

Education

Ph.D. in Computer Science

November 2021 – August 2025 (est.)

Max-Planck-Institute for Informatics, Saarland University

Saarbrücken, Germany

- › *Dissertation*: “Synthesizing 3D Human Interactions in Multi-Entity Environments”
- › *Supervisors*: Prof. Dr.-Ing. Philipp Slusallek, Prof. Dr. Christian Theobalt

M.S. in Visual Computing

ECTS: 1.5 (\equiv 3.8/4.0 US GPA)

2018 – 2021

Saarland University

Saarbrücken, Germany

- › *Thesis*: “Synthesizing Compositional Animations from Textual Descriptions Using an RNN-based Hierarchical Two-Stream Network”
- › *Supervisor*: Prof. Dr.-Ing. Philipp Slusallek

M.Sc. (Hons.) in Computer Science

GPA: 8.31 (\equiv 3.8/4.0 US GPA)

2014 – 2016

St. Xavier’s College

Kolkata, India

- › *Thesis*: “Real-Time Sign Language Processing System using Perceptrons”
- › *Supervisor*: Dr. Asoke Nath

B.Sc. (Hons.) in Computer Science

GPA: 7.81 (\equiv 3.8/4.0 US GPA)

2011 – 2014

St. Xavier’s College

Kolkata, India

Professional Experience

AI Researcher

August 2021 – Present

German Research Center for AI (DFKI), Saarbrücken

Supervisor: Prof. Dr.-Ing. Philipp Slusallek

- › Working with generative AI models for multimodal understanding (LLM, VLM), audio processing, computer vision, 3D scene understanding.
- › Collaborating in multi-domain EU projects related to human-human and human-object actions and interactions for large-scale visual generation and understanding, including pedestrian simulations, motion understanding, 3D character-scene interactions via LLMs, and character interactions in VR.
- › Supervising Masters’ thesis on related projects in multi-person pose forecasting.

ML Research Intern

June 2024 – September 2024

Snap Research, New York City, NY, USA

Supervisors: Dr. Chuan Guo, Dr. Bing Zhou

- › Developed a multimodal generative AI method for two-person interactions based on audio and visual signals. [Keywords: Generative AI, Multimodal Content Understanding, Two-Person Motion Generation, Tokenization, Masked Transformer]

ML Research Intern

June 2023 – April 2024

Google Inc., Mountain View, CA, USA

Supervisors: Dr. Avisek Lahiri, Dr. Vivek Kwatra

- › Developed a deep learning generative model for facial expression synthesis from audio inputs, with particular attention on lip synchronization for automated visual dubbing. [Keywords: Speech to Facial Motion, Face Landmark Detection, Speech Processing, Audio Encoding, CNNs, RNNs, Transformers]

AI Research Assistant

May 2019 – July 2021

German Research Center for AI (DFKI), Saarbrücken

Supervisor: Prof. Dr.-Ing. Philipp Slusallek

- › Worked with the Agents and Simulated Reality group on natural human motion simulation based on multimodal cues, including natural language instructions and objects to interact with.

- › Developed one of the first deep learning-based text-to-motion synthesis models capable of generating motions for complex instructions, published in the premier International Conference on Computer Vision (A* tier) with 180+ citations.

Teaching Assistant (Image Processing & Computer Vision) April 2020 – September 2020
Saarland Informatics Campus, Germany *Professor: Dr. Joachim Weickert*

- › Delivered tutorials, graded assignments and exams for a graduate-level course on Image Processing and Computer Vision.

Visiting Scholar on Computer Vision February 2018 – July 2018
Indian Statistical Institute, Kolkata, India *Supervisor: Dr. Dipti Prasad Mukherjee*

- › Developed machine-learning-based computer vision techniques to estimate the constituents of coal samples from the microscopic petrographic images. [Keywords: *Decision Trees, Multi-class Classifiers, Support Vector Machine, Computational Imaging, Pattern Recognition*]

Signal Processing Research Associate March 2017 – November 2017
Indian Institute of Science, Bengaluru, India *Supervisor: Dr. Gurunath Gurralla*

- › Developed an adaptive algorithm for accurate, real-time estimation of grid signal frequencies at high sampling rates. [Keywords: *Fourier Transform, Frequency Estimation, Adaptive Bandpass Filter, Discrete-Time Energy Separation Algorithm*]

Professional Service

Invited Speaker

- › HiGraphics Workshop: Synthesizing 3D Human Interactions in Multi-entity Environment 2024
- › SIGGRAPH Frontiers Workshop: Expressive Avatar Interactions in Online Co-Experiences 2023

Conference Program Committee and Reviewer

- › AAAI Annual Conference on Artificial Intelligence (AAAI) 2024, 2023
- › ACM SIGGRAPH 2025, 2023
- › ACM SIGGRAPH Asia 2023
- › ECVA European Conference on Computer Vision (ECCV) 2024
- › Eurographics 2025
- › IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2025
- › IEEE International Conference on Virtual Reality and 3D User Interfaces 2023
- › Pacific Graphics 2023
- › UBICOMP International Symposium on Wearable Computers (ISWC) 2023

Journal Reviewer

- › Computer Graphics Forum 2025
- › IEEE Transactions on Multimedia 2025

Conference Volunteer

- › IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC) 2017

Achievements and Awards

- › **Research Grant from Snap Inc. 2024-25.** Acquired research grant for projects on multimodal Generative AI incorporating audio and visual signals for multi-person interactions.
- › **Best Master's Project Award 2016.** From St. Xavier's College, Kolkata, India, for the project "Real-Time Sign Language Processing System".
- › **Certificate of Merit 2013.** From St. Xavier's College, Kolkata, India, for securing the first position in semesters III and IV of Bachelor of Science in Computer Science.

- **Vidyasagar Scholarship 2009.** From the Government of West Bengal, India, for securing second rank among girls in Kolkata in the statewide 10th standard examination.

Skills

Machine Learning	PyTorch, PyTorch-3D, Torchvision, Tensorflow, TensorBoard, NumPy, Pandas, Keras, SciPy, Scikit-learn, NLTK
Generative AI	MLP, RNN, CNN, GAN, cVAE, vector-quantization, diffusion, transformer, tokenization, LLM, VLM, GPT, data-collection, data annotation
Programming	Python, Java, C, C++, Jax, Powershell, MATLAB, SQL
Softwares	Visual Studio, Motion capturing, Containerization (Kubernetes, Docker), \LaTeX , Github, MS-Office (Word, PowerPoint, Excel)
Graphics	OpenGL, WebGL, Unity, Blender, Matplotlib, Seaborn, OpenCV, Open3D
Web Design	Angular JS, HTML5, Markdown, CSS, JavaScript
Languages	Bengali (native), English (fluent), Hindi (fluent), German (beginner)

Scientific Publications

- [11] **Ghosh, A.**, Zhou, B., Dabral, R., Wang, J., Golyanik, V., Theobalt, C., Slusallek, P., Guo, C. “DuetGen: Music Driven Two-Person Dance Generation via Hierarchical Masked Modeling”. SIGGRAPH 2025.
- [10] [Under Review] Sui, K., **Ghosh, A.**, Hwang, I., Wang, J., Guo, C. “A Survey on Human Interaction Motion Generation”. arXiv preprint:2503.12763, March 2025.
- [9] Li, Z., Cheng, K., **Ghosh, A.**, Bhattacharya, U., Gui, L., Bera, A. “SimMotionEdit: Text-Based Human Motion Editing with Motion Similarity Prediction”. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2025.
- [8] **Ghosh, A.**, Dabral, R., Golyanik, V., Theobalt, C., Slusallek, P. “ReMoS: 3D Motion-Conditioned Reaction Synthesis for Two-Person Interactions”. European Conference on Computer Vision (ECCV), pp 418–437. 2024. [Poster]
- [7] **Ghosh, A.**, Dabral, R., Golyanik, V., Theobalt, C., Slusallek, P. “IMoS: Intent-Driven Full-Body Motion Synthesis for Human-Object Interactions”. EUROGRAPHICS Computer Graphics Forum, Vol. 42, No. 2, pp. 1-12. 2023. [Oral]
- [6] **Ghosh, A.**, Cheema, N., Oguz, C., Theobalt, C., Slusallek, P. “Synthesis of Compositional Animations from Textual Descriptions”. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp 1396-1406. 2021. [Poster]
- [5] **Ghosh, A.**, Cheema, N., Oguz, C., Theobalt, C., Slusallek, P. “Text-Based Motion Synthesis with a Hierarchical Two-Stream RNN”, ACM SIGGRAPH 2021 Posters, Article No. 42, pp. 1-2. 2021.
- [4] Bansal, S., **Ghosh, A.**, Seelamantula, C. S., Gurralla, G., Ghosh, P. K. “Adaptive Frequency Estimation Using Iterative DESA with RDFT-based Filter”. IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), pp. 1-6. IEEE. 2017. [Oral]
- [3] Dasgupta, A., Seth, D., **Ghosh, A.**, Nath, A. “Real Time Sign Language to Braille Interfacing System”. 7th International Conference on Communication Systems and Network Technologies (CSNT), pp. 371-375. IEEE. 2017. [Oral]
- [2] Agarwal, S., **Ghosh, A.**, Nath, A. “Green Enterprise Computing - Approaches Towards A Greener IT”. International Journal of Innovative Research in Advanced Engineering (IJIRAE), pp. 1-5. 2016.

- [1] Seth, D., **Ghosh, A.**, Dasgupta, A., Nath, A. “Real Time Sign Language Processing System”. International Conference on Smart Trends in Information Technology and Computer Communications (SmartCom), pp. 11-18. Springer Singapore. 2016. *[Oral]*