion Challenge (AVEC) Workshop, 2018.
- Shizhe Chen, Jia Chen, Qin Jin, Alexandar Hauptmann. Video Captioning with Guidance of Multimodal Latent Topics. ACM Multimedia, 2017.
- Qin Jin, Shizhe Chen, Jia Chen, Alexandar Hauptmann. Knowing Yourself: Improving Video Caption via In-depth Recap. ACM Multimedia, 2017.
- Shizhe Chen, Qin Jin, Jinming Zhao and Shuai Wang. Multimodal Multi-task Learning for Dimensional and Continuous Emotion Recognition. ACM Multimedia Audio-Visual Emotion Challenge (AVEC) Workshop 2017.

- Shizhe Chen, Jia Chen, Qin Jin. Generating Video Descriptions with Topic Guidance. International Conference on Multimedia Retrieval (ICMR) 2017.
- Jia Chen, Qin Jin, Shiwan Zhao, Shenghua Bao, Li Zhang, Zhong Su, Yong Yu. Boosting Recommendation in Unexplored Categories by User Price Preference. ACM Transactions on Information Systems (TOIS) Volume 35 Issue 2, October 2016.
- Qin Jin, Jia Chen, Shizhe Chen, Yifan Xiong. Describing Videos using Multi-modal Fusion. ACM Multimedia, 2016.
- Jia Chen, Qin Jin, Yifan Xiong. Semantic Image Profiling for Historic Events: Linking Images to Phrases. ACM Multimedia 2016.
- Shizhe Chen, Qin Jin, Multi-modal Conditional Attention Fusion for Dimensional Emotion Prediction, ACM Multimedia 2016.
- Yifan Xiong, Jia Chen, Qin Jin, Chao Zhang. History Rhyme: Searching Historic Events by Multimedia Knowledge. ACM Multimedia 2016.
- Xirong Li, Yujia Huo, Qin Jin, Jieping Xu. Detecting Violence in Video using Subclasses. ACM Multimedia, October, 2016.
- Shizhe Chen, Xinrui Li, Qin Jin, Shilei Zhang, Yong Qin. Video Emotion Recognition in the Wild Based on Fusion of Multimodal Features. International Conference on Multimodal Interaction (ICMI) 2016.
- Guankun Mu, Haibing Cao, Qin Jin. Violent Scene Detection Using Convolutional Neural Networks and Deep Audio Features. Chinese Conference on Pattern Recognition (CCPR) 2016.
- Shizhe Chen, Yujie Dian, Xinrui Li, Xiaozhu Lin, Qin Jin(*), Haibo Liu, Li Lu. Emotion Recognition in Videos via Fusing Multimodal Features. Chinese Conference on Pattern Recognition (CCPR) 2016.
- Xirong Li, Qin Jin. Improving Image Captioning by Concept-based Sentence Reranking. Pacific-Rim Conference on Multimedia (PCM), September 2016.
- Qin Jin, Junwei Liang, Xiaozhu Lin. Generating Natural Video Descriptions via Multimodal Processing. Interspeech 2016.
- Qin Jin, Junwei Liang. Video Description Generation using Audio and Visual Cues. International Conference on Multimedia Retrieval (ICMR) 2016.
- Jia Chen, Qin Jin, Shenghua Bao, Junfeng Ye, Zhong Su, Shimin Chen, Yong Yu. Exploitation and Exploration Balanced Hierarchical Summary For Landmark Images. IEEE Transactions on Multimedia (TMM), Volume:17 Issue:10, 2015.
- Jia Chen, Qin Jin, Yong Yu, Alexander G. Hauptmann, Image Profiling for History Events on the Fly. ACM Multimedia 2015.
- Shizhe Chen, Qin Jin. Multi-modal Dimensional Emotion Recognition using Recurrent Neural Networks. ACM Multimedia Audio/Visual Emotion Challenge and Workshop 2015.
- Shimin Chen and Qin Jin, Persistent B+-Trees in Non-Volatile Main Memory, VLDB, Hawaii, USA, 2015.
- Qin Jin□, Xirong Li□, Haibing Cao, Yujia Huo, Shuai Liao, Gang Yang, Jieping Xu. RUCMM at MediaEval 2015 Affective Impact of Movies Task: Fusion of Audio and Visual Cues. MediaEval Workshop 2015, Wurzen, Germany.

- Xirong Li, Qin Jin, Shuai Liao, Junwei Liang, Xixi He, Yujia Huo, WeiYu Lan, Bin Xiao, Yanxiong Lu, Jieping Xu. RUC-Tencent at ImageCLEF 2015: Concept Detection, Localization and Sentence Generation, CLEF working notes, 2015.
- Jia Chen, Min Li, Qin Jin, Yongzhe Zhang, Shenghua Bao, Zhong Su, Yong Yu, Lead Curve Detection in Drawings with Complex Cross-Points, Neurocomputing, 2015.
- Qin Jin, Junwei Liang, Xixi He, Gang Yang, Jieping Xu, Xirong Li, Semantic Concept Annotation for User Generated Videos Using Soundtracks. International Conference on Multimedia Retrieval (ICMR) 2015.
- Qin Jin, Chengxin Li, Shizhe Chen, Huimin Wu, Speech Emotion Recognition With Acoustic And Lexical Features, in Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia, 2015.
- Junwei Liang, Qin Jin, Xixi He, Gang Yang, Jieping Xu, Xirong Li, Detecting Semantic Concepts In Consumer Videos Using Audio, in Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia, 2015.
- Jia Chen, Qin Jin, Shiwan Zhao, Shenghua Bao, Li Zhang, Zhong Su, Yong Yu. Does Product Recommendation Meet its Waterloo in Unexplored Categories? No, Price Comes to Help, SIGIR 2014.
- Junwei Liang, Qin Jin, Xixi He, Xirong Li, Gang Yang, Jieping Xu. Semantic Concept Annotation of Consumer Videos at Frame-level Using Audio. Pacific-rim Conference on Multimedia (PCM) 2014.
- Shizhe Chen, Qin Jin, Xirong Li, Gang Yang, Jieping Xu. Speech Emotion Classification using Acoustic Features. International Symposium on Chinese Spoken Language Processing (ISCSLP), 2014.
- Jia Chen , Qin Jin , Weipeng Zhang , Shenghua Bao , Zhong Su , Yong Yu , Tell Me What Happened Here in History , ACM International Conference on Multimedia, 2013.

社会兼职

Associate Editor, ToMM

Area Chair of ACM Multimedia 2018, 2020

Special Session Chair of APSIPA-ASC 2016, 2019.

Member of CCF, ACM, IEEE, ISCA

荣誉获奖

CVPR 2020 ActivityNet Large Scale Activity Recognition Challenge (ANET) Dense Captioning Events in Videos Task (Winner)
The End-of-End-to-End A Video Understanding Pentathlon @CVPR 2020 (Rank 2nd)
Outstanding Method Award in VATEX Video Captioning Challenge @ ICCV 2019
2019, 之江杯全球人工智能大赛视频内容描述生成 (第一名)
CVPR 2019 ActivityNet Large Scale Activity Recognition Challenge (ANET) Temporal Captioning Task (Winner)
2019 TRECVID (Video to Text Description) Grand Challenge (Rank 1st)
2019 Audio-Visual Emotion Challenge @ ACM Multimedia 2019 (Winner)
CVPR 2018 ActivityNet Large Scale Activity Recognition Challenge (ANET) Temporal Captioning Task (Winner)
2018 TRECVID (Video to Text Description) Grand Challenge (Rank 1st)
2018 Audio-Visual Emotion Challenge @ ACM Multimedia 2018 (Winner)
2017 TRECVID (Video to Text Description) Grand Challenge (Rank 1st)
Best Grand Challenge Paper Award at ACM Multimedia 2017
2017 ACM Multimedia (Video to Language) Grand Challenge (Rank 1st)
2017 Audio-Visual Emotion Challenge @ ACM Multimedia 2017 (Winner)
2016 IBM SUR Award
2016 ACM Multimedia (Video to Language) Grand Challenge (Rank 1st)
2016 Audio-Visual Emotion Challenge (AVEC) (Rank 2nd)
2016 MediaEval Movie Emotion Impact Challenge (Rank 1st)
2016 Chinese Multimodal Emotion Challenge (MEC) (Rank 2nd)
2016 NLPCC Chinese Weibo Stance Detection (Rank 1st)
2015 ImageCLEF (Image Sentence Generation) Evaluation (Rank 1st)
2015 Outstanding Bachelor Thesis Advisor of Renmin University of China

Copyright © 2015 School of Information, Renmin University of China. All rights reserved.
版权所有中国人民大学信息学院

地址：北京市海淀区中关村大街59号
邮编：100872