

Muhammad Hamza Mughal

Curriculum Vitae

✉ mmughal@mpi-inf.mpg.de
🌐 m-hamza-mughal.github.io

Education

- since 2024 **Ph.D.**,
Max-Planck Institut für Informatik & Universität des Saarlandes, Germany
Advisor: Prof. Christian Theobalt & Prof. Vera Demberg
Member of RTG 2853 "Neuroexplicit Models for Language, Vision and Action"
- 2021–2023 **M.Sc. Visual Computing**,
Universität des Saarlandes, Germany
Advisor: Prof. Christian Theobalt
- 2015–2019 **B.E. Electrical Engineering**,
National University of Sciences and Technology, Pakistan
Advisor: Asst. Prof. Muhammad Shahzad

Publications

- *Retrieving Semantics from the Deep: an RAG Solution for Gesture Synthesis*
M. Hamza Mughal, Rishabh Dabral, Merel C.J. Scholman, Vera Demberg, Christian Theobalt
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2025.
- *ConvoFusion: Multi-Modal Conversational Diffusion for Co-Speech Gesture Synthesis*
Muhammad Hamza Mughal, Rishabh Dabral, Ikhsanul Habibie, Lucia Donatelli, Marc Habermann, Christian Theobalt
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2024.
- *MoFusion: A Framework for Denoising-Diffusion-based Motion Synthesis*
Rishabh Dabral, **Muhammad Hamza Mughal**, Vladislav Golyanik, Christian Theobalt
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2023.
(Highlight - top 10% of selected papers)
- *Assisting UAV Localization Via Deep Contextual Image Matching*
Muhammad Hamza Mughal, Muhammad Jawad Khokhar, Muhammad Shahzad
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol.14, 2021.

Experience

- Oct 2022 – **Research Assistant**, *Max Planck Institute for Informatics, Saarbrücken*
- Dec 2023 *Visual Computing and Artificial Intelligence Department (D6)*
Researching novel methods of Multi-modal Motion and Gesture synthesis. I contributed to Motion Synthesis Framework which used text and audio conditionings to synthesize realistic motions for text-to-motion & music-to-dance synthesis task. Supervised by Dr. Rishabh Dabral & Prof. Christian Theobalt
- Aug 2021 – **Visiting Researcher (DAAD Scholar)**, *Technical University of Munich*
- Nov 2021 *Chair of Data Science in Earth Observation, TUM Germany*
Conducted research on Multi-modal Representation Learning using Self-supervised Learning (SSL) methods including non-contrastive and contrastive optimization procedures. Supervised by Prof. Muhammad Shahzad & Prof. Xiaoxiang Zhu

- Sep 2020 – **Machine Learning Lead**, *Scribe Audio Inc.*
 Aug 2021 Islamabad, Pakistan
 Responsibilities included modifying, implementing and training text-to-speech algorithms like GST-Tacotron, TransformerTTS, TortoiseTTS & GradTTS to achieve control over emotion in Expressive Speech Synthesis. My tasks also involved training and improving neural vocoders like WaveRNN, DiffWave & HiFiGAN for robust melspectrogram to sound wave conversion.
- Sep 2019 – **Computer Vision Engineer**, *VisionX LLC*
 Sep 2020 Islamabad, Pakistan
 Designed and implemented algorithms for 3D Pose Estimation using PoseNet on 2D images to enable Motion Capture in Augmented Reality. Moreover, my work mainly included designing, implementing and deploying custom deep learning models for tasks like Human Attributes Classification and Human Activity Recognition for Retail Analytics.
- Jun 2019 – **Computer Vision Intern**, *VisionX LLC*
 Sep 2019 Islamabad, Pakistan
 Tested and deployed 2D & 3D facial landmark detection algorithms like 3D-FAN and devised algorithm for face pose detection in 3D world coordinate system. Moreover, I also worked on imposing 3d models of objects on 3D facial landmarks.
- Sep 2018 – **Research Assistant**, *National University of Sciences and Technology*
 Sep 2019 Machine Vision & Intelligent Systems Lab, NUST, Islamabad
 Worked on researching and implementing classical computer vision and deep learning based methods for UAV localization, which was published in IEEE-JSTARS. Supervised by Prof. Muhammad Shahzad.

Technical Skills

| | | | |
|--------------------------|--|-------------------|--|
| Languages | Python, C++, SQL | Frameworks | Pytorch, Tensorflow, Kubernetes |
| Augmented Reality | Apple ARKit - RealityKit | IT | Managed GPU Servers (Docker): NVIDIA HGX/DGX A100 |
| Cloud | AWS: SageMaker, Kinesis, Lambda, Fargate; GCP: BigQuery, Firestore; Azure: Azure Kubernetes Cluster, Functions | | |

Languages

| | |
|----------------|----------------|
| English | TOEFL: 110/120 |
| German | A1.2 |