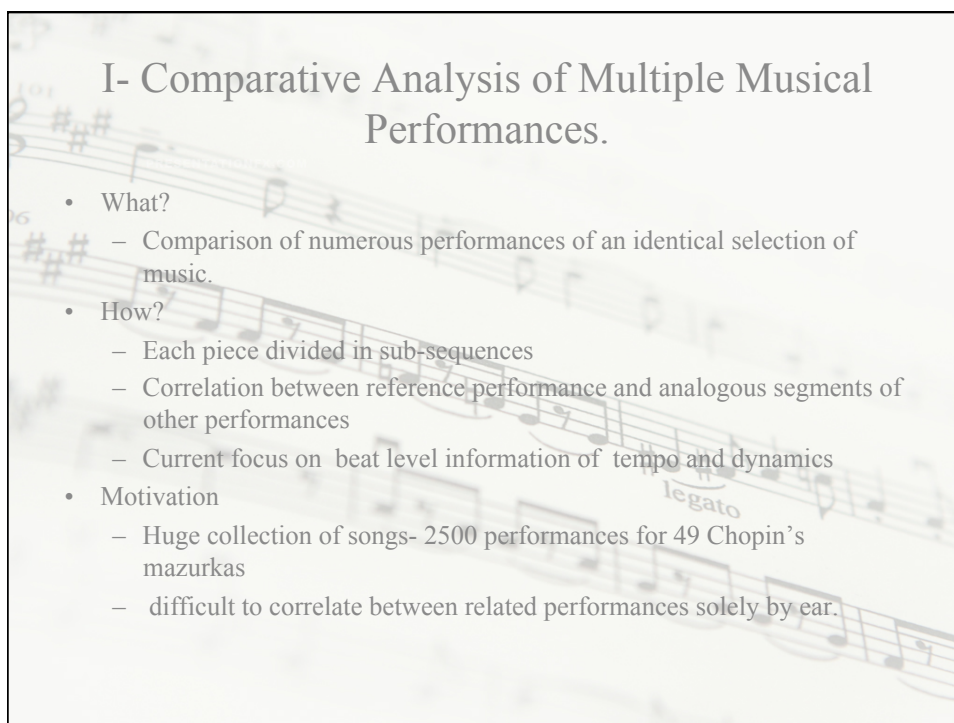


1. COMPARATIVE ANALYSIS OF MULTIPLE MUSICAL PERFORMANCES

2. HYBRID NUMERIC/RANK SIMILARITY METRICS FOR MUSICAL PERFORMANCE ANALYSIS.

Craig Stuart Sapp
CHARM, Royal Holloway, University of London

Presented by
Ravi Nitin Balajee
USC EE 675 / ISE 575
Feb 23, 2010



I- Comparative Analysis of Multiple Musical Performances.

- What?
 - Comparison of numerous performances of an identical selection of music.
- How?
 - Each piece divided in sub-sequences
 - Correlation between reference performance and analogous segments of other performances
 - Current focus on beat level information of tempo and dynamics
- Motivation
 - Huge collection of songs- 2500 performances for 49 Chopin's mazurkas
 - difficult to correlate between related performances solely by ear.

Mazurka (in Polish, *mazurek*)

- stylized Polish Folk Dance in triple meter,
- lively tempo that has a heavy accent on the second beat.
- Several compositions available, but best from Chopin's solo piano pieces
 - Great deal of repetition
 - groups of measures or entire sections or of an initial theme.

Source: en.wikipedia.org
http://www.usc.edu/dept/polish_music/dance/mazur.html

Raw Data

- Tempo and loudness
 - Our old friend- Sonic Visualizer with audio analysis plugins
- Raw power calculated for comparisons of musical dynamics
 - Power calculated for the audio signal every 10ms is sampled at each beat location.
 - Smoothed with exponential smoothing filter
 - Applied in forward and reverse directions to center smoothed data around original time location.
- Loudness level information
 - From smoothed power value about 70ms after the onset is used.
- Pearson Correlation applied

Scape plot-an example

The slide displays two triangular matrices. The left matrix is a 6x6 grid of letters: Row 1: ABCDEF; Row 2: ABCD BCDEF; Row 3: ABC BCD CDEF; Row 4: AB BC CD DE EF; Row 5: A B C D E F; Row 6: A B C D E F. The right matrix is a 6x6 grid of numbers: Row 1: 5,3; Row 2: 5,6 5; Row 3: 5 5,3 4,8; Row 4: 5 4,3 5 5,7; Row 5: 6,5 4 3,5 6,5 6; Row 6: 7 6 2 5 8 4.

- Human-readable format
- Six beats A-F
- Next level-grouped by sequential pairs
- Averaging in each cell
- Any operation can be performed

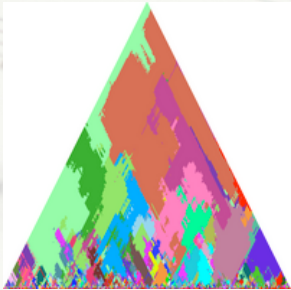
PRESENTATIONEX.COM

Steps

- Choose one performer to be the reference for a particular plot.
- For each cell in the scape plot, measure the correlation between the reference performance and all other performances which yields the highest correlation value.
- Color the cell with a unique hue assigned to that highest correlating performance.

PRESENTATIONEX.COM

Bacha, Abdel Rahman El Mazurka, Op. 24, No. 2 Full tempo curve




25.4%	Falvay 1989
13.8%	Nezu 2005
8.7%	Malcuzyński 1951
6.5%	Malcuzyński 1961
5.9%	Hatto 1997
5.9%	Fliere 1977
5.2%	Indjic 2001
5.0%	BenOr 1989
2.6%	Perlemuter 1992
2.5%	Osinska 1989
1.3%	Lushtak 2004
1.1%	Fiorentino 1962
1.1%	Magin 1975
1.0%	Wasowski 1980

<http://mazurka.org.uk/ana/pcor-perf/>
Interesting Rubinstein's performance:
<http://mazurka.org.uk/ana/pcor-perf/rubinstein/index-noavg.html>

Dynascapes

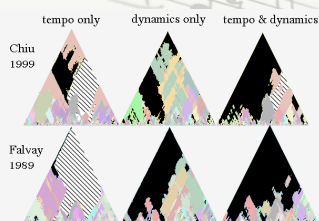
- Beat level amplitude measurements
- Less unique to a single individual performer- loudness defined in compositions
- Greater likelihood of random patterns.
- Example: Uninsky's performance in C# minor, 63/3



- Recordings 40 years apart and closest to his own dynamic interpretation than to any of the other performers.
- Difficult to ascertain if the match is relevant due to the limited range of dynamics between the performances.

Two Papers by Craig Sapp

- Both time and dynamics processed in the same plot
 - Likely to show a direct relation between performance rather than a random occurrence
- Time series of each feature is interleaved
- Tempo $t = (t_1, t_2, t_3, \dots, t_n)$; Dynamics $d = (d_1, d_2, d_3, \dots, d_n)$;
- Joint feature sequence: $J = (J_{t,1}, J_{d,1}, J_{t,2}, \dots, J_{d,n})$;
- To mix unrelated data, the Standard deviation and mean should be equal .



- Using concepts such as correlation at multiple timescales, performance “noise floor” leads to measurements that are more refined than correlation alone.
- Hybrid numeric/rank similarity metrics for musical performance analysis
- Comparing the measurements with plain correlation values helps to identify the same Chopin mazurka played by same artist out of a collection of various other pianists.

Beat level tempo

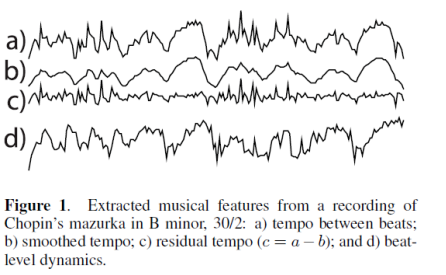


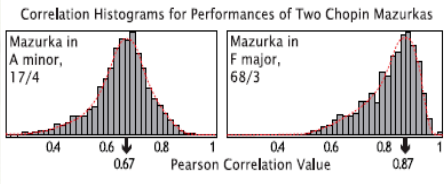
Figure 1. Extracted musical features from a recording of Chopin's mазurka in B minor, 30/2: a) tempo between beats; b) smoothed tempo; c) residual tempo ($c = a - b$); and d) beat-level dynamics.

Dynamics measurements:
Starting from comparison method of correlation, series of intermediate similarity measurements (S1-S3) are used to derive a final measurement technique(S4).

PRESENTATIONEX.COM

Definitions

- **S₀- Type -0 score-** Top level correlation between dynamics curves:
- Pearson Coefficient or 'r' value calculated.
- Value range: -1.0 to 1.0→identical match
 - 0 indicates no predictable linear relation between the sequences.
- Example:



Correlation Histograms for Performances of Two Chopin Mazurkas

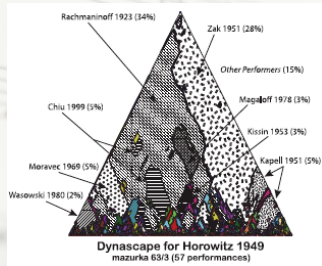
Mazurka	Mode
Mazurka in A minor, 17/4	0.67
Mazurka in F major, 68/3	0.87

Complex composition
Mode=0.67

Simpler composition
with fewer options.
Mode=0.87

• **S₁- Type-1 score:**

- S₀ causes variability on correlation distributions.
- Hence scape-plots developed.
 - Nearest-neighbor performances in terms of correlation at all timescales for a particular reference performance
- Fraction of area each target performance covers in a reference performer's scape-plot.



Target	S ₀	R ₀	S ₁
Rac23	0.60	3	0.34
Zak37	0.64	1	0.28
Mor69	0.59	4	0.05
Chi99	0.49	20	0.05
Kap51	0.51	17	0.05
Mag78	0.62	2	0.03
Kis93	0.52	15	0.03
Was80	0.58	5	0.02

(reference plot)

- Vladimir Horowitz performance.
- S₀(Zak) > S₀(Rac) → Zak-best correlation for entire feature sequence.
- Zak's music (though 28%) matches well towards end of music.

• **S₂- Type-2 score:**

- Scape plots sensitive to *Hatto Effect*:
 - When performance is completely identical to the reference performance, the scape plot likely to have a solid-color.
 - S₁ score of all other performances will be approximately 0 regardless of how similar they are to the reference performance.
 - Therefore, S₁ scores not useful in viewing similarities to other performances which are hidden behind close neighboring performances.
- One solution: remove the best match from scape plot to view/calculate the next best match.
- S₂ calculations.
 - Q- reference performance
 - Closer performances have lesser shadows on reference
 - In this case, remove 'b', 'a' becomes visible
 - Keep continuing till S₂ ranks decrease towards bottom match

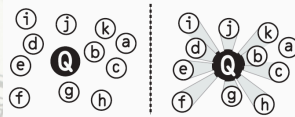
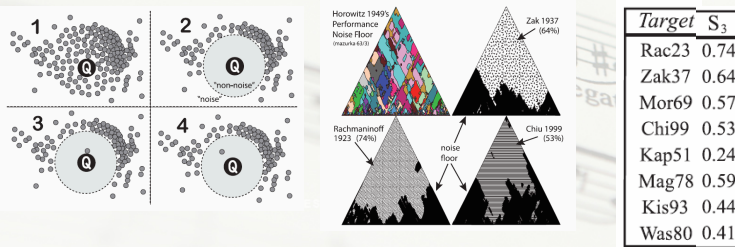


Figure 4. Schematic of nearest-neighbor matching method used in comparative timescales.

Target	S ₂
Rac23	0.34
Zak37	0.34
Mor69	0.13
Chi99	0.20
Kap51	0.08
Mag78	0.27
Kis93	0.09
Was80	0.11

• **S₃- Type-3 score:**

- Lesser *Hatto effect*:
 - Continuing S₂ calculations, problem occurs with Chiu's region overlapping with Moravec's performance.
- Difficult to determine which performance matches best according to S₂ scores. Hence S₃ scores employed.
- All performances ranked approximately by similarity to reference using S₀ or S₂ calculations.
 - 2 groups- Better matched performance placed over poorly matched performance (performance "noise-floor")
- All non-noise performances are removed from search database. (No.2)
- Non-noise performances re-introduced separately along with noise-floor performances and scape-plot is generated. (No.3,4)
- The coverage area is the S₃ similarity measurement with respect to query performance.



• **S₄- Type-4 score:**

- Refinement of S₃ scores.
- Unlike correlation values, S₃ scores are not symmetric.
 - (ie) score from A to B is not same as B to A.
- Geometric mean is used to mix S₃ score with reverse-query score (S_{3r})
 - When the 2 scores are not similar, the final score is penalized.

$$S_3 = A \rightarrow B$$

$$S_{3r} = B \rightarrow A$$

$$S_4 = (S_3, S_{3r})^{1/2}$$

Target	S ₀	R ₀	S ₁	S ₂	S ₃	S _{3r}	S ₄	R ₄
Rac23	0.60	3	0.34	0.34	0.74	0.82	0.78	1
Zak37	0.64	1	0.28	0.34	0.64	0.60	0.62	2
Mor69	0.59	4	0.05	0.13	0.57	0.54	0.55	3
Chi99	0.49	20	0.05	0.20	0.53	0.54	0.53	4
Kap51	0.51	17	0.05	0.08	0.24	0.17	0.20	22
Mag78	0.62	2	0.03	0.27	0.59	0.37	0.47	7
Kis93	0.52	15	0.03	0.09	0.44	0.23	0.32	11
Was80	0.58	5	0.02	0.11	0.41	0.55	0.47	6

- R₀ - similarity ranking by sorting S₀
- R₄ - similarity ranking by sorting S₄

Evaluation

- Usefulness of this technique
 - Identify the recordings of same performer among the collection
 - Presumably, pianists tend to play similar to their performance over time rather than performances of other artists.
- Rubinstein performance matching
 - 3 recordings of the entire mazurka cycle
- Other performers

PRESENTATIONEX.COM

Rubinstein Performance Matching

Query	Target	Tempo			T _a			Dynamics			TD					
		R ₁	R ₂	R ₃	R ₁	R ₂	R ₃	R ₁	R ₂	R ₃	R ₁	R ₂	R ₃			
Rub39	Rub52	2	3	12	3	8	8	3	2	1	6	8	3	3	12	
Rub39	Rub66	1	1	1	1	1	1	2	3	2	4	4	2	1	2	
Rub52	Rub39	6	9	2	27	31	12	3	3	2	28	23	6	12	9	2
Rub52	Rub66	2	2	1	3	2	1	2	2	1	2	2	1	2	2	1
Rub66	Rub39	3	4	2	2	3	1	3	6	5	15	8	6	5	5	3
Rub66	Rub52	1	2	1	3	2	2	2	2	1	1	3	2	1	2	1
<i>Ranking</i>		R ₁	2.5		6.5		2.5		9.3		4.0					
<i>Averages</i>		R ₁	3.5		7.8		3.0		8.0		4.2					
<i>(by feature)</i>		R ₁	1.5		4.2		2.0		4.2		1.7					
<i>Overall Averages:</i>		R ₁	= 4.97			R ₂	= 5.32			R ₃	= 2.70					

Table 3. Rankings for 17/4 Rubinstein performances. Shaded numbers indicate perfect performance of a similarity metric.

- 3 reference performances
 - 1- overall similarity
 - 2- Second best match
- No perfect matching pairs in Dynamics
 - Contains measurement noise
 - Varying of dynamics over time than tempo
 - Combination of both
- More prominent interpreter of Chopin's piano music
- Tends to vary his performance interpretation more than other pianists
 - Difficult case to match with other evaluations.

PRESENTATIONEX.COM

Evaluation of other performers

Mazurka	Query	Target	T		T _s		T _d		D		TD	
			R ₀	R ₄	R ₅	R ₄	R ₀	R ₄	R ₀	R ₄	R ₅	R ₄
17/4	Cze49	Cze49b	1	1	1	1	1	1	3	1	1	1
17/4	Cze49b	Cze49	1	1	1	1	1	1	7	1	1	1
63/3	Fri23	Fri30	1	1	1	1	1	1	1	1	1	1
63/3	Fri30	Fri23	1	1	1	1	1	1	1	1	1	1
17/4	Hor71	Hor85	2	1	13	1	2	1	1	1	2	1
17/4	Hor85	Hor71	1	1	1	1	1	1	1	1	1	1
30/2	Fou78	Fou05	2	1	1	1	13	17	2	2	2	1
30/2	Fou05	Fou78	1	1	1	1	2	6	3	2	2	1
63/3	Uni32	Uni71	1	1	1	1	5	1	1	1	1	1
63/3	Uni71	Uni32	1	1	1	1	1	1	1	1	1	1

- S4 performs perfectly on identifying the performance by same pianist(except Fou)
- Case of Hor71 and Uni

PRESENTATIONEX.COM

Results

- S0 and S4 effective than other metrics in identifying similar performances
- Precise meaning of color in plot is difficult to pin down
 - Hard to determine how relevant the best matches between performances are.
 - Only a step towards identifying significant relations
- Tempo and dynamics only considered
 - Focuses attention in specific areas
 - Other important features important in characterizing a performance discarded
 - Legato and staccato-difficult to extract
 - Note timings, voicing, pedaling and articulation are not considered.
 - Attack differences between the left and right hands for pianists.

PRESENTATIONEX.COM