# Kwang Bin Lee

Daejeon, South Korea | klee166@kaist.ac.kr | kwangaji.github.io

## Education

**KAIST, Graduate School of Culture Technology**, Daejeon, South Korea 2023 – Present
M.S. in Culture Technology (Advisor: Sung Hee Lee)

- Master's Thesis: *Scene-agnostic Hierarchical Bimanual Task Planning via Visual Affordance Reasoning*
- GPA: 4.2 / 4.3

**Johns Hopkins University**, Baltimore, MD, USA 2017 – 2019
M.S. in Computer Science (coursework nearly completed; not conferred due to family circumstances)

- GPA: 3.64 / 4.0

**Johns Hopkins University**, Baltimore, MD, USA 2012 – 2017
B.S. in Computer Science with Honors (Advisor: Vladimir Braverman)

- GPA: 3.52 / 4.0

## Research Experience

**Graduate Researcher (Motion Team)**, KAIST, Daejeon, South Korea 2023 – Present

- Contribute to the Telepresence & Motion research team, spanning text understanding, motion processing, motion generation, and spatial reasoning.
- Lead development of Unity-based telepresence infrastructure, including motion data processing (BVH I/O, playback, retargeting) and planning pipelines integrated with LLMs and retrieval-augmented generation (RAG).
- Support team research initiatives by co-authoring technical sections for funded grant proposals on dance motion generation and embodied virtual agent frameworks.

**Research Intern** Johns Hopkins University, Baltimore, MD, USA 2017 - 2018

- Developed a Unity-based virtual block-building game to evaluate infants' spatial reasoning, replicating a study from the Language and Cognition Lab at Johns Hopkins University (advised by Dr. Anand Malpani).
- Implemented an XML-based logging system for structured behavioral data capture and downstream analysis.

**Summer Intern**, KAIST, Daejeon, South Korea 2015

- Reconstructed and post-processed Kinect-scanned 3D scenes for research use, applying MeshLab to fill holes and correct mesh artifacts.
- Enhanced reconstructed scenes by adding geometric properties and metadata within Unity.

## Publications

**Scene-Agnostic Hierarchical Bimanual Task Planning via Visual Affordance Reasoning** Under Review

***Kwang Bin Lee***, Jiho Kang, Sung-Hee Lee

- As first author, developed a unified framework that translates high-level task instructions into coordinated, affordance-aligned bimanual manipulation sequences across unseen 3D environments.
- Designed three modules—Visual Point Grounding, Bimanual Subgoal Planning, and Interaction-Point–Driven Bimanual Prompting—that integrate affordance cues, spatial layout, and hand-state constraints to structure subgoals and generate valid two-handed action sequences.

**Situated Embodied XR Agents via Spatial Reasoning and Prompting** IEEE ISMAR 2025 Demo

Jihun Shin, Hyeonjin Kim, Eunseong Lee, DongHwan Shin, ***Kwang Bin Lee***,
Taehei Kim, Hyeshim Kim, Joonsik An, Sung-Hee Lee

- Built an XR agent prototype that combines spatial reasoning with language and motion planning for context-aware interaction in a walkable augmented environment.
- Integrated a GPT-4o real-time API server with Unity C#, enabling dialogue and motion generation grounded in

scene metadata.

- Designed prompt engineering pipelines using JSON-encoded scene context to align dialogue and motion outputs for consistent agent behavior.

**Anonymous Submission** Under Review

- As a co-author, built an auto-annotation pipeline for the BABEL motion dataset, converting high-level motion labels into body-part-specific text descriptors for text-guided motion editing.

**Real-time Target-aware Part-wise Translation of Upper-body Gestures to Virtual Avatars in Dissimilar Environments** In Preparation

Jiho Kang, Jihun Shin, ***Kwang Bin Lee***, Sung-Hee Lee

- Built a Unity-based motion capture system to synchronously record gaze, hand, and finger movements.
- Developed an IK-based motion refinement pipeline for an MR real-time telepresence system.
- Contributed to research discussions and writing on deep neural network integration for interactive systems.

## Projects

**Pioneer Research Center Project: Embodied AI Agents for Temporal–Spatial Interaction (STEAM Research Grant)** 2024 – Present

- Contributed as a researcher to a nationally funded project, *Temporal–Spatial Interaction*, focused on developing embodied agents capable of recording and replaying actions across time and space to enable telepresence across locations; a collaboration among four KAIST labs and one Yonsei University lab.
- Designed a motion framework for the *Temporal* team, enabling semantically aligned responses to user interactions in mixed reality.
- Helped set up a motion capture system to collect full-body CPR motion data, processed BVH files, and implemented a recording module to construct a motion capture dataset supporting experimental studies and real-time playback by integrating an OptiTrack interface with Unity.
- Integrated a LangChain-based Python server with Unity for real-time motion streaming and agent control.

**Lyric Transcription in Noisy Environments** 2024

- Designed a lyric transcription model for real-world settings (e.g., café or background music) by fine-tuning the Whisper-medium model on a noise-augmented JamendoLyrics dataset with MUSAN noise samples, improving robustness in low-SNR conditions.
- Developed and deployed a fine-tuning framework with integrated dataset preprocessing and evaluation pipelines, ensuring reproducible training and inference.

**Reflexion-based Multi-Agent System for Game Accessibility in Older Adults** 2025

- Designed a Reflexion-based multi-agent system to adapt FPS gameplay for older adults in real time, using a training-free reinforcement approach to address sparse gameplay data.
- Conducted evaluation sessions with ten older adults, analyzing accessibility outcomes in performance, comfort, and engagement.

## Work Experience

**Administrative Assistant**, Mapo Tax Office (Military Service) 2020 – 2022

- Completed mandatory military service by serving as an administrative assistant at the Mapo District Tax Office.
- Supported daily operations including document processing, tax filing assistance, and data entry.
- Assisted staff and citizens with administrative inquiries, ensuring accurate and timely service delivery.

## Teaching Experience

**Teaching Assistant** Spring 2025

KAIST, Graduate School of Culture Technology

- Natural Language Processing

**Course Assistant** 2014, 2017–2018

Johns Hopkins University

- Computer Vision, Randomized and Big Data Algorithms, Introduction to Algorithms, Intro to Java

## Certificates

**Coursera: Deep Learning Specialization (Andrew Ng)** 2023

- Completed a five-course specialization covering neural networks, CNNs, sequence models, and deep learning best practices.
- Gained hands-on experience with TensorFlow for model building and optimization.

**Upstage: Large Language Models Certificate** 2024

- Studied LLM fundamentals (architectures, tokenization, transformer internals) and applied parameter-efficient fine-tuning with LoRA.
- Built a medical RAG platform (LangChain + Oracle DB) that delivered region-specific clinical insights (e.g., mosquito-borne risks, prescribing trends); project selected as a finalist in the Upstage Course Hackathon.

## Relevant Coursework

**Mathematics & Foundations:** Linear Algebra; Differential Equations; Discrete Mathematics, Intro to Algorithms; Randomized and Big Data Algorithms; AI for Programming; Principles of Programming Languages; Data Structures

**Vision & Graphics:** Computer Vision; Image Processing and Analysis I; Computer Graphics; FFT in Graphics and Vision; Computer Graphics Survey; Vision as Bayesian Inference

**Human-Centered Systems:** Augmented Humans

## Skillset

**Programming:** Python, PyTorch, Unity (C#), C, C++, Git, LaTeX, MATLAB

**AI/ML:** Fine-tuning Large Language Models, Vision–Language Models, Deep Learning Frameworks

**Tools:** OptiTrack with Manus (Motion Capture), LangChain, Sourcetree, Overleaf, Blender, MeshLab