[x] look up papers on singing transcription

[x] pick the one that seems easiest

[x] read it

Methods:

* attention network (nishikimi)
* pitch-time curve: <https://ieeexplore.ieee.org/abstract/document/6837431>
* review article – includes evaluation measures, says mulder 2004 (what do they do?) is the best; dataset is available: <https://riuma.uma.es/xmlui/bitstream/handle/10630/8372/298_Paper.pdf?sequence=1>
* mulder auditory model:
  + <https://ieeexplore.ieee.org/document/1326812>
  + <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1326812&tag=1>
  + based on
  + <http://ismir2002.ircam.fr/proceedings/02-FP04-3.pdf>
* probabilistic model:
  + <https://www.isca-speech.org/archive_open/archive_papers/sapa_04/sap4_40.pdf>
  + review says it was bad
* pitch extraction
  + auditory model: <https://asa.scitation.org/doi/pdf/10.1121/1.402840>
  + libraries:
    - <https://aubio.org/>
    - praat with parselmouth: <https://stackoverflow.com/questions/32595404/pitch-detection-in-python>
    - neural network based: <https://pypi.org/project/crepe/>

[ ] make checklist based on paper

notes on Mulder paper (also use Martens which theirs is based on):

* the paper is a full search system, I only care about their “front end”
* Front end outline:
  + cochlear processing – not clear what they’re doing here
    - ampex –speech analysis pitch extractor
    - look for premade libraries
    - <https://hal.archives-ouvertes.fr/hal-00923967/document>
  + Pitch extraction (math??)
  + Note segmentation
    - Generate starting boundaries at minima
    - Label notes as rest, low-frequency or high-frequency
    - If the segments are similar, eliminate boundary
  + make new boundaries for legato notes (don’t understand?)

automatically transcribe when singing

* available apps are currently paid or bad
* according to assignment, can be a single mode recognizer that explores a modality not explored in the miniprojects

solved problem but could still learn from it

automatically transcribe and also select chords

transcription of some novel area?

reproduce a paper

<http://eita-nakamura.github.io/articles/Nishikimi_etal_SingingTranscriptionUsingAttention_ICASSP2019.pdf>

<http://www.ee.columbia.edu/~dpwe/papers/BrownD98-binsynth.pdf>

v1. record a sequence of pitches (detect change in pitch) with time information

v2. detect rhythm information (time signature, etc.) and create score. could have user tap for beat?

v3. additional features

* detect multiple pitches
* record multiple parts
* automatically suggest chords