SUTUDY ON SKIN PROPAGATION OF SPEECH USING THREE DIMENSIONAL DISPLACEMENT MEASUREMENT

Masashi Nakayama, Shunsuke Ishimitsu, Satoshi Iijima
3-9-12, Ozuka-higashi, Asaminami, Hiroshima, Japan
email: masashi@hiroshima-cu.ac.jp
Yoshiyasu Futawatari
FK Bldg. 3rd floor, 1-8-6, Hamamatsucho, Minato, Japan

Many researchers focus on speech and body-conducted speech including body-conducted speech to clear propagation mechanisms for making innovative interfaces. Especially, skin propagation of speech on a face skin analyzed by Laser Doppler Vibrometer generally. In this study, it is presented that sound measurement on skin propagation during vocalization using three dimensional displacement measurement. At this experiment, each subject vocalized 7 vocabularies that are 5 Japanese vowels, 1 syllabic nasal and 1 silence. Then, speech, body-conducted speech and skin displacements are measured for analysis with Spectrum and its envelope. Speech is recorded at about 20 cm from a mouth with a microphone, and body-conducted speech is recorded at upper left of upper lip with an accelerometer. The displacements are captured from upper lip using the 3D displacement measurement. As the results, direction and its attenuation depend on each axis and a transfer function can be obtained that converts between the dimensions of acceleration and displacement at the same measurement location on the lip.

Keywords: Skin-conduction, Speech, Three Dimensional Displacement Measurement

1. Introduction

A measurement using a laser vibrometer or an accelerometer has been investigated to clarify skin and body propagation of a human body and skin during speech vocalization and singing. Many research topics are investigating for measuring and analysing the device. H. Tabatabai et al. experimented to measure the signals which are muscle tension, tremor, heart sound and voice with a single point of vibrometer [1]. T. Kitamura measured with a vibrometer when 3 Japanese singers vocalized /a/ Japanese vowel with fixed fundamental frequency [2]. And, Y. Avargel and I. Cohen proposed speech enhancement with a vibrometer information, and then the frequency characteristics improved with Spectral Gain Modification and the information [3]. In this study, we try to measure skin propagation during the vocalization by sound vibration measurement target using three dimensional (3D) displacement measurement that can capture 3D displacement at a human face. At this time, speech, body-conduction speech and skin displacement are simultaneously measured, and the relationship with each signal is discussed.
2. 3D dimensional measurement

To measure skin vibration on human face, scanning measurement with a laser Doppler meter is generally used. Because this measurement uses the reflection of laser at a skin, it is a one-dimensional displacement measurement in the vertical direction against this device. However, considering the complexity of human skin changes, multidimensional measurement is required. Therefore, by using 3D displacement measurement, we try to measure a skin surface vibration and propagation. In 3D displacement measurement, it is possible to image the surface to be measured using a camera ability of high-speed imaging installed at different angles and obtain 3D displacement measurement at an arbitrary point within the imaging range. 3D displacement measurement has been used in measurements for test and examination of advanced materials and parts [4], however a vibration measurement for human skin has not been attempted, so the method expects to measure 3D displacement measurement for human.

3. Signal recording for skin-propagated sound

In general, during a vocalization, airborne sound emitted from a lip and nostrils, and body-conducted sound as solid propagation such as skin and the conduction are also propagated. In this study, an experiment compare with speech, body-conducted speech and 3D displacement measurement to find the differences in frequency characteristics respectively.

3.1 Experimental setup

A speech was measured by a microphone, a body-conducted speech by an accelerometer, and a skin propagation by 3D displacement measurement. Figure 1 shows a scene of sound recording. This experiment was performed in a space, a part of the office limited to use environment and equipment used in the 3D displacement measurement. Participants vocalized a speech while sitting condition on a chair, and the vocabulary choose as 7 conditions including 5 Japanese vowels, 1 syllabic nasal and 1 silence. Under the experimental conditions of Table 1, the microphone is installed at a position of about 30 cm from the mouth, and the accelerometer is placed on an upper where is employed as the measuring location [5]. The cameras for the 3D displacement measurement focused on an opposite side of upper lip for the accelerometer. It is necessary to measure for 4th or 5th formants such as about 4 to 5 kHz and more, however ability of sampling frequency on the 3D measurement has a trade-off between the frequency and resources for measurement such as luminance and memory for capturing the images. Considering the trade off, experiment is employed on 10 kHz sampling with 0.2 s recording.

![Figure 1: Signal recording with 3D displacement measurement](image)
Table 1: Experimental setup

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Model name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software</td>
<td>Istra4D, Dantec Dynamics</td>
</tr>
<tr>
<td>3D measurement</td>
<td>Q-400, Dantec Dynamics</td>
</tr>
<tr>
<td>Camera</td>
<td>VEO410L, Phantom</td>
</tr>
<tr>
<td>Lens</td>
<td>NIKKOR 105mm, Nikon</td>
</tr>
<tr>
<td>Lighting</td>
<td>Halogen &amp; LED lighting</td>
</tr>
<tr>
<td>Microphone</td>
<td>MI-1431, Ono Sokki</td>
</tr>
<tr>
<td>Accelerometer</td>
<td>NP-2110, Ono Sokki</td>
</tr>
<tr>
<td>Amplifier</td>
<td>SR-2200, Ono Sokki</td>
</tr>
</tbody>
</table>

3.2 Result and discussion

Figures 2 and 3 show speech and body-conducted speech. In the speech, the fundamental frequency and 1st to 4th formants were observed on Spectrum and its envelope on the speech, and frequency characteristic of body-conducted speech measured with up to about 1 kHz. The differences become a factor which is decreasing of sound quality and dismissing of 3rd and higher formants. In general, vocabulary can recognize with 1st and 2nd formants, so there is no problem for hearing of vocabulary in spoken words and sentences of the speech. Figure 4 shows x and y axes of 3D displacement signal measured on the upper lip. A vowel is vocalized sound with vocal cord, repetition of the fundamental wave is seen in the waveforms of both axes can be observed on the spectrum. Attenuation of about 10 dB occurs up to about 500 Hz rapidly after the fundamental frequency, and the attenuation of the gain can be confirmed in the order of speech, body-conducted speech, and skin-propagation. In addition, since the magnitude differs depending on the axis, it can be seen that there is a bias in the vibration direction respectively. From these result, the following two outcomes are considering by this experiment. This bias is means to expect the dependency for axis on human skin, so this is one of important outcome as a necessity for measurement with 3D displacement measurement. In addition, it is considered that a transfer function can be obtained that converts between the dimensions of acceleration and displacement at the same measurement location such as an upper lip, because of the differences was only dimension between displacement and accelerometer as signal measurement.
4. Conclusion and future work

In order to confirm the possibility of skin measurement using 3D displacement measurement, speech, body-conducted speech and skin propagation on upper lip were measured. Attenuation of about 10 dB was observed up to about 500 Hz on the skin surface via body-conducted speech. Especially, there are two findings, direction and attenuation depend on each axis and a transfer function can be obtained that converts between the dimensions of acceleration and displacement at the same measurement location on the lip. In the future, we will discuss to find that estimation of transfer functions that convert displacement and acceleration and measurement with 3D displacement with z-axis added will be performed as truly 3D dimensional measurement. In addition, the vibration using 3D for more precisely analysis try to visualize detailed skin propagation and vibration.

Acknowledgement

I would like to thank Mr. K. Tomioka and Mr. M. Sato, CORNES Technology Limited for the opportunity to measure 3D displacement, and also thank Mr. G. Kuwabara, Photron Limited for supporting of experiment with high-speech camera.

REFERENCES