2002

School of Electrical Engineering & Telecommunications

Music for Hearing Implant

Author: Adrian Kuo Ching LEE

UNSW Supervisor: A/Prof W H Holmes

Co-Supervisor: A/Prof H McDermott, Department of Otolaryngology, The University of Melbourne

Introduction

Users of cochlear implants experience difficulties in listening to music. Improved processing of music signals for hearing implants requires an accurate and robust estimation of the musical pitch. This thesis presents a comparative study of two different signal processing strategies for estimating the fundamental frequencies (F0) of almost-periodic input signals.

Fundamentals behind cochlear prostheses

The primary motivation for initial research is to provide auditory speech communication to the profoundly hearing impaired. In the case of sensori-neural deafness, many hair cells in the cochlea may be damaged or missing. In some cases, although populations of hair cells are diminished, some cochlear neurones survive. The principle of cochlear prostheses is to generate sensations of sound by directly applying electrical stimulation to the residual auditory neurones.

Design of Signal Processing Strategies

The two F0 estimators investigated have a nominal compass of 2 octaves, which, in musical terms, extends from A2 (110Hz) to A-flat4 (415.3Hz). Their outputs are to be quantised to the 12 chroma defined by the equal tempered musical scale based on the western tonal tradition, although there is no explicit assumption that the signals are derived from musical instruments instead of from speech.

Results

The following table is a summary of the two F0 estimators' performance at SNR = 0dB for sampled signals with quasi-stationary steady-state waveforms.

INSTRUMENTS Pitch / Accuracy (%)		FLUTE G3	FLUTE C4	VIOLIN G3	VIOLIN C4	TRUMPET C3	TRUMPET G3
Sifting Algorithm	Pitch + Height	92%	94%	56%	92%	40%	98%
	Pitch only	92%	94%	96%	92%	92%	98%
Sinusoidal Coding Analysis	Pitch + Height	93%	95%	93%	88%	93%	98%
	Pitch only	93%	95%	96%	88%	98%	98%

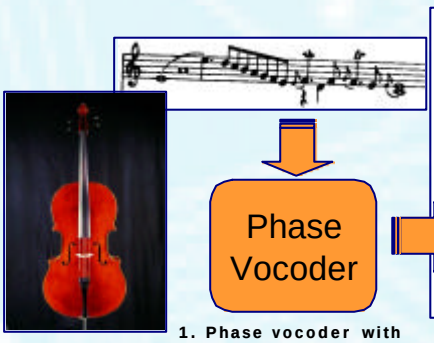
The sifting algorithm is sometimes affected by pitch halving and /or doubling. Both approaches are able to track the vibrato of the coloratura (sung vowel /æ/ at pitch F4 [349.2Hz]).

Conclusion

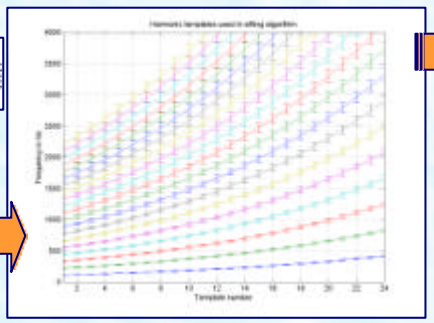
Both F0 estimators are found to be good pitch estimators. The sinusoidal coding analysis substantially reduces the possibility of pitch halving or doubling. However, with the current hearing implant technology, the sifting algorithm is favoured due to its simpler computational requirements. Nonetheless, it is hoped that the F0 estimates generated from either of the above mentioned strategies can improve not only pitch perception for the cochlear implant users, but also provide them extra cues for tonal language understanding.

Acknowledgements

I would like to thank Department of Otolaryngology, The University of Melbourne, for their financial support.



Phase Vocoder



Pitch Estimate

Pitch controls pulsatile rate

Images provided courtesy of Cochlear Ltd

1. Phase vocoder with sifting algorithm

- This strategy, proposed by The University of Melbourne, is based on frequency analysis using a phase vocoder as a spectrum analyser.
- Outputs of the phase vocoder bandpass filters are sieved according to a template based on the harmonics of each candidate F0.
- Total power collected in each template is calculated, and the results for each candidate F0 are compared to select the best estimate of the pitch.

Sinusoidal Coding Analysis

$$r(w_0) = \sum_{k=1}^{K(w_0)} \bar{A}(kw_0)\left\{ \max_{w_l \in L(kw_0)} [A_l D(w_l - kw_0) - \frac{1}{2}\bar{A}(kw_0)] \right\}$$

Pitch Estimate

Pitch controls pulsatile rate

2. Sinusoidal coding analysis

- This strategy is based on the sinusoidal coding method, proposed by McAulay and Quatieri.
- Sets of amplitude, frequency and phase estimates, $\{A_l, \omega_l, \phi_l\}$, are generated for each harmonic constituent of the signal by minimising the mean-square difference between the signal and the estimated waveform.
- The above function relies on the SEEVOC algorithm. It was intended mainly for speech transmission at low bit rates.
- Amplitude spectrum of musical tones can vary considerably from that of speech, and thus some novel modifications are made, which include:
  - § An extra criterion introduced to the SEEVOC peak-searching algorithm in order to capture pitch intervallic movements.

UNSW

ENGINEERING @ UNSW