Ziyue Xiang (Alan Xiang)

Senior Researcher @ Tencent Americas

Ph.D., Electrical and Computer Engineering, Purdue University, West Lafayette

Sunnyvale, California, USA
ziyue.alan.xiang@gmail.com

ziyue-alan-xiang
alanshawn.com
github.com/xziyue
UKO6iCgAAAAJ
0000-0001-6054-5801

Language: Mandarin, English, Cantonese
Pronoun: He/His

Last Update: May 12, 2026

Contents

Summary I
Education I
Experience I
Selected Projects I
Skills II
Publications II
Patents III
Membership and Certifications V
Professional Service V

Summary

I have experience in audio/video/image compression, signal processing, computer vision, and deep learning. I am currently working on developing next generation international video coding standards. My Ph.D. thesis is focused on digital media forensics, with topics including media metadata analysis, bitstream analysis, deepfake detection, and media manipulation detection.

Education

Ph.D. in Electrical and Computer Engineering *Aug. 2020–May. 2024*
Purdue University *West Lafayette, IN, USA*
- Advisor: Prof. Edward J. Delp

MSc. in Computer Science *Aug. 2018–May 2020*
Syracuse University *Syracuse, NY, USA*

BSc. in Information and Computing Science *Sept. 2014–June 2018*
Sun Yat-sen University *Guangzhou, Guangdong, China*
- Best undergraduate thesis award

Experience

Senior Researcher *Jan. 2024–Now*

Applied Research Intern *June 2023–Dec. 2023*
Tencent Americas *Palo Alto, CA, USA*
- Developing next generation international video coding standards beyond Versatile Video Coding (VVC)

Teaching Assistant *Jan. 2023–May 2023*
Purdue University *West Lafayette, IN, USA*
- Teaching assistant for *ECE-27920: Vertically Integrated Projects (VIP)* instructed by Prof. Edward J. Delp

Research Assistant *Aug. 2020–Dec. 2023*
Purdue University *West Lafayette, IN, USA*
- Using computer vision, signal processing, and machine learning techniques to conduct media forensics research

Research Assistant *Aug. 2018–July 2020*
Syracuse University *Syracuse, NY, USA*
- Developing algorithms to automatically analyze the integrity of scientific figures at Prof. Daniel E. Acuna's SOS+CD lab

Teaching Assistant *Jan. 2020–May 2020*
Syracuse University *Syracuse, NY, USA*
- Teaching assistant for *CIS-375: Discrete Mathematics* instructed by Prof. Andrew C. Lee

Selected Projects

Video Coding Standard Beyond Versatile Video Coding (VVC)
- Submitted JVET Proposals: JVET-AM0220, JVET-AO0199, JVET-AP0082
- Proposed Counter-based Temporal Probability Initialization (CTPI) that achieved coding gain over ECM-19.0

Contents

Summary I
Education I
Experience I
Selected Projects I
Skills II
Publications II
Patents III
Membership and Certifications V
Professional Service V

Analysis of MP4 Container Metadata Using Machine Learning

- In project [6], a custom MP4 container metadata parser was developed with the ability to export the tree structure and interpret vendor-specific data formats
- Employed feature engineering techniques to convert the MP4 metadata tree into feature vectors and used random forest classifiers to analyze the MP4 containers
- In a followup project [2], **Self-supervised Learning** strategies were used to train a **Graph Neural Network** for MP4 metadata analysis

Analysis of H.264 Bitstream Using Deep Learning

- In project [5], the **openh264** H.264 decoder was modified to extract encoding parameters such as macroblock type, motion vectors, quantization parameters from the bitstream
- Analyzed the H.264 encoding parameters with a **Vision Transformer**-based **Siamese Neural Network** to determine the capturing device of a given video bitstream

Analysis of MP3 Bitstream Using Deep Learning

- In project [4], the **minimp3** MP3 decoder was modified to extract MP3 encoding parameters such as MDCT coefficients, scalefactors, Huffman table indices from the bitstream
- The MP3 encoding parameters were analyzed by a **Convolution Neural Network+Transformer Neural Network** hybrid network to determine which parts of the MP3 audio are compressed more than once; the compression inconsistency can be used for audio manipulation detection

Synthetic (Deepfake) Audio Detection

- Modified MP3 decoder (**minimp3**) and AAC decoder (**fdk-aac**) to extract encoding parameters from audio bitstreams
- In project [13], the AAC scalefactors were used with **Transformer Neural Networks** for synthetic speech detection
- In project [3], the MP3 encoding parameters were used to compute the spectrogram efficiently, which was analyzed by popular neural network architectures such as **ResNet**, **EfficientNetV2**, **MobileNetV3** for synthetic speech detection
- In project [12], a **Variational Autoencoder** was used to extract **Disentangled Representation** for synthetic speech signals, which allows interpretable synthetic speech detection

Skills

- Programming languages: Python, C++
- Machine learning frameworks: TensorFlow, PyTorch, scikit-learn
- Image/audio/video processing algorithms
- The ability to modify and improve video/audio/image codec software such as *openh264*, *fdk-aac*, *ECM*, *VTM*, and *libav*
- Experience in computer graphics (OpenGL, ray tracing, simulation)
- Experience in compiler design (front-end, LLVM IR, optimization)
- Experience in computer vision (object detection, semantic segmentation, homography estimation, camera calibration, binocular vision, etc.)

Publications

[1] Multimedia Forensics Using Metadata
Z. Xiang
Ph.D. Thesis, Purdue University, 2024

[2] MTN: Forensic Analysis of MP4 Container Metadata Using Graph Neural Networks
Z. Xiang, A. K. S. Yadav, P. Bestagini, S. Tubaro, E. J. Delp
IEEE/CVF Computer Vision and Pattern Recognition Workshops (CVPRW), 2023

Contents

Summary I
Education I
Experience I
Selected Projects I
Skills II
Publications II
Patents III
Membership and Certifications V
Professional Service V

[3] Extracting Efficient Spectrograms From MP3 Compressed Speech Signals for Synthetic Speech Detection
Z. Xiang, A. K. S. Yadav, P. Bestagini, S. Tubaro, E. J. Delp
ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec), 2023

[4] Forensic Analysis and Localization of Multiply Compressed MP3 Audio Using Transformers
Z. Xiang, P. Bestagini, S. Tubaro, E. J. Delp
International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2022

[5] H4VDM: H.264 Video Device Matching
Z. Xiang, P. Bestagini, S. Tubaro, E. J. Delp
International Conference on Pattern Recognition Workshops (ICPRW), 2022

[6] Forensic Analysis of Video Files Using Metadata
Z. Xiang, J. Horváth, S. Baireddy, P. Bestagini, S. Tubaro, E. J. Delp
IEEE/CVF Computer Vision and Pattern Recognition Workshops (CVPRW), 2021

[7] Scientific Image Tampering Detection Based on Noise Inconsistencies: A Method and Datasets
Z. Xiang, D. E. Acuna
arXiv preprint arXiv:2001.07799, 2020

[8] Compression robust synthetic speech detection using patched spectrogram transformer
A. K. S. Yadav, **Z. Xiang**, K. Bhagtani, P. Bestagini, S. Tubaro, E. J. Delp
arXiv preprint arXiv:2402.14205, 2024

[9] Generation of synthetic echocardiograms using video diffusion models
A. O. Pellicer, A. K. S. Yadav, K. Bhagtani, **Z. Xiang**, Z. Pizlo, I. Gradus-Pizlo, E. J. Delp
IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), 2024

[10] PS3DT: Compression Robust Synthetic Speech Detection Transformer
A. K. S. Yadav, **Z. Xiang**, K. Bhagtani, P. Bestagini, S. Tubaro, E. J. Delp
IEEE International Conference on Machine Learning and Applications (ICMLA), 2023

[11] FGSSAT: unsupervised fine-grain attribution of unknown speech synthesizers using transformer networks
K. Bhagtani, A. K. S. Yadav, **Z. Xiang**, P. Bestagini, E. J. Delp
57th Asilomar Conference on Signals, Systems, and Computers, 2023

[12] DSVAE: Interpretable Disentangled Representation for Synthetic Speech Detection
A. K. S. Yadav, K. Bhagtani, **Z. Xiang**, P. Bestagini, S. Tubaro, E. J. Delp
IEEE International Conference on Machine Learning and Applications (ICMLA), 2023

[13] ASSD: Synthetic Speech Detection in the AAC Compressed Domain
A. K. S. Yadav, **Z. Xiang**, E. R. Bartusiak, P. Bestagini, S. Tubaro, E. J. Delp
International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2023

[14] An Overview of Recent Work in Multimedia Forensics
K. Bhagtani, A. K. S. Yadav, E. R. Bartusiak, **Z. Xiang**, R. Shao, S. Baireddy, E. J. Delp
International Conference on Multimedia Information Processing and Retrieval (MIPR), 2022

[15] Deepfake Detection Using Multiple Data Modalities
H. Hao, E. R. Bartusiak, D. Güera, D. M. Montserrat, S. Baireddy, **Z. Xiang**, S. K. Yarlagadda, R. Shao, J. Horváth, J. Yang
Handbook of Digital Face Manipulation and Detection: From DeepFakes to Morphing Attacks, 2022

[16] Sat U-net: a Fusion Based Method for Forensic Splicing Localization in Satellite Images
J. Horváth, **Z. Xiang**, E. D. Cannas, P. Bestagini, S. Tubaro, E. J. Delp
Multimodal Image Exploitation and Learning, 2022

[17] Estimating a Null Model of Scientific Image Reuse to Support Research Integrity Investigations
D. E. Acuna, **Z. Xiang**
arXiv preprint arXiv:2003.00878, 2020

Patents

(1) Blended prediction construction with adaptively derived weights
R. Chernyak, L.-F. Chen, **Z. Xiang**, S. Liu, Y. Wang, Y.-U. Yoon, B. Wang
US Patent App. 19/308,173, 2026

(2) Cross-component prediction in multi-partition prediction mode
L.-F. Chen, R. Chernyak, **Z. Xiang**, B. Wang, S. Wenger, Y. Wang
US Patent App. 19/313,726, 2026

(3) VQ based transform coefficient compensation
Y. Wang, R. Chernyak, L.-F. Chen, S. Liu, **Z. Xiang**, B. Wang, Y.-U. Yoon
US Patent App. 19/306,879, 2026

(4) Merge and block vector refinement for intra template matching prediction
Y.-U. Yoon, B. Wang, R. Chernyak, L.-F. Chen, M. Xu, **Z. Xiang**, S. Liu
US Patent 12,556,688, 2026

(5) Adaptive spatial scanning for non-separable transforms
Y. Wang, Y.-U. Yoon, L.-F. Chen, B. Wang, R. Chernyak, **Z. Xiang**, S. Liu
US Patent App. 19/294,171, 2026

Contents

Summary I
Education I
Experience I
Selected Projects I
Skills II
Publications II
Patents III
Membership and Certifications V
Professional Service V

(6) Method and apparatus for intra prediction mode
B. Wang, S. Wenger, R. Chernyak, Y.-U. Yoon, L.-F. Chen, M. Xu, **Z. Xiang**
US Patent App. 19/266,130, 2026

(7) Techniques for in-loop filtering
L.-F. Chen, B. Wang, S. Wenger, Y. Wang, Y.-U. Yoon, M. Xu, R. Chernyak, **Z. Xiang**
US Patent App. 19/266,145, 2026

(8) Decoder side gradient-based intra mode derivation on non-adjacent reference line for chroma components
B. Wang, R. Chernyak, L.-F. Chen, **Z. Xiang**, Y. Wang, M. Xu, S. Liu
US Patent App. 19/266,054, 2026

(9) Implicit control derivation method for linear and non-linear models in video compression
R. Chernyak, L.-F. Chen, S. Liu, M. Xu, B. Wang, Y.-U. Yoon, **Z. Xiang**
US Patent App. 19/301,913, 2025

(10) Decoder quantization shifting offset derivation
M. Xu, R. Chernyak, B. Wang, L.-F. Chen, Y.-U. Yoon, **Z. Xiang**, M. Tang
US Patent App. 19/301,915, 2025

(11) Signaling and filtering of template and flexible partition split
R. Chernyak, B. Wang, L.-F. Chen, **Z. Xiang**, S. Liu, Y. Wang, Y.-U. Yoon
US Patent App. 19/216,658, 2025

(12) Techniques for transform kernel set selection/derivation
B. Wang, Y.-U. Yoon, S. Liu, Y. Wang, L.-F. Chen, R. Chernyak, **Z. Xiang**
US Patent App. 19/212,522, 2025

(13) Multi-template decoder side intra mode derivation
Y.-U. Yoon, B. Wang, L.-F. Chen, R. Chernyak, Y. Wang, **Z. Xiang**, T. Liu
US Patent App. 19/182,446, 2025

(14) Chain-based motion vector and chain-based motion vector predictor derivation
L.-F. Chen, B. Wang, T. Liu, R. Chernyak, M. Xu, Y. Wang, **Z. Xiang**, S. Liu
US Patent App. 19/182,500, 2025

(15) Intra mode information based on non-conventional intra predictor
B. Wang, R. Chernyak, L.-F. Chen, Y.-U. Yoon, M. Xu, **Z. Xiang**, Y. Wang
US Patent App. 19/184,907, 2025

(16) Cross-component prediction and geometric partition weight adaption in geometric partition mode
L.-F. Chen, B. Wang, Y.-U. Yoon, R. Chernyak, M. Xu, Y. Wang, **Z. Xiang**
US Patent App. 19/182,375, 2025

(17) Selective transform set in video coding
B. Wang, R. Chernyak, L.-F. Chen, Y.-U. Yoon, M. Xu, **Z. Xiang**, Y. Wang
US Patent App. 19/175,818, 2025

(18) Intra predictor and intra mode coding
B. Wang, R. Chernyak, L.-F. Chen, Y.-U. Yoon, M. Xu, **Z. Xiang**, Y. Wang
US Patent App. 19/177,348, 2025

(19) Quantization compensation through learned inverse transform kernel for video coding
Y. Wang, R. Chernyak, M. Xu, **Z. Xiang**, B. Wang, T. Liu, L.-F. Chen, S. Liu
US Patent App. 19/175,854, 2025

(20) Decoder side quantization shifting offset prediction
M. Xu, R. Chernyak, L.-F. Chen, B. Wang, Y.-U. Yoon, **Z. Xiang**, S. Liu
US Patent App. 19/248,384, 2025

(21) Reconstruction of the transform coefficients in video coding
M. Xu, R. Chernyak, T. Liu, B. Wang, **Z. Xiang**, L.-F. Chen, Y.-U. Yoon, S. Liu
US Patent App. 19/172,097, 2025

(22) Adaptive clipping processes for video and image compression
R. Chernyak, **Z. Xiang**, S. Liu, B. Wang, M. Xu, L.-F. Chen, Y. Wang
US Patent App. 19/172,501, 2025

(23) Method and apparatus for advanced coefficients coding
M. Xu, Y.-U. Yoon, B. Wang, R. Chernyak, L.-F. Chen, S. Liu, **Z. Xiang**
US Patent App. 19/078,159, 2025

(24) Intra prediction based on transform domain equation
Z. Xiang, B. Wang, L.-F. Chen, Y.-U. Yoon, M. Xu, Y. Wang, R. Chernyak
US Patent App. 19/078,164, 2025

(25) Merge candidate construction
L.-F. Chen, B. Wang, Y.-U. Yoon, R. Chernyak, **Z. Xiang**, S. Liu, M. Xu
US Patent App. 19/070,333, 2025

(26) Transform kernel derivation in inter-coded block with intra prediction mode information
B. Wang, R. Chernyak, L.-F. Chen, Y.-U. Yoon, M. Xu, **Z. Xiang**, Y. Wang
US Patent App. 19/072,848, 2025

Contents

Summary I
Education I
Experience I
Selected Projects I
Skills II
Publications II
Patents III
Membership and Certifications V
Professional Service V

(27) Multi-quantizer method for image and video compression
M. Xu, R. Chernyak, L.-F. Chen, Y.-U. Yoon, B. Wang, S. Liu, **Z. Xiang**
US Patent App. 19/047,514, 2025

(28) Inverse pre-filter for image and video compression
R. Chernyak, L.-F. Chen, B. Wang, M. Xu, S. Liu, **Z. Xiang**, M. Tang
US Patent App. 19/048,596, 2025

(29) MPM candidate derivation improvement by using intra template-matching
B. Wang, L.-F. Chen, R. Chernyak, Y.-U. Yoon, **Z. Xiang**, M. Xu, S. Liu
US Patent App. 19/039,636, 2025

(30) Intra mode coding based on template
B. Wang, R. Chernyak, L.-F. Chen, Y.-U. Yoon, M. Xu, S. Liu, **Z. Xiang**
US Patent App. 19/045,257, 2025

(31) Video coding with training-based coding tool
B. Wang, R. Chernyak, L.-F. Chen, Y.-U. Yoon, X. Zhao, M. Xu, S. Liu, **Z. Xiang**
US Patent App. 19/023,018, 2025

(32) Adaptive template expansion
Y.-U. Yoon, B. Wang, **Z. Xiang**, L.-F. Chen, R. Chernyak, M. Xu, S. Liu
US Patent App. 18/923,652, 2025

(33) Adaptive clipping in models parameters derivations methods for video compression
R. Chernyak, B. Wang, L.-F. Chen, **Z. Xiang**, Y.-U. Yoon, M. Xu, S. Liu
US Patent App. 18/923,643, 2025

(34) Fine-grained intra prediction fusion
Z. Xiang, B. Wang, R. Chernyak, Y.-U. Yoon, L.-F. Chen, M. Xu, S. Liu
US Patent App. 18/923,634, 2025

(35) Direction-adaptive region-based prediction combination
Y.-U. Yoon, S. Liu, R. Chernyak, B. Wang, L.-F. Chen, M. Xu, X. Zhao, **Z. Xiang**
US Patent App. 18/923,624, 2025

Membership and Certifications

- Member of Institute of Electrical and Electronics Engineers (IEEE)
- Member of Moving Picture Experts Group (MPEG), ISO/IEC JTC 1/SC 29/WG 4
- Member of Joint Video Experts Team (JVET), ISO/IEC JTC 1/SC 29/WG 5

Professional Service

- Reviewer for *Journal of Visual Communication and Image Representation*
- Reviewer for *Conference on Vision and Intelligent Systems*
- Reviewer for *Journal of Testing and Evaluation*
- Reviewer for *Forensic Science International: Digital Investigation*
- Reviewer for *ACM Multimedia*
- Reviewer for *International Conference on Acoustics, Speech, & Signal Processing (ICASSP)*
- Reviewer for *Conference on Computer Vision and Pattern Recognition (CVPR)* Media Forensics Workshop
- Reviewer for *International Conference on Pattern Recognition (ICPR)*
- Reviewer for *Information Processing and Management*
- Reviwer for *IEEE Transactions on Very Large Scale Integration Systems (IEEE TVLSI)*
- Active member of the LaTeX developer community; authored LaTeX3 language tutorial that is recommended by the LaTeX project and has over 30,000 online views; contributed CTAN packages such as *luaprogtable* and *smart-eqn*