WEN WEN

CS Ph.D. Candidate

🌐 Google Scholar   📞 (+86)18820427417

🔗 Linkedin   ✉️ winwinwenwen77@gmail.com

EDUCATION

Ph.D.	City University of Hong Kong - Computer Science	*2021 – 2025 (Expected)*
M.A.Sc.	Sun Yat-sen University - Applied Statistics	*2019 – 2021*
B.Sc.	Sun Yat-sen University - Statistics	*2015 – 2019*
Visiting	University of California, Berkeley - Math	*2017 Fall*

EXPERIENCE

Google YouTube - Student Researcher - South Bay, CA, USA   *Jun 2024 – May 2025*

- Designed and developed a novel MLLM architecture with dual vision encoders to independently analyze perceptual quality at both video context and pixel distortion granularities. A language decoder subsequently reasons about their interplay. This design yields robust quality scores, interpretable quality descriptions, and enhanced sensitivity to pixel distortions, such as fine-grained compression artifacts. The model achieved state-of-the-art or equivalent cross-dataset performance on VQA benchmarks with superior robustness to pixel-level distortions.
- Devised methodologies to assess the impact of frame pre-processing and robust sampling strategies on MLLM capabilities for short-form video quality perception. Subsequently, introduced a lightweight, learning-based ensemble method that adaptively integrates MLLM predictions with state-of-the-art blind VQA (BVQA) models, achieving superior generalization. Analysis of content-aware ensemble weights revealed video characteristics underrepresented by existing BVQA models, suggesting future research directions.
- Constructed an MLLM by integrating vision encoders with a large language model and developed a corresponding parameter-efficient fine-tuning pipeline.

ByteDance Multimedia Lab - Intern - Shenzhen, Guangdong, China   *Apr 2023 – May 2024*

- Developed a modular BVQA model featuring a base quality predictor, a spatial rectifier, and a temporal rectifier. These components address visual content/distortion, spatial resolution, and frame rate changes, respectively. Extensive experiments on professionally-generated and user-generated content video databases demonstrated superior or comparable performance to existing methods. The model's modularity also facilitates analysis of video quality databases regarding their spatial and temporal complexity.
- Developed a lightweight model to validate the necessity of combining traditional handcrafted image and video features for detecting how spatial and temporal downsampling affect video quality.

Alibaba Cloud - Intern - Hangzhou, Zhejiang, China   *Dec 2019 – Aug 2020*

- Solved practical industrial control problems by applying multi-objective optimization, resulting in one patent.

Guangfa Securities - Intern - Guangzhou, Guangdong, China   *Nov 2018 – May 2019*

- Developed recommendation algorithms by analyzing large-scale user data using Spark and SQL, leading to one patent.

SELECTED PUBLICATION

Under Review	**CP-LLM: Context and Pixel Aware Large Language Model for Video Quality Assessment** Wen Wen, Yaohong Wu, Yue Sheng, Neil Birkbeck, Balu Adsumilli and Yilin Wang
ICASSP 2025	**An Ensemble Approach to Short-form Video Quality Assessment Using Multimodal LLM** Wen Wen, Yilin Wang, Neil Birkbeck, and Balu Adsumilli
CVPR 2024	**Modular Blind Video Quality Assessment** Wen Wen, Mu Li, Yabin Zhang, Yiting Liao, Junlin Li, Li Zhang, and Kede Ma
CVPR 2024	**Learned Scanpaths Aid Blind Panoramic Video Quality Assessment** Kanglong Fan, Wen Wen, Mu Li, Yifan Peng, and Kede Ma
TPAMI 2024	**Analysis of Video Quality Datasets via Design of Minimalistic Video Quality Models** Wei Sun, Wen Wen, Xiongkuo Min, Long Lan, Guangtao Zhai, and Kede Ma
TCSVT 2024	**Perceptual Quality Assessment of Virtual Reality Videos in the Wild** Wen Wen, Mu Li, Yiru Yao, Xiangjie Sui, Yabin Zhang, Long Lan, Yuming Fang, and Kede Ma

OTHER

Skills	Python, PyTorch, TensorFlow, JAX, SQL
Service	Reviewer of top-tier conferences including NeurIPS, ICIP and ICME