2025 IEEE World Haptics Conference (WHC)
Suwon Convention Center, Suwon, Korea
July 8 ~ 11, 2025

Preliminary Collision Analysis for Safe Human-Robot Collaboration

Jieun Jang, Taejun Kwon and Saekwang Nam*

I. Introduction

The integration of collaborative robots in manufacturing facilities has increased collision risks between humans and robots, posing significant safety threats to workers [1]. To ensure safety, workers typically set high collision detection sensitivity, particularly in environments with frequent human-robot interactions such as food manufacturing. However, this heightened sensitivity often causes unintended emergency stops triggered by the robot's operational vibrations.

To address this issue, we conducted a study to distinguish between task-induced vibrations generated by the robot's operations and human collision vibrations during collaborative tasks. In particular, we measured vibrations in the form of acoustic signals using a microphone [2] to distinguish between the robot's operational vibrations (Fig. 1(a))–occurring while the robot arm, holding a basket with food items, shakes off the frying oil after lifting the basket from a deep fryer–and the vibrations transmitted to the robot arm due to collisions with parts of a worker's body (Fig. 1(b)). After systematic labeling of the three types of acoustic events (task-induced vibrations (TV), Collisions with a human(COL) and no-vibration situations (NV)), we analyzed their characteristics by applying Short-Time Fourier Transform (STFT)-based spectrogram analysis [3].

II. Materials and Methods

A. Experimental Setup

For the testbed, we utilized deep frying equipment combined with a Doosan Robotics A0509 collaborative robot arm (Fig. 1(a)). The robot arm is mounted on top of the deep frying equipment and is designed to repeatedly deep fry food items using a basket. To measure the acoustic signals generated by vibrations, we attached a microphone at the point on the robot arm where the signals could be effectively transmitted [4], as shown in Fig. 1(c).

The acoustic signals transmitted to the surface of the robot arm were recorded using a microphone and stored on a Raspberry Pi 5. To ensure compatibility, an Adafruit I2S MEMS Microphone was used and wired directly to the Raspberry Pi [5]. Additionally, to minimize external noise interference and maximize the recording of surface vibrations from the robot arm, the microphone element was wrapped with insulating tape [6]. Finally, the insulated microphone was glued to the location on the robot arm where the transmission of vibrations from the three different types of acoustic events could be most effectively captured.

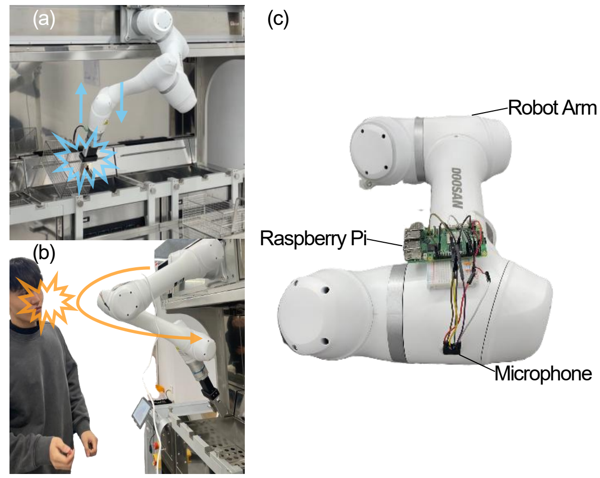

Fig. 1. Experimental Setup: (a) Task-induced vibrations, (b) Collisions with a human, and (c) Sensor installation on robot arm

B. Data Collection

The robot arm performs a pre-programmed operation for deep frying food. During a single operation, the robot arm shakes the basket containing the food approximately 51 times, which is considered task-induced vibrations. To incorporate collision with a human, a worker wearing protective gear made strong contact at various locations on the robot arm with a fist during the middle of the operation. As a result, each operation includes approximately 51 instances of task-induced vibrations and 5 to 8 instances of intentional human collision vibrations. The measured acoustic signals are recorded as 48 kHz WAV files, with each operation generating about 100 seconds of data. This experimental protocol was repeated 50 times to produce a large dataset.

C. Data Refinement

Each raw WAV file obtained from a single deep frying operation comprises three distinct categories of acoustic events: task-induced vibrations caused by the robot's basket-shaking motion, human collision vibrations, and no-vibration events. Initially, we converted the raw acoustic signals into spectrograms using the Short-Time Fourier Transform (STFT) to represent the power distribution of frequencies over time, enabling the identification of differences in the acoustic power spectrum patterns between different vibration

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. RS-2023-00242528) and by the IITP (Institute for Information & Communications Technology Planning & Evaluation)-ITRC (Information Technology Research Center) grant funded by the Korea government (MSIT) (No. IITP-2025-RS-2024-00437756).

Jieun Jang, Taejun Kwon, and Saekwang Nam are with Graduate School of Data Science, Kyungpook National University, Daegu 41566, South Korea. e-mail: s.nam@knu.ac.kr

*Corresponding author

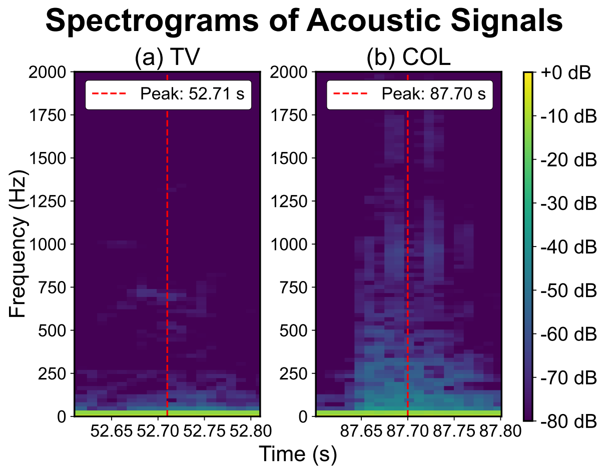

Fig. 2. Spectrograms of acoustic signals: (a) Task-induced vibrations (TV) and (b) Collisions with a human (COL). Red dashed lines indicate peak moments.

types. For instance, the pattern of task-induced vibrations from basket shaking produces a spectrogram as shown in Fig. 2(a), whereas the spectrogram resulting from collisions with a human is depicted in Fig. 2(b).

Subsequently, we systematically annotated the spectral patterns through manual labeling based on visual inspection. Task-induced vibrations were labeled as ”1” and those related to collisions with a human were labeled as ”2”. For instance, the red-dashed vertical line in Fig. 2(a) is labeled as ”1”, while the one in Fig. 2(b) is labeled as ”2”. The remaining spectral segments over time, where no-vibration activity occurred, were assigned a label of ”0.”

D. Conversion of Spectrogram Data from Time Domain to Frequency Domain

As there are clear differences in the acoustic power spectrum among the three acoustic categories, we further quantified these differences by transforming the time-domain spectrogram data to the frequency domain. Specifically, amplitude distributions across frequency were extracted from 22 spectrogram data points for each category. The results for task-induced vibrations (labeled as ”1”) are represented by sky-blue colored lines in Fig. 3, with their mean distribution shown in blue. Collisions with a human (labeled as ”2”) are plotted in light-orange and orange.

III. Preliminary Results and Disscussion

The frequency responses exhibited distinct distribution differences among the three acoustic categories, as shown in Fig. 3. Task-induced vibration demonstrated significantly lower variability than collisions with a human, with the light-blue shaded area showing a narrower amplitude range than the light-orange area. This consistency reflects the robot’s repetitive basket-shaking motion, while the larger variation in human collisions results from varying impact locations and magnitudes. No-vibration situations showed minimal amplitude variation, representing baseline conditions.

Operational vibrations exhibited a characteristic peak around 700 Hz, corresponding to the metal basket-gripper

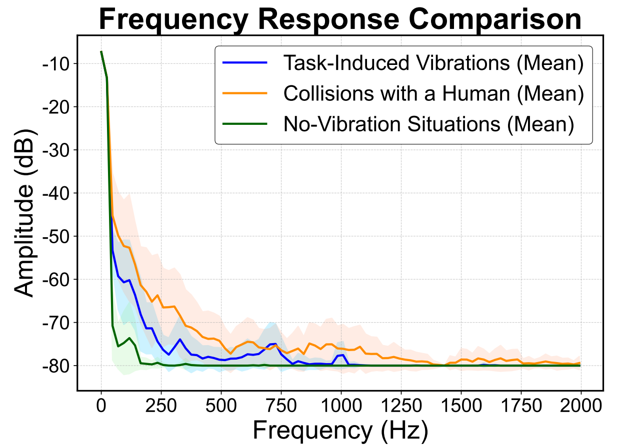

Fig. 3. Frequency response comparison by collision types (0-2000Hz). Shaded areas indicate ±1 standard deviation.

contact. Conversely, collisions with a human showed higher amplitude values in the low-frequency range due to the soft tissue-metal collision characteristics, resulting in stronger mid-to-low frequency components.

IV. Future Work

The frequency response analysis revealed distinct patterns among the three acoustic categories (task-induced vibrations, collisions with a human, and no-vibration situations), indicating the feasibility of classification using a one-dimensional neural network approach [7]. The systematically labeled frequency-domain data, representing the amplitude distribution across frequencies, will serve as the training dataset for classifying the collision type.

V. Ethics and Acknowledgment

This experiment adhered to the principles of the Declaration of Helsinki. Given its non-invasive nature and simple measurement methods, it was exempt from formal IRB approval. This experiment was conducted in collaboration with ITCOBOT.

References

[1] J. Vasic and A. Billard, ”Safety issues in human-robot interactions,” in 2013 IEEE ICRA, pp. 197-204.
[2] B. Khojasteh, Y. Shao, and K. J. Kuchenbecker, “Robust surface recognition with the maximum mean discrepancy: Degrading haptic-auditory signals through bandwidth and noise,” IEEE Trans. Haptics., vol. 17, no. 1, pp. 58-65, 2024.
[3] M. J. Yang, J. Cho, H. Chung, K. Park, and J. Kim, “Touch classification on robotic skin using multimodal tactile sensing modules,” in 2023 IEEE ICRA, pp. 9917-9923.
[4] X. Fan, D. Lee, Y. Chen, C. Prepscius, V. Isler, L. Jackel, H. S. Seung, and D. Lee, “Acoustic collision detection and localization for robot manipulators,” in 2020 IEEE/RSJ IROS, pp. 9529-9536.
[5] J. Llave, A Low-Cost Robotic System to Convey Object Texture via Vibro-Tactile Stimulation. California State University, Long Beach, 2021.
[6] M. D. Rao, “Recent applications of viscoelastic damping for noise control in automobiles and commercial airplanes,” J. Sound Vib., vol. 262, no. 3, pp. 457-474, 2003.
[7] J. H. Sim, J. Yoo, M. L. Lee, S. H. Han, S. K. Han, J. Y. Lee, S. W. Yi, J. Nam, D. S. Kim, and Y. S. Yang, “Deep learning model for cosmetic gel classification based on a short-time fourier transform and spectrogram,” ACS Appl. Mater. Interfaces, vol. 16, no. 20, pp. 25825-25835, 2024.