Pitch Perception Using a Semitone Mapping to Improve Music Representation in Nucleus Cochlear Implants

Sherif A. Omran[1,2], Michael Büchler[1], Waikong Lai[1], Norbert Dillier[1]

[1]Laboratory of Experimental Audiology, ENT Department, University Hospital Zurich
[2]Institute of Neuroinformatics – University and ETH Zurich, Switzerland

Keywords

Virtual Channels, Semitone Mapping, Music, Pitch ranking, Complex Tones

Introduction

Cochlear Implants (CIs) are devices that aim at restoring speech perception for patients with profound or complete hearing loss. Current speech coding strategies are not necessarily optimal for music representation. Some patients report that music is perceived as an unpleasant noise while others enjoy it. Music is a complex sound consisting of notes having harmonic overtones. The majority of musical instruments generate fundamental frequencies below 1kHz (Pierce 1983). Clinically used mapping strategies compress the lower part of the frequency range [250Hz-1kHz] and thereby distort e harmonic structure. Another reason for the poor reproduction of musical sounds with current cochlear implants is the limited number of electrode channels which are mapped to a wide frequency range [250Hz-6kHz]. This causes the fundamental frequencies of adjacent musical notes to be often mapped to the same stimulation channel, especially in low frequency ranges below 500Hz.

The Semitone (Smt) mapping strategy presented here aims to improve music representation for CIs by increasing the frequency resolution. The approach consists of mapping the fundamental frequency of each semitone in a selected restricted frequency range to a separate electrode. Two versions of the Smt mapping were evaluated: the first one (Smt-LF) covers low to mid frequencies [125Hz-1.56kHz] and the second (Smt-MF) covers mid to high frequencies [440Hz-6kHz]. Smt mapping was first investigated by Kasturi and Loizou (2007) for the Advanced Bionics 12-channel cochlear implant using a frequency range of [300-600 Hz].

Hypothesis

Smt mapping can be tested by first processing music signals with an Acoustic Model (AMO) (Laneau, Moonen et al. 2006) and presenting them to normal hearing (NH) subjects. In addition, Busby and Plant (2005) reported that stimulating dual electrodes causes the perception of an intermediate pitch. This concept is used to extend the available 22 electrodes to 43 channels providing an increase in frequency representation and thereby hopefully improving the subject's performance.

Smt mapping is expected to give a better music representation due to preserving the harmonic structure of overtones. This would lead to an increase in discrimination between two notes whose fundamental frequencies are a few semitones apart if compared with the standard (Std) mapping.

Methods

Musical tones were processed using Std, Smt-MF and Smt-LF maps and were presented pairwise to NH subjects in a pitch ranking test. Their task was to choose the item which was perceived higher in pitch. Three groups of semitone distances were used: 1, 3 and 6 semitones. Each has two references (D, G#) forming six pairs (D-D#, D-F, D-G#, G#-A, G#-B, G#-D^{+1}) of tones, where the two tones D and G# are references, while D^{+1} is in the next octave. A pair of tones is randomly selected and presented in random order to the subject. The input audio signal was first band passed, then processed with the AMO. Loudness was roved between ±6dB to minimize loudness cue dependencies. A stimulation width parameter of 1 mm was assumed for the AMO. Six NH subjects were tested in a quiet environment.

Additionally, a Melody Contour Identification (MCI) test (Galvin, Fu et al. 2007) was performed, where five notes from nine patterns (rising, rising flat, rising falling, flat rising, flat, flat falling, falling, falling flat, falling rising) were played, each having the base note "A" from octaves three, four and five with semitone chromatic scale intervals [1-5].

Stimuli

Synthetic complex tones were prepared composed of fundamental frequencies and four harmonic overtones as musical notes. The four harmonic overtones had a successive 20% decrease in amplitude to avoid any change in the relative intensities of overtones that may lead to pitch reversals (Omran, Büchler et al. 2007). To avoid temporal and loudness cues, all tones had the same temporal structure and their perceptual loudness was equalized using Adobe Audition (Riley 2004). Each tone had a duration of 500 msec including 30 msec on/off ramp with a pause of 500 ms between each successive note. Musical tones from the 3rd, 4th and 5th octaves with fundamental frequency ranges [130-246Hz], [261-493Hz] and [523-987Hz] respectively were used for testing (Pierce 1983).

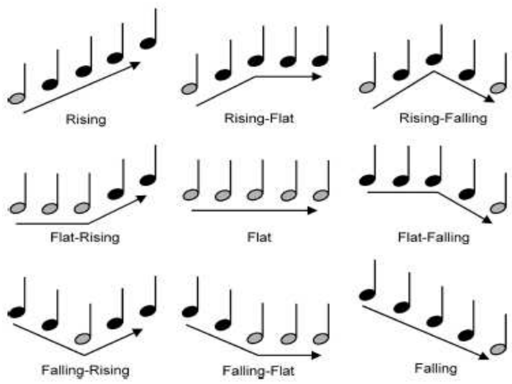
Procedure

A training session was first conducted to estimate the subjects' ability to perform this test. They were acoustically presented with 43 channels with different frequency to channel maps (Std, Smt-MF and Smt-LF) using a calibrated loudspeaker at 65 dBA at a distance of 1.5m in front of the subject. MACarena software (Lai and Dillier 2002) was used to present a randomly selected pair of tones from a group of semitone pairs (Table 1), and the subject had to choose the one higher in pitch. All tones were processed with the AMO. Each tone pair was presented a total of 8 times.

	Semitones				Semitones		
Ref	1	3	6	**Ref**	1	3	6
D3	D#3	F3	G#3	**G#3**	A3	B3	D4
D4	D#4	F4	G#4	**G#4**	A4	B4	D5
D5	D#5	F5	G#5	**G#5**	A5	B5	D6

Table 1: Pitch ranking tones and references

NH subjects conducted the MCI test, where a melody contour was acoustically presented and the subjects' task was to choose the correct one (rising, rising flat, rising falling, flat rising, flat, flat falling, falling, falling flat, falling rising). The notes were played using a calibrated loudspeaker at 65 dBA with a roving of ± 1 dB to minimize pitch shifts due to loudness effect (Zwicker and Fastl 1990).



Results

Figures 1-3 show the pitch ranking results of the Std, Smt-MF and Smt-LF maps with both D and G# references in Octaves 3, 4 and 5 using 1, 3 and 6 Semitones respectively. The vertical axis shows the results, where values between [-50 0] indicate a pitch reversal. Results between ±12.5 (2 errors in 8 trials) were considered to be at or below the chance level ($p < 0.05$) as indicated by the gray shaded area. The red points are outliers.

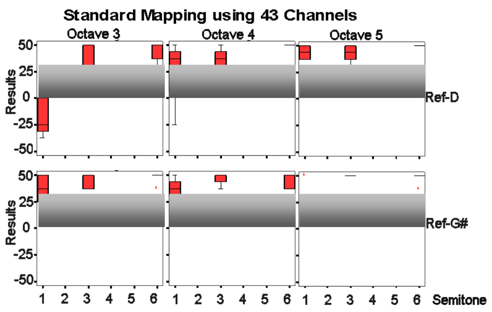


Figure 1: Results of Std mapping.

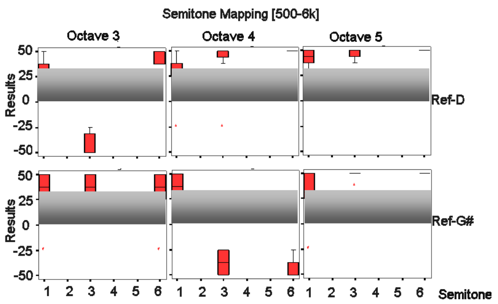


Figure 2: Results of Smt-MF mapping.

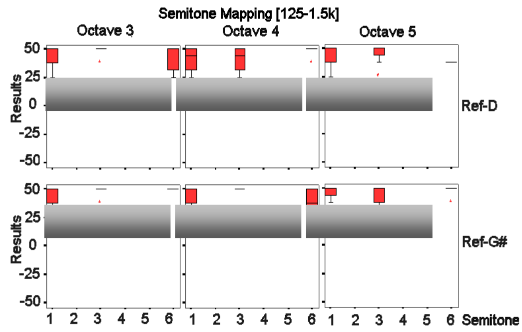


Figure 3: Results of Smt-LF mapping.

Comparing figures 1 and 2, one could see that pitch reversals occur more often with Smt-MF mapping (G# reference in octave 4 with 3 and 6 semitones and with the D reference in octave 3 with 3 semitones). Minor improvements were seen with Smt-MF (octave 3 with 1 and 6 semitones for the D reference and octave 4 with 1 and 3 semitones for the D reference).

Comparing figures 1 and 3, in octave 3 with both references (D and G#), in general the results are higher than the Std mapping. Smt-LF showed an improvement in the median and the variance (in octave 4 with 1 semitone the results improve with the D reference, and with the G# reference the median of the distribution is higher, and with 3 semitones the results are improved with reference G#. In octave 5 the variance increases with 3 semitones with the D reference.

Figure 4 shows the results from the MCI test for different semitone intervals with different maps. Each 3x3 matrix has nine cells representing the percent correct responses of the melody pattern test. The darker the cell, the higher the percentage correct was achieved. There is a trend for an improvement with Smt mapping compared to Std. However, Smt-LF produced a discrepancy at octave 5 because high frequencies (and therefore higher harmonics) are outside of the mapped frequency range. Similarly, Smt-MF produces worse results at octave 3 where low fundamental frequencies are outside of the mapped frequency range.

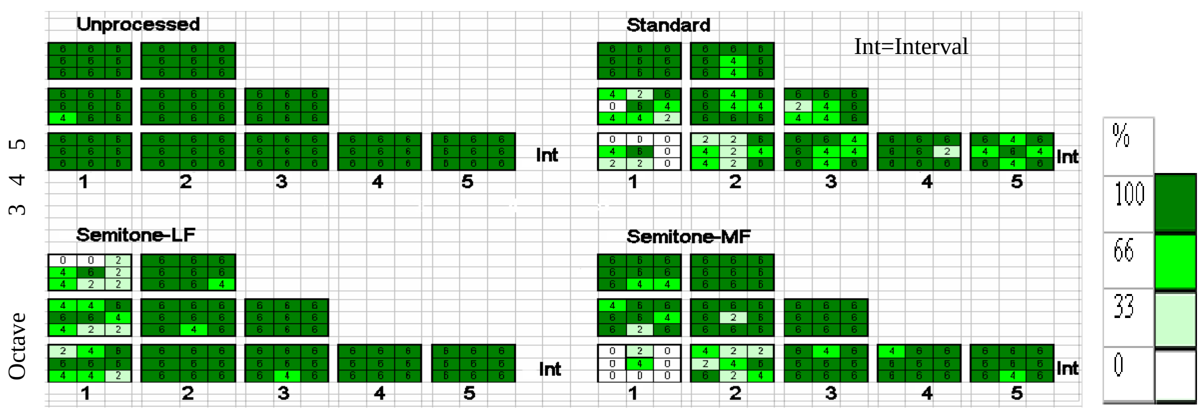


Figure 4: Melody Contour Identification test results.

Discussion

The aim of the study was to check if the Smt mapping would improve music representation. The results show that the tone pairs processed by the Std map are more easily discriminated for larger semitone intervals. Six semitones are easier to discriminate than three and three are easier than one. The results with the Smt-MF mapping are generally worse than the Std mapping due to its likelihood to produce pitch reversals, while the results with the Smt-LF mapping are generally better or at least similar to the results from the Std mapping. Furthermore, no pitch reversals were observed with this map. A disadvantage would be its smaller frequency range compared to the Smt-MF mapping.

Further investigations were performed with the MCI. Results of this test showed that Smt provided an improvement compared to the Std mapping. However, Smt-LF and Smt-Mf generate ambiguities at octave 5 and 3 because high and low frequency overtones were outside the respective frequency mapping range. In summary, the Smt-LF mapping yielded improved pitch ranking. In melody contour identification there was a trend of improvement with Smt-LF in comparison with the Std mapping.

Conclusion

Semitone mapping is believed to be a useful strategy for improving music representation as it increases the frequency representation and allows resolving the harmonic overtones. However, pitch reversals occur when using the Smt-MF map at low frequencies because the fundamental frequencies below 440 Hz are not considered in the map. This situation is somewhat improved using the Smt-LF which starts as low as 125 Hz and maps the processed sounds to higher frequencies than tonotopically expected which might, on the other hand, affect speech representation.

Acknowledgment

Swiss National Science Foundation. Grant Number 320000-110043.

References

Pierce J (1983). The science of musical sound, Library of congress.

Zwicker E and Fastl H (1990). Psychoacoustics: Facts and Models. Berlin, Germany, Springer-Verlag, 1990.

Lai W and Dillier N (2002). MACarena: a flexible computer-based speech testing environment. 7th International Cochlear Implant Conference 2002, Manchester.

Riley R (2004). Audio Edition with Adobe Audition, PC Publishing.

Busby P A and Plant K L (2005). "Dual electrode stimulation using the nucleus CI24RE cochlear implant: electrode impedance and pitch ranking studies." Ear Hear **26**(5): 504-11.

Laneau J M Moonen, et al. (2006). "Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants." J Acoust Soc Am **119**(1): 491-506.

Galvin J J I, Fu Q J, et al. (2007). "Melodic Contour Identification by Cochlear Implant Listeners." Ear and Hearing **28**(3): 302-319.

Kasturi K S and Loizou P C (2007). Effect of filter spacing and correct tonotopic representation on melody recognition: Implications for cochlear implants. 2007 ARO

Omran S A, Büchler M., et al. (2007). Pitch Ranking of Complex Tones using a Model of the Virtual Channels in the Nucleus Freedom System. EFAS 2007, Heidelberg - Germany, German Society for Audiology (Deutsche Gesellschaft für Audiologie).