Optimizing virtual channel selection

Kalyan S. Kasturi and Philip C. Loizou

Department of Electrical Engineering, University of Texas at Dallas

kalyan@utdallas.edu, loizou@utdallas.edu, http://www.utdallas.edu/~loizou/cimplants/

Introduction

Several studies reported that a number of subjects can perceive distinct pitches between pairs of adjacent electrodes via simultaneous delivery of current to pairs of electrodes.

Given processor constraints (e.g., power consumption, stimulation rate, etc) we need to know the minimal number of virtual channels needed to improve speech understanding. We would also like to know how to best select and place the virtual channels along the electrode array.

The answers to these questions can not be easily obtained with cochlear implant listeners due to the large number of confounding factors.

Cochlear implant simulations were used to assess the effect of frequency location and number of virtual channels on sentence recognition.

Methods

Subjects: Nineteen normal hearing listeners participated in this experiment. All subjects were native speakers of American English.

Material: The test material consisted of IEEE sentences.

Twenty sentences were used per condition.

Signal Processing

Speech material was first low-pass filtered using a sixth order elliptical filter with cut-off frequency 6000 Hz. Filtered speech was passed through a pre-emphasis filter with a cut-off frequency of 2000 Hz.

This was followed by band-pass filtering into *N* different frequency bands using sixth-order Butterworth filters. The filters were designed to span the frequency range from 300 Hz to 5500 Hz in a logarithmic fashion.

Corresponding to each channel a sinusoid was generated with frequency set to the center frequency of the channel and with amplitude set to the root-mean-squared (rms) energy of the channel envelope estimated every 4 msecs and finally the channel outputs were summed.

TABLE. The 3-dB frequency boundaries of the 4 bands with the corresponding center frequencies (Hz) of each band.

Channel	Lower Frequency (Hz)	Upper Frequency (Hz)	Center Frequency (Hz)
1	300	621	460
2	621	1285	953
3	1285	2658	1971
4	2658	5500	4079

Virtual channels were implemented by adding an extra sine wave corresponding to the frequency (place) of the virtual electrode.

Conditions examined:

(a) Number of virtual channels: 1 and 2

(b) Frequency location of virtual channels:

i. Low-frequency (LF) 460-953 Hz
ii. Mid-frequency (MF) 953-1971 Hz
iii. High-frequency (HF) 1971-4079 Hz

(c) Number of virtual channels (i.e., resolution) between pairs of electrodes: 1 - 4

(d) Selection of subset of virtual channels based on local maximum (within LF, MF or HF regions) or global maximum (across all regions) of virtual channel amplitudes.

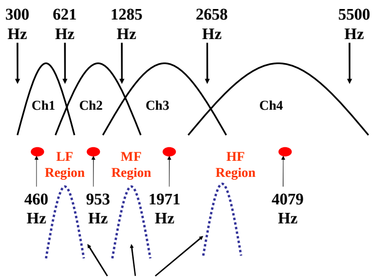


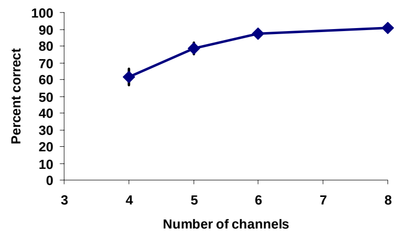
Figure 1

Experiment 1: Baseline

In this experiment, sentence recognition was assessed as a function of number of channels varying from 4 to 8.

These results served as a baseline for comparing the performance gain obtained with the use of virtual channels.

Baseline performance

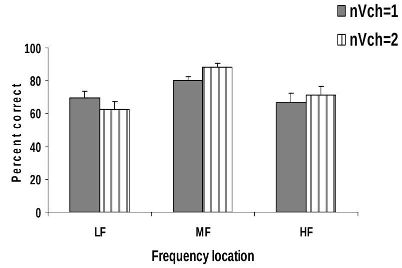


Experiment 2: Effect of frequency location of virtual channels

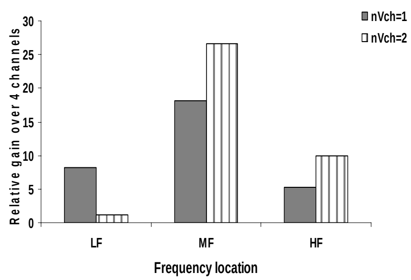
In this experiment, the effect of frequency location and number of virtual channels on sentence recognition was investigated.

In one condition a single virtual channel was added in either of three different frequency locations (LF, MF or HF) and sentence recognition was quantified in each case (see Figure 1).

In another condition, two virtual channels were added in either of the three different frequency locations and performance was quantified.



Relative intelligibility gain (% correct) over 4 channels



Analysis and Discussion

Two-way ANOVA indicated a significant effect ($p<0.0005$) of frequency location, a non-significant effect ($p=0.602$) of number of virtual channels and a significant interaction ($p=0.001$).

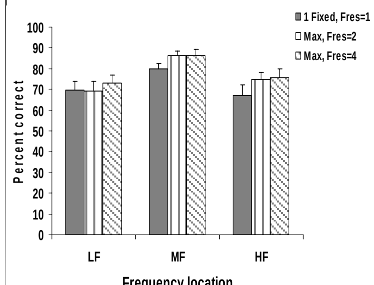
For the 2-Vchannel condition, post-hoc tests (Fisher's LSD) indicated that performance in MF region was significantly higher ($p<0.005$) than performance in LF and HF regions. Performance in the MF region with two virtual channels was better than that with one virtual channel ($p=0.007$).

The placement (frequency location) of virtual channels is extremely important for speech recognition.

Experiment 3: Effect of frequency resolution on maximum selection

In this experiment we assessed the effect of peak-picking virtual channel amplitudes within each of three frequency regions (LF, MF, HF), while varying the frequency resolution from 1-4 virtual channels.

In the condition, for instance, when the frequency resolution was four, the maximum was selected among the four virtual channels available within each frequency region. The total number of stimulating channels was five (4 + 1 local max) in all conditions and for all frequency resolutions.



Analysis and Discussion

Two-way ANOVA indicated a non-significant ($p=0.38$) effect of frequency resolution, a significant effect ($p=0.003$) of frequency location and a non-significant interaction ($p=0.453$).

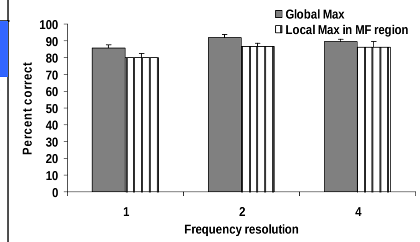
When it comes to maximum selection within each frequency region (LF/MF/HF), frequency resolution (i.e., the number of virtual channels between pairs of electrodes) is not important.

Experiment 4: Global maximum selection versus local maximum selection

In this experiment, the use of global maximum selection (across all frequency bands) was investigated while varying the frequency resolution (number of Vchannels) from 1 to 4.

In the condition, for instance, where the frequency resolution was four, the maximum was selected across all 12 Vchannel (=4x3) amplitudes in the three regions (LF, MF, HF).

The total number of stimulating channels was five (4 + 1 global max) in all conditions and for all frequency resolutions.



Analysis and Discussion

Statistical analysis indicated that when the frequency resolution is one or two (Vchannels) performance with global maximum is significantly ($p<0.005$) better than that with local maximum in the MF region.

There is no significant difference in performance between global and local max selection when the frequency resolution is four.

Better performance is obtained when using global maximum selection rather than local maximum selection of Vchannel amplitudes.

Conclusions

- Results indicated that the place of virtual channel stimulation (i.e., the frequency region) had a significant effect on word recognition. Adding virtual channels in the low or high frequency regions yielded no significant benefit on word recognition.
- Significantly higher scores were obtained when virtual channels were introduced in the mid-frequency region possibly due to a better F2 representation. In one condition, adding one additional virtual channel in the mid-frequencies produced performance equivalent to that of 8 channels.
- The present results suggest that having access to a small number (1-2) of virtual channels, particularly in the mid-frequency region, can bring significant benefits to speech understanding.
- The addition of 1-2 virtual channels if placed and selected appropriately can potentially *double* the number of effective channels of frequency information transmitted by current implant devices.