

GHAZANFAR ALI, PH.D.

ASSISTANT PROFESSOR | EMBODIED AI, MULTIMODAL INTERACTION AND DIGITAL HUMANS

End-to-end research leadership across datasets, models, evaluation, real-time systems, human studies and publication

Seoul, South Korea | ghazanfar.6785@gmail.com | [LinkedIn](#) | [ORCID](#) | [Google Scholar](#)

ACADEMIC PROFILE

Assistant Professor in the Department of AI at Gachon University and an end-to-end AI systems researcher working at the intersection of embodied intelligence, human motion, multimodal interaction, computer vision, and HCI. Research contributions span co-speech gesture and facial animation, controllable physical personas, live LLM agents, motion tokenization and video-to-motion ingestion, grounded RAG agents, diffusion-based style transfer, and AR/MR interaction. Record includes 17 journal, conference, poster, and live-demonstration publications, two research awards, and systems validated through controlled benchmarks and human studies with up to 113 participants.

Research themes: Embodied and multimodal AI; digital humans and conversational agents; human motion modeling; co-speech gesture and facial behavior; controllable personas; efficient retrieval and contrastive learning; RAG and LLM agents; computer vision; AR/VR/MR; human-centered evaluation.

ACADEMIC APPOINTMENTS

Assistant Professor

Mar 2026 - Present

Department of AI, Gachon University | South Korea

- Lead research on live conversational agents, scalable monocular-video motion estimation, and persona modeling for LLM-based embodied agents.
- Direct two PhD students and own research delivery from concept and architecture through datasets, implementation, evaluation, demonstration, user study, and paper development.
- Designed and implemented a real-time LiveAgent platform integrating a custom LLM chatbot, ASR, TTS, gesture generation, and facial/body animation across Unity, desktop, and web clients.

InnoCORE Postdoctoral Researcher

Sep 2025 - Feb 2026

LLM Center, Korea Advanced Institute of Science and Technology (KAIST) | South Korea

- Developed the major components of ControllableGesture, a text-only framework for authoring persistent physical personas through explicit gesture and affect controls.
- Built the data/control pipeline and shared retrieval-generation architecture; subsequent refinements are preparing the system for an expert HCI study.

Postdoctoral Researcher

Sep 2023 - Aug 2025

Korea Institute of Science and Technology (KIST) | South Korea

- Led research on efficient, personalized co-speech gesture generation, grounded embodied agents, multimodal behavior synthesis, and interactive digital-human systems.
- Originated ConGRets and co-developed RIDGE, the Joseon Dynasty RAG agent, lightweight upper-face animation, and related evaluation frameworks.
- Translated research models into real-time Unity demonstrations and publication-grade benchmark and user-study pipelines.

Research Assistant

Sep 2017 - Aug 2023

Korea Institute of Science and Technology (KIST) | South Korea

- Built language-driven animation systems spanning data curation, representation learning, motion retrieval, real-time character integration, and human-subject evaluation.
- Developed embodied-agent applications for mixed reality, healthcare communication, automated previsualization, no-code authoring, and multilingual interaction.
- Mentored approximately seven graduate researchers and interns as a senior PhD researcher.

EDUCATION

Ph.D. in AI-Robotics

Sep 2017 - Aug 2023

KIST School, University of Science and Technology (UST), South Korea

Dissertation: Scalable Hybrid Approach of Co-speech Text-to-Gesture Generation for Interactive Digital Humans. Advisor: Dr. Jae-In Hwang.

B.E. in Computer Engineering

Sep 2009 - Jan 2014

National University of Sciences and Technology (NUST), Pakistan

Final-year project: Design and development of a driving simulator for autonomous vehicles.

CURRENT RESEARCH LEADERSHIP AND SYSTEMS

LiveAgent: Real-Time Conversational Digital Human Platform

Research lead | Working demo

A reusable embodied-agent platform for demonstrations, downstream research, and large-scale HCI studies.

- Architected the full stack: Python communication and algorithm server; socket/HTTP interfaces; custom LLM chatbot; API-based ASR/TTS; gesture, face, and body animation; Unity, desktop, and Node.js/Three.js clients; session management and logging.
- Streams and parallelizes LLM, TTS, and gesture work with server/client queues; generates each two-second gesture segment in under 50 ms and reaches approximately two-second first-word latency with a fast LLM.

MotionBridge and MotionStage

Originator, architect and primary developer

A paired research program for accessible video-to-motion ingestion and synthetic motion-data generation on consumer hardware.

- MotionBridge maps monocular video observations into a frozen body/hand/root motion-token vocabulary for clean, complete, retargetable 4D motion without expensive per-frame geometry optimization.
- MotionStage is the companion synthetic-data specification: arbitrary skeletal motion to video plus exact 3D/2D joints, camera parameters, occlusion labels, masks, and coherent finger/body motion.
- Completed a three-pass review of 368 works and framed evaluation around throughput, usable motion-hour cost, yield, labor, and consumer-GPU accessibility.

Motion Evaluation Suite and GQNet

Lead developer | In development

A dual-purpose evaluation program for co-speech gesture and text-to-motion/action generation.

- Designed a seven-axis suite spanning realism, synchrony, semantic appropriateness, diversity, reference accuracy, physical artifacts, and efficiency, with approximately 35 metrics and human-correlation meta-evaluation.
- GQNet v1 learns quality sensitivity from injected noise/artifacts; Arm 2 learns human-perceived quality directly from unmodified generated motion and human ratings.
- The resource benchmark evaluates seven popular methods and exposes the latency and memory costs of recent motion-generation pipelines.

Behavioral Imposter Test (BIT)

Study design lead | IRB under review

An embodied-persona fidelity instrument that measures how quickly users detect a mismatch in lexical style, knowledge, voice, or gesture behavior during live interrogation.

- Delivered the LiveAgent infrastructure and designed the controlled-persona protocol: eight matched personas and 48 single-channel controlled builds.
- A PhD student is adapting the platform for a crowdsourced study; preparation and IRB review are in progress.

ControllableGesture

Major components implemented | Expert study preparation

Text-only controllable co-speech behavior for persistent physical personas, including handedness, gesture amplitude/speed, and excited or relaxed styles.

- Implemented major model components and a control store spanning 99,022 two-second units, 55 hours, and 30 speakers; expert HCI participants will author personas and evaluate adherence during free-form conversation.

SELECTED ESTABLISHED RESEARCH CONTRIBUTIONS

ConGRets: Scalable, Low-Latency Co-Speech Gesture Generation

Originator and first author | Under review

A text-only contrastive retrieval system with zero-shot motion-style adaptation and fair decoded-frame evaluation against generative methods.

- Processes 1,000 text queries in under 500 ms on CPU (approximately 0.5 ms/query) with a memory footprint below 100 MB; achieved FGD 84.62, 99% speaker-style accuracy, and 89% content-to-cluster accuracy.
- Uses a global motion-only speaker encoder, BERT-mini text backbone, shared latent space, curated gesture units, overlap-aware segmentation, and runtime blending.

RIDGE: Rule-Infused Deep Learning

Lead author | Published 2025

A hybrid LLM-assisted rule and deep-learning architecture for semantically appropriate co-speech gesture generation.

- Improved Gesture Cluster Affinity to 0.73 versus 0.61 and 0.52 baselines; LLM-extracted rules reached approximately 80% overlap with expert annotation.
- Introduced the GCA semantic-coherence metric and trained across approximately 300 hours of in-the-wild video plus 70 hours of motion capture.

Multilingual Gesture Synthesis and GestureCLR

Lead research contribution

A cross-lingual, real-time text-to-gesture pipeline combining contrastive pose matching, motion clustering, translation, word timing, and Unity playback.

- GestureCLR reached 97% accuracy on high-noise 2D poses versus 18% for a cosine baseline; expanded the motion library from 56 to 2,035 units and the rule base from 84,000 to 210,000 mappings.
- A 51-participant study found improvements in naturalness, human-likeness, temporal and semantic consistency, and social presence without degradation from translation.

Grounded Embodied Agents for Historical and Medical Communication System design, implementation and evaluation
RAG-backed interactive agents that combine evidence-grounded responses with voice, lip synchronization, expressions, and gesture.

- Joseon Dynasty agent: 49.6 million-character corpus; benchmark scores of 88.7 factual accuracy, 92.2 reliability, and 91.1 reasonableness; 23-50 point gains over comparison systems.
- Medical agent: improved answer F1 from 0.492 to 0.779; supported 48 expressions and six emotions; 113-participant study found greater social presence than text/audio slides.

Visual Computing, AR/MR and Automated Previsualization

Research systems and publications

End-to-end systems spanning screenplay understanding, real-time 3D/VR generation, mobile AR perception, and wearable mixed-reality agents.

- ASAP generated storyboards, 3D previz, and VR scenes from screenplays; text-to-action accuracy reached 93% on simple and 87% on complex sentences.
- Silhouette Mesh used a 2.4M-parameter mobile segmentation model (0.971 mIoU) and improved spatial awareness, presence, and naturalness in a 24-participant study.

MANUSCRIPTS AND RESEARCH IN PROGRESS

Under review: ConGRets: Contrastive Learning for Scalable, Low-latency Co-speech Gesture Generation with Zero-shot Speaker Style Adaptation. First author and research lead.

Full draft, internal review: What Does a Motion-Hour Cost? Video-to-Motion Ingestion, Motion Tokenisation, and the Unreported Economics of Motion Corpora. First author.

Planned after study completion: ControllableGesture expert-study paper; Behavioral Imposter Test study; GQNet and human-validated motion-evaluation suite.

SPONSORED AND INSTITUTIONAL RESEARCH PROGRAMS

Interactive Virtual Avatar Technology Development

KIST Flagship Research Program | Jan-Dec 2024

Led the gesticulation subsystem for an end-to-end lifelike digital-human platform.

Development of Text-based Character Animation Synthesis System

Korea Creative Content Agency (KOCCA) | Jun 2021-Dec 2022

Curated motion-text data and developed deep learning and retrieval modules for language-driven gesture, action, and facial behavior.

Non-face-to-face Monitoring Technology for Medical Staff

NST Korea | Jan 2021-Dec 2023

Developed embodied conversational interfaces within a multi-group telehealth research program.

Digital Experience of Botanical Gardens and Zoos through Augmented Reality

KOCCA | Sep 2017-Dec 2020

Developed HoloLens-based agents integrating spatial mapping, gaze, speech, object recognition, cloud AI, and expressive animation.

TEACHING, MENTORING AND RESEARCH LEADERSHIP

- Current research leadership: direct technical leadership for two PhD students working on LiveAgent downstream studies and evaluations.
- Graduate mentoring: guided approximately seven graduate researchers and interns while serving as a senior PhD researcher at KIST.
- Lab Engineer, University of Central Punjab (Feb 2014-Jul 2017): planned and supervised hands-on computer-programming laboratories serving approximately 600 undergraduate students.

HONORS, ACADEMIC SERVICE AND DEMONSTRATIONS

Honors and Awards

- Best Paper Award, 37th International Conference on Computer Animation and Social Agents (CASA), 2024.
- Academic Excellence Award, UST-KIST School, 2023.

Academic Service

- Journal reviewer: Cognitive Computation; Scientific Reports; Discover Computing; Discover Artificial Intelligence.
- Program Committee Member: 8th International Workshop on Robotics Software Engineering (RoSE'26).

Selected Demonstrations

- Presenter, ASAP, SIGGRAPH Asia 2023 Real-Time Live!; research demonstrations for KOCCA Korea and additional institutional venues.

TECHNICAL EXPERTISE

AI and machine learning: Python, PyTorch, PyTorch Lightning, transformers, BERT/Sentence-BERT, contrastive learning, retrieval systems, RAG, LLM integration, diffusion models, clustering, representation learning, benchmarking.

Motion, vision and graphics: motion capture processing, motion tokenization, 2D pose estimation, OpenPose, facial/body animation, blendshapes, gesture synthesis, animation blending, Unity3D, Blender, ARCore, Microsoft HoloLens.

Interactive systems and HCI: Node.js, Three.js, JavaScript, Python servers, sockets, HTTP, ASR/TTS APIs, session/logging infrastructure, real-time deployment, controlled user studies, crowdsourcing, statistical evaluation.

PEER-REVIEWED PUBLICATIONS

Journal Articles

2026. Hwang Youn Kim, **Ghazanfar Ali**, Jeongha Lee, and Jae-In Hwang. "Lightweight Speech-Conditioned Upper-Face Animation for Virtual Agents via Emotion-Liveness Composition." *IEEE Access*, 2026.
2025. **Ghazanfar Ali**, Woojoo Kim, Muhammad Shahid Anwar, Jae-In Hwang, and Ahyoung Choi. "Expanding Multilingual Co-Speech Interaction: The Impact of Enhanced Gesture Units in Text-to-Gesture Synthesis for Digital Humans." *IEEE Access*, 13, 145144-145157.
2025. **Ghazanfar Ali**, Hwang Youn Kim, and Jae-In Hwang. "RIDGE: Rule-Infused Deep Learning for Realistic Co-Speech Gesture Generation." *Computer Animation and Virtual Worlds*, 36(4), e70034.
2025. Jeong Ha Lee, **Ghazanfar Ali**, and Jae-In Hwang. "A Retrieval-Augmented Generation System for Accurate and Contextual Historical Analysis: AI-Agent for the Annals of the Joseon Dynasty." *Computer Animation and Virtual Worlds*, 36(4), e70048.
2025. Hanseob Kim, **Ghazanfar Ali**, Bin Han, Hwang Youn Kim, Jieun Kim, Hyemin Shin, Gerard Jounghyun Kim, and Jae-In Hwang. "ASAP for Multi-Outputs: Auto-generating Storyboard and Pre-visualization with Virtual Actors Based on Screenplay." *Multimedia Tools and Applications*, 84(20), 22377-22400.
2024. Hwang Youn Kim, **Ghazanfar Ali**, and Jae-In Hwang. "Enhancing Doctor-Patient Communication in Surgical Explanations: Designing Effective Facial Expressions and Gestures for Animated Physician Characters." *Computer Animation and Virtual Worlds*, 35(3), e2236. Best Paper Award, CASA 2024.
2021. Hanseob Kim, **Ghazanfar Ali**, Andreas Pastor, Myungho Lee, Gerard J. Kim, and Jae-In Hwang. "Silhouettes from Real Objects Enable Realistic Interactions with a Virtual Human in Mobile Augmented Reality." *Applied Sciences*, 11(6), 2763.
2020. **Ghazanfar Ali**, Myungho Lee, and Jae-In Hwang. "Automatic Text-to-Gesture Rule Generation for Embodied Conversational Agents." *Computer Animation and Virtual Worlds*, 31(4-5), e1944.

Conference Papers, Posters, and Live Demonstrations

2026. Jeongha Lee, Yujin Kim, **Ghazanfar Ali**, Suhyun Kim, and Jae-In Hwang. "Through Van Gogh's Eyes: Global Style Transfer with Diffusion Model." *European Conference on Computer Vision (ECCV)*, 2026.
2025. Jeongha Lee, **Ghazanfar Ali**, and Jae-In Hwang. "RAG-based AI-Agent for Contextualized Analysis of High-Density Historical Records: Application to the Annals of the Joseon Dynasty." *IEEE ISMAR-Adjunct*, 893-894.
2025. Hwang Youn Kim, **Ghazanfar Ali**, and Jae-In Hwang. "LUFA: Lightweight Upper-Face Animation for VR/MR Avatars." *IEEE ISMAR-Adjunct*, 841-842.
2022. **Ghazanfar Ali**, and Jae-In Hwang. "Improving Co-Speech Gesture Rule-Map Generation via Wild Pose Matching with Gesture Units." *SIGGRAPH Asia Posters*, 1-2.
2022. Hanseob Kim, **Ghazanfar Ali**, Bin Han, Hwang Youn Kim, Jieun Kim, and Jae-In Hwang. "ASAP: Auto-generating Storyboard and Previz." *SIGGRAPH Asia Real-Time Live!*, 1.
2022. Hanseob Kim, Jieun Kim, **Ghazanfar Ali**, and Jae-In Hwang. "No-Code Digital Human for Conversational Behavior." *SIGGRAPH Asia Posters*, 1-2.
2021. Hanseob Kim, **Ghazanfar Ali**, and Jae-In Hwang. "ASAP: Auto-generating Storyboard and Previz with Virtual Humans." *IEEE ISMAR-Adjunct*, 316-320.
2021. Hanseob Kim, **Ghazanfar Ali**, Seungwon Kim, Gerard J. Kim, and Jae-In Hwang. "Auto-generating Virtual Human Behavior by Understanding User Contexts." *IEEE VR Abstracts and Workshops*, 591-592.
2019. **Ghazanfar Ali**, Hong-Quan Le, Junho Kim, Seung-Won Hwang, and Jae-In Hwang. "Design of Seamless Multi-modal Interaction Framework for Intelligent Virtual Agents in Wearable Mixed Reality Environment." *International Conference on Computer Animation and Social Agents (CASA)*, 47-52.

ADDITIONAL PROFESSIONAL LEADERSHIP

Web Development Advisor and Team Lead

Jul 2015 - Present

Penumbra Digital | Remote

- Lead web-development delivery from requirements and design collaboration through implementation, deployment, and post-launch services.
- Oversee WordPress and project-specific JavaScript stacks for confidential B2B, corporate, NGO, and technology clients across Pakistan, the Middle East, and the United States.

REFERENCES

Available upon request.