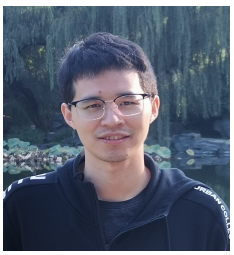
MingLang Qiao

Curriculum Vitae

Multimedia Communication Computing Lab
Room 207, IRC Building
Beihang University, China.
+86 178 0100 4742
minglangqiao@buaa.edu.cn
minglangqiao.github.io/
Third-year Ph.D. student



BIOGRAPHY

2018 – Now **Ph.D.**, *Dept. of Electronic Information Engineering*, Beihang University, Beijing, China.
GPA: 3.62/4.0
- Group: Multimedia Communication Computing (MC2) Lab (*See Home Page*)
- Major: Information and Communication Engineering with advisor, Prof. Mai Xu (*See Home Page*)
- Research interests:
  - **Computer Vision:** *Low-level Vision, Multi-task Learning, Multi-modal Learning*
  - **Perception:** *Visual Attention Modeling, Salient Object Detection, Quality Assessment*
  - **Video Coding:** *Perceptual Video Coding, Video Quality Enhancement*

2014 – 2018 **B.E.**, *Dept. of Electronic Information Engineering*, Beihang University, Beijing, China.
- Rank: 34/291

Scholarships

2021 Academic Scholarship of Beihang University ***2nd Prize***

2020 Academic Scholarship of Beihang University ***1st Prize***

2018 Admission Scholarship of Beihang University for First-year Postgraduate Student

2017 National Encouragement Scholarship

– *Directly awarded by the National Ministry of Education.*

2017 Outstanding Science and Technology Scholarship

Working Experience

2021 – Now QoE-Oriented Perceptual Coding of UGC Video with *Alibaba Cloud*

– *Alibaba Innovative Research Programme (See Home Page)*

– *Advisor: Doctor Yunjin Chen (See Home Page)*

2020 – 2021 QoE-Oriented Transcoding for E-Commerce Images on *Taobao* of Alibaba

– *Alibaba Innovative Research Programme*

2019 – 2020 Perceptual Transcoding of Live Video with *Momo Inc.* (*See Demo*)

– *The research results have been put into commercial use in Momo live video platform*

PUBLICATIONS

(∗: Corresponding Author. †: Co-first author.)

The publications are also listed in my *Google scholar page*.

Papers

[1] [Pre-print] **Minglang Qiao**, Yufan Liu, Mai Xu*, Xin Deng, Bin Li, Weiming Hu, Ali Borji. *Joint Learning of Visual-Audio Saliency Prediction and Sound Source Localization on Multi-face Videos.* International Journal of Computer Vision (**IJCV**, Under-review), 2021. *[Paper]* *[Dataset]*

[2] **Minglang Qiao**†, Yufan Liu†, Mai Xu*, Bing Li, Weiming Hu, Ali Borji. *Learning to Predict Salient Faces: A Novel Visual-Audio Saliency Model.* Proceedings of the European Conference on Computer Vision (**ECCV**), 2020. *[Paper]*

[3] **Minglang Qiao**, Mai Xu, Zulin Wang, Ali Borji. *Viewport-Dependent Saliency Prediction in 360° Video.* IEEE Transactions on Multimedia (**TMM**), 2020. *[Paper]* *[Dataset]*

[4] Mai Xu*, Yuhang Song, Jianyi Wang, **Minglang Qiao**, Liangyu Huo, Zulin Wang *Predicting Head Movement in Panoramic Video: A Deep Reinforcement Learning Approach.* IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**), 2018. *[Paper]* *[Code]* *[Dataset]*

[5] Lai Jiang, Mai Xu, Tie Liu, **Minglang Qiao**, Zulin Wang *Deepvs: A Deep Learning Based Video Saliency Prediction Approach.* Proceedings of the European Conference on Computer Vision (**ECCV**), 2018. *[Paper]* *[Code]* *[Dataset]*

Patents (A total of 9)

[6,..,13] Yuhang Song, Xianglong Kong, Haochen Wang, .., **Minglang Qiao**, Dajun Cui. *A Power/Joint/Servo/Connector Module Device in a Modular UAV System.* 4 Invention & 4 Utility Model Patents. Patent Numbers: ZL201620302692.X/.../201610224853.2. *[Patents]*

[14] **Minglang Qiao**, Haochen Wang, Jianyi Wang, Jun Wang, Zhongshen Shun, Songyang Zhang, Weiyang Qun. *A hand gesture recognition method based on distance-velocity feature.* Patent Numbers: CN107024685A. *[Patent]*

Experience

Research Experience (Completed)

2020 – 2021 **Joint Learning of Visual-Audio Saliency Prediction and Sound Source Localization on Multi-face Videos**.

- Main Works:
  - *Established a database consisting of 300 video sequences with human eye fixations captured in visual-audio condition, which can be used for multi-modal perception task*
  - *Investigated the influence of audio on human attention, and explored the factors that impact sound source localization*
  - *Proposed a model based on CNN and graph attention network to jointly learn the tasks of visual-audio saliency prediction and sound source localization*
- *Presented an application of saliency prediction in perceptual video compression, and the maximum increase of EWPSNR on multi-face video is about 2 dB over the conventional x264*

2019 – 2020 **Learning to Predict Salient Faces: A Novel Visual-Audio Saliency Model**.

- Main Works:
  - *Analyzed how human attention is affected by multiple factors including face and sound, as well as the difference of human attention between visual-only and visual-audio condition*
  - *Proposed a multi-modal network for human attention prediction in visual-audio condition*
- Improved performance of saliency prediction, outperforming state-of-the-art visual-audio saliency prediction approaches

2018 – 2019 **Viewport-Dependent Saliency Prediction in 360° Video**.

- Main Works:
  - *We found that the distribution of human fixations is influenced by the objects and location of the corresponding viewport in 360° video*
  - *First attempt to introduce the distribution bias into perception modeling of viewport in 360° videos*
  - *Built a multi-task deep neural network for viewport saliency prediction in 360° video*
- In experiments, our method performed considerably better than other state-of-the-art methods for viewport saliency prediction over 360° video

2017 – 2018 **Predicting Head Movement in Panoramic Video: A Deep Reinforcement Learning Approach**.

- Main Works:
  - *Constructed a new panoramic video database that consists of head movement positions of 58 subjects across 76 panoramic video sequences*
  - *Proposed an offline deep reinforcement learning approach (DHP) to predict head movement (HM) maps.*
  - *Developed an online-DHP approach to predict the HM position of one subject with the online manner*
- Experiments validate that proposed approach is effective in both offline and online prediction of HM positions for panoramic video.

2017 – 2018 **Deepvs: A Deep Learning Based Video Saliency Prediction Approach**.

- Main Works:
  - *Established an eye-tracking database that consists of 538 videos with diverse content, and analyzed the factors that influence human attention in videos*
  - *Proposed a novel object-to-motion CNN structure to predict intra-frame saliency, which integrates both objectness and object motion*
  - *Introduced a saliency structured ConvLSTM network with the center bias dropout and sparsity-weighted loss, to learn the saliency transition across inter-frames*

2017 – 2018 **Perceptual Transcoding of Live Video with Momo Inc**.

- Main Works:
  - *Developed a computation-efficient model for real-time saliency prediction of live video, which runs about 100 FPS on GPU*
  - *Implemented a region-of-interest based video compression approach in x264, which is guided by the predicted saliency result*
  - *Developed a real-time enhancement method to improve the visual quality of video before compression*
  - *Conducted several subjective quality assessment experiments to evaluate and optimize the proposed perceptual coding scheme*
- Saved bit-rate by 30% - 40% (over the conventional x264) for live video, with similar or better visual quality compared with the video before transcoding.
- Commercial use on Momo live video platform, covering millions of users

Research Experience (On-going)

2021 – Now **Simultaneous Detection and Ranking of Salient Objects**.

- Main Works:
  - *Developed a multi-task architecture to simultaneously detect, segment and rank multiple salient objects*
  - *Employed the encoder of Transformer and scene-graph to capture the relations between different salient objects*

2021 – Now **QoE-Oriented Perceptual Coding on UGC Video with Alibaba Cloud**.

- Main Works:
  - *Quality of Experience (QoE) oriented coding for user generated content (UGC) video*
  - *Target at compressing GUC video with better visual quality and lower bit-rate, which is critical for the transcoding service on Alibaba cloud*
  - *Based on the features of human visual system, including attention mechanism and masking effects* (e.g., JND)

HONORS & AWARDS

2020 Excellent Academic Paper Award of Beihang University

2017 National **1st** Prize, National Undergraduate Electronics Design Contest

– *Top 2.4%, among 14, 400+ .*

2016-2017 "Lanqiao Cup" National Software and Information Technology Contest ***2nd & 3rd Prize***
- *Sponsored by the National Ministry of Industry and Information*

2017 Second Prize on the 27th "Fengru Cup" Innovation Contest
- *The top innovation competition in Beihang University*

2017 Honorable Winner in 33th COMAP's Mathematical Contest in Modeling

2017 Second Prize on the Fengru Cup "Nokia" Innovation Contest

2016 First Prize on the 26th "Fengru Cup" Innovation Contest

2016 First Prize on the Beihang Electronic Innovation Contest

TEACHING

2021 **Introduction to Machine Learning**, *Department of Computer Science*, Beihang University, China.
- Role: Teaching Assistant & Tutor
- Lecturers:
  - *Prof. Mai Xu* (*See Home Page*)

2021 **Comprehensive Innovation: Digital Communications**, Beihang University, China.
- Role: Teaching Assistant & Tutor

SKILLS

Programming: Python, Matlab, C, C++, Java (basic), Cuda (basic)
Platform: Caffe, Pytorch, Tensorflow, Theano, Linux, Altium Designer, C51
Word processing: LaTex, Microsoft Office, Adobe Illustrator