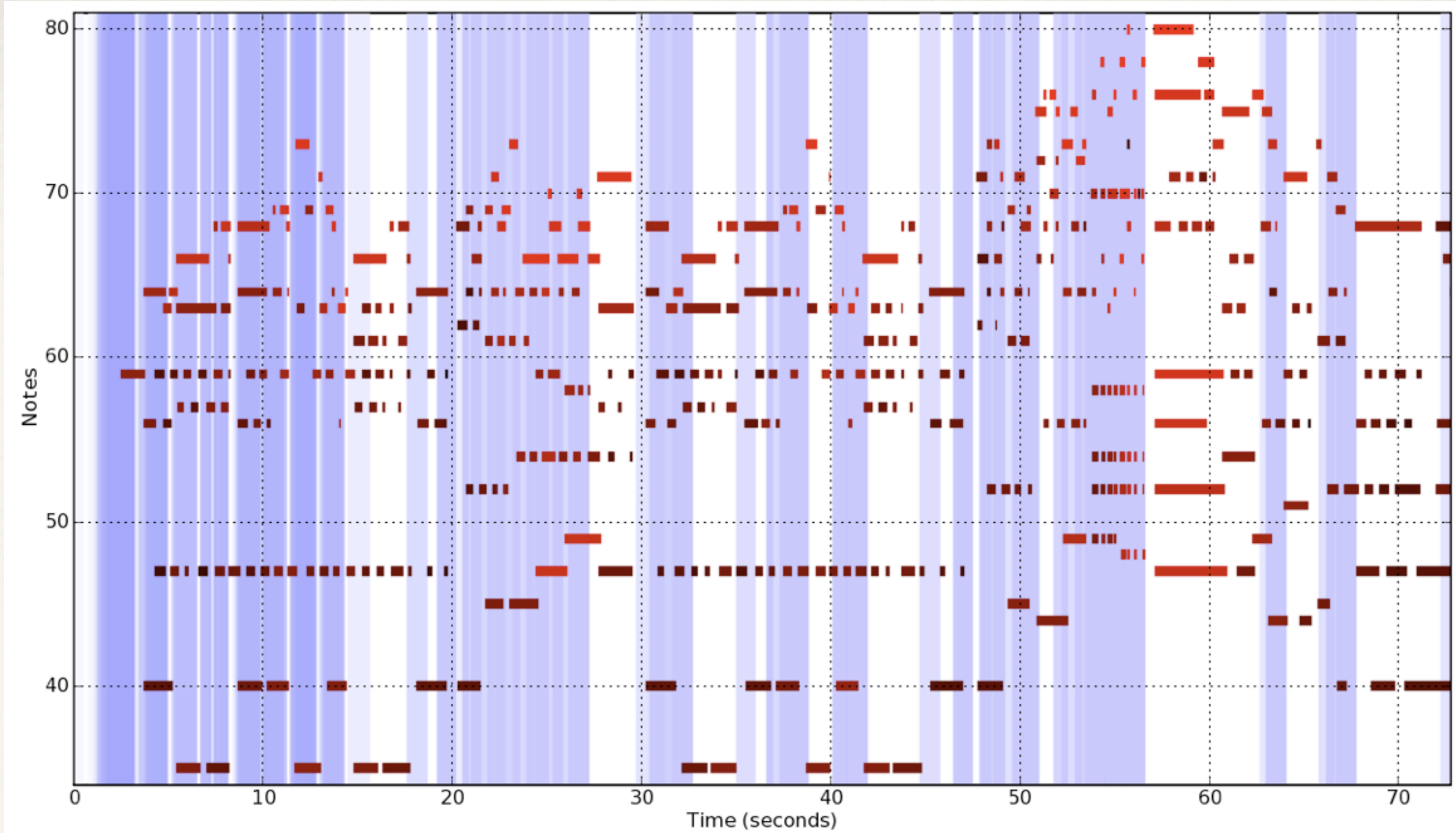


Expressive Performance

MUMT 621: Music Information Acquisition, Preservation, and Retrieval

April 2, 2009

Performance & Music Psychology

Rule-based Performance Models

Modeling Performance Data

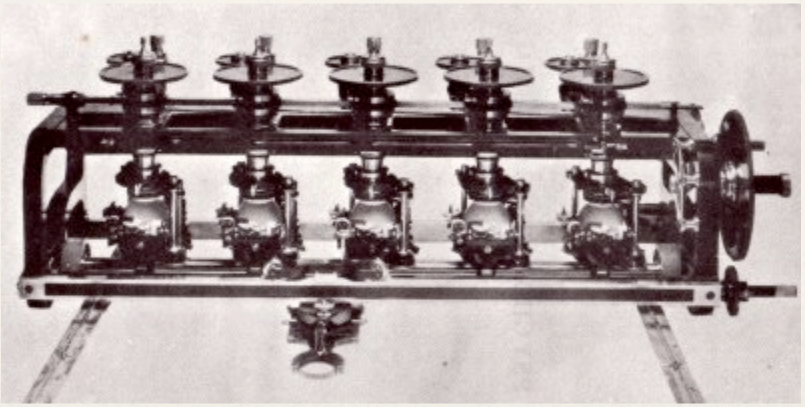
Automatically Extracting Performance Data

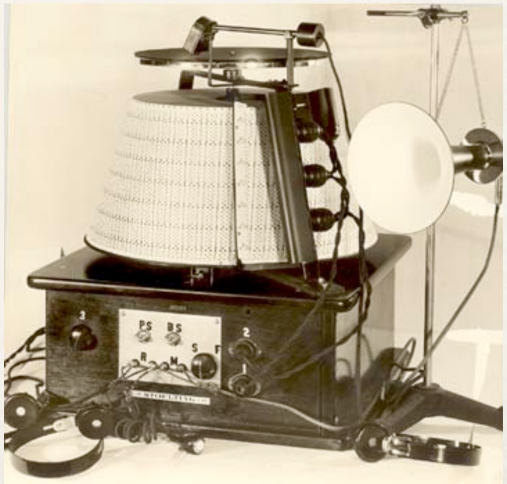
Study of Intonation Practices

Performance & Music Psychology

Performance & Music Psychology

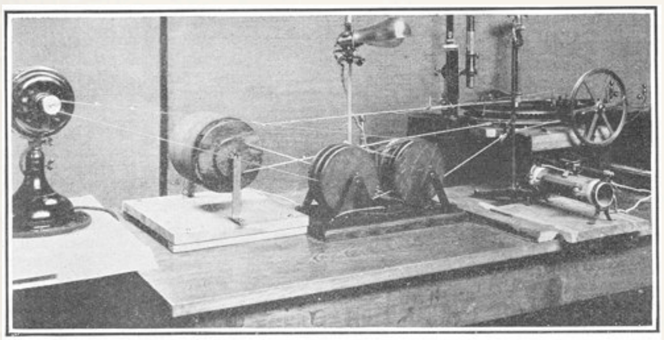
- Carl Seashore (1938) studied timing, dynamics, intonation, and vibrato in pianists, violinists, and singers
  - Equipment: piano rolls, films of the movement of hammers during performance, phono-photographic apparatus





Performance & Music Psychology

- Interest in empirical performance analysis diminished between the Second World War and 1980's, in part due to its labouriousness



Wave recorder for use with disk phonograph; the lever, acting like a pantograph, traces the waves on a revolving smoked drum

The tonoscope for analyzing the pitch of the tones on a disk phonograph record

Performance & Music Psychology

- The resurgence in interest in the late 1970s/early 1980s coincided with
  - a movement by musicologists away from equating scores with music
  - an increased interest by cognitive psychologists in music.

- Ingemar Bengtsson and Alf Gabrielsson (1980) undertook a number of systematic experiments on musical rhythm in performance

Bruno Repp

- 1992 Experiment
  - performed extensive evaluations of Beethoven's and Schumann's piano music
  - found that the degree of ritardando could be consistently related to the hierarchical levels of the phrase
  - observed that the higher the structural level the more pronounced the ritardandi were
  - demonstrated the differences in pianists' styles

Bruno Repp

- 1997 Experiments
  - Experiment 1 - Schumann's "Träumerei"
    - 10 performances by piano graduate studentsA
    - An averaged performance
  - Experiment 2 - Chopin's Etude in E major
    - 9 performances by graduate student pianists
    - 15 commercial recordings of famous pianists
    - Averaged performances of each group and a combined average of all of the recordings
    - 3 performances using the first three principal components of the expert performances

Performance & Music Psychology

- Repp (1997) Music Examples

Schubert	Chopin
Example One	Example One
Example Two	Example Two
Example Three	Example Three
	Example Four
	Example Five

Bruno Repp

- Repp (1997) Results
  - Experiment One
    - Examples: Performance 1, Average, Performance 10
    - Rankings: Performance 10, Average, Performance 3
  - Experiment Two
    - Examples: Component 2, Student 1, Expert Average, Student Average, Expert 11
    - Rankings: Expert Average, Expert 11, Student 1, Student 3, Student 9 Student 2, Student Average
- All of the audio files are available at: http://www.haskins.yale.edu/misc/REPP/AP.html

Popularity of the Piano

- Factors
  - the large amount of solo repertoire
  - the instrument's percussive nature
  - the ease with which one can acquire accurate, minimally intrusive performance measurements from a pianist via MIDI technology
  - the feasibility of using specially equipped pianos to measure performance data


Bösendorfer

Popularity of the Piano

- Issues with MIDI-based studies
  - require a MIDI-rigged piano
  - typically done in a lab environment
  - precision is limited for other instruments
- Music Information Retrieval techniques allow for extraction of performance data from recorded signals
- We'll return to this later...

Rule-based Performance Models

KTH model

- Developed at the Royal Institute of Technology in Stockholm
- Set of performance rules to predict aspects of timing, dynamics, and articulation
  - ostensibly based on the local musical context
- “analysis-by-synthesis” approach
  - pros: models one kind of performer-listener interaction
  - cons: relies heavily on the performer as only a small number of examples were presented

KTH model

- DURATION-CONTRAST Rule
  - modifies the ration between sequential notes to emphasis difference in their durations
  - quality control parameter (k)
    - 1 - full effect
    - 0 - no effect
    - -1 - reverse the effect
  - problems: several rules influence the duration of the note, which makes this rule dependent on these

KTH model: Empirical Evaluation

- To produce a predictive model the parameters of the model need to be tuned
    - analysis-by-synthesis approach presents some basic suggestions
- Empirical evaluation of recorded performances are also necessary
    - Sundberg et al. (1991) determined the perceptual threshold for the k values
    - Friberg (1995) used a greedy search method to fit parameters to the PHRASE ARCH rule based on the first nine measures a single piece
    - Sundberg et al. (2003) fitted PHRASE ARCH k values manually to a single performance of a Mozart sonata movement

KTH model: Empirical Evaluation

- Empirical evaluations outside of KTH
  - Zanon and De Poli (2003a; 200b) tested both fixed and time varying k values
  - Gabrielsson and Juslin (1996) related model to emotional colourings
  - Juslin et al. (2002) developed a comprehensive computation model of expressive performance
    - G - generative KTH model
    - E - Juslin's earlier work on emotional models
    - R - random variability
    - M - analogies to physical motion

Todd model

- Developed by Neil Todd from late 1980s to early 1990s
- Structure-level models of expressive timing and dynamics
- "analysis-by-measurement" approach
  - empirical evidence obtained directly from measurements of human performances
  - assumptions:
    - direct link between musical structure and performance
    - relationship can be modeled with a single rule
  - pros: appeals to a theoretical framework to assess musical context (Lerdahl and Jackendoff 1983)
  - cons: overly simplistic ("the faster, the louder")

Todd model

- “the faster, the louder”
- Intensity is proportional to the squared tempo
- Used a recursive look-ahead procedure to allow the hierarchical grouping in the music to control the instantaneous tempo
- Leads to increased dynamics and tempo at the middle of phrases and reduced dynamics/slowing down at points of stability, such as phrase boundaries
  - this is modeled at each level of the piece’s hierarchy

Todd model: Empirical Evaluation

- Todd (1992) compared the model's output with tempo and dynamic curves of one or two performances of a small number of pieces by Haydn
- Windsor & Clarke (1997) used regression analysis of several performances generated by Todd's model against two repeated human performances
  - residuals: idiosyncrasies of human performance not explained by the model
- Clarke & Windsor (2000) had human listeners evaluate performances generated by Todd's model

Mazzola model

- Developed by Guerino Mazzola and colleagues in Zurich
- Mathematical music theory and performance model
- Applies analysis and performance components
    - computer-aided analysis tools for musical structure
        - each aspect implemented in a Rubbette (plugin)
    - performance is generated with the Rubettes
    - uses "Stemma/Operator" theory for mapping

Mazzola model

- MetroRUBETTE
  - (inner) metrical analysis
  - result is different than Lerdahl and Jackendoff-esque (outer) metrical analysis
  - used linear mapping between metrical weight and tone intensity to generate a performance
- Empirical Evaluation
  - not evaluated against real performances

Modeling Performance Data

Widmer model

- Developed at Vienna
- Multi-level model of expressive timing and dynamics
- Uses large amounts of empirical data extracted from a performance to train a machine learning model to predict local, note-level expressive deviations and higher-level phrasing pattern
  - applies inductive machine learning and data mining techniques

Widmer: Note level model

- Inductive rule learning algorithm to learn note-level rules for timing, dynamics, and articulation
  - i.e., how the performer will play a particular note
  - complementary to higher-level manipulations
- Training method
  - recordings of 13 Mozart piano sonatas by one performer
  - each note melody described by 29 attributes
  - computer learned a set of 17 simple classification rules

Widmer: Multi-level model

- Attempts to account for the role of tempo, dynamics, and articulation in shaping abstract structures, such as motifs, groups, and phrases
- Assumptions
    - expressive timing or dynamics gestures can be reasonably approximated quadratic curves
    - a multi-level performance can be represented as a linear combination of these shapes at different hierarchical levels
    - similar phrases will be played similarly by different pianists

Widmer: Multi-level model

- Inputs to the system
    - example performances by musicians
    - hierarchical phrase analysis of the music
    - tempo and dynamics curves
- System fits quadratic approximation functions to the curves associated with each phrase
- Predicts elementary expressive shapes for similar phrases in different pieces
- Can be combined with the note-level model, such that the note-level model compensates for the "residuals"

Widmer: Empirical Evaluation

- Widmer and Tobudic (2003a) tested the predictive performance of the multi-level model by measuring how closely the tempo and dynamics curves of the new performances matched those predicted by the model
    - results were better than chance and mechanical performances
- Tobudic and Widmer (2003b) optimized the case-based learning algorithm and used first-order logic and structural similarity to model the phrases' hierarchical context
    - produced some quantitative improvements

Automatically Extracting Performance Data

Applicable MIR techniques

- Beat tracking
- Music Alignment
  - Real time score following (online)
  - MIDI-score alignment (offline)

Beat Tracking

- Works okay for percussive instruments, e.g., piano, for some tasks
    - though not accurately enough to produce reliable data
- Generally fails for non-percussive instruments
- MIREX 2007 Onset Detection Evaluation

Real-time score alignment

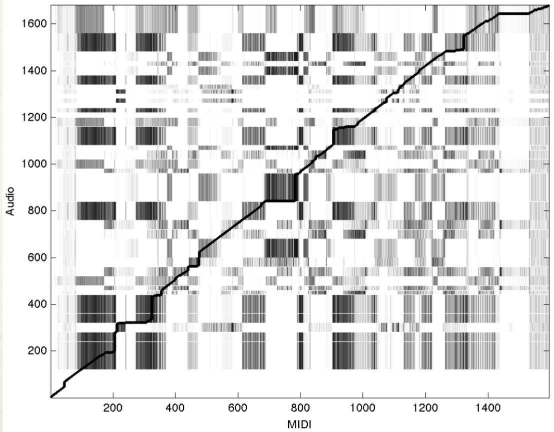
- Hidden Markov Models
  - Cano, Loscos, Bonada (1999)
  - Orio and Dechelle (2001)
  - Schwarz, Orio, and Schnell (20040
- MIREX 2008 Real-time Audio to Score Alignment Evaluation

Real-time score alignment

- Christopher Raphael's Music Plus One (2004)
  - Uses Graphical Models
  - Task I : Listen
    - Inputs: sampled acoustic signal musical score
    - Output: Time at which notes occur
  - Task 2 : Play
    - Inputs:output from Listen module musical score rehearsal data from musician performances of accompaniment
    - Output: Music accompaniment in real time

Dynamic Time Warping (Offline)

- Orio and Schwarz (2001) - peak structural distance
- Hu et al. 2003 - chromagrams, also considered pitch and MFCCs
- Turetsky and Ellis 2003 - a combination of spectral power, first order differences between channels, first order difference in frequency, and noise suppression
- Kurth et al. 2007 - chroma



Similarity matrix with the DTW path indicated in black

Study of Intonation Practices

Goals

- To reliably extract perceived pitch information from *a cappella* recordings of *a cappella* singing voice (solo and in ensembles)
- To develop a model of the observed intonation tendencies

INTONATION IN SOLO VOCAL PERFORMANCE

Original

INTONATION IN SOLO VOCAL PERFORMANCE

Original

Quantized

Potential Applications

- Relate to studies of expressive performance
- Assess whether there is an observable relationship between intonation and music theories dealing with musical attraction and tension
- Develop a predictive model for intonation practices
- Generate more accurate digital re-creations

Project Overview

- Experiment uses recordings of Schubert's 'Ave Maria' and a composed melody
- Recordings are analyzed with signal processing techniques
- Estimated perceived fundamental frequency of each note is related to its musical context

Experiment

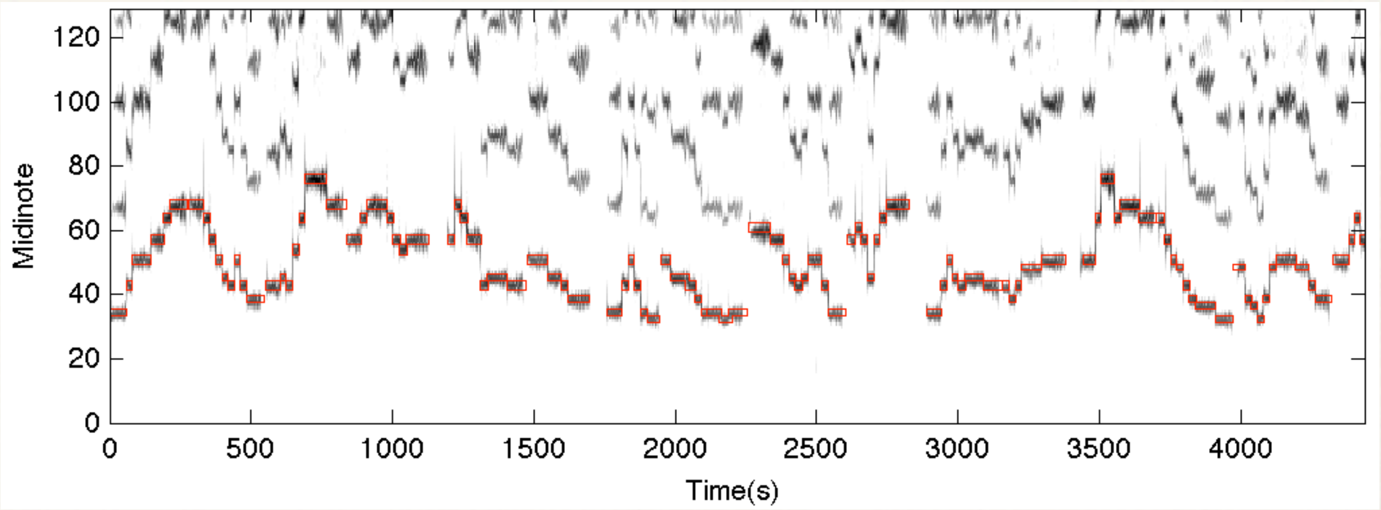
- Subjects
  - 6 professional sopranos without perfect pitch
  - 6 undergraduate sopranos with perfect pitch
  - 6 undergraduate sopranos without perfect pitch
- Experimental material
  - Schubert's 'Ave Maria' - a cappella and with accompaniment
  - A composed melodic exercise


A - ve Ma - ri - - - a, Gra - ti - a ple na Ma - ri - a gra ti - a
4
ple - - - na, Ma - ri - - - a gra - ti - a ple - na A - ve A - ve Do - mi-
6
nus, Do - mi - nus te - cum Be - ne - dic - ta tu in mu - li - e - ri - bus et
8
be - ne - di - - - - - ctus, et be - ne-dic - tus fru - ctus ven - tris, ven - tris
10
tu - i Je - - - - - sus. A - ve - Ma - ri - - - a!


10
19
28
37
46
55
65

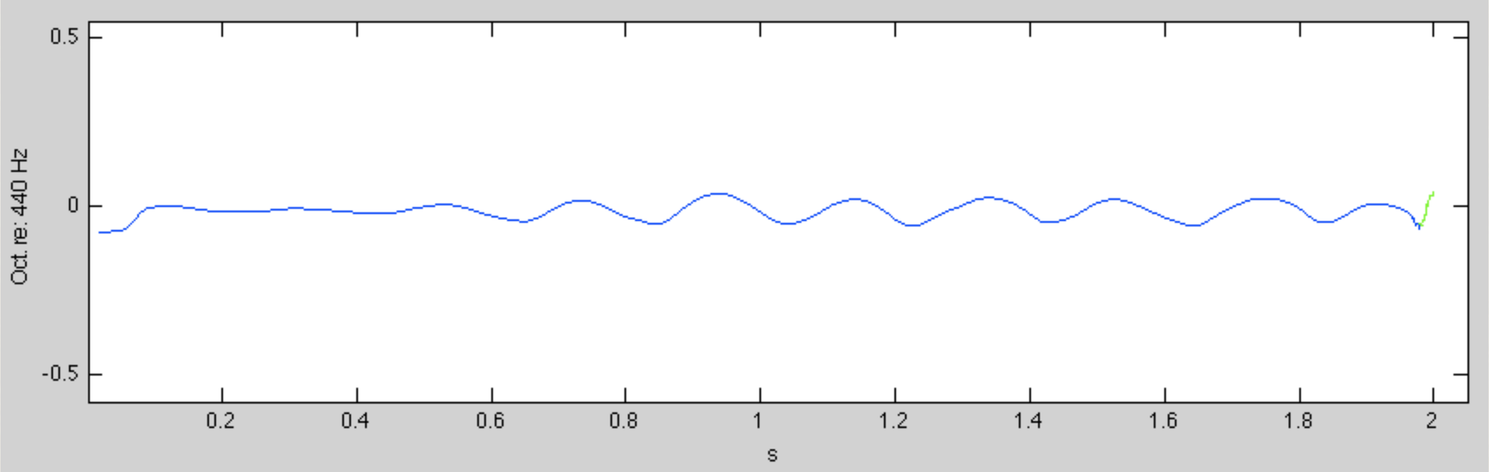
Signal Processing

- Note onsets and offsets are determined by first aligning a MIDI version of the score to the audio using Dynamic Time Warping



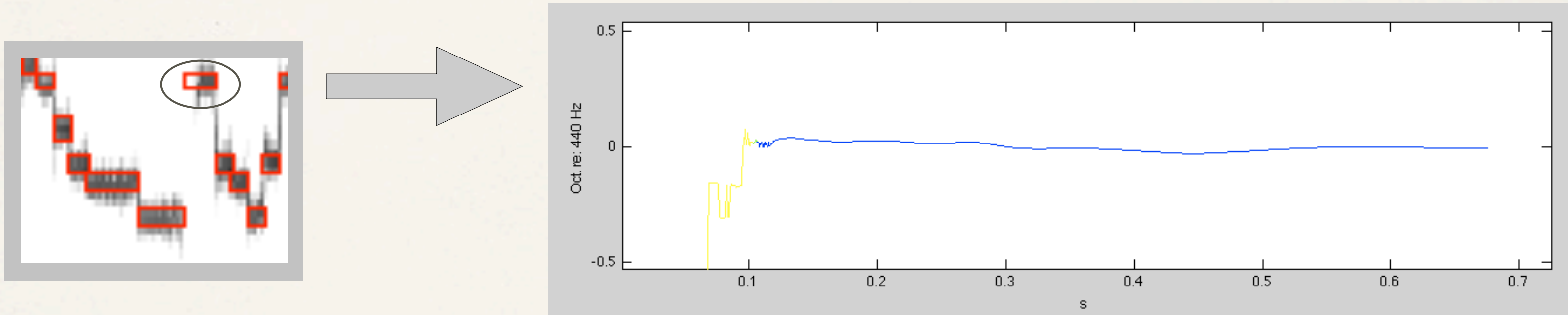
Signal Processing

- Once the onsets and offsets have been determined, fundamental frequency estimation is done with Alain deChevegine's YIN algorithm (2002)
- The perceived pitch over the duration of the sustain portion of the note was calculated as the geometric mean

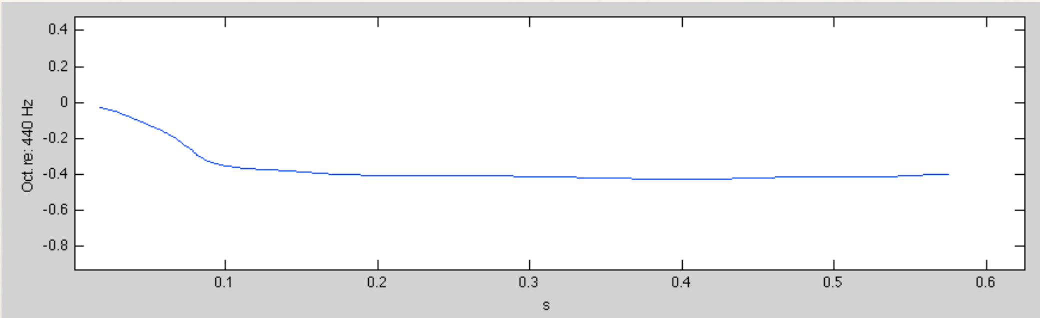


Challenges

Problems with the alignment algorithm



Singers 'sliding' into notes



a cappella vs accompanied

Self-consistency

	a cappella	Accompanied
Undergrad	14.9	19.0
Pro	33.74	10.2

Intra-singer consistency

a cappella	Accompanied
21.31	16.43

Leading-tone and Tonic

Leading-tone and Tonic

Leading-tone and Tonic

Leading-tone and Tonic

	Mean	Standard Deviation
Leading tone to tonic (72 instances)	90	15.1
Tonic to leading tone (84 instances)	-89.5	15.5
Eb to D in different keys (24 instances)	-84.3	10
D to Eb in different keys (6 instances)	99.1	6.4

Polyphonic Ensembles

- the seventh of the home key functions differently in a piece that modulates
- B is the leading tone of C Major and the mediant of G Major
  - in C Major it is an unstable pitch that generally resolves to the tonic (C)
  - in G major it is a stable pitch
- even within a single chord there are potential tuning conflicts
- in the context of a G Major chord in C major, B is both the leading-tone of the key and the third of the chord
  - it is commonly held that leading-tones are tuned sharp
  - theories of sensory consonance suggest that a vertical major third will be tuned flat

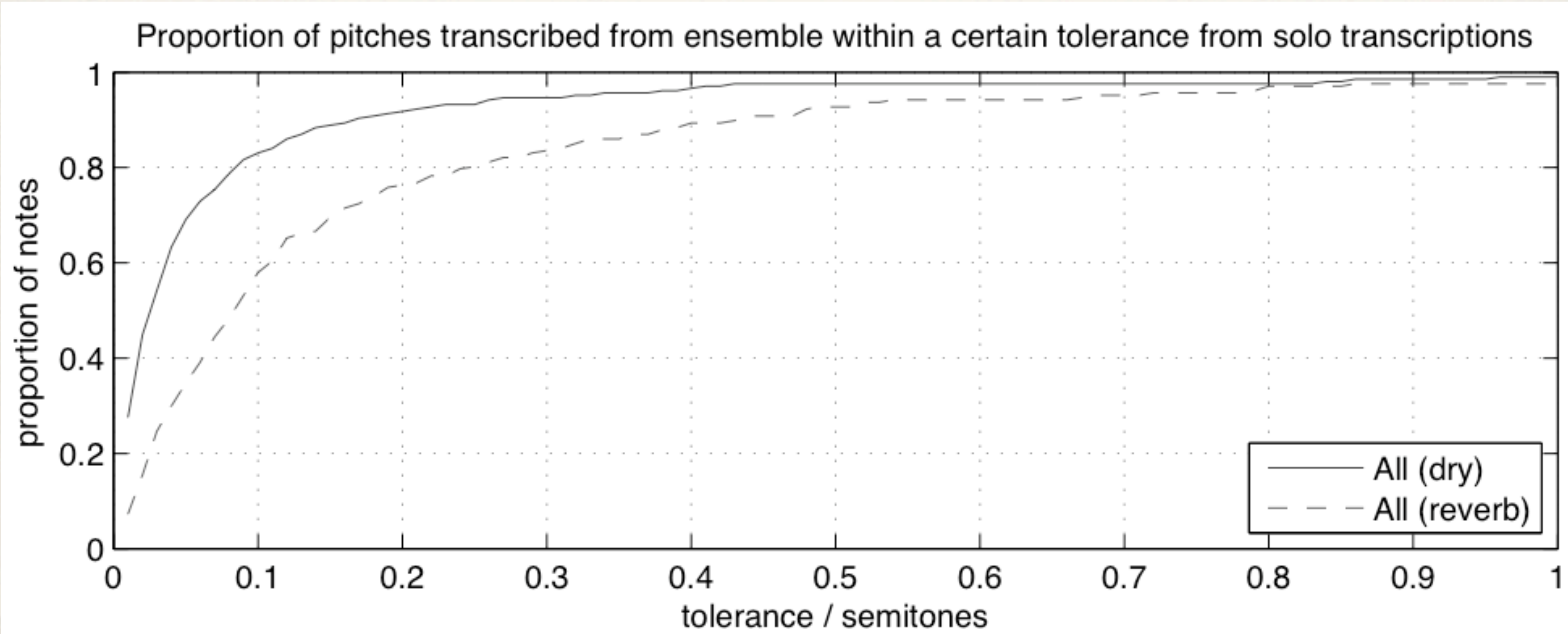
Extracting Pitch Data for Ensembles

- Robust polyphonic transcription is still an unsolved problem
- However, there is a workaround when a score is available:
  - align the MIDI score to audio
  - use the MIDI score to guide the signal processing analysis to estimate accurate frequency information

Extracting Pitch Data for Ensembles

- Can use an instantaneous frequency estimation technique to create a special spectrogram with more exact frequency information than a standard spectrogram

- Aligned MIDI file indicates the frequency-range and time-span for each expected note



Current accuracy rate of pitch estimation

Future Work

- Improve alignment and pitch estimation accuracy
- Once accurate pitch estimates have been obtained for a number of recordings have been the collected data can be modelled
    - Short-term goal is to find if any generalities exist
    - Longer-term goal is to develop a theory of vocal intonation practices

References

Bengtsson, I., and A. Gabrielsson. 1980. Methods for analyzing performance of Musical rhythm. *Scandinavian Journal of Psychology.* 21. 257-68.

Clarke, E.F., & Windsor, W.L. (2000). Real and simulated expression: A listening study. Music Perception, 17, 277–313.

Friberg, A. (1995). Matching the rule parameters of Phrase Arch to performances of “Träumerei:” A preliminary study. In: A. Friberg, & J. Sundberg (Eds.), Proceedings of the KTH Symposion on Grammars for Music Performance (pp. 37–44). Stockholm, Sweden: Department of Speech Communication and Music Acoustics.

Gabrielsson, A., & Juslin, P.N. (1996). Emotional expression in music performance: Between the performer's intention and the listeners experience. Psychology of Music, 24, 68–91.

Juslin, P.N. (2003). Studies of music performance: A theoretical analysis of empirical findings. In: R. Bresin (Ed.), Proceedings of the Stockholm Music Acoustics Conference (SMAC'03),August 6–9, 2003(Vol. II, pp. 513–516). Stockholm, Sweden: Department of Speech, Music, and Hearing, Royal Institute of Technology.

Repp, B. H. 1992. Diversity and commonality in music performance: an analysis of timing microstructure in Schumann's 'Träumerei'. *Journal of the Acoustical Society of America.* 92: 2546–68.

Seashore, C. E. 1938. *Psychology of music.* New York: Dover Publications.

Repp, B.H. (1992). Diversity and commonality in music performance: an analysis of timing microstructure in Schumann's “Träumerei”. Journal of the Acoustical Society of America, 92, 2546–2568.

Stamatatos, E., & Widmer, G. (2002). Music performer recognition using an ensemble of simple classifiers. In: F.V. Harmelen (Ed.), Proceedings of the 15th European Conference on Artificial Intelligence (ECAI'2002), July 21–26, 2003, Lyon(pp. 335–339). Amsterdam: IOS Press.

Sundberg, J., Friberg, A., & Bresin, R. (2003). Attempts to reproduce a pianist's expressive timing with Director Musices performance rules. Journal of New Music Research, 32, 317–325.

References

Sundberg, J., Friberg, A., & Frydén, L. (1991b). Threshold and preference quantities of rules for music performance. Music Perception, 9, 71–92.

Todd, N.P.M. (1992). The dynamics of dynamics: A model of musical expression. Journal of the Acoustical Society of America, 91, 3540–3550.

Widmer, G. (2002). Machine discoveries: A few simple, robust local expression principles. Journal of New Music Research, 31, 37–50.

Widmer, G. (2003). Discovering simple rules in complex data: A meat-learning algorithm and some surprising musical discoveries. Artificial Intelligence, 146, 129–148.

Widmer, G., Dixon, S.E., Goebl, W., Pampalk, E., & Tobudic, A. (2003). In: search of the Horowitz factor. AI Magazine, 24, 111–130.

Widmer, G. & W. Goebl. (2004). Computational models of expressive performance: The state of the art. Journal of New Music Research 33, 203-216.

Widmer, G., & Tobudic, A. (2003). Playing Mozart by analogy: Learning multi-level timing and dynamics strategies. Journal of New Music Research, 32, 259–268.

Windsor, W.L., & Clarke, E.F. (1997). Expressive timing and dynamics in real and artificial musical performances: Using an algorithm as an analytical tool. Music Perception, 15, 127–152.

Zanon, P., & De Poli, G. (2003a). Estimation of parameters in rule systems for expressive rendering in musical performance. Computer Music Journal, 27, 29–46.

Zanon, P., & De Poli, G. (2003b). Time-varying estimation of parameters in rule systems for music performance. Journal of New Music Research, 32, 295–315.

Zanon, P., & Widmer, G. (2003b). Recognition of famous pianists using machine learning algorithms: First experimental results. In: Proceedings of the 14th Colloquium on Musical Informatics (CIM'2003) 84–89. Florence, Italy.