Real-Time Multimodal Tactile Sensor with Visual and Auditory Feedback

Hyosung Kim, Junhui Lee, and Saekwang Nam
Graduate School of Data Science, Kyungpook National University, Republic of Korea

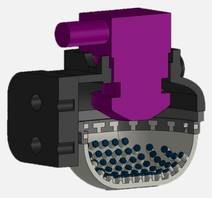
Introduction

- **Background**
  - **Tactile sensing**: fundamental to how humans interact with their environment, and equipping robots with analogous capabilities is critical for achieving **dexterous manipulation** [1].
  - **Vision-Based Tactile Sensors (VBTS)**: utilizing internal cameras to generate **high-resolution tactile data** from physical deformations. A prominent example is the **TacTip** [2].
  - **Human skin**: perceives complex tactile information through **slow-adapting (SA)** receptors, which detect **sustained pressure and shape**, and **fast-adapting (FA)** receptors, which detect **subtle vibrations and texture** [3].
- **Problems**
  - VBTS operates at 30 - 60 times per second on average.
  - While this is sufficient to perform SA receptors, it is inadequate for the response speed required to perform FA receptors.
- **Objectives**
  - Develop a **multimodal tactile sensor** that **integrates heterogeneous sensors** to perform the roles of both SA receptors and FA receptors to address these limitations.
  - Specifically, it integrates the high-resolution **spatial information** provided by **existing VBTS** with the **vibration information** detected by a high sensitivity **microphone**.

Materials & Methods

- **Hardware Configuration**

- Jetson Orin NX (Host System)
- **GPIO 40-pin** connector for **sensor connectivity**.

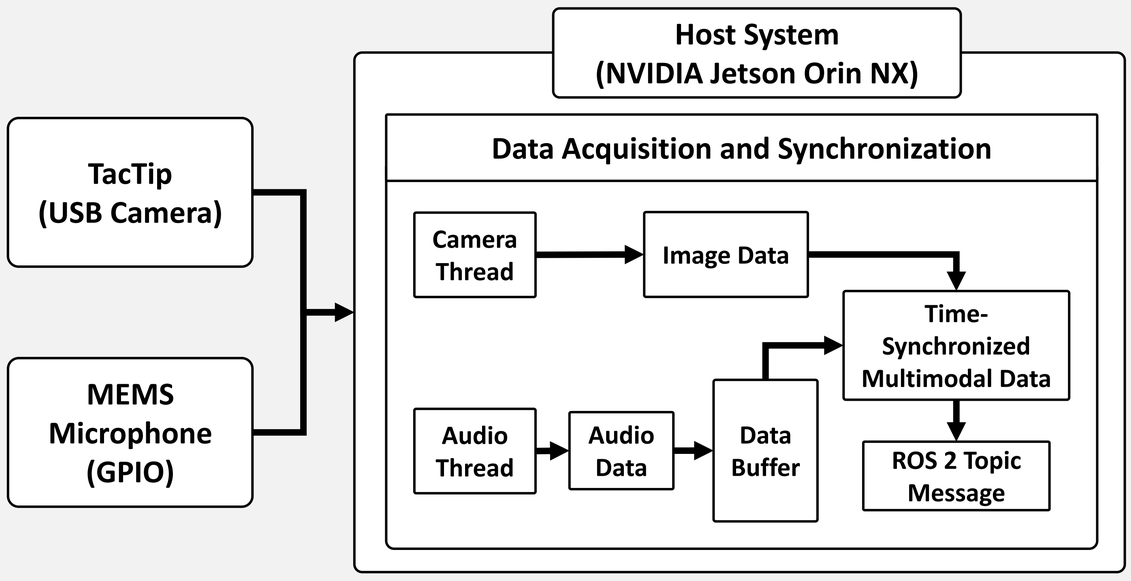


- TacTip (VBTS)
- Soft, deformable skin with an internal array of pins.

- USB Camera
- Internal camera of the TacTip.

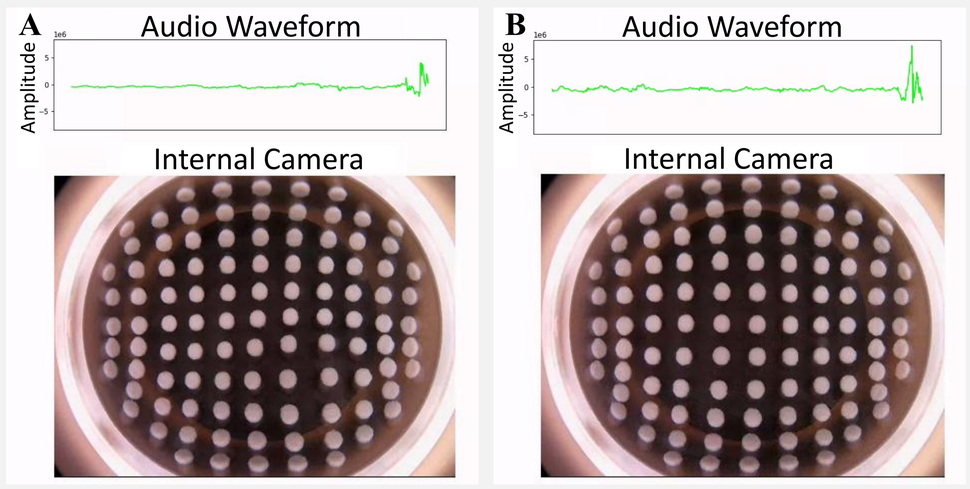
- SPH0645 (Microphone)
- Attached to the bottom of the TacTip to detect **vibration**.

- **Data acquisition and synchronization pipeline**
- All data acquisition and integration processes within the system are efficiently handled within a single ROS 2 node.

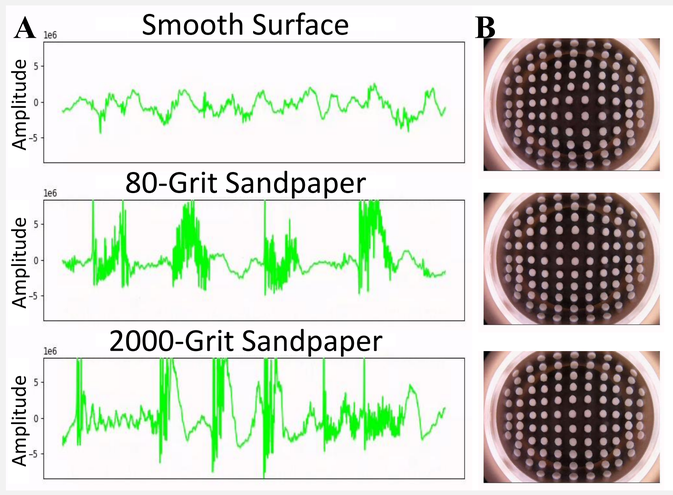


Experiments & Results

- **Experiment 1: Transient Contact Detection**
  - Objective: Compare the temporal precision of the vision-only and multimodal methods in detecting a transient tap event.
  - Results: microphone can capture and differentiate transient events with high fidelity.



- **Experiment 2: Surface Texture Discrimination**
  - Objective: Verify whether audio data can distinguish between different surface textures during a slip event.
  - Results: The audio stream exhibited unique vibration signatures corresponding to each texture.



Conclusions

- **Conclusions**
  - Presented and validated a multimodal tactile sensor that fuses a vision-based SA channel with a microphone-based FA channel on a single, ROS 2-driven embedded platform.
  - The system achieves superior temporal resolution by detecting transient contact events from auditory signal.
  - It successfully discriminates between visually-indistinguishable surfaces by identifying their unique vibrational signatures.
- **Future Works**
  - Key limitation of our system: the microphone is susceptible to ambient noise from sensor movement and the surrounding environment.
  - To address this, integrating a secondary, reference microphone to specifically measure and subtract ambient vibrations, thereby isolating the true contact signal. This enhancement is expected to yield more robust multimodal data, contributing to the overarching goal of achieving human-like robotic dexterity.

[1] M. Raibert, K. Blankespoor, G. Nelson, and R. Playter, “BigDog, the rough-terrain quadruped robot,” in Proc. 17th IFAC World Congress, 2008, pp. 10822-10825.
[2] B. Ward-Cherrier et al., “The TacTip family: Soft optical tactile sensors with 3D-printed biomimetic morphologies,” Soft Robotics, vol. 5, no. 2, pp. 216-227, 2018.
[3] R. S. Johansson and J. R. Flanagan, “Coding and use of tactile signals from the fingertips in object manipulation tasks,” Nature Reviews Neuroscience, vol. 10, no. 5, pp. 345-359, May 2009.
This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2023-00242528, RS-2024-00436182) and by the IITP (Institute for Information & Communications Technology Planning & Evaluation)-ITRC (Information Technology Research Center) grant funded by the Korea government (No. IITP-2025-RS-2024-00437756) and the other IITP grant funded by the Korea government (MSIT) (No. RS-2025-02263277).

Contact: s.nam@knu.ac.kr