Segmenting sounds: simpler version

LSS 31 3 2019.

This work is based on work I did a long time ago, but updated a bit since then. Original references are [Sound Segmentation Using Onsets and Offsets](http://www.cs.stir.ac.uk/~lss/Papers/JNMR1994.pdf), Smith L.S.

Journal of New Music Research, 23, 1, 11-23, March 1994.

Interestingly, re-reading my old papers (yes, one forgets what one write after more than 20 years) I see that in a later paper, for NIPS 1995 (*Onset based sound segmentation*, pp 729--735, in Touretzky D.S., Mozer M.C., Hasselmo M.E. (eds) Advances in Neural Information Processing Systems 8 (Proceedings of the 1995 Conference), MIT Press, 1996), I added a layer of integrate-and-fire neurons. I haven’t included that approach here, but see the document NotSoSimpleSegementation.doc. But I digress: the aim of this document is to explain what this simple system does.

The overall concept here (i.e. in findsegments\_1.m) is:

1. Bandpass the incoming sound signal into a number of bands, using the gammatone filter-bank. One can adjust the number of bands, the lowest and highest frequencies of the centres of each band, and the bandwidth of the bands. These are set in *findsegments\_all*, with default values, over-ridable using the varargin mechanism. See 'mincochfreq', 'maxcochfreq', 'n\_erbs'.
2. Apply the onset-offset filter to each band. This is achieved by
   1. Rectifying the signal
   2. Smoothing the rectified signal (convolution)
   3. Applying the onset-offset filter (convolution)

See lines 101-103 of findsegments\_1.m. Note that the rectified signals are summed prior to smoothing or the application of the onset/offset filter. If one was going to use onset neurons, one would not do this, but would smooth and then apply the onset offset filter to each band. This would give a set of onset/offset signals, one for each band. (I must have done this in the 1990’s but I’ve no idea where the code is).

1. Use the onset/offset signal generated to do the segmentation. Currently, this is done by
   1. Finding the onset peaks (peaks of the positive signal) and the offset peaks (peaks of -1 \* onsetoffset signal).
   2. Use these to segment the signal in a slightly ad-hoc way.
   3. The segments are returned as an n by 2 array (or as a null array if none are found).

In addition, in the directory stimuli\_2019, there are a few files with speech and noise. The system works well for narrowband noise (here, a 440Hz sine), but less well for wideband noise (here, white noise). You should try out the system with these.