Title
A haptic-inspired audio approach for structural health monitoring decision-making

Authors
Mao, Zhu
Todd, Michael
Mascareñas, David

Publication Date
2015-03-23

DOI
10.1117/12.2084464

Peer reviewed

A Haptic-Inspired Audio Approach for Structural Health Monitoring Decision-Making

Zhu Mao[1], Michael Todd[2*] and David Mascareñas[3]
[1]Department of Structural Engineering, University of California San Diego, La Jolla CA 92093-0085, email: zmao@ucsd.edu; [2]Department of Structural Engineering, University of California San Diego, La Jolla CA 92093-0085, email: mdtodd@ucsd.edu; [3]Engineering Institute, Los Alamos National Laboratory, Los Alamos, NM 87545, email: dmascarenas@lanl.gov

Abstract

Haptics is the field at the interface of human touch (tactile senstation) and classification, whereby tactile feedback is used to train and inform a decision-making process. In structural health monitoring (SHM) applications, haptic devices have been introduced and applied in a simplified laboratory scale scenario, in which nonlinearity, representing the presence of damage, was encoded into a vibratory manual interface. In this paper, the "spirit" of haptics is adopted, but here ultrasonic guided wave scattering information is transformed into audio (rather than tactile) range signals. After sufficient training, the structural damage condition, including occurrence and location, can be identified through the encoded audio waveforms. Different algorithms are employed in this paper to generate the transformed audio signals and the performance of each encoding algorithms is compared, and also compared with standard machine learning classifiers. In the long run, the haptic decision-making is aiming to detect and classify structural damages in a more rigorous environment, and approaching a baseline-free fashion with embedded temperature compensation.

Keywords: haptics, ultrasonic guided waves, audio encoding, damage detection, damage localization, structural health monitoring

1. INTRODUCTION

Human decision-making ultimately relies upon data from the five usual senses (vision, aural, tactile, taste, and smell). Compared to the state-of-the art computational algorithms, the human perspective and decision-making power are typically more adaptive and robust, which is not often well-exploited in terms of interpreting structural health monitoring (SHM) data. Among the five traditional senses people have, haptics refers to the tactile sensing to the response of force/pressure, vibration, texture, and/or thermal changes to the human body, and haptic engineering is the utilization of the sense of touch for a variety of control, learning, and identification purposes. Specifically in the context of SHM, haptics has been introduced to demonstrate how nonlinear impacts to a structure is identified via a smart glove embedded with a vibro-transducer array that can map SHM signals to modulated tactile pulses. Different patterns of tactile stimuli are identified by human, thus damage types and locations will be classified [1].

Beyond the obvious sense of vision, another sense that is exploited by humans for decision-making is hearing, and by processing the audio information, people identify various types of objects and phenomena. For instance, people may recognize other's voice although what they hear was not exactly heard before. Facilitated by sound, people may also be able to tell the moving direction of an object, as well as whether a bottle if full while being filled. In engineering applications, people have long history of utilizing audio stimuli to extract useful information, among which the railway applications are one of the most successful. In order to inspect the integrity of locomotive systems, wheel tapping has been conducted for over a hundred years, to make sure there are no cracks on train wheels and to ensure that axle boxes are not overheating. Inspired by all these capabilities of sense of touch and hearing, this paper converts ultrasonic guided waves scattered from moveable magnet on a metallic surface at different locations into audio range signals to determine whether human auditory perception is superior for identifying damage occurrence, locations, and/or types.

Based on the ultrasonic guided wave detection technologies, Figure 1 shows how the audio haptic encoding is processed. As a tone burst of excitation is generated by the mounted piezoelectric actuator, the responses at the same or different locations will be recorded, and subtracted by the corresponding signal at a baseline condition. The subtracted residual contains the time-of-flight information, no matter it is obtained from the pulse-echo or pitch-catch mode. After appropriate audio-encoding, the classification of different damage locations will be conducted in this work.

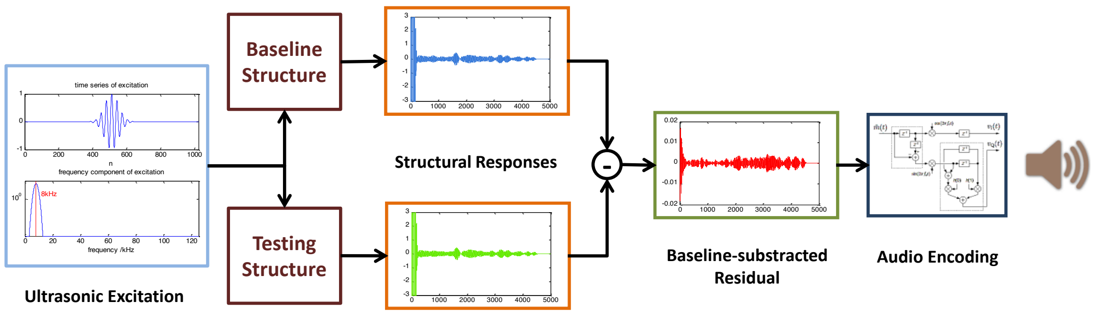


Figure 1: Flow of ultrasonic guided wave signal encoding for audio-haptic damage detection

A plate beam Figure 2 demonstrates briefly how the waveforms are different from each other. Actuators and sensors are placed near the two end of the beam, and multiple damage locations are marked. Since the damage placement is symmetric, there will be 5 distinct damage cases in total. Magnet is placed on each of the 5 locations, as Figure 2 illustrated, and the magnet will change the scatter patter of the ultrasonic response, therein induce non-trivial residual signals.

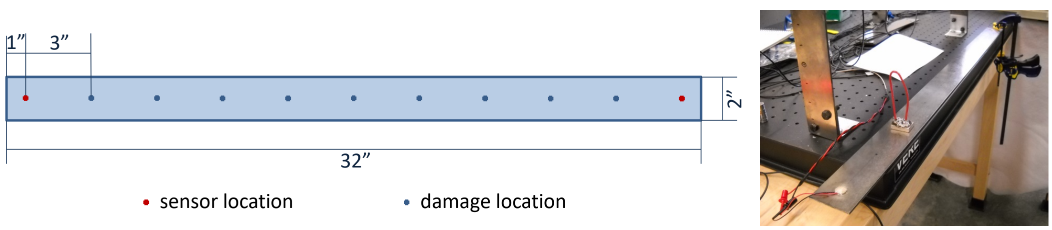


Figure 2: Test structure and damage locations.

Figure 3 plots the baseline-subtracted residuals for undamaged condition and the other 5 damaged conditions. The undamaged condition in upper left is under undamaged condition, and the residual is more like random noise, rather than the informative wave packets in the other 5 subplots. Comparing the 5 damaged residuals, there is obvious difference between those cases. However, considering the uncertainty existed in the evaluation process, such as temperature change and operational variability, the patterns of each damaged case may not be fully recognized to facilitate a future classification. Despite the state-of-the-art feature extraction techniques applied in SHM, such as first arrival analysis, audio decision-making treats the whole waveform as a complete set. Obviously, detecting damage is very straightforward, and it would not take lots of effort to obtain an optimized encoding algorithms. Yet the classification between cases is not an easy task. To make ultrasonic waves audible, the naïve implementation is to stretch the signal to a slower rate, therefore the original signal will be at lower frequency and proportionally longer in time. This stretch strategy will only make the sound track audible, but not help people making easier classification decisions. Because the tone of all the time series have the same frequency content, and there will only be slight differences in the magnitude which are not significant enough to be differentiated.

In the rest of this paper, two different encoding algorithms will be presented, namely a chord encoding and a melody encoding. Both algorithms will be based on the pitches existed on a piano keyboard. In section 2, the chord encoding will be introduced, in which ultrasonic signals are acquired from a beam with magnet moving at certain locations, and different damage conditions will output different chords. In section 3, melodies are generated by transferring the guided wave signals acquired from a metal plate. In section 4, several human subjects are trained and all the audio tracks are classified, and the performances of the two approaches are compared in this section as well.

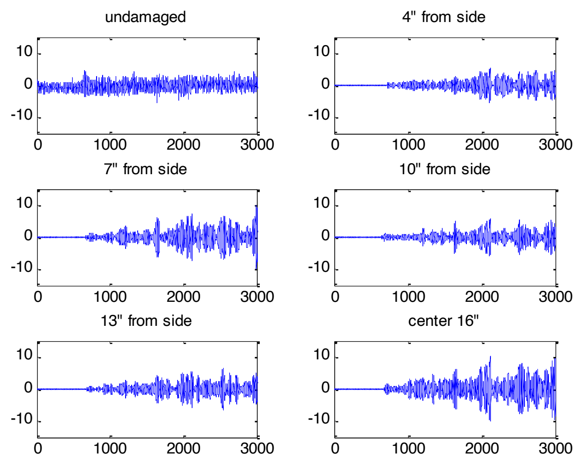


Figure 3: Residual signal for pitch-catch guided wave detection

2. CHORD ENCODING

As abovementioned, the difference of time series from multiple damage conditions are not dramatically significant, so encoding the original series to a more sensitive and specific sound tracks becomes important. The pitches (only single tone) from piano keyboard will be employed in this study, as they are the most "human friendly" sound. Equation (1) is the first thrust of transforming the time series into a weighted sum of all 88 pitches on a keyboard, denoted as *w*(*t*),

$$w(t) = \sum_{i=1}^{88} a_i \sin(2\pi f_i t), \tag{1}$$

where f_i is the linear frequency of the i-th pitch on the keyboard, given by Equation (2):

$$f_i = 2^{\frac{k-49}{12}} \times 440 \text{ Hz}\cdot \tag{2}$$

In a selected pack of waveform, 88 local maxima are determined, and each of the magnitudes is put in Equation (1) as the weight (a_i) of that specific tone. As Figure 4 illustrated, all the highlighted dots are lined up with the associated tone frequency, and their positions on a piano keyboard. By this means, the sound track series obtained in Equation (1) forms a steady 88-pitch chord lasting in time duration *t*.

However, the waveform envelope of the ultrasonic response has a slow pattern in time domain, therefore the weight of certain key will tend to be similar to the keys nearby. This causes a lot of information redundancy and too many tones are not more helpful for human to obtain a good quality of training and easy haptic decision. In addition, the similar values of close keys, regarded as equally loud pitches, will cause uncomfortableness because of physiological reasons. The sound track will be overwhelmed with dissonant chords, such as those major and minor seconds, and because of those uncomfortable intervals, no sustainable training can be applied onto human.

Instead of picking every local peak, Figure 5 demonstrates the peak-picking on waveform envelope. In a certain time window, the highest several (4 in this paper) peaks on the envelope are selected, and selecting the number of peaks is a trade-off between the human hearing capability and the complexity of the encoded audio signals. More peaks will induce a more diversified audio pool, yet too many possible choices of chord will make the training and testing on human subject way harder. Moreover, to address the issue of dissonant intervals, only C, E and G will be included in lieu of the entire 88 notes on the keyboard. All combination of C, E and G form a C major chord and its various inversions, which are all friendly/harmonic to a human's sense of hearing. The selected time range is divided into 16 bins, and mapped to all C, E and G pitches across 5 octaves, and Figure 6 highlights the pool of notes that would be possibly included in the chord with their names and Hertz values. The highest four peaks in the waveform envelope is adopted as the weight to the corresponding bins that those peaks fall in, and all other bins will not be included in the chord.

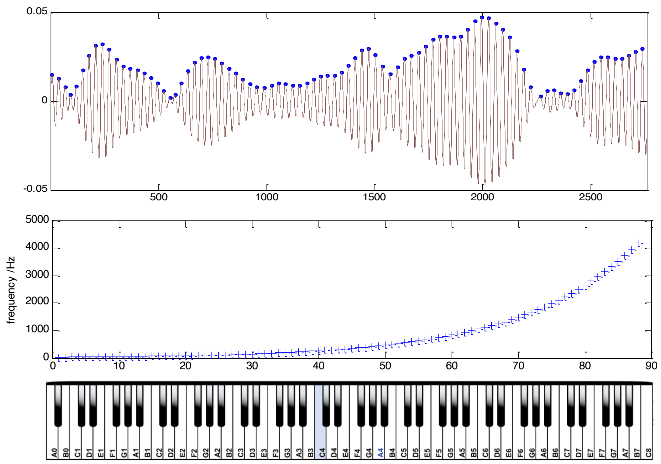


Figure 4: Encoding ultrasonic series into combination of keyboard notes

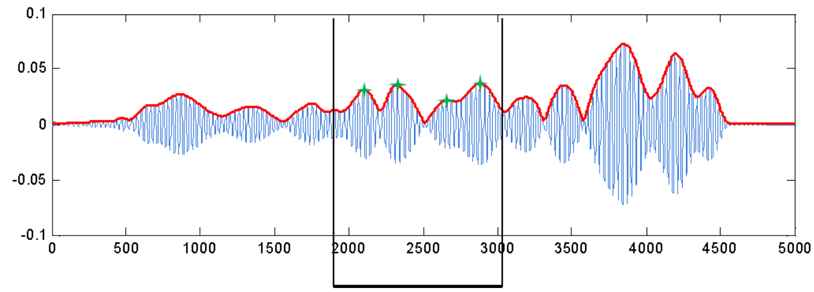


Figure 5: Encoding ultrasonic series via envelope peaks

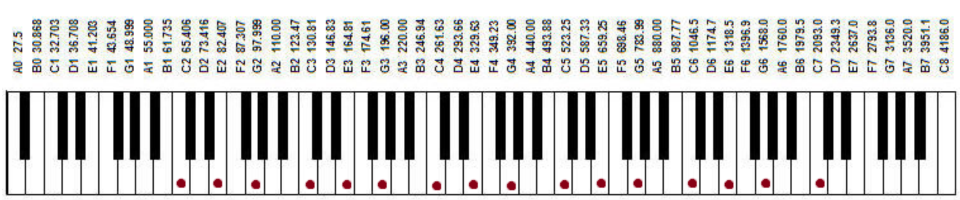


Figure 6: Selected notes for audio encoding on the keyboard

To implement the optimized audio encoding for damage identification, a plate-like structure is sketched in Figure 7, and the sensor locations and 6 damaged locations are marked in the figure. On the right-hand-side of Figure 7, the received waveforms under pulse-echo and pitch-catch modes are both plotted and compared, when there is no damage. For the purpose of sanity checking, the time of flight from both mode are found on the waveforms and they are nicely consistent with each other as the direct path of the two modes are equally long.

The encoded audio sound tracks are 4-tone sinusoids and the power spectra of the 6 damaged cases are potted in Figure 8. The power spectra clearly illustrate the pitch content of the encoded chords, as well as the weight of each single tone.

This encoding algorithm will be applied to multiple realizations to generate a training set and also applied to the unknown testing cases for identifying the location of damages as described in Figure 7.

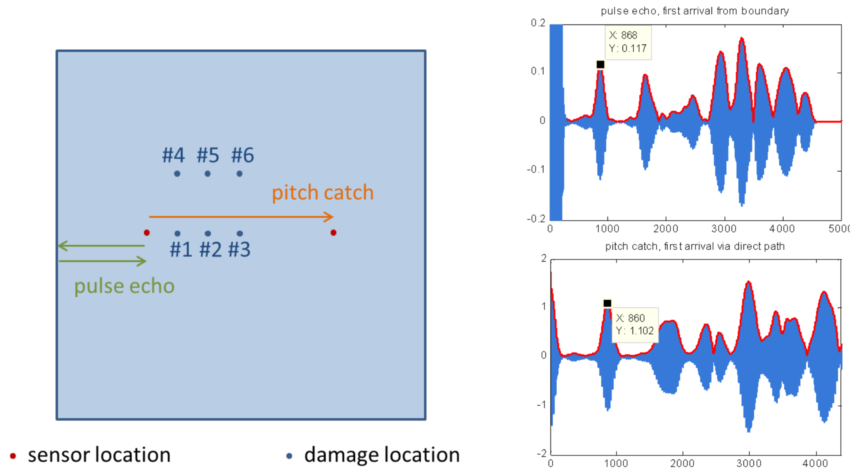


Figure 7: Selected notes for audio encoding on the keyboard

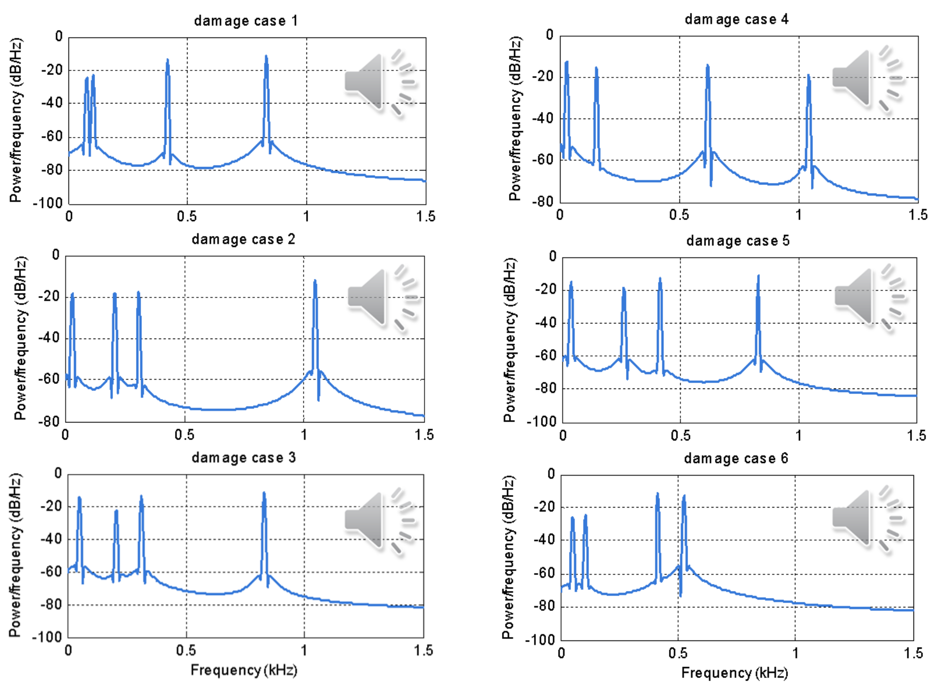


Figure 8: Power spectra of the encoded chords for each damage case

3. MELODY ENCODING

In parallel to the chord encoding algorithm, a different methodology will be introduced in this section. When a chord sound track is generated, the pitch combination does not change along the time dimension, nor does the volume. That is to say, the encoded chord signal is stationary. This helps and simplifies the training. However, to certainty complexity degree, it may be better if the change along time is meaningful and can be utilized to carry SHM information. A melody encoding is studied in this section as Figure 9 illustrates.

As abovementioned, both frequency content and the volume affect the hearing perception. In the chord encoding, the sound tracks have fixed frequency contents and constant volume. Figure 9 shows the algorithm of encoding with time-variant frequency contents and volumes. The wave pack is discretized in both the horizontal (time) and vertical (magnitude) axis. In the time axis, the entire wave form is divided into different bulks. The maximum value of each segment will be mapped to the pre-defined four bins along magnitude axis, representing four different tones. The length of each segment will be adopted to define the time duration of the encoded signal. In the end, the sinusoid in each time segment, with corresponding frequency, will be multiplied by the envelope of that wave bulk. As Figure 9 demonstrates, there will be a pitch transition at the end of each bulk, and the volume at the transition point is the lowest.

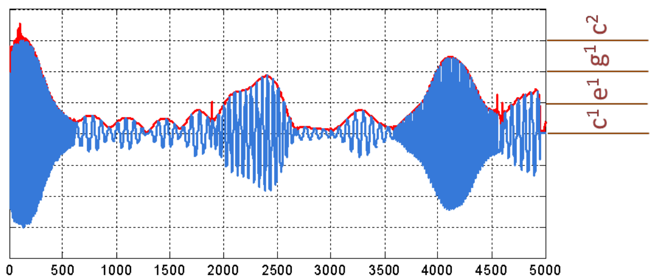


Figure 9: Encoding ultrasonic series into melody

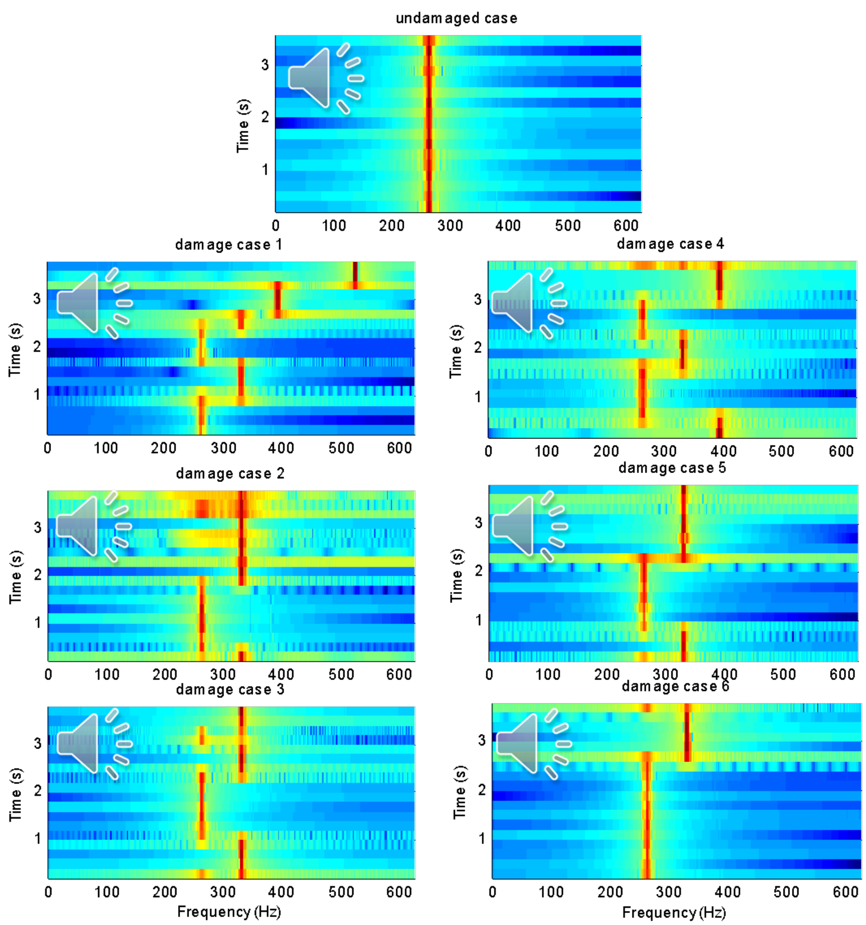
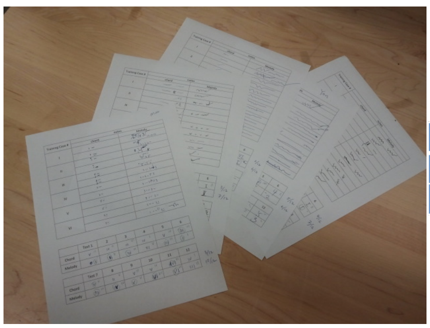


Figure 10: Spectrogram of the encoded melodies for baseline and each damage case

This melody encoding algorithm enables changes in the time domain, and in certain ways, the shape of the original waveform is transformed with respect to the frequency and length of each tone bust. Figure 10 shows the spectrogram of each encoded soundtrack. When there is no damage, the residual waveform has low magnitude and a noise-like characteristic, so that the encoded soundtrack is a constant C^1. When there are different damages, the spectrograms illustrate different patterns of melody.

4. HUMAN SUBJECT TESTING AND CONCLUSION

Human subject testing is conducted adopting the two audio-haptic encoding algorithms introduced in section 2 and 3. There are 24 tests implemented covering all the 6 damaged locations. These 24 tests are split into 12 trainings and 12 testing cases; therefore the number of making correct classification versus the total testing number is adopted to evaluate the audio-haptic decision-making performance. In Figure 11, the photo of test sheets and results are given. Among all the 6 test subjects, their background is diversified, especially for their music training experience.



	#1	#2	#3	#4	#5	#6
Chord	3/12	2/12	5/12	4/12	3/12	4/12
Melody	7/12	6/12	6/12	7/12	10/12	12/12

Figure 11: Human subject testing results

Based upon the testing observations and results, several conclusions are summarized as below:

- Damage detection is very straightforward, and every test subjective is able to tell the occurrence of damages, no matter what type the damage is.
- Generally speaking, the melody algorithm facilitates damage classification better than the chord encoding, with respect to a higher correct classification rate on average. People are in general better trained in daily life with melody, but not so much with chord.
- For chord encoding, there are several test results barely greater than random guess (2/12), and that indicates the ambiguity existed in audio signal generation as well as the capability limitation among the test subjects. The best result is 5 correct out of 12.
- For melody encoding, the correct rate is always greater than 50% among all the test subjects, and this approach outperforms chord encoding a lot. Test subject #6 has many years of instrument training, and classifies all the 12 tests correctly.
- Time of training is essential to the classification result. Test subject #5 is the algorithm designer and has much longer time exposed to the training cases, therefore does 10 correct out of 12.
- The correct classification rate is highly dependent on individual test subject. Test subject #2 has the lowest rate among all test subjects for both approaches, and this might be related with the individual audio perception.

Acknowledgement

This research was supported by the research grant (UD130058JD) of the Agency for Defense Development of the Korean government and by the Leading Foreign Research Institute Recruitment Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning (2011-0030065).

Reference

[1] D. Mascareñas, C. Plont, C. Brown, M. Cowell, N. J. Jameson, J. Block, S. Djidjev, H. Hahn, and C. Farrar, “A vibro-haptic human–machine interface for structural health monitoring” Struct. Health Monit., vol. 13, no. 6, pp. 671–685, Nov. 2014.

[2] Yoon MJ, Yu KH. Psychophysical Experiment of Vibrotactile Pattern Recognition at Fingertip. In SICE-ICASE International Joint Conference ; 2006; Busan, Korea.

[3] Vallbo AB, Johansson RS. Properties of cutaneous mechanorecpetors in the human hand related to touch sensation. Human Neurobiology. 1984; 3(1).

[4] Myle K, Binseel M. The Tactile Modality: A Review of Tactile Sensitivity and Human Tactile Interfaces. Army Research Laboratory; May 2007. Report No.: ARL-TR-4115.