論 文 の 要 旨

題 目　　A Study on Vibration Source Localization Using High-speed Vision

（高速ビジョンを用いた振動源定位に関する研究）

氏 名　　姜 明俊

Localizing a target vibrating source robustly is a significant task in the various fields and most of the existing vibration source localization methods are realized by using the acoustic cues emitted by the vibrating source. However, sound source localization has its inherent bottleneck in accuracy because of the low directivity in sound propagation. Another vibration source localization concept is vision-based tracking that uses appearance features of target, such as color histograms, object shapes, and other landmarks. Most of the video cameras in used in existing tracking methods are restricted to standard video signal formats (e.g, NTSC 30 fps or PAL 25 fps), which are modeled around the human eye's properties. As a result, they have been implicitly designed on the basis of appearance-based image features, and dynamics-based image features, such as high-speed vibration excited at 100 Hz or higher on the audio frequency range, which cannot be observed with the naked eye, have never been considered.

To solve the problem, I proposed a concept of vision-based vibration source localization to extract vibration image regions by using pixel-level digital filters in a high-frame-rate (HFR) video. An image sensor is regarded as a collection of photo sensors, the number of which corresponds to the pixel number of the image sensor, and the image intensity at every pixel in an image can be considered as a time sequential signal.

The temporal periodic changes in image intensity can be observed at pixels of vibration sources in images, depending on their vibration frequencies in the audio frequency range, when the frame rate of a vision system is sufficiently high to allow vibration measurement.

Thus, an HFR vision system can localize spatiotemporal changes in image intensities as a vibration distribution by implementing digital filters, which is one of the basic operations in acoustic signal processing for the analysis of sound and vibration dynamics at all the pixels in images in order to pass signals in a specific band of frequencies for identifying and inspecting their vibration frequency and other properties.

A) Robustness Analysis of Vibration Features Against Appearance Changes in High-Frame-Rate Videos

We investigate the effect of appearance variations on the detectability of vibration feature extraction with pixel-level digital filters for HFR videos. In particular, we consider robust vibrating object tracking, which is clearly different from conventional appearance-based object tracking with spatial pattern recognition in a high-quality image region of a certain size.

For HFR videos of a rotating fan located at different positions and orientations and captured at 2000 or 300 frames per second with different lens or exposure time settings, we verify how many pixels are extracted as vibrating regions with pixel-level digital filters. The effectiveness of dynamics-based vibration features is demonstrated by examining the robustness against changes in aperture size and the focal condition of the camera lens, the apparent size and orientation of the object being tracked, and its rotational frequency, as well as complexities and movements of background scenes and motion blurs in captured videos.

Tracking experiments for a flying multicopter with rotating propellers are also described to verify the robustness of localization under complex imaging conditions in outside scenarios.

B) Real-Time Vibration Source Tracking Using High-Speed Vision

By applying pixel-level digital filters to clipped region-of-interest (ROI) images, in which the center position of a vibrating object is tracked at a fixed position, we reduce the latency effect on a digital filter, which may degrade the localization accuracy in vibration source tracking. Pixel-level digital filters for 128×128 ROI images, which are tracked from 512×512 input images, are implemented on a 1000 fps vision platform that can measure vibration distributions at 100 Hz or higher. Our tracking system allows a vibrating object to be tracked in real time at the center of the camera view by controlling a pan-tilt active vision system. We present several experimental tracking results using objects vibrating at high frequencies, which cannot be observed by standard video cameras or the naked human eye, including a flying quadcopter with rotating propellers, and demonstrate its performance in vibration source localization with sub-degree-level angular directivity, which is more acute than a few or more degrees of directivity in acoustic-based source localization.

In future work, we intend to improve these points toward more universal vibration source localization under more extreme conditions and accelerate the computational speed for real-time processing of HFR video. And we plan to improve the algorithm for more universal vibration source localization with automated tuning of parameters for digital filters, ROI scale adjustment, and other additional processes as well as the performance of the system by using GPUs for enlarging the spatial resolution of the ROI region and accelerating the visual sampling rate for higher frequency vibration detection; and expand our system for practical applications, such as simultaneous surveillance of flying drones with accurate localization.