CHARACTERIZING SINGING VOICE FUNDAMENTAL FREQUENCY TRAJECTORIES

Johanna Devaney, Michael I. Mandel, and Ichiro Fujinaga

CIRMMT Centre for Interdisciplinary Research in Music Media and Technology — McGill — Schulich School of Music / École de musique Schulich

SUMMARY

- This paper uses the 1st and 2nd Discrete Cosine Transform (DCT) coefficients as approximations of slope and curvature in singing voice fundamental frequency (F_0) trajectories
- These characterizations are useful for describing similarities in the evolution of the fundamental frequency in different notes
- Such descriptors can be applied in the areas of performance analysis and singing synthesis

CALCULATING SLOPE AND CURVATURE

- OPTION 1: Calculate slope as rise over run of endpoints
  - *Problem*: highly sensitive to noise in endpoints
- OPTION 2: Derive slope and curvature from a second-order polynomial
  - *Problem*: all terms dependent on polynomial order
- OPTION 3: Approximate slope and curvature from DCT coefficients
  - *Benefit*: orthogonal basis, coefficients do not depend on truncation and provides polarity for each coefficient

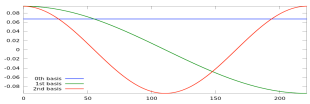
DISCRETE COSINE TRANSFORM

- We are using the type-II DCT for this research

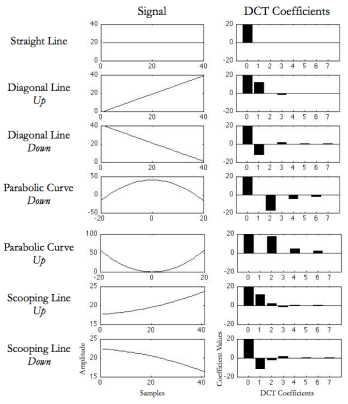
$$y(k) = w(k)\sum_{n=0}^{N-1} x(n)\cos\frac{k(2n+1)p}{2N} \qquad k = 0,1,2$$

$$\text{where } w(k) = \begin{cases} \frac{1}{\sqrt{N}} & k = 0 \\ \sqrt{\frac{2}{N}} & 1 \le k \le 2 \end{cases}$$

- The figure to the right is a visualization of the 0th, 1st, and 2nd DCT basis functions



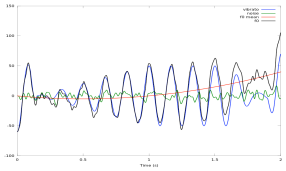
- The figure to the right shows seven examples of DCT coefficients for simple signals.
  - The plots on the left are the original signals: straight line, diagonal line, parabolic curve, and scooping line
  - The bar graphs on the right are the values for the raw DCT coefficients 0–7



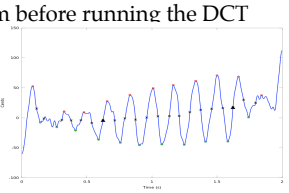
- To compute slope and curvature independent of signal length, the DCT coefficients are scaled
  - $y(0)/N^{1/2}$ is proportional to the mean of signal, in cents
  - $y(1)/N^{3/2}$ is proportional to the slope, in cents/second
  - $y(2)/N^{5/2}$ is proportional to the curvature, in cents/second^2

EVALUATION

- The DCT was evaluated on 945 test signals (7 x 3 x 3 x 3 x 5)
  - 7 types of slope and curvature
  - 3 types of note transition effects
  - 3 combinations of vibrato rate and depth
  - 3 different note lengths (1000, 2000, and 3000 ms)
  - 5 random draws of low-pass noise
- In the figure to the right, the black line shows the F_0 trajectory composed of the sum of the mean trajectory (red), vibrato (light blue), and low-pass noise (green)



- Ran vibrato elimination algorithm before running the DCT
- In the figure to the right, the red and green circles show the estimated peaks and troughs, the x's show the midpoints, and the triangles show the start and end of the vibrato section.



RESULTS

- 3 different types of note transitions, 315 test signals in each
- Tables show the absolute difference from the ground truth

Slope	Ground Truth	Type 1 (315)		Type 2 (315)		Type 3 (315)	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
Straight line	0	0.8	*0.9*	0.8	*1.1*	0.7	*0.9*
Diagonal line, up	10	0.9	*1.5*	0.8	*0.9*	0.9	*1.1*
Diagonal line, down	-10	0.6	*0.8*	0.8	*0.9*	0.8	*1.0*
Parabolic shape, up	-1.0	0.7	*0.8*	0.8	*0.9*	0.9	*1.3*
Parabolic shape, down	1.1	0.8	*1.1*	0.6	*0.7*	0.7	*1.0*
Scooping line, up	7.6	0.8	*1.2*	0.8	*0.9*	0.7	*0.9*
Scooping line, down	-7.5	0.9	*1.3*	0.9	*1.0*	0.8	*0.9*

Curvature	Ground Truth	Type 1 (315)		Type 2 (315)		Type 3 (315)	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
Straight line	0	1.1	*1.5*	1.6	*3.3*	1.2	*2.1*
Diagonal line, up	0	1.2	*2.0*	1.8	*3.4*	1.6	*3.2*
Diagonal line, down	0	1.0	*1.4*	1.1	*1.8*	1.2	*1.8*
Parabolic shape, up	-2.6	1.3	*1.9*	1.2	*2.4*	1.6	*2.8*
Parabolic shape, down	2.6	1.7	*3.3*	0.9	*1.3*	0.9	*1.3*
Scooping line, up	1.3	1.6	*2.5*	1.2	*1.7*	1.0	*1.8*
Scooping line, down	-1.3	1.5	*3.8*	1.1	*2.1*	1.6	*2.5*

CONCLUSIONS

- DCT-based estimates of slope and curvature are robust to vibrato, jitter, and transition effects for a range of note shapes
- Slope estimates are more robust than curvature
- Overall running DCT on F_0 trajectories provides a good summary of pitch-related slope and curvature characteristics

Fonds de recherche Société et culture Québec — Social Sciences and Humanities Research Council of Canada / Conseil de recherches en sciences humaines du Canada — Canada