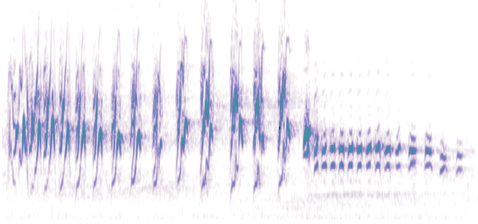
Heterodyne chirp analysis of bird vocalisations

Dan Stowell and Mark D Plumbley
Centre for Digital Music, QMUL
Queen Mary University of London
centre for digital music

The Problem

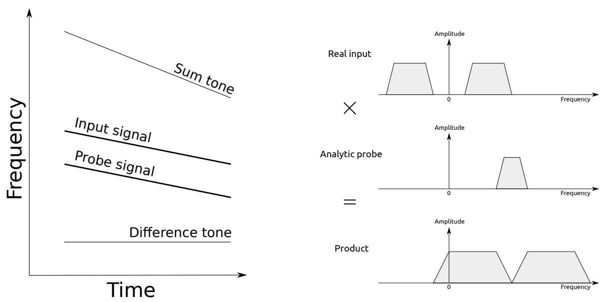
- ***Birdsong is highly nonstationary***
(fast pitch modulations)



- Spectrogram, MFCC and LPC are typical in birdsong analysis, yet all assume local stationarity. (***risk:*** loss of information)
- Nonstationary analyses exist (wavelet, chirplet, etc) but rarely used - why? Computational complexity? Downstream analysis complexity?
- We describe an ***efficient, simple*** chirplet analysis and demonstrate it can improve species classification.

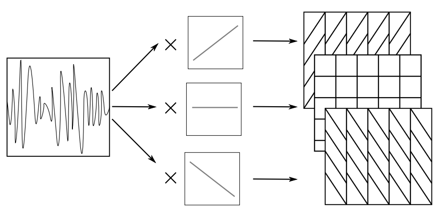
Heterodyne chirp analysis

Multiplying two chirps of the same slope produces a ***stationary difference tone***.

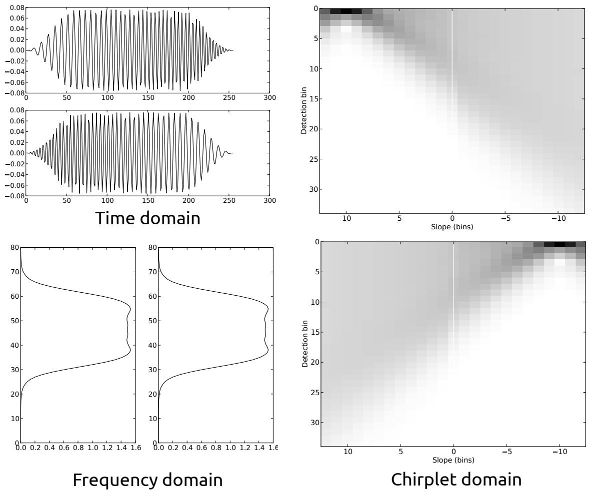


We can use this to detect chirp-like segments of audio efficiently.

For each input frame:
heterodyne (multiply) with a small dictionary of chirp atoms
(only ***one*** atom needed per slope value)
then take FFT. *Highly parallelisable.*



Example: analysing two synthetic chirps



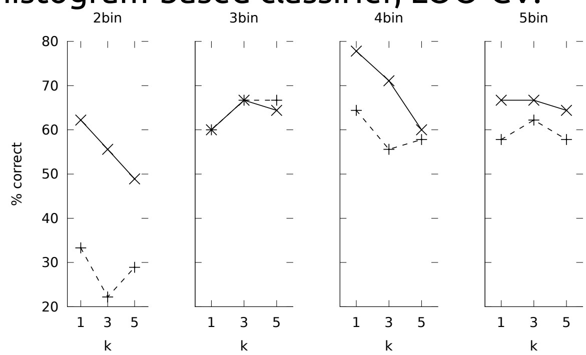
Experiment

Species classification

9 species, 45 amateur recordings (from xeno-canto), varied duration and noise.

Binomial name	Common name	Num
Chloris chloris	Greenfinch	4
Cyanistes caeruleus	Blue tit	4
Erithacus rubecula	European robin	7
Parus major	Great tit	4
Periparus ater	Coal tit	4
Phylloscopus trochilus	Willow warbler	6
Pica pica	Magpie	4
Turdus merula	Blackbird	5
Turdus viscivorus	Mistle thrush	7

Analysed with ***no segmentation***, histogram-based classifier, LOO-CV:



Chirp (solid) outperforms FFT (dashed).

Current work

- Comparison against other chirplet recovery techniques (e.g. sparse representations)
- Tracking multiple birds in audio (polyphonic recognition)

dan.stowell@eecs.qmul.ac.uk
http://tinyurl.com/dan-qmul