Investigating the MP3000 coding strategy for music perception

Waikong Lai, Norbert Dillier

ENT Department, University Hospital, CH-8091 Zürich, Switzerland

Keywords: MP3000 coding strategy, music perception, electrodogram

Introduction

With Nucleus speech processors, input acoustic signals are divided into up to 22 frequency bands or channels, corresponding to the number of active intracochlear electrodes, and up to *n* of these 22 channels with the most energy are then selected for encoding and presentation on the corresponding intracochlear electrodes. The number *n* is referred to as the number of selected maxima.

Due to the nature of the energy distribution of the input signals, the standard ACE coding strategy from Nucleus tends to select clusters of channels. Stimulation on adjacent channels will produce interactions and refractory effects due to forward masking will result in some degree of redundancy. The MP3000 coding strategy (Noguiera et al. 2005) attempts to eliminate some of these redundancies, resulting in less channels *m* ($<n$) being used to transmit the same amount of information.

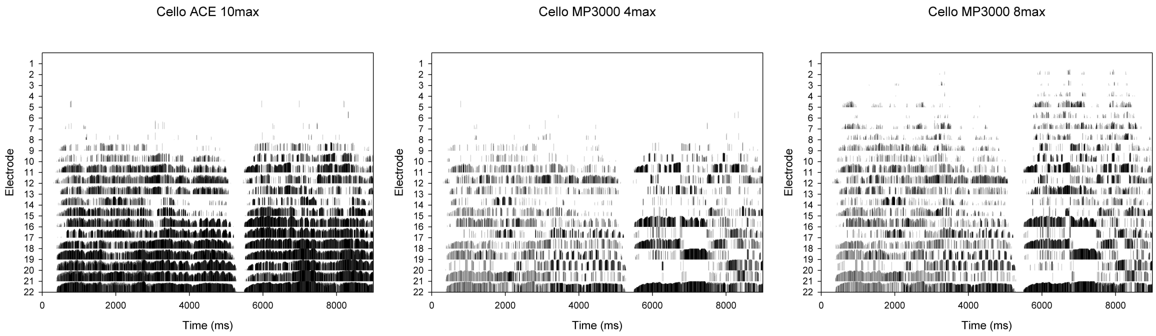

Figure 1: Electrodograms of ACE (10 maxima) and MP3000 (4 and 8 maxima respectively) outputs showing the reduction in stimulation activity between ACE 10max and MP3000 4max, as well as the additional high-frequency spectral information present with MP3000 8max.

On the other hand, the MP3000 coding strategy can be used to increase the effective amount of spectral information presented if a similar number of selected maxima is specified (see Figure 1). Arguably, the transmitted spectral information could also be achieved with the ACE strategy by increasing the number of selected maxima. However, this could be limited by the Nucleus CI24 implant's maximum overall stimulation rate of 14400 pulses per second (pps). For instance, at 1200 pps, *n* cannot exceed 12. Here, the MP3000 coding strategy would in theory be able to increase the effective amount of spectral information being presented without having to lower the stimulation rate.

This pilot study investigates whether the MP3000 coding strategy would be beneficial for music perception as compared to the corresponding ACE strategy. Instrument identification and pitch ranking tests were conducted, as well as a sound quality judgement questionnaire of various music pieces. To better understand how the musical sounds were being processed, electrodographic analysis was also carried out.

Method

2 subjects MB and GZ, both experienced adult CI users, took part in this pilot study. MP3000 maps at 1200 pps with 8 and 4 maxima were compared against the subject's standard map (ACE 1200 pps with 10 maxima in both cases) in 2 separate sessions two weeks apart. The MP3000 maps were optimised to ensure that the overall loudness was similar to that of the ACE map.

In the first session, an Instrument Identification test with 8 music instruments (trumpet, trombone, flute, clarinet, violin, cello, guitar and piano) from 4 instrument families (brass, woodwind, bowed string and plucked/struck string respectively) playing the same melody was conducted. The test material consisted of original acoustic recordings of professional musicians from the Zürcher Opernhaus Orchester. In addition, an adaptive pitch difference limens test involving clarinet tones was also carried out.

A questionnaire with an accompanying CD of 13 selected pieces of various genres (classical-orchestral, pop, jazz, vocal and solo instruments) which the subject could complete at home by comparing both ACE and MP3000 maps in the ensuing 2 weeks, was then issued. The questions all involved a 9 point quality judgement scale ranging from very poor to very good. The subjects were not specifically requested to use the MP3000 map exclusively in this period.

In the second and final session 2 weeks later, the Instrument Identification test was repeated and a pitch ranking test using clarinet tones was also carried out.

Electrodograms of the CI speech processor output with the various maps were made using a custom Python program together with the Nucleus Implant Communicator (NIC) research software.

Results

The Instrument Identification test results from both sessions were combined and summarised below.

Subject	MB	GZ
ACE 10max	70.8 %	85.4 %
MP3000 4max	66.7 %	87.5 %
MP3000 8max	52.1 %	79.2 %

All results except subject MB with the MP3000 8max condition were significantly above chance. In general, neither subject had difficulty identifying the instrument family (eg. brass instruments were never wrongly identified as plucked/struck string instruments), although GZ had some difficulty distinguishing between trombone and cello with both MP3000 maps but not the ACE map. Within a given instrument family (eg. trumpet and trombone, or violin and cello), more confusions were observed.

Neither the pitch difference limens test, nor the pitch ranking test showed any difference between all three maps. The pitch ranking test showed some pitch reversals which could possibly be accounted for after examining the electrodograms, but it was also difficult to determine if the same cues were involved in the judgement of each pair of tones.

The questionnaire results in Figure 2 showed that music enjoyment was fair to good with the cochlear implant for both subjects, although both subjects showed a preference (MB: marginally, GZ: clearly) for their standard ACE 10max map. The MP3000 maps did not confer any advantages over the subjects' more familiar ACE maps.

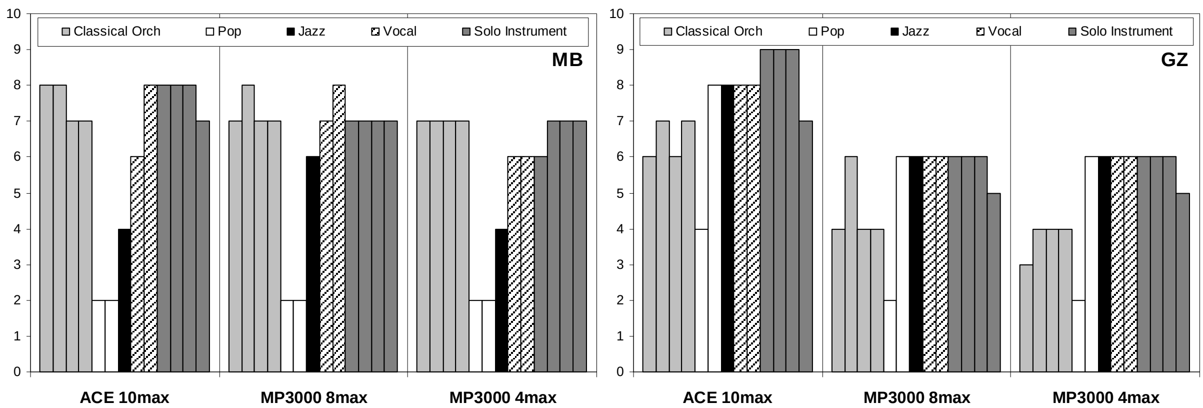

Figure 2: Questionnaire results. The 9 point response scale has a minimum of 1 and a maximum of 10.

Discussion

Instrument families differ mainly in their timbre, and there was generally little difficulty using timbre cues to identify the instruments according to their family. Within a given family, the instruments differ in their tonal or pitch ranges, and it is here that the subjects have greater difficulties. It would seem that timbre cues are more easily used than the pitch cues for identifying the instruments. Figure 3 shows electrodograms for the Trombone and the Cello with ACE 10max and MP3000 8max maps. GZ's difficulty in discriminating between these two instruments with the MP3000 maps is unlikely to be due to the spectral (pitch) content, and more likely due to similarities in the temporal content (timbre) of the signals, although this claim is difficult to substantiate with mere electrodograms.

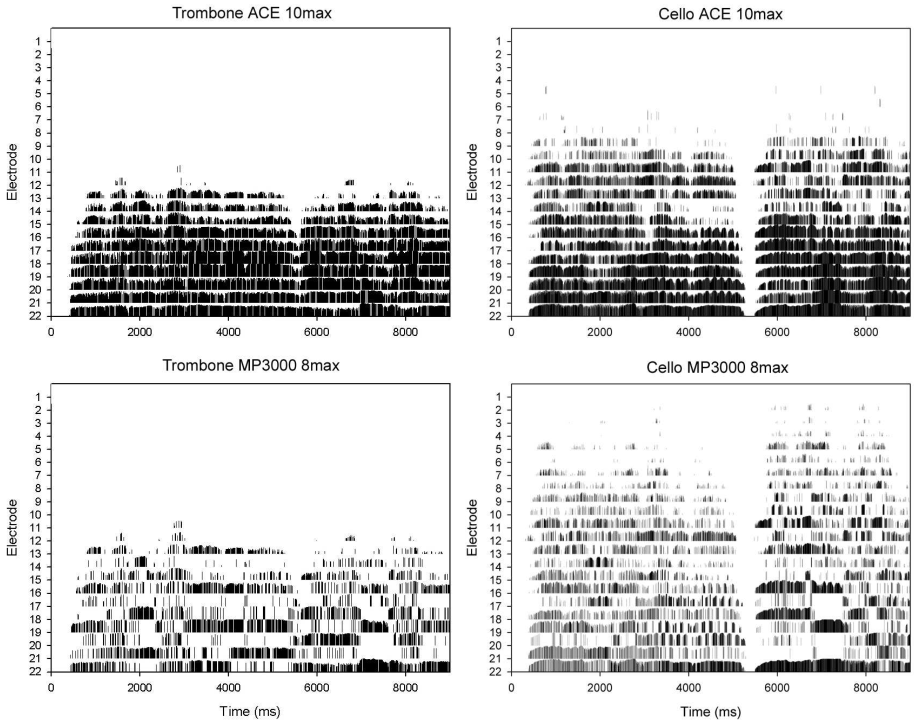

Figure 3: *GZ had difficulties distinguishing between the Trombone and the Cello with the MP3000 8max map (lower figures) despite the differences in spectral content.*

In retrospect, pitch ranking and discrimination (eg difference limens) tests were difficult to carry out using real (complex) musical tones as changing a single parameter could simultaneosly affect multiple cues in the music signal. On the other hand, oversimplified signals (e.g. pure tones) which can be better controlled are probably not representative of music sounds. As such, great care has to be exercised in the choice and design of the test material for these kind of tests. Electrodographic analysis of the clarinet tone signals was useful for checking the test signals for irregularities which could then affect the task, such as pitch reversals.

The preference of both subjects for their familiar ACE map when listening to the music pieces does not necessarily mean that the ACE coding strategy is superior to the MP3000 coding strategy. The MP3000 maps might not be properly optimized, and the subjects' familiarity with their ACE maps may have also produced a bias. Nevertheless, there was no spontaneous improvement in music quality despite the additional spectral information presented with the MP3000 8max map. One possible advantage of the ACE coding strategy is that the higher density of the encoded stimulus activity may provide more fluid percepts, and this might be important for the transmission of temporal cues in the music signals.

Conclusion

The present study did not reveal any advantage of MP3000 over ACE in terms of music perception, although familiarity with the subjects' own ACE maps may have been an important factor. Complex tones from real instruments have too many interrelated parameters to control to be well suited for adaptive testing, while oversimplified test material (e.g. pure tones) are probably not representative of real music listening. Objective electrodographic analysis of musical signals is useful for gaining insights into the nature of musical sounds presented to the cochlear implant.

References

Noguiera W, Büchner A, Lenarz T, Edler B (2005) A psychoacoustic "NofM"-type speech coding strategy for cochlear implants. EURASIP Journal of Applied Signal Processing 18, 3044-3059.

Acknowledgement

This work was supported by a research grant from Cochlear AG, Basel and research grant 320000-110043 from the Swiss National Science Foundation.